



HAL
open science

Apports et difficultés d'une collecte de données à l'aide de récepteurs GPS pour réaliser une enquête sur la mobilité

Thi Huong Thao Pham

► **To cite this version:**

Thi Huong Thao Pham. Apports et difficultés d'une collecte de données à l'aide de récepteurs GPS pour réaliser une enquête sur la mobilité. Economies et finances. Université Paris-Est, 2016. Français. NNT : 2016PESC0060 . tel-01429677

HAL Id: tel-01429677

<https://theses.hal.science/tel-01429677>

Submitted on 9 Jan 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.



UNIVERSITE PARIS - EST
École doctorale Organisations, Marchés, Institutions (OMI)

Thèse présentée en vue de l'obtention du grade de Docteur en Sciences Economiques, par :

Thi Huong Thao PHAM

**Apports et difficultés d'une collecte des données
à l'aide du récepteur GPS
pour réaliser une enquête sur la mobilité**

Sous la direction de Jean-Loup MADRE
Encadrée à l'Ifsttar par Philippe MARCHAL et Jimmy ARMOOGUM

Soutenue le 02 mai 2016

Composition du jury :

Gabriel Dupuy, Professeur à l'Université Paris 1, Panthéon Sorbonne, (rapporteur)

Mario Cools, Maître de conférences à l'Université de Liège (rapporteur)

Pierre Taillant, Ingénieur économiste à l'ADEME (examineur)

Jean-Loup Madre, Directeur de recherches émérite, à l'IFSTTAR (directeur de thèse)

Remerciements

Mes travaux de Thèse, présentés dans ce manuscrit, n'auraient pu aboutir sans l'aide et la présence de nombreuses personnes, que je tiens à remercier très sincèrement ici.

Tout d'abord un grand et très chaleureux Merci à mon directeur de thèse Jean Loup MADRE, et mes encadrants Jimmy ARMOOGUM et Philippe MARCHAL. Merci pour votre aide, vos conseils, vos encouragements, pour m'avoir guidé, remotivé parfois et surtout pour m'avoir fait confiance. Merci d'avoir partagé vos vastes connaissances et longues expériences, vos efforts et votre passion de la recherche. Je ne peux pas oublier la présence de Jimmy dans mon audition de demande de bourse de l'IFSTTAR, la préparation de présentation de conférence en Corse avec Philippe, les conseils et correction dans le rapport ADEME avec Jean Loup et Jimmy.

Je remercie Shuning YUAN, Pierre Olivier FLAVIGNY, Thanh Tu NGUYEN, Sophie ROUX de d'avoir partagé leurs expériences sur les enquêtes GPS et la programmation du post-traitement des données GPS.

Je remercie aussi Monsieur Jean Louis COSTER, chef du bureau des statistiques de la multimodalité du Ministère de l'Ecologie, du Développement durables et de l'Energie, de m'accueillir au SOeS pour récupérer et traiter les données GPS.

Je suis également très reconnaissant envers les membres de mon jury de thèse. Les questions soulevées au cours de la présentation affinent et complètent les perspectives à donner à ce travail. De plus, les échanges établis avec mes rapporteurs avant la soutenance m'ont permis de compléter et d'éclaircir certains points de mon manuscrit.

Un grand merci à Laurent pour ses conseils de communication et soutiens pour avancer ma thèse et à Katia, Sophie, Quang Nguyễn, Claire, Toky, Hoai Thu pour leur amitié et encouragement au cours de cette thèse.

Je remercie affectueusement à mes parents, ma famille pour leur soutien, leurs encouragements et leur patience, avec une pensée particulière à pour Huy Quang et Van Anh.

Le dernier mot, tout simplement personnel, est pour Trung Tinh pour sa présence à mes côtés, sa patience et sa compréhension, malgré les errances de mon humeur, les nombreux coups de bourre.

SOMMAIRE

Remerciements	3
SOMMAIRE	5
Résumé	9
Introduction	12
Organisation de la thèse	13
Glossaire	16
1 Objectifs des enquêtes de mobilité en France.....	17
1.1 Pourquoi une enquête sur la mobilité ?.....	18
1.2 Nouvelles utilisations des enquêtes de mobilité	20
1.3 Limites des enquêtes classiques	21
CONCLUSION	22
2 Les nouvelles technologies peuvent-elles améliorer les enquêtes classiques ?.....	24
2.1 Apport des nouvelles technologies pour observer la mobilité.....	24
2.1.1 Le GPS	25
2.1.2 GSM.....	27
2.1.3 Wi-Fi.....	28
2.1.4 Méthode hybrides	28
2.2 Les différentes sources d'erreurs avec les nouvelles technologies	29
2.3 Analyse de l'utilisation des nouvelles technologies dans les enquêtes de déplacement	30
2.3.1 Utilisation en complément d'une enquête ménages transports.....	30
2.3.2 Utilisation pour une étude de faisabilité du remplacement de l'enquête nationale de déplacement	31
2.4 Analyse du pilotage d'une enquête par GPS en Angleterre.....	32
CONCLUSION	34
3 Présentation de l'Enquête Nationale sur les Transports et les déplacements (ENTD) et du protocole de collecte par GPS utilisé en France	35
3.1 Présentation générale de l'ENTD 2007-2008	36
3.1.1 Le déroulement de l'enquête et les thèmes abordés.....	37
3.1.2 Organisation de la composante GPS dans l'ENTD 2007-2008.....	41
3.2 Elaboration du logiciel de déchargement des traces GPS	45
3.3 Analyse des enregistrements GPS en fin de deuxième visite	45
3.4 Les récepteurs GPS utilisés dans l'ENTD 2007-2008.....	47
3.4.1 Choix d'un suivi de personnes et non pas de véhicules.....	47
3.4.2 Sélection des récepteurs-enregistreurs GPS, caractéristiques et paramétrage	47
3.5 Panorama des trois sources de données à confronter dans le volet GPS de l'ENTD 2007-2008	50
4 Acceptabilité, acceptation et biais des enquêtes de transport réalisées à l'aide de récepteur GPS.....	51

4.1	Introduction	51
4.2	Taux d'acceptabilité d'une enquête réalisée à l'aide de GPS	51
4.3	Profils des personnes qui acceptent les enquêtes avec GPS	52
4.4	Comparaison des estimations entre l'ENTD 2007-2008 et le sous-échantillon « GPS »	55
4.4.1	Redressement du sous-échantillon des individus volontaires	55
4.4.2	Comparaison en moyenne	56
4.4.3	Et en distribution	60
4.5	Meilleure description de la mobilité ou une mobilité plus grande ?	61
4.6	Conditions complémentaire pour l'acceptabilité de l'enquête GPS.....	64
4.7	Acceptation d'une enquête réalisée à l'aide de GPS	66
CONCLUSION	67
5	Analyse des traces brutes collectées par GPS	69
5.1	Algorithme Decoutrace	69
5.1.1	Couche donnée GPS	70
5.1.2	Couche "Point d'Intérêt" (POI)	86
5.2	Evaluation de l'algorithme DECOUTRACE par le court questionnaire complémentaire CAPI-GPS	87
CONCLUSION	87	
6	Validation des résultats par confrontation avec les données du questionnaire classique de l'ENTD 2007-2008.....	89
6.1	Panorama de l'appariement des traces GPS et déplacements quotidiens	89
6.2	Analyse des déplacements quotidiens/ traces GPS manquants	91
6.3	Identification des déplacements quotidiens non appariés avec des trace GPS	93
6.3.1	Effets des caractéristiques du déplacement	95
6.3.2	Effets des caractéristiques de l'enquêté	99
6.3.3	Raisons de l'oubli de décrire des déplacements quotidiens.....	100
6.4	Identification des traces GPS sans déplacements quotidiens	101
6.4.1	Caractéristiques des traces GPS non appariées.....	103
6.4.2	Effets des caractéristiques de l'enquêté	106
6.5	Identification des traces GPS appariés avec des déplacements.....	109
6.5.1	Comparaison des heures de départ et d'arrivée.....	110
6.5.2	Comparaison des durées entre déplacement quotidien et trace GPS	112
6.5.3	Comparaison des distances parcourues entre déplacement quotidien et trace GPS.....	114
6.5.4	Comparaison des distances linéaires entre déplacement quotidien et trace GPS	115
6.5.5	Comparaison des modes de déplacement imputés sur les traces GPS avec la description des déplacements quotidiens	115
7	Conclusion générale	117
7.1	Les principaux résultats	117
7.2	Recommandations.....	122
	Le récepteur GPS.....	122
	La formation de l'enquêteur	122
	Qualité de l'interview	123
	Appui et motivations	124
	Contrôle de l'utilisation du GPS – Rappel des données manquantes à cause de l'oubli du GPS	124
	Disponibilité de l'appareil GPS.....	124

	Post-traitement des données GPS.....	124
	Questionnaire complémentaire CAPI-GPS.....	125
	Enquête permanente.....	125
7.3	Prospective.....	126
8	Annexes.....	127
8.1	Annexe 1 : Notice GPS.....	127
	Contenu du « pack GPS » fournis aux enquêtés.....	127
	Fonctionnement du récepteur GPS.....	128
	Quelques conseils d'utilisation.....	128
	Consignes importantes.....	128
	Autres remarques.....	128
	Matériel et documents fournis aux enquêteurs.....	129
	Principe de la chaîne GPS pour l'enquêteur.....	130
8.2	Annexe 2 : Imputation des modes de traces GPS à partir des données complémentaires CAPI-GPS.....	131
8.3	Annexe 3 : Identification des déplacements quotidiens par CAPI non trace GPS.....	134
8.4	Annexe 4 : Identification des traces GPS non appariées avec déplacement quotidien par CAPI.....	138
9	Bibliographie.....	143

Résumé

Les méthodes de collecte de données sur la mobilité basée sur les nouvelles technologies ont évolué au cours des dernières décennies. Le suivi de la mobilité par GPS pour la recherche sur les comportements de déplacement se diffuse, principalement en raison de la couverture mondiale et de la précision du système GPS. Le sous-échantillon d'environ 957 volontaires qui ont accepté de porter un GPS pendant au moins une semaine dans le cadre de l'Enquête Nationale Transport et Déplacements (ENTD) 2007-08 est la première expérience nationale de ce type dans le monde.

Cette thèse présente l'intérêt des nouvelles technologies pour les enquêtes de mobilité, leurs avantages et leurs biais. D'abord, en étudiant l'acceptabilité et les biais des enquêtes de mobilité à l'aide de GPS, les résultats confirment que les enquêtés volontaires pour une enquête GPS sont plus consciencieux et décrivent mieux leur mobilité. Ensuite, nous décrivons le profil des enquêtés qui acceptent l'enquête GPS. L'une des raisons principales du faible taux de l'acceptation du GPS dans l'ENTD 2007-2008 est la disponibilité de GPS. Donc pour l'accroissement du taux de disponibilité des GPS, nous avons besoin de calculer les temps et durées des visites de l'enquête pour réduire le taux d'immobilisation forcé des GPS.

Indépendamment de ce problème d'acceptabilité du GPS par l'enquêté, un défi dans le traitement a posteriori des données collectées par GPS est la mise au point de méthodes permettant de combler les données manquantes et de reconstituer de manière automatisée des séquences continues, à la fois dans l'espace et dans le temps. Nous présentons l'algorithme de post-traitement et les résultats du logiciel de post-traitement des données GPS de l'ENTD 2007-2008. La validation est effectuée en comparant avec les données collectées par les méthodes classiques, ainsi qu'avec le CAPI-GPS qui fournit des informations supplémentaires sur la fiabilité de l'appareil et des caractéristiques plus détaillées (mode, motif...) pour les déplacements réalisés un jour tiré au sort.

Ensuite, nous comparons les descriptions de la mobilité obtenues par deux méthodes, questionnaire classique de l'enquête, d'une part et traces GPS d'autre part, pour les jours couverts par ces deux instruments d'observation. Un panorama de l'appariement des traces GPS et des déplacements quotidiens est effectué en s'appuyant sur leur chronologie de manière automatique mais également par un contrôle manuel. En identifiant les caractéristiques des déplacements quotidiens et traces GPS non appariés, nous estimons les raisons de leur non-appariement.

Cette recherche montre que l'enquête à l'aide de GPS peut être utilisée avec succès pour compléter les enquêtes de transport classiques, mais qu'il est encore prématuré d'imaginer la substitution complète des enquêtes classiques de mobilité par l'enquête GPS.

Mots-clés : enquête de transports, mobilité, nouvelles technologies, collecte de données, GPS, post-traitement, acceptabilité, acceptation, CAPI- GPS, traces GPS, déplacement quotidien.

Abstract

Travel survey methods based on new technologies have evolved in the past few decades, shifting from limited experiments to large-scale travel surveys. GPS-based data collection methods have become particularly popular in travel behavior research, mainly because of the worldwide coverage and the accuracy of the GPS system. We have taken the opportunity of the French National Travel Survey (FNTS) to have the first nationwide experience with embedding such a “GPS package” in a traditional survey, with a sub-sample of approximately 957 voluntary respondents.

This thesis begins by presenting a review of interest in new technologies for mobility surveys, advantages and limitations. Prior to going into the field with the processing of GPS data, it is important to understand the likely acceptability of GPS device among FNTS respondents – specifically whether a sufficient proportion would be willing to use a GPS device while completing a travel diary and, if not, why not. The findings confirmed that the voluntary respondents with GPS receiver are more conscientious and better describe their mobility. We find the profile of respondents who accept the GPS survey. One of the main reason for low GPS acceptance rate in the FNTS 2007 is the availability of GPS. For increasing the GPS availability rate, we need to calculate the time and duration of the survey visits to reduce the GPS immobilization rate.

A challenge in the GPS data post-processing is the development of methods to fill GPS missing data and to reconstitute automatically continuous sequence, both in space and time. Then, the architecture of GPS data post-processing for the FNTS 2007-08 is described. After, the results of the post processing software is calibrated by comparison with conventional methods, such as CAPI-GPS providing for a few trips additional information on the reliability of the device and on more detailed characteristics (e.g. mode and purpose).

Next, we provide an overview of comparing the descriptions of mobility obtained by two methods CAPI and GPS in the days covered by these two observation tools. The automatic matching of GPS traces to daily trips based on their chronologies and the manual control of this matching are performed. By identifying the characteristics of unmatched GPS traces and daily trips, we expect the reasons for their mismatches.

Finally, we evaluate the contributions and challenges of data collection using GPS device. This study shows that the GPS survey can be used successfully to complete the conventional transport surveys, but it is still too early to predict the complete substitution of conventional survey by the GPS mobility survey.

Keywords: transport survey, mobility, new technologies, data collection, GPS, post processing, acceptability, GPS traces, daily trip.

Introduction

A travers le monde, les Enquêtes Nationales Transports ont pour objectifs majeurs, d'une part de décrire la mobilité et de fournir une photographie à une date donnée des comportements de mobilité, d'autre part de permettre de caler des modèles de prévision de la demande de transport.

Les enquêtes de transport sont le principal outil permettant d'étudier le comportement de déplacement. Cependant, les modes de déplacement sont de plus en plus complexes et les modèles de prévision exigent des données plus riches, ce qui impose une charge plus élevée sur les enquêtés et provoque un accroissement substantiel du taux de non-réponse pour les enquêtes traditionnelles.

Pour ces raisons, de nouvelles solutions sont à l'étude. Depuis le milieu des années 90, dans le domaine de la collecte de données sur les comportements de mobilité, l'attention s'est concentrée sur le potentiel des systèmes de localisation tels que les systèmes par satellite (GPS), par les antennes de téléphonie mobile (GSM) ou par les stations de radio (RDS). L'intérêt des méthodes expérimentées avec ces différents systèmes réside d'une part dans la possibilité de recueillir des données plus précises de manière automatique ou semi-automatique, et d'autre part dans la capacité d'étendre la période d'observation par rapport aux méthodes classiques.

Dans ce contexte, les enquêtes basées sur le GPS sont de plus en plus fréquentes et fiables, grâce à plusieurs études menées à partir de la fin des années 1990 et au développement technologique des appareils GPS. Lors des premières utilisations du GPS pour le recueil de données sur la mobilité, la majorité des expériences se sont concentrées sur le suivi des véhicules, notamment pour des raisons de facilité d'alimentation électrique des récepteurs GPS sur de longues périodes, mais également parce que les GPS du milieu des années 90 n'offraient pas une précision suffisante ou bien ne permettaient d'estimer aucune position lorsqu'ils étaient utilisés dans des conditions difficiles comme celles que l'on peut connaître en transport en commun, en milieu urbain.

Leur taille et leur poids était également un frein pour les faire accepter par des enquêtés sur une période relativement longue. Par ailleurs, le traitement a posteriori de données « véhicules », s'il n'est pas immédiat, s'avère beaucoup plus aisé que celui de données « piétons », lesquels se déplacent assez souvent en dehors des réseaux routiers généralement identifiés dans les bases de données cartographiques. Néanmoins, même dans les années 1990, quelques expériences ont été mises en place pour un suivi de personnes. Des expériences réussies ont été menées plus récemment aux Etats-Unis (Wagner, 1997), au Canada, au Japon, en Australie ou en Europe dans le cadre d'enquêtes effectuées sur des échantillons relativement petits, généralement à un niveau local / régional. A ce jour, la collecte de données au moyen de GPS portable sous forme passive a

été appliquée seulement dans quelques enquêtes expérimentales en Europe à Paris et Londres (Flavigny et al., 2000) ou pour un suivi individuel aux Pays-Bas (Kalf et Saris, 1997).

La qualité des données collectées par ces dispositifs est plus précise que les données recueillies avec des méthodes d'enquête traditionnelles. En outre, les GPS imposent une charge moindre sur les enquêtés en réduisant leur interaction avec les enquêteurs. Plutôt que ce dialogue parfois difficile, les enquêtes de transport avec GPS traitent les données a posteriori pour récupérer des informations de transport avec un grand degré de précision, avec peu ou pas de sollicitation des enquêtés. Cela permet de mener des enquêtes plus longues sans perdre la qualité des données à cause de la fatigue des enquêtés.

Malgré ces avantages, plusieurs études ont souligné la possibilité de biais liés à cette nouvelle technologie, par exemple, l'acceptabilité / l'acceptation par les enquêtés, la confidentialité des données, les erreurs de l'utilisateur (GPS oublié, non rechargé, non allumé), les difficultés de traitement à posteriori de données GPS brutes, etc.

Sur un sous-échantillon de la dernière Enquête Nationale Transports et Déplacements (ENTD) Française 2007-08, nous avons collecté les traces GPS de la mobilité durant une semaine pour 957 volontaires

Cette thèse s'appuie notamment sur l'analyse, le traitement et la comparaison de données collectées par trois différents instruments de l'ENTD 2007-2008 : le questionnaire classique, les traces GPS brutes et le questionnaire complémentaire (CAPI-GPS), pour étudier la possibilité d'une transition vers la généralisation de l'utilisation des nouvelles technologies dans les enquêtes sur la mobilité des personnes.

Organisation de la thèse

Ce manuscrit de thèse est organisé en sept parties. Dans la première partie, nous présentons la littérature portant sur les enquêtes de mobilité en France. Nous répondons à la question « Pourquoi une enquête sur la mobilité ? » au regard des enjeux sociétaux et économiques avec des applications multiples. Suite à l'évolution de plus en plus variée de la mobilité des personnes, de nouvelles utilisations des enquêtes de mobilité sont exigés. Cette partie donne ensuite un bref panorama des limites des enquêtes classiques.

Le développement des nouvelles technologies telles que GPS, GSM, Wi-Fi, etc. offrent des apports importants pour collecter des données de transport, mais ont aussi des erreurs inévitables. Tenant compte de leurs avantages et inconvénients, la seconde partie analyse l'utilisation des nouvelles technologies dans les enquêtes de déplacement en France et dans le monde. A côté des test

positifs, des expérimentations négatives ont toutefois également été menées, et nous rappelons l'enquête nationale de transport 2003-2004 en Angleterre pour laquelle les Britanniques ont renoncé à utiliser des traceurs GPS pour leur enquête nationale.

La troisième partie est consacrée à la présentation de l'Enquête Nationale sur les Transports et Déplacements et au protocole de collecte par GPS. Nous rappelons l'élaboration du logiciel de déchargement des traces GPS et l'analyse des enregistrements GPS en fin de deuxième visite, les récepteurs GPS utilisés dans l'ENTD 2007-2008. La fin de cette partie dresse un panorama général des trois sources de données à confronter dans le volet GPS de l'ENTD 2007-2008.

La faisabilité des enquêtes de mobilité à l'aide de récepteur utilisant le GPS à grande échelle dépend (mais pas uniquement) de l'acceptabilité d'une telle technologie par les personnes sélectionnées. Dans la quatrième partie, nous proposons de caractériser l'acceptabilité et de déterminer si elle est ignorable (c'est-à-dire indépendante des comportements de mobilité que l'on souhaite observer) ou non ignorable (Armoogum, 2002). Nous comparons les estimations sur l'ensemble de l'ENTD et sur l'échantillon des individus qui sont susceptibles d'accepter le principe du suivi par GPS (les deux échantillons ayant été redressés de la même manière). Nous vérifions l'hypothèse de Stopher et al. (2010) selon laquelle les individus qui répondent à une enquête de mobilité à l'aide de récepteur GPS sont plus consciencieux et décrivent mieux leur mobilité. Nous proposons des recommandations pour augmenter le taux d'acceptabilité de l'enquête GPS. Nous vérifions également si la disponibilité d'un GPS influence le taux d'acceptation de l'enquête GPS.

Après nous être concentrés sur l'acceptabilité et les biais de l'enquête GPS, en nous appuyant sur les données brutes GPS de l'enquête nationale de transports, nous développons l'algorithme de découpage des traces (Decoutrace) qui est mentionné initialement dans la thèse de Yuan (Yuan, 2010). Grâce au court questionnaire complémentaire CAPI-GPS, nous pouvons construire un modèle probabiliste conditionnelle pour imputer les modes de transports des traces GPS et l'évaluer en comparant les résultats entre les deux sources de données fournies par les GPS et par ce CAPI-GPS.

Un problème habituel dans les enquêtes sur la mobilité quotidienne conduites avec une méthode classique est la sous-estimation du nombre de déplacements et l'inexactitude de la déclaration de la durée et de la distance des déplacements. Plusieurs méthodes ont été développées pour dépasser ces limites. L'utilisation du récepteur GPS dans une enquête transport fournit une opportunité pour estimer les différences quant au nombre, à la durée, à la distance, aux heures de départ et d'arrivée des déplacements enregistrés par ces deux méthodes. La sixième partie présente les résultats globaux de ce travail. Est tout d'abord proposé un panorama de l'appariement automatique des traces GPS et déplacements quotidiens en s'appuyant sur leur chronologie et par un contrôle manuel. Une première analyse des déplacements quotidiens et traces GPS manquants non appariés

est réalisée. Ensuite, en s'appuyant sur les effets des caractéristiques du déplacement quotidiens/traces GPS et de l'enquête, nous identifions des déplacements sans traces GPS ; et aussi des traces GPS sans déplacements quotidiens et essayons de trouver les raisons de ces non-appariements entre les déplacements quotidiens et les traces GPS. L'analyse est complétée en comparant les heures de départ, d'arrivée, le mode imputé aux traces GPS avec celui déclaré pour les déplacements quotidiens qui sont appariés.

Au-delà de ces six parties, le mémoire s'achève comme il se doit par une conclusion générale, retraçant les apports de ce travail, esquissant quelques recommandations et pistes de prolongement.

Glossaire

Acceptabilité : réponse positive à la question de l'acceptation du suivi par GPS par l'enquêté.

Acceptation : prise en charge de l'appareil proposé par l'enquêteur.

Activité : une interaction continue avec l'environnement physique, un service ou une personne, dans un même cadre socio-spatial pertinent pour l'analyse. Elle comprend toute période d'attente avant ou pendant l'activité.

Appropriation : utilisation de l'appareil sans l'oublier et en le rechargeant régulièrement, ce qui devrait produire des traces valides.

CAPi: Computer Assisted Personal Interview. Enquête en face-à-face assistée par ordinateur; IPAO (Interview Personnelle Assistée par Ordinateur) pour les québécois.

CATi: Computer Assisted Telephone Interview. Enquête par téléphone assistée par ordinateur.

Démarrage à Chaud : moins de deux heures à partir du dernier positionnement.

Démarrage à Froid : Récepteur GPS utilisé pour la première fois peut avoir la perte d'informations de l'éphéméride ; la position initiale du récepteur peut s'être éloignée de plus de 1 000 kilomètres.

Démarrage à Tiède : plus de deux heures à partir du dernier positionnement.

Déplacement : une séquence de trajets continue entre deux activités.

DOP (Dilution Of Precision : atténuation ou diminution de la précision) : La précision de la position obtenue dépend, toutes choses égales par ailleurs, de la géométrie du système : si les satellites visibles se trouvent tous dans un cône d'observation de faible ouverture angulaire, la précision sera évidemment moins bonne que s'ils sont répartis régulièrement dans un large cône. Les effets de la géométrie du système de mesure sur la précision sont décrits par ce paramètre.

Durée de perte du signal : on distingue trois catégories "courte", "longue" et "très longue" selon l'effet observé. Si dans la période de perte du signal, le statut (mouvement, arrêt, changement de la vitesse) de l'enquêté varie peu ou que ce changement n'influence pas l'algorithme de découpage, on dit que la période est courte, par exemple quelques secondes perdues sur autoroute. Si dans la période où les signaux sont perdus, on ne peut plus être certain du statut (mouvement, arrêt, variation significative de la vitesse) de l'enquêté, on dit que la période est longue, par exemple une trentaine de minutes perdues dans le métro. S'il y a des déplacements ou activités totalement perdus, on dit que la période de perte du signal est très longue.

DRMS : est la racine carrée de la moyenne des carrés des erreurs sur les deux axes horizontal et vertical.

Etat de mouvement : on définit trois catégories : en mouvement, à l'arrêt, un ou plusieurs déplacement(s) manquant(s).

SIG : un Système d'Information Géographique (GIS en anglais) capable d'organiser et de présenter des données alphanumériques spatialement référencées, ainsi que de produire des plans et des cartes.

GPS : Global Positioning System – un système de géo-localisation fonctionnant au niveau mondial. Il s'agit du principal système de positionnement par satellites.

#GNSS : Global Navigation Satellite Systems – GNSS englobe l'ensemble des systèmes de navigation par satellite. Les systèmes GPS, Glonass, Galileo, Beidou sont tous des GNSS. #

HDOP : Horizontal Dilution Of Precision, coefficient d'affaiblissement de la précision sur la composante horizontale du résultat de positionnement instantané.

1 Objectifs des enquêtes de mobilité en France

Les enquêtes nationales transport existent dans un assez grand nombre de pays, principalement en Europe, mais avec des objectifs, fréquences, échantillonnages, méthodologies et des protocoles d'enquêtes différents entre les pays ce qui limite la comparaison des résultats, faute d'harmonisation des données et méthodes.

A notre connaissance, seuls les pays suivants ont réalisé des enquêtes couvrant la totalité du territoire national et portant à la fois sur la mobilité locale et de longue distance : Autriche, Allemagne, Belgique, Danemark, Espagne, Finlande, France, Grande-Bretagne, Italie, Norvège, Nouvelle Zélande, Suède, Suisse, USA. Il y a seulement cinq pays suivants qui maintiennent encore des enquêtes nationales en continu :

- Les Pays-Bas depuis 1978 ;
- L'Angleterre depuis 1999 ;
- L'Allemagne avec le panel depuis 1994 ;
- La Nouvelle-Zélande depuis 1994 ;
- Le Danemark depuis 2006.

L'enquête Nationale Transport française est réalisée environ tous les douze à quinze ans et la dernière date d'il y a 7 ans. La question de son renouvellement se pose pour actualiser l'observation des comportements de transports des résidents français. Pour y répondre, dans un premier temps, nous avons besoin de préciser les objectifs et enjeux de la réalisation d'une enquête sur la mobilité. D'autre part, la mobilité des personnes est de plus en plus variée dans le temps et dans l'espace : les données sur la mobilité évoluent, et nous sommes face à de nouvelles opportunités et de nouveaux défis dans l'utilisation des enquêtes sur la mobilité. Face à la complexité des comportements de transport, les enquêtes classiques ne sont pas parfaites et comportent plusieurs limites potentielles. Pour améliorer l'observation des mobilités complexes, nous présentons quelques pistes de méthodologies s'appuyant sur les nouvelles technologies dans cette partie.

1.1 Pourquoi une enquête sur la mobilité ?

L'observation de l'usage des véhicules sur les routes s'appuie en premier lieu sur les systèmes d'information des transports (comptages et capteurs). Ils nous donnent une image du volume du trafic mais ne fournissent aucune information sur les personnes en déplacement, ni sur les motifs ou sur l'enchaînement de ces déplacements (Armoogum et al., 2010). Une enquête sur les déplacements collecte des informations sur les liens entre motifs, modes, origines et destinations ou temps consacré aux déplacements mais aussi des informations démographiques, socio-économiques ainsi que sur les habitudes de déplacements des individus d'un ménage. Ces informations sont nécessaires en matière d'aménagement et d'aide à la décision dans le domaine des transports.

La France dispose de bons outils d'observation de la mobilité :

- France entière, les Enquêtes Nationales sur les Transports (ENT) sont riches en informations mais lourdes. En conséquence, elles sont réalisées à un rythme de plus en plus espacé : la dernière, celle de 2007-08, est la cinquième de sa lignée (1966-67, 1973-74, 1981-82 et 1993-94) ;
- sur les zones urbaines, les Enquêtes Ménages Déplacements (EMD) ont une périodicité de renouvellement de 7 à 10 ans. Mais toutes les villes ne mesurent pas la mobilité de leurs habitants, qui correspond d'ailleurs à une part décroissante de leur trafic interne (le transit et les trafics d'échange, notamment avec les couronnes péri-urbaines, prennent une importance croissante).

L'objectif de ces enquêtes est la connaissance des déplacements des ménages résidant en France et de leur usage des moyens de transport tant collectifs qu'individuels. Elles reprennent d'ailleurs les définitions et les principes établis lors des précédentes éditions afin d'assurer la mesure des évolutions structurelles.

Quand les enquêtes sur la mobilité sont renouvelées, elles permettent des comparaisons dans l'espace et dans le temps, éventuellement sur une longue période. Elles peuvent notamment servir à estimer des tendances en matière d'évolution des trafics, globalement et par mode, à décrire les pratiques de mobilité individuelle, et à mesurer des évolutions dans la demande de transports et les comportements, dont les causes peuvent être économiques, sociales et/ou urbanistiques, en relation avec l'évolution de l'offre de transport.

Les applications sont multiples, avec des enjeux sociétaux et économiques : une enquête ménage déplacements est un moyen de collecte d'informations riches sur la mobilité des personnes et leur contexte socio-économique le plus large. Elle fournit des résultats non seulement au niveau national mais aussi au niveau régional et des zones d'étude. Elle permet de répondre à différents enjeux :

En matière de sécurité routière : l'information sur les kilométrages réalisés en voiture selon le type de route utilisée (voirie urbaine, autoroute, etc.), tant pour les déplacements de la veille et du dernier week-end que pour les voyages à longue distance, est utile notamment pour des applications en matière de sécurité routière (mesure de l'exposition au risque d'accident de la route pour des groupes de population).

En matière environnementale : l'enquête permet de mesurer les effets de la mobilité sur la consommation d'énergie et sur les émissions polluantes (globales ou locales). Via l'estimation des distances parcourues (véhicules-kilomètres ou voyageurs – kilomètres), on peut suivre l'émission de polluants, la consommation d'énergie. L'enquête par sa vision nationale de la mobilité urbaine et périurbaine due à l'extension des villes, permet de tirer un bilan au niveau national des modifications structurelles de la mobilité.

Au regard des enjeux économiques, les informations de l'enquête permettent de comprendre les pratiques de mobilité et leurs dimensions essentielles : le nombre de déplacements, les modes et le temps qui y est consacré. En se basant sur ces caractéristiques, nous pouvons analyser les comportements de la population à partir de données démographiques. On peut par exemple étudier les dynamiques et les conséquences de l'équipement et de l'usage automobiles (Gallez, 1994) ; (Berri, 2005) ; (Collet, 2007) pour ensuite effectuer une projection à long terme des évolutions des parcs automobiles (Gallez, 1994). Ces recherches permettent ensuite de mettre en perspective temporelle la progression de la motorisation des ménages et l'évolution du contexte économique (revenu, prix). Il sera donc possible, à partir des données d'enquête, de mesurer, par exemple, l'influence du taux d'équipement automobile ou de l'accès aux moyens de transport sur la mobilité. Elles peuvent aussi être utilisées pour calculer le coût des déplacements via les caractéristiques des véhicules ou les tarifs utilisés dans les transports en commun. Jusqu'en 1981, les ENT comportaient une partie "dépenses". Pour ne pas trop charger le questionnaire, on préfère maintenant traiter les dépenses par appariement avec les enquêtes Budget de Famille et se concentrer sur une description plus fine des déplacements, par exemple par GPS en tenant compte de la baisse des prix du matériel, et de ses améliorations techniques.

En ce qui concerne le domaine du marketing, la plupart des études reposent sur des données référencées au lieu de résidence principal. Pourtant, les gens passent la majeure partie de leur temps éveillé en dehors de leur lieu de résidence et utilisent ce temps pour travailler, mais également pour consommer et de plus en plus pour faire leurs achats. Il est donc maintenant important de raisonner en terme "d'économie présentielle" en s'appuyant sur les données fournies par les enquêtes de mobilité (Potier et Terrier, 2007). Pouvoir suivre les individus, connaître leurs trajets, identifier les moments où ils sont à proximité des points de vente, revêt donc aujourd'hui une importance capitale.

Quant aux enjeux de trafic, les données de mobilité, surtout celles issues d'enquêtes ménages déplacements, permettent d'estimer non seulement la demande de transport en fonction de l'offre de l'infrastructure mais aussi l'état du trafic pour ensuite modéliser par exemple le choix modal ou le comportement des conducteurs en matière de sécurité routière.

Concernant les enjeux de la planification des investissements d'infrastructure, la conception puis l'évaluation de projets de transport nécessitent une bonne connaissance des déplacements et par conséquent des données sur les besoins et pratiques de mobilité. Cette connaissance permet ainsi d'estimer la clientèle d'une ligne de bus, de métro ou tramway ou d'une infrastructure routière. C'est sur la base de données sur la mobilité et de la composante économique du coût de transport pour l'utilisateur que des estimations financières de tarification peuvent être élaborées, pour aboutir aux simulations financières des projets.

Le secteur des transports est à l'origine d'une partie importante des rejets de polluants et de gaz à effet de serre. L'enquête sur la mobilité permet de fournir des données de base afin d'évaluer ces effets (CETE N-P et INRETS, 2006). Les thèmes de recherche qui en découlent sont les effets de la mobilité sur la consommation d'énergie ainsi que l'estimation des émissions de polluants dans l'air (qualité de l'air et réchauffement climatique). Ces recherches sont un outil d'aide à la décision dans l'élaboration de la politique nationale et locale et permettent d'appréhender les modifications souhaitées des comportements de mobilité en vue de satisfaire les engagements internationaux et nationaux sur la réduction des émissions polluantes.

En conclusion, les données de mobilité sont utilisées dans l'évaluation de projets de transports en terme socio-économique pour la planification des infrastructures et pour les études de prospective. Elles contribuent aussi à évaluer les émissions de gaz à effet de serre, de polluants locaux et de consommations énergétiques liées aux transports. Elles fournissent une base de données au service de la collectivité et sont utilisables pendant plusieurs années entre deux collectes. A plus court terme, certaines enquêtes mesurent l'audience de l'affichage et d'autres permettent aux autorités organisatrices de répartir les recettes (notamment Versement Transport) des réseaux de transport en commun inter-connectés.

1.2 Nouvelles utilisations des enquêtes de mobilité

Aujourd'hui, la mobilité des personnes est de plus en plus variée dans le temps et l'espace. Cette évolution peut s'expliquer notamment par la fragmentation spatiale, le travail à temps partiel ou à domicile, la spécialisation de la main-d'oeuvre, une augmentation du nombre de familles avec plusieurs actifs, la plus grande diversité dans la composition des ménages et aussi par une tendance à l'augmentation du nombre de voitures par ménage.

D'autre part, les besoins de données sur la mobilité évoluent, de l'estimation la plus précise des trafics pour le dimensionnement des infrastructures (matrices Origine-Destination, etc.), vers une meilleure compréhension des comportements individuels, non seulement au cours d'une journée « ordinaire » mais tout au long de l'année. En effet, la compréhension de la mobilité peut répondre à un besoin économique (comme dans la mesure de l'audience des médias extérieurs, notamment l'affichage) et les centres d'intérêt se déplacent vers des enjeux sociaux et environnementaux, car le secteur des transports est confronté à un défi redoutable par rapport aux objectifs du développement durable, puisque sa part continue à augmenter dans la consommation de produits pétroliers et dans l'émission de gaz à effet de serre.

Face à la complexité accrue des comportements et en tenant compte de l'évaluation des besoins en données, actuellement, la recherche se concentre de plus en plus sur l'enchaînement de la mobilité d'un individu pour chaque jour et chaque semaine, et sur les relations entre la structure spatiale globale (lieux de travail, d'achat, de loisirs, etc.) et les détails des lieux réellement visités au cours des déplacements (Kizel (2003) ; Mat et Timmermans (2004); Golo (2000); Barnet et Sarmiento (1998)). Un déplacement n'est plus traité comme un comportement autonome, mais comme faisant partie d'un ensemble complexe de mobilité. En conséquence, il est fondamental d'être en mesure de recueillir des données précises sur la localisation de l'origine et de la destination, le type de destination, la longueur, la durée, l'heure de départ et d'arrivée et les modes pour tous les déplacements quotidiens.

1.3 Limites des enquêtes classiques

Traditionnellement, les données de mobilité sont collectées au moyen d'enquêtes sur papier ou par téléphone. Les personnes « recrutées » sont invitées à décrire leur comportement de mobilité d'une journée ou à reconstruire leur comportement sur un ou plusieurs jours précédents. Il est démontré que les données recueillies par ces méthodes présentent systématiquement des écarts par rapport au comportement réel. Lors d'un recueil classique de données sur la mobilité, on demande à la personne sélectionnée de décrire les déplacements qu'elle a effectués la veille de l'interview (et parfois au cours du dernier week-end). Ces modes d'observation ne sont pas parfaits et comportent plusieurs types d'erreurs potentielles :

- I. Non-réponse totale : l'interviewé omet par exemple de décrire certains déplacements (Brog et al. 1980 ; Brog et Meiburg, 1981 ; Armoogum et Madre, 1997 ; Armoogum, 2002),
- II. Non-réponse partielle : l'interviewé répond correctement sur le nombre de déplacements mais ne renseigne pas certaines variables (motif, mode de transport, etc.) (Zmud et Arce, 1997 ; Armoogum et Madre, 1997),

- III. Arrondi : par exemple, heures de départ ou d'arrivée données à 5 ou 10 mn près, ce qui peut entraîner une grande incertitude quand on calcule un temps de transport par différence (Armoogum et Madre, 1997 ; Yamamoto 2009),
- IV. Sous-estimation : l'interviewé sous-estime les petits déplacements ainsi que ceux qui ne commencent pas ou ne se terminent pas au domicile et les conducteurs de voiture sous-estiment leurs temps de parcours (Ettema et al. 1996 ; Stopher, 1992),
- V. Surestimation : les personnes qui voyagent avec les transports publics ont tendance à surestimer les temps (Ettema et al. 1996 ; Stopher 1992).

En raison de leur coût et des moyens mobilisés, les enquêtes classiques n'ont lieu qu'avec une faible périodicité (jusqu'à une dizaine d'années), ce qui est relativement long si des ruptures apparaissent dans les comportements. Ces types d'enquête donnent une image réductrice de la mobilité du point de vue des individus (Gascon et Madre, #2010#). Tout se passe comme si leur comportement d'un jour de semaine était "reproductible", ce qui on le sait n'est pas réaliste et le devient de moins en moins en raison des désynchronisations des rythmes de vie, qui se traduisent notamment par des ouvertures tardives des commerces voire le dimanche pour certains (Golob et Meurs (1986) ; Axhausen et al. (2002) ; Bhat et al. (2004)).

Améliorer l'observation des mobilités complexes est maintenant possible puisque :

- D'une part, les méthodologies d'enquête ont évolué : nous sommes passés de l'ère PAPI (Paper and Pencil interview) à l'ère CATI ou CAPI (Computer Assisted Telephone Interview ou Personal Interview) ;
- D'autre part, les nouvelles technologies comme la localisation automatique par satellite (GPS mais aussi à moyen terme Galileo), par les antennes de téléphonie mobile (GSM) ou celles des stations de radio (RDS) ou encore par les infrastructures Wi-Fi, permettent d'améliorer considérablement la précision du cadre temporel (heures de départ, d'arrivée...) et spatial (distances et lieux d'activité) des déplacements.

CONCLUSION

Nous avons considéré les applications et enjeux sociétaux et économiques de l'enquête de mobilité au niveau national, mais aussi au niveau régional et des zones d'études. Les données de mobilité sont utilisées dans l'évaluation de projets de transports en terme socio-économique pour la planification des infrastructures et pour les études de prospective. Elles contribuent aussi à évaluer les émissions de gaz à effet de terre, de polluants locaux et de consommations énergétiques liées aux transports.

Les enquêtes de transport sont le principal outil permettant d'étudier le comportement de déplacement. Aujourd'hui, de plus en plus de variation de la mobilité dans le temps et l'espace

demande l'évolution des données collectées sur la mobilité, la haute précision de l'estimation des trafics et la meilleure compréhension des comportements individuels non seulement au cours d'une journée, mais tout au long de l'année.

Cependant, les modes de déplacement sont de plus en plus complexe et les modèles de prévision exigent des données plus riches, ce qui impose une charge plus élevée sur les enquêtés et provoque un accroissement substantiel du taux de non-réponse pour les enquêtes traditionnelles. Pour ces raisons, de nouvelles solutions sont à l'étude.

2 Les nouvelles technologies peuvent-elles améliorer les enquêtes classiques ?

La méthode de collecte classique contient plusieurs erreurs potentielles et les nouvelles technologies sont prometteuses pour améliorer la qualité des données de mobilité. Cette partie présente les apports mais aussi les différentes sources d'erreurs des nouvelles technologies dans l'enquête de transport. En fait, de nombreuses études sur l'application du GPS dans les enquêtes de transport ont été menées. Il y a des résultats favorables de l'utilisation de GPS en complément d'une enquête ménage de transport. On est en train d'étudier la faisabilité de l'utilisation du GPS en remplacement de l'enquête nationale de déplacement traditionnelle. A côté des tests positifs, toutefois, sur la base d'un test négatif, les Britanniques ont renoncé à utiliser des traceurs GPS pour leur National Travel Survey 2003-2004.

2.1 Apport des nouvelles technologies pour observer la mobilité

Dans un premier temps, la méthode PAPI (Paper and Pencil Interview) a été utilisée pour toutes les enquêtes ménages déplacements par des entretiens en face-à-face, par appel téléphonique ou par courrier. Des questions sur les déplacements d'une ou plusieurs journées étaient posées à une personne de référence du ménage ou à chacun de ses membres. Pour chaque déplacement, on demandait l'adresse à l'origine et à la destination, le lieu, le motif, le mode, le nombre de personnes accompagnant le répondant, etc. Avec cette méthode, l'enquêté peut facilement fournir des informations imprécises ou même erronées, omettre des déplacements, arrondir la durée et la distance parcourue ou omettre des lieux d'origine et de destination.

Pour remédier à ces erreurs dans les données sur la mobilité, à partir des années quatre-vingts, on a remplacé la méthode manuelle par la méthode assistée par ordinateur : d'abord le CAPI (Computer Assisted Personal Interview), le CATI (Computer Assisted Telephone Interview) le CASI (Computer Assisted Self Interview ou enquête auto-administrée) ou CAWI (Computer Assisted Web Interview). Parmi ces méthodes, chacune a des avantages et inconvénients. La méthode CATI permet d'allier souplesse, rapidité d'exécution, et possibilité de rappeler automatiquement les individus à des moments différents de la journée selon un protocole précis. Cependant, du fait de la multiplication de l'utilisation de cette technique par un certain nombre d'institutions, il existe un risque de voir diminuer la qualité des données et la participation à ce type d'enquête. D'autre part, la durée des interviews est courte et il n'est pas possible d'utiliser les images. Dans l'Enquête Nationale de Transports et Déplacements 2007-2008, la méthode CAPI a été retenue.

Il y a eu de nombreuses études sur les méthodes classiques et les possibilités de transition vers des méthodes automatisées de collecte de données sur la mobilité. Des études américaines ont examiné

les possibilités d'utilisation de Systèmes d'Information Géographique (SIG) dans les enquêtes sur les déplacements (Abdel-Aty et al., 1995), Greaves, 1997) ainsi que deux projets d'études sur l'utilisation de l'ordinateur de poche et de l'Internet afin de collecter des données sur la longue distance (Polak et al., 1999).

Plus récemment, on a expérimenté le recueil de données de mobilité par repérage satellitaire GPS sur un sous-échantillon de l'ENTD en France. Au cours de la dernière décennie, les technologies mobiles ont commencé à être utilisées pour suivre la circulation des véhicules, les expéditions de fret, le transport de marchandises et la mobilité des personnes. Cette section passe en revue les principales caractéristiques de ces technologies, notamment la téléphonie mobile (GSM), le système de positionnement global (GPS, et bientôt GALLILEO, le projet européen de système de positionnement par satellites), les assistants numériques personnels (PDA) et le recours à des ordinateurs portables ainsi que les systèmes de communications avancés (Wi-Fi) et la combinaison des méthodes précédentes (méthodes hybrides).

2.1.1 Le GPS

Le GPS est un système de radio positionnement par satellites développé par l'armée américaine (US Navy) conçu au début des années 70 par les militaires pour leur besoin propre. Il remplaça le système Transit qui fut conçu dans les années 50 et fonctionna jusqu'au milieu des années 1980. Le système donne la position et la vitesse d'un mobile : à tout instant, en tout endroit, dans un système de référence mondial (appelé WGS84) avec une précision inférieure à 100 m, et le temps à une microseconde près. Il permet de déterminer les coordonnées géographiques d'un point situé n'importe où dans le monde 24 h sur 24 h.

Les coordonnées d'un point sont obtenues en mesurant, à l'aide de signaux radio, sa distance par rapport à un certain nombre de satellites (dont les orbites sont quasi circulaires, de rayon 26 500 km, de période environ égale à 12 heures, et situés dans six plans inclinés de 55° par rapport à l'équateur) et en appliquant le principe de triangulation. Quatre satellites sont nécessaires pour établir les coordonnées x, y et z de ce point.

Pour des raisons de sécurité, le département américain de la défense a dans le passé brouillé les signaux des satellites. Du fait de ce brouillage, la position d'un point déterminée à l'aide d'un appareil de terrain ne pouvait être connue qu'à 100 m près.

La technique de correction différentielle, élaborée par les ingénieurs civils, permettait de réduire les effets du brouillage pour arriver à une précision décimétrique. Cette technique nécessite la présence d'un récepteur fixe de position connue. A tout moment, la différence de position entre les

coordonnées réelles et les coordonnées mesurées permet de déterminer le décalage en x, y et z. Ce décalage à un moment t étant connu, la correction à appliquer à la mesure effectuée par un récepteur mobile de terrain à un même moment t est alors connue.

Depuis le 1er mai 2000, le brouillage, ou « selective availability » (SA), a été désactivé, mettant à disposition de tous des données avec une précision inférieure à 20 m. La technique de correction différentielle permet toujours d'améliorer ces résultats, grâce à la mise en relation des données du récepteur fixe avec celles des données de terrain. Ainsi, il est possible d'obtenir une précision inférieure à 10 m.

Le GPS a été déclaré opérationnel en février 1994 par le congrès américain et ses possibilités ont fait que la communauté civile scientifique s'est intéressée à lui. Il est aujourd'hui sous le contrôle du département de la défense (DoD) et du département des transports (DoT) des Etats Unis. Ce système opérationnel depuis au moins 30 ans offre une solution à presque tous les besoins de positionnement et de navigation, et présente une utilisation simple sur le terrain. Les récepteurs sont donc passifs. Le GPS a apporté une révolution dans les méthodes de relevé topographique, en géodésie et en navigation, voir pour le guidage sur route. Aujourd'hui tous les nouveaux réseaux géodésiques nationaux sont réalisés par des méthodes spatiales et en particulier par GPS. Comme le récepteur doit être en vue des satellites, le GPS ne fonctionne pas très bien voire plus du tout dans certains lieux comme dans un tunnel, dans le métro ou à l'intérieur des bâtiments. Les dispositifs GPS permettent en général d'obtenir une localisation dans les zones rurales plutôt meilleure que dans les zones urbaines (rues étroites, canyons urbains, ...), et sont relativement peu affectés par la densité de population.

En complément du système GPS, nous connaissons le système européen EGNOS - European Geostationary Navigation Overlay Service soit Service Européen de Navigation par Recouvrement Géostationnaire - qui utilise un système inspiré du GPS différentiel, mais étendu à un ensemble de stations dont les données sont corrélées entre elles. Le service EGNOS est double :

- Un service ouvert et utilisable par tous (Open Service), améliorant la précision donnée par le GPS d'origine et sa disponibilité ;
- Un service à information garantie (Safety Of Life), utilisable notamment par l'aviation civile.

Le service EGNOS diffuse en effet des données de corrections du signal GPS permettant non seulement de compenser les effets de propagation au travers de l'ionosphère, mais surtout l'erreur résiduelle sur les axes verticaux et horizontaux, en garantissant cette information pour le service Safety Of Life sur les 150 prochaines secondes, avec une très grande probabilité compatible avec les systèmes embarqués d'aide à la navigation aérienne.

L'infrastructure EGNOS est basée sur un réseau de 34 stations terrestres de référence qui reçoivent les deux fréquences de satellites GPS. Situées à des endroits connus, elles sont équipées d'horloges atomiques, et permettent de déterminer le trajet des signaux des systèmes de positionnement des États-Unis GPS et, historiquement, du russe GLONASS.

Ces stations de référence terrestres, réparties sur le territoire européen ainsi que quelques sites distants, constituent un maillage serré qui complète la triangulation obtenue à partir des satellites du GPS. EGNOS permet de localiser les usagers avec une précision de 1,5 mètre. EGNOS est devenu officiellement disponible pour l'aviation à partir de mars 2011. Ses signaux sont très utiles pour la sécurité du transport aérien.

La technologie GPS est maintenant utilisée dans de nombreux domaines : une des applications en transport est comme un outil de suivi de la mobilité des personnes. Le service ouvert EGNOS est disponible depuis le 1er Octobre 2009 pour ceux qui sont équipés d'un récepteur GPS compatible le WAAS6.

2.1.2 GSM

Le GSM (Global System Mobile) constitue au début du 21ème siècle le standard de téléphonie mobile le plus utilisé en Europe. Il s'agit d'un standard de téléphonie dit « de seconde génération » (2G) car, contrairement à la première génération de téléphones portables, les communications fonctionnent selon un mode entièrement numérique.

Le GSM s'appuie sur la triangulation et la puissance du signal des BTS (Base Transceiver Station ou en français, station de transmission de base) pour déterminer la position. La précision de positionnement de GSM dépend de la densité des BTS. Il parvient donc à fournir les informations les plus précises dans les zones urbaines denses, où les stations de base sont peu espacées, mais ne fournit que des positions imprécises dans des zones reculées où les stations de base sont très largement espacées, lorsqu'elles existent. Les dispositifs GSM présentent l'avantage de fonctionner dans certains endroits où le GPS n'est pas en mesure de fournir une position. Pourtant, même dans les zones urbaines denses, la précision du positionnement d'un appareil GSM n'est généralement pas inférieure à ± 40 mètres, le plus souvent au-delà de 100 mètres dans les zones suburbaines (Varshavsky et al., 2006). En outre, le GSM ne donne pas d'information sur les déplacements courts intérieurs à une cellule couverte par une antenne, et surtout il peut fournir une information trompeuse s'il change d'antenne de référence alors que le récepteur n'a pas bougé.

2.1.3 Wi-Fi

Indépendamment du GSM et du GPS, un autre système de positionnement pouvant trouver une application dans les enquêtes sur la mobilité des personnes mérite d'être présenté : le positionnement par Wi-Fi. Le Wi-Fi est l'abrégié de Wireless Fidelity et le nom commercial de la norme IEEE 802.11b, une technologie de réseau local Ethernet sans fil (en anglais : WLAN, pour Wireless Local Area Network / le mot "wireless" signifie "sans fil"), basé sur la fréquence 2.4 GHz.

Il s'agit donc d'une technologie de transmission de l'information par ondes radio électromagnétiques, de fréquence 2.4 GHz (Giga Hertz), entre différents ordinateurs ou appareils. Des logiciels ou kits de développement permettent d'utiliser de plus en plus le Wi-Fi dans des applications de localisation. Le dispositif qui peut être utilisé pour cela est un PC portable, un Tablet-PC, un Smartphone, ou des étiquettes RFID. Plus la densité des signaux Wi-Fi dans une zone donnée est importante, meilleur sera le positionnement. Pour les enquêtes de mobilité, en utilisant un véhicule équipé en même temps par un GPS et un récepteur de signal Wi-fi, il est possible d'atteindre une précision de 13 à 20 mètres dans les zones urbaines denses, diminuant à 30 à 40 mètres dans les quartiers de banlieue (Chiou et al. 2009)(Lim et al. 2008). Comme dans le positionnement par GSM, les résultats dépendent de la densité des stations de base. Le Wi-Fi semble être une bonne méthode pour la précision dans les zones urbaines denses. Cependant, la qualité de l'enquête dépend du taux d'acceptabilité par les enquêtés et elle n'est pas assurée dans les zones rurales où la densité des stations de base est faible.

2.1.4 Méthode hybrides

La combinaison de nouvelles technologies de localisation nous permettra d'obtenir les meilleurs résultats car elle pourra diminuer les inconvénients de chaque technologie. Par exemple, nous pouvons profiter de la capacité du GPS de se localiser précisément dans les espaces ouverts et utiliser le Wi-Fi ou GSM en milieu urbain dense. Le GPS a pu être combiné dans le passé à titre expérimental avec un PDA, un ordinateur portable ou un téléphone mobile (Murakami and Wagner, 1999) (Doherty et al., 2001). De même, le GSM peut être utilisé comme dispositif unique, mais peut également être intégré à un PDA, ou à d'autres appareils de communication, ou même être combiné avec un récepteur GPS (Hato, 2006), (Itsuno and Hato, 2006). Une autre méthode hybride est la combinaison de GPS, Wi-fi et GSM : on a recensé les données brutes des stations de Wi-fi, des satellites et des stations de téléphone mobile avec les algorithmes avancés de localisation. Un téléphone portable équipé de ce logiciel peut être localisé dans tous les endroits.

En générale, la combinaison du GPS avec les autres nouvelles technologies est prometteuse. Le développement de nouvelle technologie et la baisse de prix du récepteur GPS ouvrent des

perspectives pour une enquête de mobilité à grande échelle. Utilisation du GPS dans les enquêtes de mobilité La méthode d'enquête sur les déplacements par interview face-à-face avec papier et crayon a été remplacée progressivement par des méthodes assistées par ordinateur telles que le CAPI, le CATI ou le CAWI, mais de nouveaux moyens pour recueillir des données sur la mobilité de manière plus précise et efficace sont encore recherchés. Le Global Positioning System peut contribuer considérablement à améliorer les données temporelles et spatiales sur les déplacements individuels en capturant exactement leurs heures de départ et d'arrivée, leurs origines et destinations, en fournissant également des indications précieuses sur le choix d'itinéraire. En s'appuyant sur ces données, la durée et la distance parcourue peuvent être facilement déduites.

2.2 Les différentes sources d'erreurs avec les nouvelles technologies

Les nouvelles technologies présentent de nombreux avantages pour les enquêtes sur les déplacements concernant principalement l'estimation de la durée, de la distance de déplacement ainsi que pour l'enregistrement des petits déplacements souvent oubliés par les enquêtés. Mais elles présentent à l'opposé des inconvénients pour les aspects qui suivent :

1. L'acceptabilité de ces enquêtes, qui peut entraîner des biais de sélection
2. Le mode de transport : les récepteurs GPS ne donnent aucune information directe sur le mode de déplacement ;
3. Le motif : on ne sait pas pour quel motif le déplacement est réalisé ;
4. Le nombre de personnes participant au même déplacement : les informations fournies par le récepteur GPS ne nous donnent aucune information sur la personne qui l'a porté, ni sur le nombre de personnes dans le même mode de transport. On ne peut pas non plus distinguer le conducteur et le passager du véhicule ;
5. Les déplacements manquants : il y a des jours sans aucune trace GPS, les causes sont multiples (la personne n'est pas sortie ; la personne a oublié le GPS ; le GPS ne fonctionnait pas, etc.) ;
6. La qualité des données : les données GPS ne sont pas toujours très précises. Il s'agit de la durée pendant laquelle les signaux sont perdus. On peut perdre le signal, par exemple quand la personne est dans le métro ou dans des quartiers très denses ou lorsque le récepteur est placé dans de mauvaises conditions (au fond d'un sac, loin d'une vitre dans les transports collectifs, etc.).

Il y a un grand nombre de recherches sur l'amélioration des données GPS pour mieux comprendre la mobilité. Deux méthodes parmi les plus utilisées sont :

- la combinaison du suivi GPS avec d'autres techniques pour obtenir des informations supplémentaires comme l'utilisation d'Internet ou du téléphone (Bohte and Maat, 2009) ; (Auld et al., 2009) ; (Doherty et al., 2006) et,
- la construction d'algorithmes pour identifier les modes et motifs de déplacements (Schuessler and Axhausen, 2009) ; (Yuan, 2010) ; (Forrest and Pearson, 2005).

2.3 Analyse de l'utilisation des nouvelles technologies dans les enquêtes de déplacement

Les nouvelles technologies utilisées dans l'enquête de déplacement sont diverses et chaque méthode a ses avantages et inconvénients. Le GPS peut permettre d'avoir une précision de localisation inférieure à 10 m mais il ne fonctionne pas très bien dans certains lieux comme dans un tunnel, dans le métro ou à l'intérieur des bâtiments. Le GSM peut fonctionner dans certains endroits où le GPS n'est pas en mesure de fournir une position mais sa précision dépend de la densité de réseau BTS. Le Wi-Fi peut fonctionner dans le milieu urbain dense et sa précision peut atteindre 13m-20m mais comme le GSM, sa précision dépend de la densité de station de base. Pour profiter des avantages des nouvelles technologies, nous pouvons utiliser une méthode hybride en combinant par exemple Wi-Fi et GSM ou GPS. Dans le cadre de cette thèse, nous nous sommes intéressés à la collecte de données par GPS.

En France, l'analyse de la mobilité par suivi GPS a pu être réalisée dans plusieurs enquêtes expérimentales :

- L'enquête sur la mobilité individuelle à Lille lancée par AFFIMETRIE, mise en œuvre par I.S.L. (Institut Sondage Lavielle) ;
- Le volet GPS dans l'enquête nationale transport et déplacements (ENTD 2007-2008) ;
- L'enquête mobilité des salariés localisés au sein du PDIE de Porte des Alpes réalisée par GPS en 2010 ;
- La recherche "Enquête Globale Transport en Ile-de-France" par association d'un GPS, d'un SIG et d'un système expert ;
- Les projets de transposition des dispositifs GPS à des pays du Sud (au Mexique dans le cadre de la coopération ECOS-NORD, au Vietnam).

2.3.1 Utilisation en complément d'une enquête ménages transports

Tel que mentionné dans la partie précédente, de nombreuses études sur l'application du GPS dans les enquêtes sur les déplacements des personnes ont été menées. Ce type d'enquête a dans

certain cas été appliqué avec un suivi GPS de type passif : on donne à l'enquêté un matériel sans aucune interface graphique, il s'agit en quelque sorte d'une « boîte noire » qui se contente d'enregistrer les lieux visités et l'heure de la visite. D'autres projets ont associé un enregistrement GPS à l'utilisation d'un PDA pour compléter le recueil. Par exemple, le projet réalisé en 1996 par FHWA Lexington (Wagner, 1997), ou encore l'étude menée au Pays-Bas en 1999 (Draijer et al., 2000). Ensuite, certains ont cherché à améliorer les données par combinaison du GPS et d'un petit questionnaire après l'enquête GPS.

L'objectif de recherche était par exemple de déterminer le taux de déplacements, qui est sous-déclarés dans une enquête ménages déplacements par téléphone (Wolf et al., 2004) ou bien de déterminer les facteurs influençant la sous-déclaration des déplacements (Zmud and Wolf, 2003) ; (Armoogum, 2000) ; (Bricka and Bhat, 2006). On peut noter aussi l'enquête réalisée à Austin, Texas en 1998 pour compléter les données de l'enquête traditionnelle (Casas and Arce, 1999) ou encore l'enquête en Californie pour l'étude des comportements des camions dans les milieux urbains et ruraux afin de calibrer la modélisation de la congestion (Wagner et al., 1998). Le suivi par GPS passif a été utilisé également dans une enquête ménages déplacements à Cincinnati aux Etats-Unis pour améliorer le protocole d'enquête.

Le récepteur GPS portable a été utilisé également dans deux projets de recherche en Australie. Dans la première, les volontaires ont été invités à porter le GPS pendant quelques jours pour valider les résultats de l'enquête par interview à domicile (Stopher et al., 2007). Dans ce cas, on a remis un GPS aux ménages dont au moins un membre a pris régulièrement les transports en commun. La deuxième utilisation de ce dispositif a pour but l'évaluation d'un programme pilote sur les changements de comportement de mobilité des ménages qui venaient de déménager (Ampt et al., 2006). Pour ces deux projets, les appareils étaient donnés aux personnes qui prenaient régulièrement les transports en commun.

2.3.2 Utilisation pour une étude de faisabilité du remplacement de l'enquête nationale de déplacement

Au-delà de l'utilisation du GPS pour la correction des données d'une enquête classique, il existe plusieurs recherches sur le développement du GPS comme instrument principal d'enquête sur les déplacements. Nous pouvons citer notamment une première enquête basée sur un suivi GPS passif pour 30 volontaires à Atlanta, Georgia (Wolf et al., 2001). Le GPS passif a également été utilisé dans le suivi de véhicules pour une étude de la tarification routière à Copenhague en 2001 et 2002 (Schönfelder et al., 2007). Dans cette étude, environ 400 voitures étaient équipées de GPS pendant une période de 10 à 12 semaines. Il y a également eu une enquête de déplacements par GPS conduite dans la ville de Québec, au Canada qui recourrait à un questionnaire électronique pour

PDA – CHASE (Computerized Household Activity Scheduling Elicitor) (Doherty et al., 2001). Cet appareil permet d'enregistrer les signaux satellitaires et il a été utilisé aussi dans l'enquête ménage en Flandre (Belgique) en 2004 (Kochan et al., 2006).

En 2002, le ministère des Transports a financé une étude pilote pour évaluer l'utilisation du GPS dans l'enquête Transports à Londres (London Transport Survey – LATS), dans laquelle 154 volontaires ont été équipés d'un appareil portable, le GeoLogger, par Geostats pendant trois jours. Au Japon, une recherche sur le GPS intégré aux téléphones mobiles pour l'enquête sur les déplacements a été menée par (Hato, 2006). Une étude de (Ohmori et al., 2006) sur les déplacements de personnes a également été réalisée pour comparer la collecte par téléphone portable équipé de GPS et par une méthode classique. Un des problèmes posés par cette méthode était la faible autonomie de la batterie, de l'ordre de 6h. D'autre part, il y avait une dépense supplémentaire associée à la transmission des données GPS pour leur exploitation.

En France, un sous-échantillon GPS a été intégré dans l'ENTD 2007-2008 (Marchal et al., 2008) en utilisant un récepteur portable avec la participation d'environ de 957 volontaires. Son objectif était de savoir si l'on pouvait conduire une telle enquête au niveau national. Cette enquête sera décrite dans le chapitre suivant.

2.4 Analyse du pilotage d'une enquête par GPS en Angleterre

Dans l'évaluation de précision de l'Enquête Transport de Ménage de Sydney 2003-2004 avec GPS, en comparant les résultats statistiques de données collectées par carnet de déplacements en papier et par GPS, Stopher a montré que le GPS peut être utilisé avec succès pour compléter les enquêtes transport avec le carnet de déplacements. De plus, le résultat d'une précision élevée pour les imputations des modes et des motifs de déplacement est positif dans l'enquête de mobilité urbaine reposant essentiellement sur le GPS (avec validation sur Internet) aux États-Unis à Cincinnati (Ohio) en 2009 (Stopher et al. 2012). Toutefois, sur la base d'un test négatif, les Britanniques ont renoncé à utiliser des traceurs GPS pour leur National Travel Survey (Department of Transport of UK, 2012).

Comme il s'agit de l'expérience la plus proche du volet GPS dans l'ENTD, nous allons la décrire en détail pour mieux comprendre la décision des britanniques. En novembre 2010, le National Centre for Social Research a fait appel à l'université de technologie d'Eindhoven pour traiter les données collectées lors d'une enquête-pilote visant à tester le remplacement du carnet de trajet papier de l'Enquête Nationale Transport de Grande-Bretagne (NTS) par un suivi GPS. Les traces GPS ont été collectées pour 874 enquêtés de 12 ans ou plus pendant une période de 7 jours via un récepteur GPS équipé d'un accéléromètre. Le GPS fournit à chaque seconde les informations de base sur les coordonnées (avec leur précision), la date et l'heure ; mais il enregistre ces données seulement si la

position horizontale est connue à moins de 10 mètres près, et ne fournit pas la vitesse instantanée. De plus, l'accéléromètre fournit des informations sur l'évolution de la vitesse sur trois axes, ce qui permet de préciser si l'appareil est en mouvement ou non, ainsi que son état (éteint, déchargé,). Les données de l'accéléromètre ont été principalement utilisées pour l'identification des différents types d'activités physiques (c.à.d., marche, course, repos, etc.). En plus des données enregistrées par cet appareil, le post-traitement est enrichi par les informations complémentaires fournies lors de la remise du GPS au sujet des activités les plus fréquentes, de la propriété ou d'un accès régulier à des voitures, motos et vélos,

Pour imputer le mode de déplacement, le post-traitement « Annotator » de données GPS et d'accélération affecte à chaque mode une probabilité. Deux méthodes peuvent être utilisées pour l'imputation : affecter à chaque déplacement un seul mode (celui qui a la plus forte probabilité) ou garder les probabilités conditionnelles de chaque mode.

Dans l'expérience britannique, les deux protocoles (GPS et carnet de trajet) ont été appliqués pendant la même période (février-mars 2011), mais à des individus différents : 1726 ont rempli le carnet de trajet classique sur 7 jours, et 874 ont porté le GPS, avec un taux d'acceptation nettement supérieur à celui observé en France, mais un peu inférieur à celui de la version papier.

Cette comparaison a posteriori a fait apparaître les différences qui suivent :

- beaucoup moins de déplacements d'après les données GPS (645 déplacements/an/personne contre 934 dans le carnet de trajet), mais 1,1 trajet par déplacement dans les deux cas ;
- un temps de parcours beaucoup plus long (544 heures/personne/an par rapport aux 328 heures extrapolées sur la base du carnet de trajet ;
- une distance parcourue également beaucoup plus longue (15429 miles par personne et par an, contre 6652 miles estimés à partir du carnet.

Le traitement des données GPS a permis d'identifier beaucoup moins de déplacements à pied que ceux qui figurent dans le carnet de déplacements (10% contre 28%). En revanche, il estime sept fois plus de déplacements en train que le carnet de déplacements (14 % contre 2%). Le nombre de déplacements par jour est réparti uniformément par rapport au jour de semaine pour les données collectées dans le carnet ; il en est de même pour les 4 premiers jours de port du GPS (16-18%), mais leur proportion tombe à 14 % le 5ème jour, 11 % le 6ème et 7% le 7ème jour. Cela serait dû aux enquêtés ayant oublié de recharger ou de porter leur récepteur GPS. Enfin, le carnet de déplacements décrit bien les pics durant la matinée et l'après-midi correspondant aux heures de pointe (08h00-09h00 et 15h00-16h00) ; et il est très surprenant que les données GPS traitées ne montrent pas les mêmes fluctuations horaires (pic entre 14h00 et 15h00).

Si l'on se réfère à ces nombreuses et importantes différences entre les résultats de l'enquête pilote GPS et les données du carnet de déplacements pour une même période, on est tenté de conclure sur cette expérience, qui a été perturbée par des problèmes matériels (disponibilité des appareils) et logiciels, que les récepteurs GPS ne fournissent pas un outil de collecte de données acceptable pour remplacer l'Enquête Nationale Transport en Angleterre, en tout cas avec la méthodologie et le programme de traitement des données GPS « Annotator » utilisé.

CONCLUSION

Les nouvelles technologies pouvant être utilisées dans des enquêtes sur la mobilité sont diverses et chacune ou la combinaison de plusieurs présente à la fois des avantages et des inconvénients. Pour profiter des avantages des nouvelles technologies, deux méthodes parmi les plus utilisées sont : d'une part, la combinaison du suivi GPS avec d'autres technologies pour obtenir des informations complémentaires comme l'utilisation d'Internet ou du téléphone, de l'accéléromètre, etc ; et d'autre part, la construction d'algorithmes de post-traitement de données brutes GPS pour identifier les modes et motifs de déplacement.

Le GPS a été utilisé pour la collecte de données de mobilité pour la première fois en 1997 à Lexington Texas USA pour un complément de l'enquête de transport, et ensuite plusieurs recherches ont été menées, notamment aux USA, en Australie, en Angleterre pour différents objectifs comme évaluer ou compléter les données d'une enquête classique. Le GPS est également utilisé pour des tests de faisabilité du remplacement des enquêtes nationales de déplacement. Nous pouvons citer :

- Le volet GPS dans l'ENTD 2007-2008 en France ;
- L'enquête à grand échelle menée à Cincinnati en 2010 ;
- L'enquête GPS à Jérusalem en 2010 ;
- Le pilote d'une enquête par GPS en Grande Bretagne en 2011 ;
- L'enquête GPS à Cleveland en 2012.

Malgré plusieurs résultats de tests favorables dans l'utilisation du GPS pour des enquêtes sur la mobilité quotidienne, les britanniques, sur la base d'un test négatif, ont conclu que les récepteurs GPS ne fournissaient pas un outil de collecte de données alternatif acceptable pour leur Enquête Nationale Transport en Angleterre qui se serait appuyé sur leur méthodologie et leur post-traitement des données GPS. Nous avons mis en évidence au moins un problème de traitement des données GPS dans ce pilote.

3 Présentation de l'Enquête Nationale sur les Transports et les déplacements (ENTD) et du protocole de collecte par GPS utilisé en France

Tous les dix ans, l'Insee entreprend avec le Ministère en charge des transports une Enquête Nationale sur les Transports (ENT). Dans les années cinquante, le champ de l'enquête était restreint à l'automobile (enquête « Parc et utilisation de l'automobile », CREDOC-Renault, 1958-59). Depuis les années soixante, l'enquête s'emploie à décrire l'ensemble des déplacements des ménages :

- Enquête sur les dépenses de transport de 1966-67 ;
- Enquête sur les dépenses de transport de 1973-74 ;
- Enquête Transport de 1981-82 ;
- Enquête Transports et Communications de 1993-94 ;
- Enquête Nationale sur les Transports et les déplacements 2007-08.

L'Enquête Nationale Transport est la seule source de données fournissant une vision globale et cohérente de tous les modes et situations de transport de personnes (articulation mobilité quotidienne et longue distance) pour l'ensemble des résidents de France métropolitaine. C'est le seul outil qui permette d'observer les comportements des habitants de toutes les tailles d'agglomérations (notamment les plus petites et les zones rurales) et d'estimer des indicateurs de mobilité tout au long de l'année (Quang Nguyen, 2014). L'enquête est notamment fiable pour l'appréhension des distances parcourues (véhicules*kilomètres ou voyageurs*kilomètres) ce qui est important pour les problématiques environnementales (émission de polluants locaux ou globaux, consommation d'énergie). Elle permet des comparaisons dans le temps (avec les enquêtes précédentes) et dans l'espace (avec les enquêtes menées à l'échelon local ou dans d'autres pays) (rapport final de l'Action COST SHANTI, 2014).

Signalons aussi que cette enquête est utilisée pour des études sur une large gamme de sujets (concurrence entre modes de transport, spécialisation en fonction des motifs, déplacements régionaux dans l'aire d'attraction des grandes villes, consommations d'énergie et environnement, déplacements touristiques, etc.).

Le Service de l'observation et des statistiques (SOeS) du Commissariat général au développement durable (CGDD) a assuré la maîtrise d'ouvrage de l'opération 2007-08 et l'INSEE en était le maître d'œuvre. L'Institut National de Recherche sur les Transports et leur Sécurité (INRETS, maintenant IFSTTAR suite à la fusion avec le LCPC) est le coordinateur scientifique de cette enquête depuis l'enquête de 1981-82. Son Département Economie et Sociologie des Transports (DEST, maintenant

laboratoire « Dynamiques Economiques et Sociales des Transports » dans le Département AME de l'IFSTTAR) s'est fortement investi dans la préparation de l'ENTD 2007-2008.

Ces enquêtes sont financées à la fois par des organismes publics et des organismes privés (Agence de l'Environnement et de la Maîtrise de l'Energie, Direction Générale des Infrastructures de Transport et de la Mer, Direction Générale de l'Aviation Civile, EDF, SNCF, Réseau Ferré de France, Association des Sociétés Françaises d'Autoroutes et d'Ouvrages à Péage, Comité des Constructeurs Français d'Automobiles, association des compagnies d'assurances, etc.).

3.1 Présentation générale de l'ENTD 2007-2008

L'objectif de ces enquêtes est la connaissance des déplacements des ménages résidant en France et de leur usage des moyens de transport tant collectifs qu'individuels. Ces enquêtes sont les seules sur la mobilité réalisée à cette échelle et qui décrivent tous les déplacements, quels que soient le motif, la longueur, la durée, le mode de transport, la période de l'année ou le moment de la journée. Elles s'intéressent aussi aux possibilités d'accès aux transports collectifs et aux moyens de transports individuels dont disposent les ménages pour comprendre leurs comportements.

Les principaux thèmes abordés sont les déplacements, de courtes et de longues distances des ménages et de leurs membres, selon les différents modes de transports. L'enquête met également l'accent sur la connaissance du parc de véhicules (voitures, motos, vélos, quads), de leur usage et sur l'accessibilité des individus aux transports collectifs (dont les abonnements et les réductions tarifaires).

Après son examen par le comité du label du Conseil National de l'Information Statistique, le 18 septembre 2006, l'enquête nationale sur les transports et les déplacements a reçu le label d'intérêt général et de qualité statistique, ainsi que le visa rendant cette opération statistique obligatoire, à l'exception de l'opération complémentaire impliquant l'usage d'un GPS, qui ne s'adresse qu'à des volontaires. L'Enquête Nationale Transports et Déplacements 2007-08 est la première enquête nationale à utiliser le GPS dans le monde.

Elle fait suite à une série d'enquêtes sur le même sujet afin d'assurer la mesure des évolutions structurelles. Comme les précédentes ENT, l'enquête 2007-08 est organisée autour des trois grands thèmes suivants :

- Description des déplacements :
 - La mobilité régulière : les déplacements habituels domicile-travail, domicile-lieu d'étude ou de garde des enfants ;

- La mobilité quotidienne : les déplacements réalisés un jour de semaine (veille du passage de l'enquêteur) et au cours du dernier week-end ;
- La mobilité à longues distances : les déplacements réalisés à plus de 80 kilomètres du lieu de la résidence principale ;
- Connaissance du parc de véhicules et de leurs usages :
 - Description des véhicules dont dispose le ménage et historique du parc du ménage (sur 3 ans) ;
 - Description des usages d'un véhicule (à 2 ou 4 roues) pendant une semaine ;
 - Détention de permis de conduire, pratique de la conduite et accidents de la circulation ;
- Accessibilité aux transports collectifs (dont abonnements et réductions tarifaires)

3.1.1 Le déroulement de l'enquête et les thèmes abordés

La phase de collecte des données de l'ENTD 2007-2008 a été réalisée du 30 mars 2007 au 27 avril 2008 auprès des ménages ordinaires de France métropolitaine. Elle s'est déroulée en six vagues successives (d'environ huit semaines chacune) de façon à éliminer les effets de saisonnalité (voyage à plus de 100 kms plus nombreux en été, etc.) qui affectent la mobilité, tout au moins privée. Le ménage reçoit l'enquêteur au cours de deux visites espacées d'environ une semaine. L'enquête est une interview en face-à-face d'une durée d'environ 2 heures au total. La collecte a été assistée par ordinateur (CAPI). Les ménages sont interrogés dans leur résidence principale.

La première visite est centrée sur les facteurs explicatifs du comportement de mobilité du ménage et des individus qui le compose :

- "Equipement" : questions sur l'équipement du ménage en moyens de transport, caractéristiques et usages des véhicules à disposition ;
- "Motilité sociale et physiologique" : la description des caractéristiques sociodémographiques des individus qui composent le ménage ; les capacités (permis) ou incapacités (gêne) à se déplacer, par exemples : le permis de conduire et la pratique de la conduite, accidents de la circulation, les abonnements et réductions dans les transports collectifs ;
- "Mobilité régulière" : les lieux habituellement fréquentés (domicile, travail, lieu d'étude, lieu de garde d'enfant, etc.) ;
- "Logement" : situation, parking, environnement du logement.

Le questionnement sur la "mobilité régulière" et sur la "motilité" s'adresse à tous les individus du ménage, avec quelques restrictions. Tout d'abord, les questions sur la mobilité régulière ne concernent que les personnes qui travaillent ou qui étudient, et les enfants en bas âge gardés à

l'extérieur du domicile. Ensuite, dans le but d'éviter des entretiens longs dans le cas de familles nombreuses, il a été prévu de n'interroger qu'un seul enfant écolier ou étudiant, sauf s'il travaille, qui est tiré au sort. Il est préférable que chaque membre du ménage puisse répondre aux questions qui le concernent mais le proxy est autorisé, c'est-à-dire qu'un membre du ménage peut répondre à la place d'un autre membre absent lors de la visite de l'enquêteur.

Le questionnement sur l'"Equipement", consacré aux véhicules à disposition du ménage est en deux parties. Il commence par un recensement de tous les véhicules, sauf les vélos d'enfant et se poursuit par leur description, qui est limitée à :

- Une voiture, véhicule utilitaire léger (VUL) ou camping-car (CC), ou parfois deux si le véhicule choisi pour recevoir le carnet est de ce type mais n'est pas le véhicule tiré au sort pour être décrit ;
- Un seul vélo tiré au sort en privilégiant ceux utilisés au cours du dernier mois ou exceptionnellement deux si le véhicule choisi pour recevoir le carnet est un autre vélo ;
- Les deux-roues à moteur, voiturettes, quads, tricycles à moteur, tracteurs agricoles utilisés depuis 12 mois (sans restriction).

Compte tenu des réponses au questionnement ci-dessus, un tirage au sort est effectué pour désigner, d'une part, l'individu "Kish" qui répondra aux questionnaires de la seconde visite et, d'autre part, un véhicule "Kish" dont l'utilisation devra être décrite pendant 7 jours dans un carnet auto-administré.

L'individu "Kish" est tiré à probabilité inégale, afin d'interroger de préférence la personne qui fait le plus de voyages à longue distance. A l'intérieur du ménage, les personnes éligibles pour être "Kish" sont celles qui ont 6 ans ou plus et qui sont joignables avant la fin de la vague. Le tirage privilégie la ou les personnes qui ont réalisé le plus grand nombre de voyages au cours des 3 mois précédents.

Le véhicule "Kish" est tiré au sort parmi tous ceux possédés par le ménage. Ce tirage vise à privilégier d'abord les deux-roues à moteur pour lesquels les informations sont rares, puis les vélos, puis les autres véhicules à moteur, à la condition qu'ils aient été utilisés au cours du dernier mois. Les véhicules qui n'ont pas été utilisés au cours des quatre dernières semaines ont une probabilité plus faible d'être tirés.

A la fin du questionnaire CAPI de première visite, le ménage reçoit un carnet de déplacements en papier sur lequel tous les utilisateurs du véhicule "Kish" doivent noter tous les déplacements réalisés avec ce véhicule pendant sept jours consécutifs (heure de départ et d'arrivée, commune de destination, kilométrage au compteur, etc.). Parallèlement, l'individu tiré au sort, s'il est majeur et

volontaire, reçoit un récepteur GPS à utiliser lors de tous ses déplacements au cours de la période séparant les deux visites.

Entre les 2 visites, le carnet de déplacements est rempli en décrivant tous les déplacements réalisés avec le véhicule "Kish" pendant 7 jours consécutifs. Dans le sous-échantillon GPS, le volontaire se déplace comme d'habitude, mais ses déplacements sont enregistrés automatiquement et passivement par GPS. Les types de données fournies sont : Numéro d'enregistrement, Date, Heure, Latitude, Longitude, Altitude, Précision de la mesure calculée en fonction de la géométrie de la constellation de satellites PDOP (Précision du positionnement 3D), HDOP (Précision horizontale), VDOP (Précision verticale), Nombre de satellites, Vitesse et Direction.

La seconde visite nécessite au minimum la présence de l'individu « Kish » tiré au sort en fin de première visite. En deuxième visite, le proxy n'est pas autorisé, sauf dans le cas d'enfant de moins de 12 ans ou de personne inapte à répondre. Exemple : un enfant de 8 ans parti en colonies de vacances (à plus de 100 km de la résidence) alors que ses parents ne sont pas partis a une assez forte probabilité d'être sélectionné comme « Kish ».

Au cours de cette deuxième visite, l'enquêteur récupère le carnet de déplacements et/ou le GPS et lance dans ce cas une application de récupération automatisée des données GPS sur son poste. Il vérifie et vérifie le carnet de déplacements (s'il y a lieu), complète éventuellement des renseignements manquants en première visite. Puis, il réalise l'interview CAPI de la personne "Kish". Ensuite, il lance éventuellement l'application de questionnaire complémentaire CAPI-GPS ou dépose un questionnaire biographique.

La seconde visite commence par le questionnaire CAPI sur la description des comportements suivants :

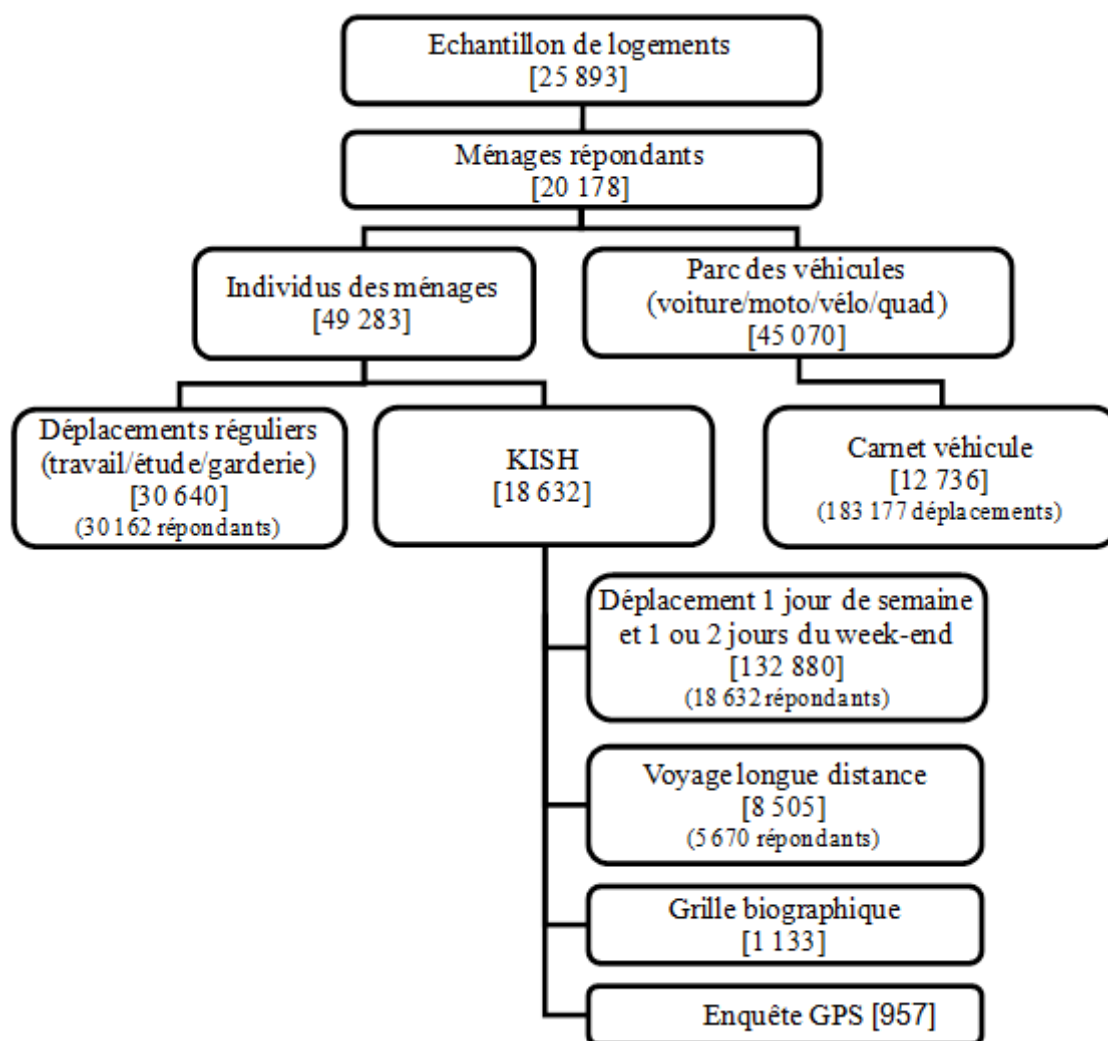
- "Mobilité quotidienne" : déplacements un jour de semaine et un jour de week-end. Il s'agit de deux journées précises de la semaine précédant la seconde visite : une journée de semaine et une journée de week-end (ou de trois journées dans les régions des Pays de la Loire et d'Ile de France où le samedi et le dimanche sont systématiquement enquêtés). Le jour de semaine à décrire est le jour le plus proche de la deuxième visite durant lequel le Kish est sorti de chez lui. C'est généralement la veille de la seconde visite. Seule une journée de week-end est décrite, sauf dans deux régions à extension : Ile de France et Pays de Loire. L'enquête Kish répond à des questions sur un jour où il a été mobile, le samedi ou le dimanche, avec une probabilité égale de tirage s'il n'est pas resté chez lui l'un de ces deux jours ;
- "Mobilité à longue distance" : voyages à plus de 100 kilomètres du lieu de résidence. La période d'observation est, au maximum, celle des trois mois (en fait 13 semaines) précédant la première visite. Mais elle peut être réduite à deux mois (en fait 9 semaines) s'il y a plus de dix voyages à

décrire. Cette description des voyages fait appel à une mémoire des enquêtés plus lointaine qui peut être source d'oubli ou d'erreur ;

- "Attitude et sécurité routière" : comportement sur la route.

Au final, l'ENTD 2007-2008 a enregistré les informations de 20178 ménages lors de la première visite. Sur ceux-là, 18632 individus Kish ont été enquêtés lors de la seconde visite. Un schéma récapitulant l'architecture de l'enquête est présenté dans la Figure 3.1. Avant d'être diffusées, les données collectées par l'enquête ont subi une étape d'anonymisation. Les informations nominatives ont tout d'abord été supprimées, mais également une partie des informations permettant la localisation géographique fine des ménages et des déplacements individuels. Pratiquement, l'identification des communes à l'origine et à la destination des déplacements a été rendue impossible, alors que les unités urbaines, les bassins de vie, les cantons, et les arrondissements ne sont renseignés que s'ils comptabilisent plus de 20000 habitants.

Figure 3.1 : Architecture de l'ENTD 2007-2008



Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

3.1.2 Organisation de la composante GPS dans l'ENTD 2007-2008

Les enquêtes de mobilité de l'avenir utiliseront probablement à grande échelle des récepteurs GPS miniatures comme ceux qui ont été utilisés dans le cadre de cette opération. Ce sont des appareils qui enregistrent les positions grâce à la technologie GPS (Global Positioning System). Grâce à ces appareils, les informations sur les heures et les distances des déplacements sont connues beaucoup plus précisément que par le protocole déclaratif habituel. Cependant, ce matériel ne permet pas de connaître le motif du déplacement, le moyen de transport utilisé, ni le nombre de personne participant au déplacement. Il ne peut pas non plus distinguer le conducteur et les passagers d'un véhicule. Un questionnaire complémentaire allégé est donc encore nécessaire.

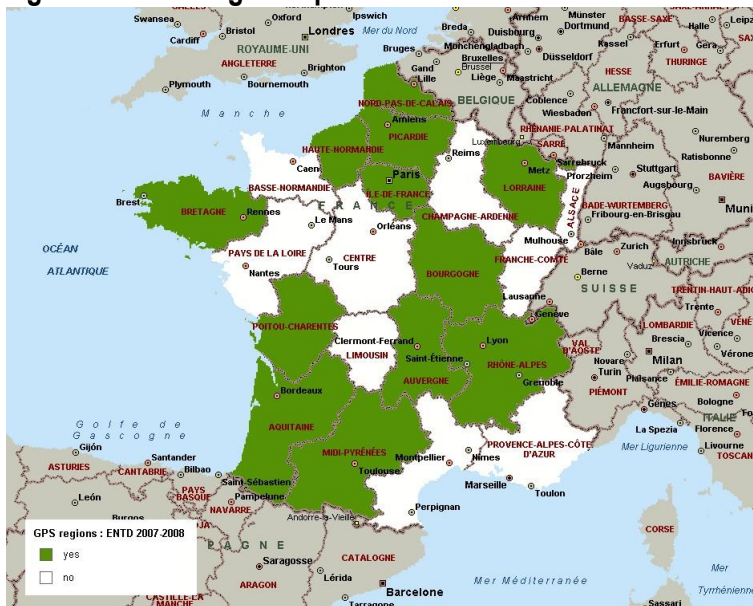
Ce changement de technique d'enquête demande à être testé sur plusieurs plans :

- Le matériel ;
- La façon dont les enquêtés le reçoivent et l'utilisent ;
- Les méthodes pour traiter les informations ainsi recueillies ;
- Les résultats statistiques, à comparer à ceux de la méthode classique.

Ce test à plusieurs niveaux est l'objectif de "l'opération GPS". Comme l'enjeu est important, l'opération requiert la participation de nombreux enquêtés mais elle reste un test. C'est pourquoi cette opération expérimentale, qui pose des problèmes de confidentialité, n'a pas le caractère obligatoire de l'ENTD et s'adresse uniquement à des personnes majeures.

Un échantillon de 957 volontaires a été équipé de récepteurs GPS pour enregistrer leurs déplacements entre les visites de l'enquêteur dans la douzaine de régions très variées qui ont accepté l'expérience et sont présentées dans la Figure 3.2 (Ile de France, Picardie, Haute Normandie, Bourgogne, Nord Pas de Calais, Lorraine, Bretagne, Poitou Charentes, Aquitaine, Midi Pyrénées, Rhône Alpes, Auvergne). La participation a été moins élevée que prévu (nous comptons sur 1500 volontaires) mais le défi technique a pu être surmonté.

Figure 3.2 : Les régions qui ont réalisé l'ENTD 2007-2008 avec GPS



Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Protocole de proposition du GPS (même s'il n'y a pas de GPS disponible) en première visite :

- Essayer d'en parler au Kish s'il est majeur, sinon continuer avec la personne de référence ;
- Poser la question à l'enquêté : le GPS est-il disponible ? Le questionnement s'arrête en cas d'indisponibilité. Un enquêteur qui ne participe pas à l'opération GPS répond systématiquement « non » ;
- Expliquer le principe de l'expérimentation et demander à la personne présente si elle est intéressée pour un test. Si elle n'est pas intéressée, le module s'arrête.

En fait, la question s'adresse au Kish ou à la personne de référence : « *Accepteriez-vous de participer à une enquête transport expérimentale avec GPS menée par l'INSEE, et d'emmener un GPS dans tous vos déplacements pendant une semaine ?* ». Il y a quatre réponses possibles :

1. Oui, sans condition,
2. Oui, à condition de pouvoir éteindre le GPS quand j'en ai envie,
3. Oui, à d'autres conditions,
4. Non.

- Expliquer les consignes d'utilisation et de recharge de l'appareil ; et reposer la question « *êtes-vous volontaire pour l'enquête GPS ?* », en cas de refus, garder le GPS, sinon, le confier et donner une notice d'utilisation qui indique également les droits relatifs à la loi sur l'informatique et les libertés.

Les trois premières étapes du protocole sont générales à toute l'enquête, que l'enquêté ou sa région soient ou non concernés par l'opération GPS, qu'il y ait ou non un GPS à distribuer. La

première partie du questionnaire nous permet le calcul du taux d'acceptabilité de l'enquête GPS. En revanche, la disponibilité du récepteur GPS et la réponse fournie dans la deuxième partie du questionnaire permet de mesurer le taux d'acceptation de l'enquête GPS.

- Prendre rendez-vous pour réaliser la deuxième visite (au minimum huit jours après la première visite, en particulier si un carnet de déplacements est à remplir) en s'assurant de la présence de la personne « Kish ».

Protocole de remise à un enquêté volontaire en première visite :

- Montrer les éléments du « pack GPS » : sac, récepteur GPS, câble allume-cigare et adaptateur secteur ;
- Noter le numéro individuel à 4 chiffres du GPS sur le dossier ménage (dans le case « remarques ») et sur la notice laissée à l'enquêté, qui est une page recto-verso ;
- Présenter les conseils d'utilisation de la notice :
 - Recharger le GPS chaque soir ;
 - Placer le GPS dans une poche poitrine ou dans le sac fourni, porté en bandoulière ;
 - Éviter de le placer au fond d'un sac ou dans le coffre d'une voiture ;
 - Essayer de se placer près d'une fenêtre dans les transports en commun ;
 - Ne pas arrêter de porter le GPS si on oublie une fois de l'allumer ou de le recharger.

Protocole pour la deuxième visite :

Il y a donc au maximum 5 modules à réaliser :

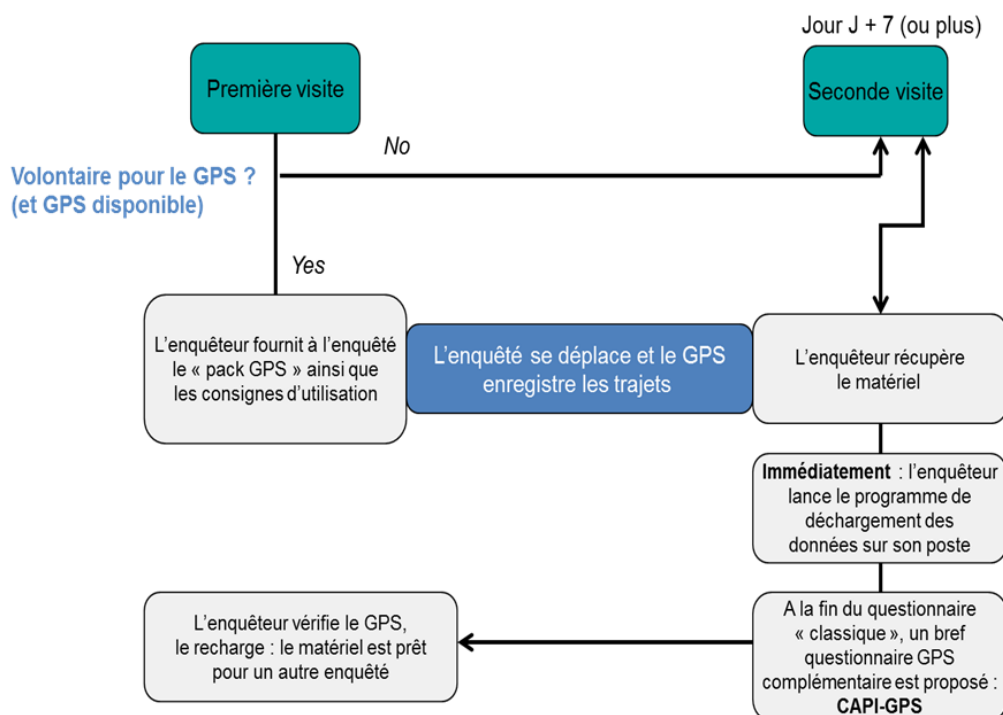
- Déchargement des données GPS : l'enquêteur récupère le GPS. Si celui-ci n'est pas rendu, l'enquêteur note la perte sur le dossier ménage, à côté du numéro de l'appareil. La perte, le vol ou la destruction du GPS sont des impondérables qui n'auront aucune conséquence ni pour l'enquêteur, ni pour l'enquêté. Il lance l'application de déchargement GPS (voir "L'élaboration du logiciel de déchargement des traces GPS" ci-dessous). L'enquêteur peut ensuite lancer l'application du questionnaire CAPI-GPS. L'application de déchargement gère la communication sans fil avec le GPS et travaille en parallèle au CAPI du questionnaire principal de l'enquête.
- Récupération et vérification du carnet de déplacements ;
- Questionnement CAPI Interrogation du KISH (comme pour les enquêtés "non-GPS") : Description des déplacements d'un jour de semaine et d'un jour de week-end (exception de deux jours de week-end dans les régions des Pays de la Loire et d'Ile de France), ainsi que celle des déplacements à longue distance ;

- Questionnement spécifique CAPI-GPS lié aux données enregistrées par le GPS :
Si le téléchargement du GPS réussit, les "traces GPS" sont copiées sur l'ordinateur de l'enquêteur, effacées du GPS et traitées par l'application. En sortie du CAPI du questionnaire principal, suivre les instructions à l'écran pour remplir le bref questionnaire CAPI-GPS complémentaire. Celui-ci demande la raison d'un jour sans trace (s'il y en a). Ensuite apparaît automatiquement un questionnement spécifique sur un ou deux jours parmi les jours entre les deux visites, distincts de ceux précédemment décrits pour la mobilité quotidienne. Le traitement immédiat des traces GPS identifie des « phases de déplacement » qui ne correspondent pas toujours aux déplacements tels qu'ils sont définis dans l'enquête, soit parce que l'enregistrement du GPS est incomplet, soit parce que deux déplacements se sont enchaînés très rapidement, soit au contraire parce que l'attente à une correspondance a été très longue.

L'intérêt de ce questionnement complémentaire est de savoir si ces phases sont bien délimitées et de connaître les caractéristiques que le GPS ne peut pas enregistrer. Après un écran de questions générales, l'enquêteur demande d'abord si le relevé du GPS correspond bien à un déplacement réalisé et non pas à une partie seulement ou à plusieurs déplacements agrégés. En cas de bonne correspondance, il demande le motif, le moyen de transport utilisé et le nombre de personnes accompagnant l'enquêté. Les questions sont répétées pour chaque déplacement identifié à partir des traces GPS d'un seul jour parmi ceux du relevé.

- Enfin, dépôt du questionnaire biographique.

Figure 3.3 : Protocole du volet GPS de l'ENTD 2007-2008



Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

3.2 Elaboration du logiciel de déchargement des traces GPS

- Couplage Bluetooth des GPS et des PC-enquêteur : le programme **MenuGPS** est le point d'entrée pour l'ensemble des applications GPS. Il permet la vérification de l'insertion de la clé USB-Bluetooth, le couplage du GPS avec le portable de l'enquêteur, puis le lancement du programme de déchargement des traces EntGPS.
- Transfert/Sauvegarde des données GPS enregistrées : le programme **EntGPS** permet de télécharger les données enregistrées par le GPS sur le portable de l'enquêteur. Cette tâche, qui peut prendre plusieurs minutes en fonction du volume de données enregistrées entre les deux visites, s'effectue en arrière-plan, et ne perturbe pas le déroulement du questionnaire "classique" CAPI de la seconde visite.
- Elaboration du programme d'interrogation GPS complémentaire (CAPI-GPS) : lorsque le questionnaire CAPI de la deuxième visite est terminé, le programme **ANATRACES** ouvre la fenêtre de Questionnement Complémentaire GPS. On y trouve des rubriques de configuration ou de test qui ne seront pas accessibles à l'enquêteur, mais qui étaient indispensables lors de la mise au point de l'outil :
 - Intérêt et aperçu du programme CAPI-GPS
 - Données externes utilisées dans le traitement
 - Analyse automatisée des traces brutes
 - Génération automatisée d'un questionnaire complémentaire

3.3 Analyse des enregistrements GPS en fin de deuxième visite

Une analyse des enregistrements, ou « traces », du GPS est réalisée pour permettre l'administration du questionnaire supplémentaire GPS. Une exploitation plus complète sera menée ensuite par les chercheurs (c'est un objectif important de cette thèse), mais ils auront le plus grand besoin des réponses à ce questionnaire supplémentaire.

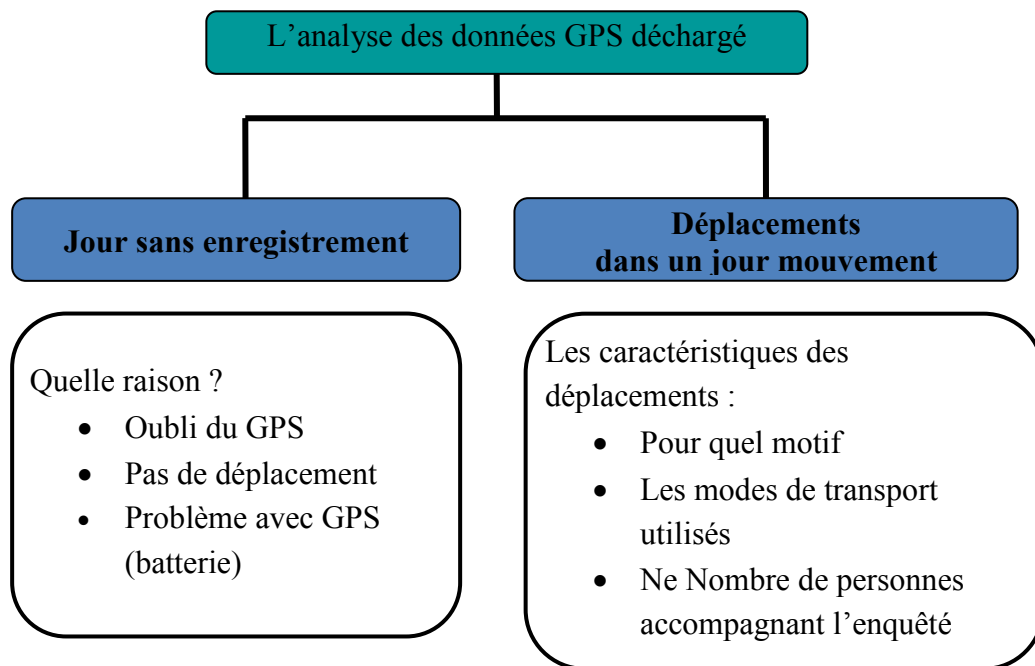
Ce questionnaire a deux parties :

- Questions générales sur la facilité d'utilisation du GPS et diagnostic éventuel sur une journée n'ayant donné aucun enregistrement. L'enquêté va répondre la question « *Avez-vous trouvé contraignant pour vous de porter le GPS avec vous dans tous vos déplacements ?* ». Il y a 3 choix de réponse : Soit OUI ; Soit NON ; Soit Ne sais pas. Si la réponse est OUI, l'enquêté va préciser pour quelles raisons.

- Retour sur une journée non vide :
 - Le programme ANATRACE reconstitue des déplacements (heures, lieux, distance, etc.). Pour chaque déplacement de la journée, il demande si cette description est complète (le questionnaire est itératif avec un maximum de 10 déplacements).
 - ✓ Si la reconstitution est bonne, quels étaient le motif et le ou les moyens utilisés pour chaque déplacement ? combien de personnes du ménage voyageaient ensemble pendant ce déplacement ? et combien de personnes extérieures au ménage ?
 - ✓ Si non, les éléments manquants sont-ils au début ou à la fin de la journée ?
 - Enfin, est-ce que des déplacements ont été oubliés dans cette journée ? Précisément, l'enquêté va répondre la question « Après ce que nous avons vu dans cette journée, reste-t-il des déplacements qui auraient dû être enregistrés ? ». La réponse soit OUI ; soit NON.

Dès que les données sont transférées sur le poste de collecte de l'enquêteur, le récepteur est vidé de toutes données individuelles. A moins d'un problème technique diagnostiqué par le programme d'analyse des traces GPS, qui demande à renvoyer l'appareil à la Direction Régionale de l'Insee concernée, le GPS pourra être proposé à un nouvel enquêté dès que la batterie aura été rechargée.

Figure 3.4: Schéma de la conception du CAPI-GPS dans le volet GPS de l'ENTD 2007-2008



Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

3.4 Les récepteurs GPS utilisés dans l'ENTD 2007-2008

3.4.1 Choix d'un suivi de personnes et non pas de véhicules

Lors des premières utilisations du GPS pour le recueil de données sur la mobilité, la majorité des expériences se sont concentrées sur le suivi de véhicules, notamment pour des raisons de facilité d'alimentation électrique des récepteurs GPS sur de longues périodes, mais également parce que la précision des données GPS n'était pas suffisante à cette époque. Leur taille et leur poids étaient également un frein pour les faire accepter par des enquêtés sur une période relativement longue. Par ailleurs, le traitement a posteriori de données « véhicules », s'il n'est pas immédiat, s'avère beaucoup plus aisé que celui de données « piétons », lesquels se déplacent assez souvent en dehors des réseaux routiers généralement identifiés dans les bases de données cartographiques.

Néanmoins, dès les années 1990, quelques expériences ont été mises en place pour un suivi de personnes (Kalfs et Saris, 1997). Plus récemment, des expériences de ce type de collecte aux Etats-Unis, au Canada, au Japon, en Australie ou en Europe dans le cadre d'enquêtes effectuées sur des échantillons relativement petits, généralement à un niveau local/ régional ont été réussies. Ces dernières années, une accélération très positive des améliorations techniques apportées aux GPS, avec des récepteurs plus petits, une meilleure précision, une plus grande capacité de stockage, ainsi qu'une importante baisse de prix nous ont permis d'envisager l'utilisation du GPS pour le suivi de personnes à grande échelle dans le cadre d'une enquête de type « Enquête Nationale Transport ».

Choix d'un suivi passif

Par « suivi passif », nous entendons un système basé sur un GPS « boîte noire », sans aucune interface graphique ni cartographie, et avec le strict minimum pour fonctionner en respectant les contraintes de confidentialité : un seul bouton arrêt-marche, qui permet à l'enquêté d'interdire l'enregistrement de certains déplacements s'il le juge souhaitable (de la même manière que dans le questionnaire classique, il peut très bien ne pas renseigner certains déplacements de son choix).

3.4.2 Sélection des récepteurs-enregistreurs GPS, caractéristiques et paramétrage

Pour l'ENTD 2007-2008 deux modèles « Blue GPS » () de Royaltek ont été choisis : initialement le RBT-3000, puis le RBT-1100. Il s'agit d'un matériel « grand public », de la taille d'une souris, qui est habituellement utilisé avec un PC portable ou un PDA, avec les logiciels de « navigation » associés. Le Blue GPS est un GPS Bluetooth : la connexion avec un ordinateur ne peut être assurée qu'avec

une liaison sans fil. Il s'agit véritablement d'une « boîte noire », il ne comporte qu'un bouton marche-arrêt et présentait donc des caractéristiques intéressantes par rapport à nos besoins. Il coûtait au moment de son achat pour test 220 euros. Deux ans plus tard, il était vendu 80 euros. Lors de son achat, c'était le seul matériel « grand public » connu à proposer la fonction « datalogger » : c.à.d. qu'il est autonome, il enregistre les « traces » même s'il n'est pas connecté à un ordinateur.

Figure 3.5: Le GPS RBT-3000 de Royaltek



Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

La capacité d'enregistrement du récepteur Blue GPS RBT-3000 est de 30000 points. En enregistrant toutes les secondes, cela correspond donc à peu près à son autonomie : il faut le recharger toutes les 8 – 10 heures. Avec la limite de capacité d'enregistrement de 8-10 heures, le récepteur GPS RBT-3000 posait un problème d'autonomie. En cas d'oubli par l'enquêteur d'éteindre le récepteur durant une période sans déplacement (par exemple au bureau), il risquait d'avoir une capacité d'enregistrement inférieure à une journée complète. Bien sûr, pour un enquêteur respectant strictement les consignes d'utilisation (assez contraignantes), et présentant des caractéristiques de mobilité « habituelles », le problème n'allait pas se poser. Mais il a été décidé d'équiper ce modèle d'une modification qui coupe le fonctionnement en cas d'immobilité prolongée.

On a ajouté un détecteur de mouvement pour économiser la batterie et permettre en toute circonstance d'obtenir plus qu'une journée complète de fonctionnement. L'appareil peut se trouver dans quatre états différents :

- Actifs éveillé, en sortie du mode repos, dès qu'un mouvement est détecté ;
- Au repos : après 10 minutes sans mouvement ;
- Soute : quand il est mobile depuis plus de 2 heures, il alterne 20 minutes de veille et 10 minutes d'activité ;
- Parking : quand il est au repos depuis plus d'une heure, il s'éveille 10 minutes par heure ;

Un premier lot de GPS de ce type a été acheté pour les besoins des tests, et les prototypes pour en étendre l'autonomie en respectant le schéma de fonctionnement décrit précédemment ont été mis au point. Au moment des demandes de devis pour l'achat de ces GPS RBT-3000 en plus grand nombre, pour l'enquête en vraie grandeur, nous avons appris que ces GPS étaient désormais retirés

de la vente : ils étaient en fait remplacés par les « Blue GPS RBT-1100 ». Ces derniers étaient plus petits, plus robustes au niveau de l'interrupteur, et surtout avec une batterie supplémentaire, ils présentaient une autonomie bien meilleure : de l'ordre de 17 heures, pour la même capacité de stockage que le RBT-3000.

Figure 3.6 : Blue GPS RBT-1100 de Royaltek



Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Bien que très différents d'aspect extérieur, ces deux modèles présentent des caractéristiques relativement proches au niveau du fonctionnement, du transfert des données, des formats utilisés et de l'adaptation des programmes pour le chargement des données. Au niveau de la précision des données collectées, du temps d'acquisition et des possibilités de se localiser dans des conditions difficiles, le second modèle est beaucoup plus performant.

Le fonctionnement de ces appareils était conforme à nos exigences : simples de manipulation avec un seul bouton, trois voyants LEDs, déchargements des données par Bluetooth. Les enquêtés peuvent facilement allumer le récepteur lorsqu'ils se déplacent et recharger quand la batterie est faible grâce aux couleurs différentes des voyants (voir "Fonctionnement du matériel" dans l'Annexe 1 "Notice du GPS"). Mais la qualité des données GPS dépend aussi du comportement des enquêtés volontaires :

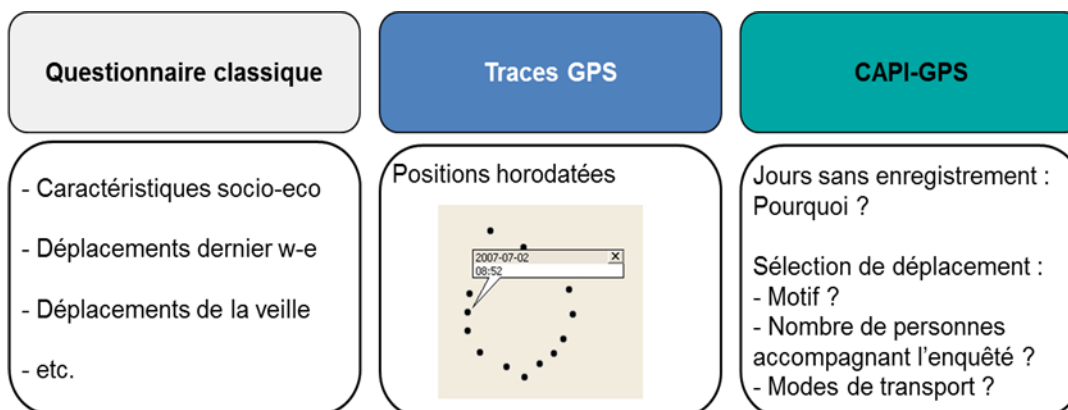
- S'ils oublient d'allumer l'appareil, les traces sont perdues. C'est aussi souvent le cas au démarrage en raison du temps mis par l'appareil pour repérer les satellites et obtenir une première localisation ;
- S'ils oublient d'éteindre l'appareil ou de recharger (il est conseillé de recharger la batterie chaque nuit), la mémoire risque d'être saturée, et ce sont les traces des derniers jours qui seront perdues.

A cause de la limite de capacité d'enregistrement, nous avons choisi la périodicité de 10 secondes afin de ne pas saturer la mémoire du GPS pour une collecte d'une semaine ou plus. Les

expériences étrangères sont souvent menées avec une fréquence d'une seconde, et on peut montrer que l'optimum se situe entre 2 et 4 secondes (Yuan, 2010).

3.5 Panorama des trois sources de données à confronter dans le volet GPS de l'ENTD 2007-2008

Figure 3.7: Trois outils de mesure des données dans l'ENTD 2007-2008



Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Nous disposons donc de trois sources pour décrire les déplacements :

- Les traces brutes GPS collectées par récepteur GPS ;
- Le questionnaire CAPI-GPS ;
- L'interrogation sur la mobilité quotidienne ;

Comme, pour ne pas trop lasser les enquêtés, on a évité d'appliquer le CAPI-GPS aux jours déjà décrits dans le cadre des déplacements quotidiens, ces sources ne peuvent être confrontées que deux à deux sur les mêmes jours d'enquête. On rapprochera donc les traces GPS enregistrées d'une part au CAPI-GPS pour certains jours, d'autre part à la description des déplacements quotidiens pour d'autres jours.

4 Acceptabilité, acceptation et biais des enquêtes de transport réalisées à l'aide de récepteur GPS

4.1 Introduction

Les enquêtes de mobilité sont nécessaires, non seulement pour faire le point et contribuer à identifier les problèmes liés à la gestion des systèmes de transport, mais également pour estimer et valider les modèles sur lesquels repose la planification. En effet, l'objectif le plus spécifique est le calibrage des modèles stratégiques destinés à prévoir l'effet à long terme des politiques de transport pour l'évaluation de projets et les diagnostics environnementaux. Pour cela, les personnes enquêtées doivent répondre à un grand nombre de questions. Cependant, pour un individu, il est difficile de se souvenir de l'ensemble de ses activités passées ainsi que de ses déplacements. En particulier, de se souvenir correctement des périodes exactes des activités (Auriat, 1996). Les technologies satellitaires basées sur les récepteurs « Global Positioning System » GPS ont avancé rapidement. Elles montrent un potentiel énorme comme instruments de mesure de la mobilité des personnes, en offrant une meilleure description du cadre spatio-temporel (Murakami, Wagner, 1999).

La faisabilité des enquêtes de mobilité à l'aide de récepteur utilisant le « Global Positioning System » (GPS) à grande échelle dépend (mais pas uniquement) de l'acceptabilité d'une telle technologie parmi les personnes sélectionnées. En effet, lors des études pilotes sur le GPS les enquêteurs ont généralement un grand degré de liberté pour « choisir » ceux qui sont prêts à coopérer et donc le taux de participation est beaucoup plus élevé que dans un échantillon tiré de façon aléatoire. Le fait d'accepter une telle enquête, est-il lié à des critères socio-démographiques ou/et aux caractéristiques de la mobilité du répondant ? Cela signifie clairement que nous devons faire attention à la méthode de redressement en étudiant le mécanisme d'acceptabilité d'une telle enquête. L'enquête fournit une description du ménage et des individus qui le composent ; l'un d'entre-eux, tiré au sort (individu « Kish »), nous a décrit sa mobilité et nous en avons profité pour lui demander s'il accepterait de porter un récepteur GPS afin de la mesurer de manière plus précise. Nous pouvons donc caractériser l'acceptabilité et déterminer si elle est ignorable (c'est-à-dire indépendante des comportements de mobilité que l'on souhaite observer) ou non ignorable (Armoogum, 2002).

4.2 Taux d'acceptabilité d'une enquête réalisée à l'aide de GPS

Dans le questionnaire principal de l'enquête nationale transports et déplacements (ENTD), nous avons posé une question sur l'acceptabilité d'un suivi de la mobilité par GPS aux individus de 18 ans ou plus. Nous n'avons considéré la réponse à cette question que pour les individus qui devaient décrire leur mobilité (soit un échantillon de 14 842 personnes).

Tableau 4.1: Acceptation d'une enquête à l'aide de récepteur GPS

Souhaitez-vous participer à une enquête sur la mobilité à l'aide d'un récepteur GPS ?	%
1. Oui, sans conditions	29,5
2. Oui, si je peux l'éteindre	5,1
3. Oui, mais sous d'autres conditions	0,8
4. Non	64,6
Ensemble de la population (de plus de 18 ans)	100,0

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Environ 2/3 des personnes de 18 ans ou plus refusent de porter un GPS (Tableau 4.1). Pour mieux comprendre les facteurs d'acceptabilité du GPS, nous avons estimé une régression logistique binaire sur l'échantillon des répondants. En effet, la régression logistique permet d'analyser une relation causale entre une variable dépendante qualitative (qui peut prendre les valeurs 0 et 1) et une ou plusieurs variables indépendantes (Gourieroux, 1989). De plus, nous avons réalisé une analyse de l'importance globale des variables utilisées lors de la régression logistique, suivant la méthodologie décrite dans (Lollivier et al., 2000). Cette méthode permet de tester la nullité simultanée de plusieurs paramètres.

4.3 Profils des personnes qui acceptent les enquêtes avec GPS

En retenant les variables significatives, les odds ratios (tableau 4.2) montrent que les facteurs qui jouent un rôle dans le fait de répondre positivement au principe d'une enquête à l'aide de récepteur GPS sont : la zone de résidence, le type de ménage, son équipement en automobiles et en ordinateurs, ainsi que le sexe, l'âge, la PCS et le nombre de déplacements réalisés la veille de la seconde visite par la personne tirée au sort. Toutefois la distance parcourue, le temps consacré à la mobilité décrit lors du deuxième entretien, l'état de santé et la taille du ménage ne semblent pas avoir d'influence.

Tableau 4.2 : Odds ratios sur l'acceptabilité d'une enquête à l'aide de récepteur GPS

Variables	Modalités	Rapport des cotes (odds ratios)	Pr > Khi-2
Sexe	Homme	1,19	<,0001
	Femme	Référence	
Age	[18 - 30[3,48	<,0001
	[30 - 40[2,63	<,0001
	[40 - 60[2,28	<,0001
	[60 - 70[1,94	<,0001
	[70 ou + [Référence	
Profession	Agriculteur	1,15	0,4231
	Artisan, Commerçant	Référence	
	Professions intellectuelles supérieures	1,69	<,0001
	Professions intermédiaires	1,52	<,0001
	Employés	1,47	0,0003
	Ouvriers	1,40	0,0020
	Retraités	1,38	0,0074
	Autres non actifs	1,45	0,0018
Type de ménage	Personne seule	1,44	<,0001
	Famille monoparentales	1,08	0,3567
	Couple sans enfant	1,28	<,0001
	Couple avec enfants	Référence	
	Autres types de ménages	1,63	<,0001
Zone de résidence	Rural	1,99	<,0001
	Couronne périurbaines AU<= 99 999 hab.	1,51	0,0061
	Pôle urbain AU<= 99 999 hab.	1,40	0,0106
	Zone multi-polarisée	2,38	<,0001
	Couronne péri-urbaine (AU 100 000 - 10 000 000 hab.)	1,91	<,0001
	Banlieue (AU 100 000 - 10 000 000 hab.)	1,67	<,0001
	Ville-centre (AU 100 000 - 10 000 000 hab.)	1,48	0,0014
	Couronne périurbaine de Paris	1,54	0,0100
	Banlieue de Paris	1,26	0,0662
	Ville de Paris	Référence	
Tranche de	[0 € - 1500 € [Référence	

revenu	[1500 € - 2400 € [1,12	0,0721
	[2400 € - 3500 € [1,27	0,0006
	[3500 € ou plus [1,26	0,0026
Motorisation du ménage	Ménage sans voiture	Référence	
	Ménage avec une voiture	1,54	<,0001
	Ménage avec deux voitures	1,64	<,0001
	Ménage avec trois voitures ou plus	1,75	<,0001
Ordinateur dans le ménage	Non	Référence	
	Oui	1,48	<,0001
Nombre de déplacement la veille de l'interview	0 déplacement	1,05	0,4385
	1 ou 2 déplacements	Référence	
	3 ou 4 déplacements	1,25	<,0001
	5 ou 6 déplacements	1,44	<,0001
	7 déplacements ou plus	1,61	<,0001

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Estrella (1998) a proposé un indicateur de la qualité de l'ajustement, il est égal à :

$$\phi_0 = 1 - \left(\frac{\text{Log } L}{\log L_0} \right)^{-\frac{2}{n} \text{Log } L_0}$$

où :

- n est, la taille de l'échantillon
- L est la valeur de la vraisemblance du modèle et
- L_0 la valeur de la vraisemblance du modèle réduit aux seuls termes constants, c'est-à-dire sans variable explicative.

Dans notre cas le $\phi_0 = 0,07$. C'est-à-dire que le modèle n'explique pas très bien le mécanisme de réponse. Puisque nous avons tenté d'introduire dans le modèle un grand nombre de variables, cela voudrait dire qu'une grande part du mécanisme de réponse serait due au hasard et qu'un redressement classique, ne privilégiant pas les dimensions que nous venons de mettre en évidence, pourrait minimiser les biais.

4.4 Comparaison des estimations entre l'ENTD 2007-2008 et le sous-échantillon « GPS »

4.4.1 Redressement du sous-échantillon des individus volontaires

Les techniques de calage sur marges ont été développées afin d'améliorer l'estimation de la variable d'intérêt, car nous réduisons la variance de notre estimateur lorsque les variables auxiliaires de calage sont corrélées avec la variable d'intérêt. Cette technique consiste à faire coïncider les marges de quelques variables de l'échantillon à celles de la population en modifiant la pondération. La méthode de calage sur marges la plus connue en matière de sondage est la 'méthode itérative du quotient' ou Raking Ratio introduite par Deming et Stephan (1940) et Stephan (1942). Deville et Särndal (1992) ont généralisé l'approche des estimateurs par régression en introduisant les estimations par « calage sur marges ». Nous pouvons ainsi corriger les différents types d'erreurs à condition d'utiliser les variables expliquant le mécanisme de la non-réponse lors du calage sur marges. Cette étape est essentielle pour assurer une bonne représentativité de l'échantillon et la comparabilité avec d'autres sources statistiques (enquêtes de l'INSEE).

A partir de l'échantillon des individus Kish de l'ENTD, nous avons exclu les individus qui ont répondu « non » à la question : Souhaitez-vous participer à une enquête sur la mobilité à l'aide d'un récepteur GPS ? Puis nous avons procédé à un redressement de cet échantillon en faisant un calage sur les mêmes marges que pour l'ENTD 2007-2008. Variables et modalité utilisées pour le redressement de l'ENTD (Armoogum et al. 2009) :

- PCS de l'individu Kish (Agriculteur ; Indépendant, profession libérale ; Professeur, instituteur actif ; Professeur, anciens instituteurs ; Employé, ouvrier actif ; Employé, ancien ouvrier ; Inactifs et chômeurs n'ayant jamais travaillé ; Individus de 6 à 15 ans) ;
- Sexe*âge de l'individu Kish (Hommes - De 6 à 24 ans ; de 25 à 34 ans ; de 35 à 49 ans ; de 50 à 64 ans ; plus de 64 ans ; Femmes - De 6 à 24 ans ; de 25 à 34 ans ; de 35 à 49 ans ; de 50 à 64 ans ; plus de 64 ans) ;
- Taille du ménage (Ménage d'une personne ; Ménage de 2 personnes ; Ménage de 3 personnes ; Ménage de 4 personnes ; Ménage de 5 personnes ou plus) ;
- Zone de résidence (Rural dans espace à dominante rurale ; Rural dans espace non à dominante rurale ; agglomération ≤ 19999 - ville centre agglomération ≤ 19999 – banlieue ; 20000 ≤ agglomération ≤ 49999 - ville centre ; 20000 ≤ agglomération ≤ 49999 – banlieue ; 50000 ≤ agglomération ≤ 99999 - ville centre ; 50000 ≤ agglomération ≤ 99999 – banlieue ; 100000 ≤ agglomération ≤ 199999 - ville centre ; 100000 ≤ agglomération ≤ 199999 – banlieue ; 200000 ≤ agglomération ≤ 1999999 - ville centre ; 200000 ≤ agglomération ≤ 1999999 – banlieue ; Ville de Paris ; Banlieue parisienne) ;

- Motorisation du ménage (0 voiture ; 1 voiture ; 2 voitures ou plus) ;
- Vague de l'enquête (6 vagues) ;
- Jour de l'enquête (lundi ; mardi ; mercredi ; jeudi ; vendredi)

4.4.2 Comparaison en moyenne

Comparons les estimations sur l'ensemble de l'ENTD et sur l'échantillon des individus qui sont susceptibles d'accepter le principe du suivi par GPS (les deux échantillons ayant été redressés de la même manière). Tous les indicateurs globaux de mobilité (nombre moyen de déplacements par jour et par personne, distance parcourue et temps consacré à se déplacer) sont significativement plus élevés chez ceux qui ont accepté le principe du suivi par GPS que dans l'ensemble de la population et les intervalles de confiance à 95% sont disjoints (Tableau 4.3).

Tableau 4.3 : Comparaison des estimations des indicateurs de Mobilité pour l'échantillon des individus qui sont susceptibles d'accepter le principe du suivi par GPS dans l'ENTD 2007-2008

	Echantillon ENTD			Sous-échantillon GPS		
	Moyenne	Borne inférieure	Borne supérieure	Moyenne	Borne inférieure	Borne supérieure
Nombre de déplacement	3,20	3,16	3,24	3,51	3,43	3,59
Budget distance	26,5	25,9	27,1	30,2	29,0	31,4
Budget temps	57,3	56,4	58,2	62,6	60,7	64,5

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Tableau 4.4 : Ventilations de la fréquence des déplacements selon l'acceptabilité du GPS

Nombre de déplacement		Echantillon ENTD			Sous-échantillon GPS		
		Moyenne	Borne inférieure	Borne supérieure	Moyenne	Borne inférieure	Borne supérieure
Sexe	Homme	3,13	3,07	3,20	3,31	3,20	3,41
	Femme	3,25	3,19	3,30	3,67	3,56	3,79
Tranche d'âge	[18 - 30[3,54	3,40	3,68	3,77	3,58	3,95
	[30 - 40[3,94	3,81	4,07	4,13	3,93	4,33
	[40 - 60[3,39	3,33	3,45	3,60	3,48	3,71
	[60 - 70[2,82	2,73	2,91	3,07	2,88	3,27
	[70 ou + [1,86	1,80	1,92	2,43	2,14	2,71
Motorisation du ménage	Ménage sans voiture	2,50	2,42	2,57	2,89	2,64	3,14
	Ménage avec une voiture	3,10	3,03	3,16	3,40	3,27	3,52

	Ménage avec deux voitures	3,52	3,45	3,59	3,74	3,62	3,87
	Ménage avec trois voitures ou plus	3,55	3,40	3,69	3,98	3,74	4,21
Zone de résidence	Rural	3,02	2,93	3,12	3,49	3,29	3,68
	Couronne périurbaines AU<= 99 999 hab.	2,99	2,85	3,13	3,40	3,17	3,64
	Pôle urbain AU<= 99 999 hab.	3,25	3,03	3,46	3,69	3,21	4,17
	Zone multi-polarisée	3,06	2,87	3,24	3,81	3,43	4,18
	Couronne péri-urbaine (AU 100 000 - 10 000 000 hab.)	3,21	3,10	3,33	3,32	3,09	3,54
	Banlieue (AU 100 000 - 10 000 000 hab.)	3,36	3,21	3,50	3,47	3,24	3,70
	Ville-centre (AU 100 000 - 10 000 000 hab.)	3,47	3,32	3,61	3,77	3,53	4,01
	Couronne périurbaine de Paris	3,04	2,89	3,18	3,21	2,93	3,50
	Banlieue de Paris	3,04	2,98	3,10	3,24	3,11	3,37
	Ville de Paris	3,05	2,94	3,15	3,24	2,97	3,51
Tranche de revenu mensuel	[0 € - 1500 €[2,68	2,62	2,75	3,12	2,96	3,28
	[1500 € - 2400 €[3,18	3,10	3,26	3,56	3,40	3,72
	[2400 € - 3500 €[3,41	3,31	3,50	3,58	3,41	3,74
	[3500 € ou plus[3,52	3,44	3,61	3,74	3,59	3,88

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Les catégories pour lesquelles la sur-mobilité en terme de nombre de déplacements de ceux qui accepteraient le GPS est la plus significative (et avec des intervalles de confiance disjoints) sont les femmes, les personnes au-delà de 70 ans, vivant dans des zones peu denses, les ménages sans voiture ou ayant plus de trois voitures et vivant dans des ménages motorisés aux revenus modestes (tableau 4.4).

Tableau 4.5 : Ventilations des distances parcourues selon l'acceptabilité du GPS

Budget distance		Echantillon ENT			Sous-échantillon GPS		
		Moyenne	Borne inférieure	Borne supérieure	Moyenne	Borne inférieure	Borne supérieure
Sexe	Homme	29,3	28,3	30,3	30,8	29,1	32,5

	Femme	24,3	23,6	25,0	29,7	28,1	31,3
Tranche d'âge	[18 - 30[29,9	27,7	32,1	30,8	27,8	33,8
	[30 - 40[34,6	32,8	36,4	39,2	36,0	42,4
	[40 - 60[29,5	28,6	30,4	31,5	29,9	33,1
	[60 - 70[22,8	21,6	24,0	26,9	23,8	30,0
	[70 ou + [10,5	9,8	11,2	14,3	11,4	17,2
Motorisation du ménage	Ménage sans voiture	10,7	10,0	11,4	12,7	10,7	14,7
	Ménage avec une voiture	23,9	23,0	24,8	26,8	25,0	28,6
	Ménage avec deux voitures	33,8	32,7	34,9	37,7	35,7	39,7
	Ménage avec trois voitures ou plus	36,4	34,2	38,6	39,4	36,0	42,8
Zone de résidence	Rural	29,7	28,3	31,1	32,6	29,8	35,4
	Couronne périurbaines AU<= 99 999 hab.	28,6	26,6	30,6	36,0	32,3	39,7
	Pôle urbain AU<= 99 999 hab.	25,6	22,3	28,9	32,7	24,3	41,1
	Zone multi-polarisée	35,6	32,5	38,7	36,7	32,0	41,4
	Couronne péri-urbaine (AU 100 000 - 10 000 000 hab.)	34,1	32,4	35,8	41,0	37,1	44,9
	Banlieue (AU 100 000 - 10 000 000 hab.)	25,7	23,9	27,5	27,5	24,3	30,7
	Ville-centre (AU 100 000 - 10 000 000 hab.)	17,8	16,4	19,2	20,7	17,9	23,5
	Couronne périurbaine de Paris	40,3	37,5	43,1	46,2	40,6	51,8
	Banlieue de Paris	24,5	23,7	25,3	25,4	23,9	26,9
	Ville de Paris	16,0	14,9	17,1	15,7	13,3	18,1
Tranche de revenu	[0 € - 1500 € [16,6	15,8	17,4	19,6	17,7	21,5
	[1500 € - 2400 € [25,2	24,1	26,3	29,4	27,1	31,7
	[2400 € - 3500 € [31,2	29,8	32,6	34,2	31,7	36,7
	[3500 € ou plus [33,5	32,2	34,8	36,0	33,6	38,4

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Les catégories pour lesquelles la sur-mobilité en terme de distance parcourue ou de temps consacré à la mobilité chez ceux qui accepteraient le GPS est la plus significative (et avec des intervalles de

confiance disjointes) sont les femmes, les personnes au-delà de 60 ans, vivant dans des zones peu denses, les ménages peu motorisés et aux revenus modestes (tableau 4.5 et 4.6).

Tableau 4.6 : Ventilations du temps consacré aux déplacements selon l'acceptabilité du GPS

Budget temps		Echantillon ENTD			Sous-échantillon GPS		
		Moyenne	Borne inférieure	Borne supérieure	Moyenne	Borne inférieure	Borne supérieure
Sexe	Homme	59,7	58,1	61,3	63,2	59,9	66,5
	Femme	55,3	54,2	56,4	62,1	59,9	64,3
Tranche d'âge	[18 - 30[65,6	62,3	68,9	66,4	62,0	70,8
	[30 - 40[65,2	62,7	67,7	68,9	65,0	72,8
	[40 - 60[61,7	60,2	63,2	64,2	61,0	67,4
	[60 - 70[53,0	51,0	55,0	59,2	54,5	63,9
	[70 ou + [33,3	31,9	34,7	45,7	39,4	52,0
Motorisation du ménage	Ménage sans voiture	52,5	50,6	54,4	60,7	54,4	67,0
	Ménage avec une voiture	55,6	54,1	57,1	61,8	58,2	65,4
	Ménage avec deux voitures	60,4	58,9	61,9	63,7	61,2	66,2
	Ménage avec trois voitures ou plus	60,9	57,7	64,1	64,3	59,5	69,1
Zone de résidence	Rural	48,4	46,1	50,7	55,3	49,5	61,1
	Couronne périurbaines AU<= 99 999 hab.	47,7	44,9	50,5	57,3	52,3	62,3
	Pôle urbain AU<= 99 999 hab.	52,4	48,0	56,8	61,5	51,8	71,2
	Zone multi-polarisée	54,6	50,6	58,6	59,3	53,4	65,2
	Couronne péri-urbaine (AU 100 000 - 10 000 000 hab.)	56,2	53,8	58,6	62,4	57,8	67,0
	Banlieue (AU 100 000 - 10 000 000 hab.)	55,9	53,1	58,7	60,3	55,3	65,3
	Ville-centre (AU 100 000 - 10 000 000 hab.)	57,2	54,5	59,9	64,4	59,2	69,6
	Couronne périurbaine	66,1	62,2	70,0	68,8	61,5	76,1

	de Paris						
	Banlieue de Paris	75,6	73,7	77,5	76,3	72,6	80,0
	Ville de Paris	78,3	74,9	81,7	79,4	72,0	86,8
Tranche de revenu	[0 € - 1500 € [46,3	44,9	47,7	52,6	49,2	56,0
	[1500 € - 2400 € [56,3	54,4	58,2	61,6	56,8	66,4
	[2400 € - 3500 € [59,4	57,5	61,3	63,8	60,6	67,0
	[3500 € ou plus [68,1	66,1	70,1	71,4	68,0	74,8

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

4.4.3 Et en distribution

C'est à partir de 3 déplacements, 20 km ou 50 minutes dans la journée que les fréquences sont plus fortes chez ceux qui accepteraient le GPS (tableaux 4.7 à 4.9).

Tableau 4.7 : Distribution de la fréquence de déplacement

Nombre de déplacement	ENTD (%)	ENTD GPS (%)
0	18,5	16,0
1	1,2	1,1
2	30,1	25,7
3	7,4	9,0
4	19,7	20,3
5	6,5	7,6
6	7,1	8,3
7	3,2	4,2
8	2,9	3,3
9	1,4	1,9
10 et plus	2,1	2,6
Total	100,0	100,0

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Tableau 4.8 : Distribution du budget distance

Budget distance	ENTD	ENTD GPS
0	18,5	16,0
]0; 10[24,1	22,5

[10; 20[16,1	15,5
[20; 30[10,6	11,0
[30; 40[7,7	8,7
[40; 50[5,8	6,3
[50; 60[4,4	4,5
[60; 70[3,1	3,7
[70; 80[2,3	2,7
[80; 90[1,7	2,4
[90; 100[1,3	1,7
100 et plus	4,5	5,2
Total	100,0%	100,0%

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Tableau 4.9 : Distribution du budget temps

Budget temps	ENTD	ENTD GPS
0	18,5	16,0
]0; 10[1,6	1,4
[10; 20[5,4	4,8
[20; 30[9,0	8,2
[30; 40[8,5	8,3
[40; 50[8,6	8,5
[50; 60[6,8	7,7
[60; 70[8,4	7,4
[70; 80[5,5	6,3
[80; 90[4,9	5,4
[90; 100[4,4	5,3
100 et plus	18,4	20,8
Total	100,0%	100,0%

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

4.5 Meilleure description de la mobilité ou une mobilité plus grande ?

Nous allons vérifier l'hypothèse de Stopher et al. (2010) selon laquelle, les individus qui répondent à une enquête de mobilité à l'aide de récepteur GPS sont plus consciencieux et décrivent mieux leur mobilité. Disposant de la réponse à la question « Souhaitez-vous participer à une enquête sur la

mobilité à l'aide d'un récepteur GPS ? » et aussi de la variable répertoriant les personnes qui ont réellement accepté le volet « GPS », nous pouvons tester si c'est le fait de porter le GPS qui fait que l'on décrit mieux sa mobilité ? Est-ce que les personnes qui ont une mobilité plus grande acceptent le plus souvent une enquête par GPS.

Pour ce faire nous avons réalisé un modèle d'analyse de la variance en mettant comme variables explicatives toutes les variables qui ont servi pour le redressement et deux variables dichotomiques :

- Personnes acceptant d'être suivies par GPS (oui/non)
- Personnes ayant participé au pilote GPS de l'ENTD (oui/non)

Finalement, les personnes qui acceptent d'être suivies par GPS ont une mobilité supérieure (+ 0,32 déplacement ; + 3.5 kilomètres et + 5,8 minutes de plus que la moyenne de la population. Tableau 4.10).

Tableau 4.10 : estimation du biais des enquêtes de mobilité par GPS

N=14 842		Fréquence de déplacement		Budget distance journalier		Budget temps journalier	
		R ² = 0,100		R ² = 0,120		R ² = 0,074	
Variables		Valeur estimée	Pr > t	Valeur estimée	Pr > t	Valeur estimée	Pr > t
	Constante	2,20	<,0001	6,64	0,0092	63,19	<,0001
Sexe	Homme						
	Femme	0,16	0,0002	- 2,67	<,0001	- 2,73	0,0051
Tranche d'âge	[18 - 30[0,86	<,0001	11,67	<,0001	18,88	<,0001
	[30 - 40[1,16	<,0001	10,89	<,0001	16,63	<,0001
	[40 - 60[0,76	<,0001	7,12	<,0001	15,60	<,0001
	[60 - 70[0,73	<,0001	7,10	<,0001	16,85	<,0001
	>70						
Profession	Agriculteur	0,73	<,0001	- 9,86	0,0001	- 4,45	0,2959
	Artisan, Commerçant						
	Professions intellectuelles supérieures	- 0,19	0,1241	- 2,99	0,0791	2,10	0,4576
	Professions intermédiaires	0,22	0,0549	- 1,38	0,3838	1,46	0,5786
	Employés	0,18	0,1036	- 9,03	<,0001	- 5,35	0,0366
	Ouvriers	- 0,11	0,3538	- 9,18	<,0001	- 7,95	0,0027
	Retraités	- 0,32	0,0098	- 13,34	<,0001	- 12,71	<,0001

	Autres non actifs	- 0,22	0,0851	- 15,10	<,0001	- 10,26	0,0003
Type de ménage	Personne seule	- 0,22	0,0024	4,18	<,0001	1,23	0,4568
	Famille monoparentales	- 0,11	0,1934	2,75	0,022	1,38	0,489
	Couple sans enfant	- 0,36	<,0001	1,33	0,0912	1,22	0,3503
	Couple avec enfants						
	Autres types de ménages	- 0,47	<,0001	0,30	0,8434	- 4,65	0,0689
Zone de résidence	Rural	- 0,13	0,2767	12,90	<,0001	- 23,65	<,0001
	Couronne périurbaines AU< 99 999 hab.	- 0,30	0,0406	9,60	<,0001	- 26,66	<,0001
	Pôle urbain AU< 99 999 hab.	0,17	0,1686	9,24	<,0001	- 19,68	<,0001
	Zone multi-polarisée	- 0,24	0,0861	16,15	<,0001	- 20,08	<,0001
	Couronne péri-urbaine (AU 100 000 - 10 000 000 hab.)	- 0,21	0,0954	13,03	<,0001	- 20,56	<,0001
	Banlieue (AU 100 000 - 10 000 000 hab.)	0,12	0,298	6,96	<,0001	- 19,17	<,0001
	Ville-centre (AU 100 000 - 10 000 000 hab.)	0,35	0,0028	1,15	0,4813	- 17,07	<,0001
	Couronne périurbaine de Paris	- 0,43	0,0117	18,91	<,0001	- 12,05	0,002
	Banlieue de Paris	- 0,24	0,0491	5,76	0,0005	- 1,81	0,511
	Ville de Paris						
Motorisation	Ménage sans voiture						
	Ménage avec une voiture	0,33	<,0001	7,47	<,0001	- 1,16	0,4511
	Ménage avec deux voitures	0,44	<,0001	11,56	<,0001	- 0,88	0,6367
	Ménage avec trois voitures ou plus	0,39	0,0001	13,67	<,0001	- 0,48	0,8384
Ménage équipé d'un Ordinateur	Oui	- 0,28	<,0001	- 1,31	0,0692	- 3,23	0,0067
	Non						
Tranche de revenu mensuel du ménage	[0 € - 1500 € [
	[1500 € - 2400 € [0,17	0,0045	3,64	<,0001	6,46	<,0001
	[2400 € - 3500 € [0,10	0,1396	5,53	<,0001	5,46	0,0007
	[3500 € ou plus [0,21	0,009	6,62	<,0001	9,22	<,0001
Volontaire pour porter un GPS	Oui	0,32	<,0001	3,53	<,0001	5,82	<,0001
	Non						
Enquêté	Oui	0,07	0,3741	1,50	0,1933	- 0,64	0,7367

GPS	Non						
-----	-----	--	--	--	--	--	--

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Plus une personne se déplace, plus sa probabilité d'accepter une enquête GPS est importante. C'est pour cela que les enquêtes auprès des personnes volontaires pour être suivi par GPS sont susceptible d'être biaisées.

4.6 Conditions complémentaire pour l'acceptabilité de l'enquête GPS

Lors de la première visite, parmi 18632 enquêtés qui ont répondu à la question « Accepteriez-vous de participer à une enquête transport expérimentale avec GPS menée par l'INSEE, et d'emmener un GPS dans tous vos déplacements pendant une semaine ? », il y a 1208 cas dans lequel l'enquêté répond « Oui, mais sous d'autres conditions ». Dans ces cas, l'enquêté va préciser sa condition complémentaire pour l'acceptabilité de l'enquête GPS. En étudiant ces conditions complémentaires, nous remarquons qu'il y a 986 enquêtés qui participerait à l'enquête GPS à condition de pouvoir éteindre le GPS quand ils en ont envie.

Tableau 4.11 : Conditions complémentaires pour l'acceptabilité de l'enquête GPS

Condition pour l'acceptabilité de l'enquête GPS	Fréquence
Choix du moment des enregistrements/éteindre GPS	986
Utilisation facile du GPS	18
GPS peu encombrant	11
Peur oubli de chargé GPS	4
Recevoir un cadeau	16
Temps d'enquête	12
Remplacement du carnet de véhicule	7
Accès aux informations fournies par le GPS	8
Peur du vol/ dédommagement	8
Anonymat	6
Peur d'être surveillé	2
Autre raison ou non réponse	130
Total	1208

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Il y a 18 enquêtés qui participeraient à condition d'une utilisation facile du GPS et 11 l'appareil est peu encombrant. La forme et la taille de l'appareil ont donc leur importance. Le GPS pourrait comporter une grande pince, alternativement à la courroie de cou, pour permettre aux répondants de le fixer à son sac ou à ses vêtements ; ceci réduirait le risque de perte des signaux lorsque le

récepteur est placé dans de mauvaises conditions (au fond d'un sac, loin d'une vitre dans les transports collectifs, etc.).

De plus, 16 enquêtés accepteraient de participer s'il y a des cadeaux ou indemnité financière. Remarquons qu'il y a 4 enquêtés qui veulent garder le GPS après l'enquête ; cela montre qu'ils n'ont pas compris que le récepteur GPS n'a aucune interface graphique, donc pas de valeur commerciale.

12 enquêtés craignent que cela ne leur prenne trop de temps et 7 enquêtés ne veulent pas remplir le carnet de véhicule. Il est haute probabilité que l'enquêté pense qu'il doit apporter GPS et remplir le carnet de véhicule pendant au moins 7 jours en même temps. En fait, il n'y a pas eu de confrontation entre les traces GPS et le carnet-véhicule, car ce n'est pas tout le temps la même personne qui est suivie. Dans le volet GPS de l'ENTD 2007-2008, l'enquêté volontaire porte l'appareil GPS tout le temps et partout pendant au moins 7 jours, tandis que le carnet de véhicule est rempli par les membres du ménage (d'autres membres).

La peur que le récepteur soit volé ou endommagé est une raison notable de refus de l'enquête GPS (8 cas). Toutefois, dans l'ENTD 2007-2008, la perte, le vol ou la destruction du GPS sont des impondérables qui n'ont eu aucune conséquence ni pour l'enquêteur, ni pour l'enquêté. Cela montre de nouveau les lacunes de l'information perçue par les enquêtés.

Parmi ceux qui n'étaient pas très ou pas du tout prêt à porter un appareil GPS, les autres raisons citées étaient l'invasion de la vie privée (6 cas pour l'anonymat et 2 cas pour se sentir surveillé) et un manque de confiance dans l'utilisation de la technologie ou l'inquiétude sur l'accès aux informations enregistrées par le GPS (8 cas). Pour persuader les enquêtés, l'enquêteur doit expliquer clairement :

- Comment les données seront utilisées et stockées en toute sécurité ;
- La façon d'utiliser le GPS, en inspirant confiance à l'enquêté ;
- Qu'ils ne seront pas tenus personnellement responsables des dommages, vol ou perte de récepteur GPS.

En plus, il sera important pour l'enquêté de se sentir soutenu au long de l'enquête, sachant qu'il pourra contacter l'enquêteur ou quelqu'un de l'INSEE ou de l'IFSTTAR via un site Web, par courriel ou par téléphone lorsqu'ils ont des questions ou rencontrent des difficultés.

Une présentation soignée des questionnaires et des documents (brochures et notices) et des conseils d'utilisation peuvent encourager les enquêtés les plus réticents vis-à-vis d'une enquête GPS.

4.7 Acceptation d'une enquête réalisée à l'aide de GPS

Lors de la première visite, la question « *Accepteriez-vous de participer à une enquête transport expérimentale avec GPS menée par l'INSEE, et d'emmener un GPS dans tous vos déplacements pendant une semaine ?* » est posé à la personne qui répond pour le ménage. Si la réponse à la question concernant l'acceptabilité de l'enquête GPS était « oui », l'enquêteur pose une question à la fin du deuxième interview à la personne qui a décrit sa mobilité « *êtes-vous volontaire pour l'enquête GPS ?* ». Dans les 12 régions où cette enquête a été réalisée, si un récepteur est disponible et si les réponses à ces deux questions sont positives, alors l'enquêté prend le GPS pour une période d'au moins 7 jours. Nous remarquons que la disponibilité de récepteur influence de manière importante le taux d'acceptation de l'enquête

Tableau 4.12 : Croisement entre l'acceptabilité et le volontariat à l'enquête GPS

	Réponse à la question « <i>êtes-vous volontaire pour l'enquête GPS ?</i> »			
	Non réponse	Oui pour être volontaire	Ne souhaite pas être volontaire	Total
Accepte une enquête par GPS -	5456	957	144	6557
refuse une enquête par GPS	12075	0	0	12075
Total	17531	957	144	18632

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Tableau 4.13 : Croisement entre la disponibilité d'un GPS et le volontariat dans le cas de l'inacceptabilité de l'enquête GPS

Dans le cas: Non Acceptabilité avec GPS				
	Pas réponse de volonté avec GPS (données manquant)	Volontaire avec GPS	Non Volontaire avec GPS)	Total
GPS Disponible	2087	0	0	2087
GPS Non Disponible	9988	0	0	9988
Total	12075	0	0	12075

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Tableau 4.14 : Croisement entre la disponibilité de GPS et le volontaire de l'enquête GPS dans le cas de l'acceptabilité de l'enquête GPS

Dans le cas: Acceptabilité avec GPS (TESTGPS = 1 ou 2 ou 3)				
	Pas réponse de volonté avec GPS (données manquant)	Volontaire avec GPS	Non Volontaire avec GPS	Total
GPS Disponible	0	957	144	1101
GPS Non Disponible	5456	0	0	5456

Total	5456	957	144	6557
--------------	------	-----	-----	------

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

On trouve 5456 cas où l'enquêté répond "oui" à la première question sur l'acceptabilité, mais comme aucun appareil n'es disponible, la seconde question n'est pas posée. C'est-à-dire que le taux d'indisponibilité de GPS est égal à 82 % (=5456/6557).

On peut donc augmenter le taux d'acceptation par les enquêtés et le nombre de suivis réalisés par l'accroissement du taux de disponibilité des GPS. Pour ce faire, on doit calculer les temps et durées des visites des enquêteurs pour baisser le taux d'immobilisation forcée des GPS. On entend par "immobilisation forcée" les cas où des GPS restent dans une Direction Régionale de l'INSEE ou à l'IFSTTAR), et ne sont proposés à aucun enquêteur GPS : soit parce qu'ils servent de "réserve", soit parce qu'aucun enquêteur GPS n'est disponible durant une certaine période dans une Direction Régionale de l'INSEE.

CONCLUSION

D'abord, en étudiant l'acceptabilité et les biais induit par des enquêtes de mobilité à l'aide de GPS, nous avons confirmé l'hypothèse de Stopher (2010) selon laquelle les enquêtés qui répondent à une enquête de transport à l'aide de récepteur GPS sont plus consciencieux et décrivent mieux leur mobilité. Nous trouvons que les personnes qui acceptent d'être suivies par GPS ont une mobilité supérieure à la moyenne de la population (+0.32 déplacement/jour ; +3.5 kilomètres/jour et + 6 minutes/jour). Ceci montre que plus une personne se déplace, plus la probabilité qu'elle accepte une enquête GPS est importante. Le profil des personnes qui acceptent l'enquête GPS est un homme jeune, exerçant une profession intellectuelle supérieure, vivant dans un ménage de structure complexe à haut revenu (plus de 2400 € par mois), équipé en ordinateur et multi-motorisé, habitant dans une zone rurale ou périurbaine d'une aire urbaine de 100000 à 10000000 habitants.

Nous pouvons augmenter le taux d'acceptation de l'enquête GPS en convaincant les enquêtés d'y participer. Les enquêteurs devront expliquer les principes de l'enquête et s'assurer que les répondants comprennent exactement ce qu'on leur demande de faire. L'enquêteur peut lever les doutes quand l'enquêté a besoin de clarifications. Les enquêteurs devront souligner qu'il s'agit d'expérimenter une nouvelle méthode d'enregistrement et de post-traitement des données GPS, en facilitant la collecte des données, en réduisant la charge de l'enquêté qui doit seulement porter chaque jour le GPS au lieu de remplir complètement un carnet de déplacements. Les enquêteurs devraient présenter également la simplicité du dispositif GPS et la précision des données collectées. Les enquêtés doivent comprendre que le dispositif GPS n'est qu'un outil de suivi passif, et qu'ils ne doivent pas se sentir surveillés. Les enquêteurs devraient insister sur l'anonymat et la confidentialité des données. Toute perte ou détérioration de l'appareil est un impondérable qui sera sans

conséquence. Les données ne seront utilisées que pour la recherche. L'enquêteur peut se référer aux instructions, montrer le GPS afin d'expliquer à l'enquêté : Comment et quand charger, transporter, allumer, éteindre l'appareil. Les notices et brochures sont utiles pour rappeler le mode d'emploi du GPS et l'objectif de l'expérimentation de cette nouvelle technologie.

Il est également important pour que les enquêtés se sentent aidé tout au long de période de collecte par la capacité à contacter l'enquêteur ou quelqu'un de l'INSEE ou l'IFSTTAR par email, téléphone ou web particulière lorsqu'ils ont des questions sur le récepteur GPS ou sur l'enquête sur les transports. Une copie des notices ou la brochure de soutien et des conseils sur l'utilisation de l'appareil GPS devrait être laissée à chaque enquêté volontaire à l'enquête GPS.

5 Analyse des traces brutes collectées par GPS

5.1 Algorithme Decoutrace

Nous pouvons imaginer qu'un point apparaît toutes les « quelques secondes » dans un espace à trois dimensions. Si l'intervalle de temps entre les points est suffisamment court et assez régulier, il ne sera pas difficile de distinguer une période de mouvement d'une période à l'arrêt. Mais, pour l'instant, on n'a qu'une chaîne de points sur « un papier blanc ».

Si l'on souhaite reconstruire l'itinéraire de la personne enquêtée, on doit projeter les points sur un plan correct afin de savoir sur quelle route il est passé, de quel lieu il est parti. Si l'on veut que ce travail puisse être réalisé automatiquement par programme, un SIG (Système d'Information Géographique) est indispensable. En utilisant un SIG, nous avons trois sources d'information qui jouent des rôles différents dans notre algorithme :

- Données GPS : Où est l'enquêté ? À quel moment ?
- SIG (Points d'Intérêt ou POI) : l'environnement et les contours au tour du point d'enregistrement, par exemple : un arrêt de bus, une station de métro ou de train, un restaurant, une station-service, un gymnase, une mairie, un hôpital, un centre commercial, etc. ;
- SIG (Réseaux) : les itinéraires empruntés, par exemple : les routes, les chemins de fer, etc. ;

Ces informations nous aident à reconstituer ce qui a eu lieu, mais parallèlement, il y a aussi des imprécisions, erreur ou omission :

- Données GPS : imprécision des coordonnées, données manquantes au démarrage ou dans de mauvaises conditions de visibilité des satellites ;
- SIG : imprécision des POI et des arcs des réseaux, mise à jour nécessaire, plusieurs activités en un même lieu (par exemple : commerce, médecin, logements, dans un même immeuble), nomenclature pas toujours adaptée à la détermination des activités et des motifs de déplacement, etc. ;
- Humain : GPS oublié, non rechargé, non allumé, mauvais positionnement du récepteur ;

Un défi pour le traitement à posteriori des données brutes collectées par GPS est de tirer le maximum des informations disponibles et de limiter les erreurs pour reconstruire les déplacements, imputer les modes et motifs de déplacement de l'enquêté de manière automatisée en séquences continues à la fois dans l'espace et dans le temps. Le traitement a posteriori des données est un sujet de recherche en cours. Récemment, certaines idées nouvelles ont été publiées (Forresr, Pearson, 2005) ; (Schuessler, Axhausen, 2009) ; (Yuan, 2010); (Nguyen, 2013). Toutefois, les questions-clés sont toujours les mêmes : comment détecter les déplacements individuels et les

activités ? Comment obtenir les modes utilisés par les enquêtés ? Comment projeter les itinéraires choisis sur le réseau ? L'algorithme DECOUTRACE se propose d'apporter des réponses à ces trois questions en extrayant un maximum d'informations des trois sources citées ci-dessus. Par exemple, pour identifier le mode: d'abord, nous découpons les enregistrements en trace ; puis quelques caractéristiques du segment, comme la vitesse moyenne, la vitesse maximale, etc. vont nous aider à identifier le mode; ensuite, les points d'intérêt (POI) autour de la fin du segment, par exemple arrêts de bus et stations de métro, peuvent amener à corriger ou confirmer le mode ; après cela, on peut repérer la ligne de bus ou l'autoroute sur un SIG ou une carte (map-matching) pour confirmation.

Par conséquent, le court questionnaire complémentaire (CAPI-GPS) et la comparaison avec les déplacements quotidiens dans les questionnaires classiques permettront l'évaluation des méthodes de traitement, et constitueront la clé de la qualité des résultats.

Nous définissons les deux couches principales suivantes pour l'algorithme DECOUTRACE :

5.1.1 Couche donnée GPS

Le découpage des données brutes de GPS en fonction du mouvement, des points d'arrêt, et une hypothèse sur le mode de transport sont les sorties souhaitées pour cette couche. Cette couche comporte les 6 étapes suivantes :

- Filtrage des données aberrantes ;
- Estimation de la vitesse réelle ;
- Lissage de données ;
- Découpage de la chaîne de données GPS en séquences de mouvement et d'arrêt ;
- Estimation des données manquantes ;
- Imputation du mode de déplacement ;
- Repérage des points d'arrêt pour identifier les lieux visités ;

5.1.1.1 Filtrage des données aberrantes

Pendant la période d'enquête, la précision du point collectée par GPS est influencée par plusieurs facteurs, la répartition géométrique des positions des satellites, l'irrégularité de leur orbite, la réflexion des signaux sur les couches de l'ionosphère et de la troposphère, etc. Le GPS est parfois utilisé dans de mauvaises conditions : le signal peut être perdu lors du passage dans un tunnel, à l'intérieur d'un bâtiment comme dans un bureau, dans un centre commercial, ou dans un bus loin de la fenêtre, dans certains trains, etc. De plus, un démarrage à froid ou un redémarrage nécessitent un laps de temps pour capter les signaux des satellites, qui nous fait très souvent perdre le début du déplacement. Dans les conditions mentionnées ci-dessus, les enregistrements seront erronés et

vont générer des résultats incorrects pour les prochaines étapes. C'est la raison pour laquelle, le filtrage de ces points est nécessaire. Nous présentons deux méthodes de filtrage de données aberrantes dont la première basée sur le nombre de satellites GPS et HDOP ; et la seconde établie sur la vitesse et l'accélération.

Méthode 1 : Nombre de satellites GPS et HDOP

Nous savons que quatre satellites sont nécessaires pour résoudre l'équation à quatre inconnues qui sont la position dans les 3 dimensions plus le décalage entre l'horloge du récepteur et celle des satellites. Afin de garder le maximum de données qui soient utiles, nous laissons l'altitude de côté, car elle peut être calculée en projetant la trajectoire sur une carte comportant des courbes de niveau. Donc, seuls trois satellites seront nécessaires.

On rappelle que l'UERE est un estimateur de l'erreur sur la mesure de la distance récepteur-satellite. HDOP est la Précision horizontale de la position obtenue par GPS. DRMS est la racine carrée de la moyenne des carrées des erreurs sur les deux axes horizontal et vertical. Les DRMS seront calculés comme dans l'équation ci-dessous (Wilson, 2009).

$$\text{DRMS}(\text{HDOP}) = \text{HDOP} \times \text{UERE}$$

L'erreur sur la distance entre deux points successifs enregistrés par GPS suit une loi de Rayleigh de paramètre 2DRMS^2 . On a calculé sa distribution :

$$\text{Probabilité (Erreur de localisation} \leq \text{Distance)} = 1 - e^{-\left(\frac{\text{Distance}}{\sqrt{2}\text{DRMS}}\right)^2}$$

Ces deux équations montrent que la vitesse, l'intervalle de temps entre deux points enregistrés, DRMS, HDOP sont des paramètres qui influencent la qualité de la position relative entre les points successifs. On trouve que la probabilité de bonne qualité des données se dégrade fortement avec l'augmentation de HDOP. On doit choisir un seuil de valeur HDOP pour décider quel point est aberrant. Cela dépend aussi de l'UERE. On doit trouver un équilibre entre la possibilité de garder un maximum de données qui soient utiles et le filtrage des données aberrantes.

Quand le paramètre DOP est disponible, le point enregistré par GPS sera bien positionné si sa valeur, c'est-à-dire la précision HDOP, est inférieure à 5 mètres et le nombre de satellites en vue est plus grand que 3.

Méthode 2 : Vitesse et Accélération

Quand le récepteur GPS ne fournit pas les informations sur la dilution de la précision (DOP) et/ou le nombre de satellites en vue, nous ne pouvons pas éliminer les points invalides en utilisant la méthode ci-dessus. Nous allons développer une méthode qui ne dépend pas du DOP ni du nombre de satellites en vue.

Nous calculons alors la vitesse moyenne entre deux points successifs en fonction de leurs coordonnées et du temps écoulé. Ainsi, nous pouvons identifier un point aberrant en comparant la vitesse moyenne entre deux segments successifs. L'idée est que la vitesse moyenne entre un point erroné et les points juste avant et juste après sera beaucoup plus élevée que celle entre deux points bien localisés.

On utilise la méthode Vitesse – Accélération dont le contenu est expliqué comme suit.

Si $\bar{v}_{(i+1,i+2)} - \bar{v}_{(i,i+1)} \leq S_v$, alors P_{i+2} est valide; si non P_{i+2} est invalide

$$\text{avec } S_v = \begin{cases} S_a \times \frac{1}{2}(T_{(i,i+1)} + T_{(i+1,i+2)}), & \text{si } \frac{1}{2}(T_{(i,i+1)} + T_{(i+1,i+2)}) \in [1, 10] \\ 100, & \text{si } \frac{1}{2}(T_{(i,i+1)} + T_{(i+1,i+2)}) \in [10, 120) \end{cases},$$

$$S_a = 10 \text{ km/h/s}, i \in [1, N]$$

Comparaison des deux méthodes

La première méthode "nombre de satellites et HDOP" supprime les points qui ont la plus haute probabilité d'être erroné, tandis que la seconde méthode "Vitesse-Accélération" élimine les points qui sont détectés comme défectueux.

La seconde méthode nécessite plus de temps de traitement que la première. Elle ne peut traiter ni le premier point, ni le dernier de la chaîne de données GPS. Elle ne peut pas travailler avec une chaîne complète de données GPS erronée ; le résultat devient imprévisible. Mais la première méthode fonctionne dans les deux cas.

Application de deux méthodes au volet GPS de l'ENTD 2007-2008

Le matériel utilisé pour collecter les données dans le sous-échantillon des volontaires avec GPS de l'ENTD 2007-2008 est le récepteur Royaltek BlueGPS avec ses deux versions successives RBT-3000, puis RBT-1100.

La différence principale entre les données collectées par ces deux récepteurs est la présence de l'information sur la précision (DOP). Le récepteur RBT-3000 ne fournit pas de DOP, tandis que le RBT-1100 la fournit.

Pour les données collectées par récepteur RBT-1100, nous utilisons les deux méthodes Satellites-HDOP et Vitesse – Accélération pour filtrer les données aberrantes. Mais avec les données obtenues des récepteurs RBT-3000, nous appliquons seulement la méthode Satellites – HDOP.

5.1.1.2 Estimation de la vitesse réelle

Nous disposons de deux mesures : La vitesse instantanée enregistrée par le GPS et la vitesse calculée à partir des coordonnées de deux points successifs. Laquelle est la plus proche de la vitesse réelle ? Nous savons que la vitesse enregistrée par le GPS est une vitesse instantanée qui ne peut représenter que la vitesse au moment où le point est enregistré. Cela signifie que c'est le meilleur choix lorsque les points sont enregistrés chaque seconde sans trou. En revanche, la vitesse calculée à partir des coordonnées de deux points successifs est la vitesse moyenne entre les deux points inscrits. Mais, en raison de l'imprécision des coordonnées, la précision de la vitesse décroît quand la distance entre deux points successifs diminue. A vitesse constante, la précision augmente avec l'intervalle de distance ou de temps entre deux points successifs. Cela signifie que si le récepteur a perdu le signal pour un temps assez long, cette vitesse moyenne sera plus proche de la vitesse réelle. Dans l'application, nous utilisons la moyenne pondérée de ces deux vitesses pour estimer la vitesse réelle.

La vitesse estimée est une moyenne pondérée de la vitesse v_i fournie par le récepteur GPS (qui représente la vitesse instantanée au moment où le point est enregistré) et de la vitesse moyenne $\bar{v}_{(i-1,i)}$ sur le segment qui est calculée en divisant la distance par la durée de deux points successifs.

$$v_{es_i} = \begin{cases} v_i, & In_i \leq In_{cf} \\ \exp\left(0.0953 \frac{In_{cf} \sqrt{In_i}}{In_i - 1}\right) v_i + \left(1 - \exp\left(0.0953 \frac{In_{cf} \sqrt{In_i}}{In_i - 1}\right)\right) \bar{v}_{(i-1,i)}, & In_{cf} \leq In_i \leq \frac{S_{Ta}}{2} \\ \bar{v}_{(i-1,i)}, & \frac{S_{Ta}}{2} \leq In_i \end{cases}$$

avec

In_i : Intervalle du temps entre deux points successifs

In_{cf} : Intervalle du temps configuré (Par défaut dans l'ENTD, $In_{cf} = 10$ secondes)

v_i : Vitesse fournie par GPS pour chaque point enregistré

$\bar{v}_{(i-1,i)}$: Vitesse moyenne entre deux points successifs

v_{es_i} : vitesse estimée

S_{Ta} : Seuil du temps d'arrêt (Dans notre cas, $S_{Ta} = 120$ secondes)

5.1.1.3 Lissage de données

Les deux méthodes de filtrage décrites ci-dessus ne peuvent pas supprimer tous les points aberrants, en particulier les petites erreurs. Ainsi, le lissage des petites erreurs est une étape nécessaire. La méthode de lissage est résumée simplement comme suit :

Si l'une des deux conditions (a), ((b) ci-dessous est satisfaite :

- (a) La vitesse estimée d'un segment est supérieure à la vitesse moyenne plus 2 fois l'écart type sur n segments avant et également supérieure à la vitesse moyenne plus 2 fois l'écart type sur n segments après ;
- (b) La vitesse estimée d'un segment est en dessous de la vitesse moyenne moins deux fois l'écart type sur n segments avant et également en dessous de la vitesse moyenne moins deux fois l'écart type sur n segments après,

alors la vitesse estimée est recalculé et égale à la moyen arithmétique entre la vitesse moyenne des n segments avant et celle des n segments après ce segment.

L'expression mathématique est donnée par l'équation suivante :

Si

$$\{(v_{es_i} > \bar{v}_{es_{n'}} + 2\sigma_{n'}) \wedge (v_{es_i} > \bar{v}_{es_{n''}} + 2\sigma_{n''})\} \vee \{(v_{es_i} < \bar{v}_{es_{n'}} - 2\sigma_{n'}) \wedge (v_{es_i} < \bar{v}_{es_{n''}} - 2\sigma_{n''})\},$$

alors

$$v_{es_i} = \begin{cases} \bar{v}_{es_{n''}}, & \text{si } n' = 0 \\ \frac{1}{2}(\bar{v}_{es_{n'}} + \bar{v}_{es_{n''}}), & \text{si } (n' \neq 0) \wedge (n'' \neq 0) \\ \bar{v}_{es_{n'}}, & \text{si } n'' = 0 \end{cases}$$

avec

Moyenne de vitesses estimées de n segments précédents :

$$\bar{v}_{es_{n'}} = \begin{cases} \frac{1}{n'} \sum_{k=i-n'}^{i-1} v_{es_k}, & \text{si } n' \neq 0 \\ 0, & \text{si } n' = 0 \end{cases};$$

Moyenne de vitesses estimées de n segments suivantes :

$$\bar{v}_{es_{n''}} = \begin{cases} \frac{1}{n''} \sum_{k=i+1}^{i+n''} v_{es_k}, & \text{si } n'' \neq 0 \\ 0, & \text{si } n'' = 0 \end{cases};$$

Ecart-type des vitesses estimé sur les n segments précédents :

$$\sigma_{n'} = \begin{cases} \frac{1}{n'} \sqrt{\sum_{k=i-n'}^{i-1} (v_{es_k} - \bar{v}_{es_{n'}})^2}, & \text{si } n' \neq 0; \\ 0, & \text{si } n' = 0 \end{cases}$$

Ecart-type des vitesses estimé sur les n segments suivants :

$$\sigma_{n''} = \begin{cases} \frac{1}{n''} \sqrt{\sum_{k=i+1}^{i+n''} (v_{es_k} - \bar{v}_{es_{n''}})^2}, & \text{si } n'' \neq 0; \\ 0, & \text{si } n'' = 0 \end{cases}$$

Nombre des segments précédents :

$$n' = \min \left\{ \begin{cases} i - 1, & \text{si } 1 \leq i \leq \left\lceil \frac{9}{In_{cf}} + 1 \right\rceil \\ \left\lceil \frac{9}{In_{cf}} + 1 \right\rceil, & \text{si } \left\lceil \frac{9}{In_{cf}} + 1 \right\rceil \leq i \leq N \end{cases}, i - i' \text{ si } \ln_{i'} > \frac{\left\lceil \frac{9}{In_{cf}} + 1 \right\rceil}{2} \right\} :$$

Nombre des segments suivants :

$$n'' = \min \left\{ \begin{cases} N - i, & \text{si } N - \left\lceil \frac{9}{In_{cf}} + 1 \right\rceil \leq i \leq N \\ \left\lceil \frac{9}{In_{cf}} + 1 \right\rceil, & \text{si } 1 \leq i \leq N - \left\lceil \frac{9}{In_{cf}} + 1 \right\rceil \end{cases}, i'' - i \text{ si } \ln_{i''} > \frac{\left\lceil \frac{9}{In_{cf}} + 1 \right\rceil}{2} \right\}$$

N: Nombre des enregistrements dans la table

v_{es_i} : vitesse estimée

In_{cf} : Intervalle du temps configuré (Par défaut dans l'ENTD, $In_{cf} = 10$ secondes)

$\lceil \cdot \rceil$: Fonction partie entière par excès.

Nous choisissons la valeur seuil du temps d'arrêt, notée S_{Ta} , égale à 120 secondes.

5.1.1.4 Découpage de la chaîne de données GPS en séquences de mouvement et d'arrêt

Chaque GPS fournit un fichier qui contient tous les points recueillis pour un enquêté pendant toute la période d'enquête. Quand la vitesse enregistrée est faible, on pense que l'enquêté est arrêté. Si la vitesse entre des enregistrements successifs est faible, la durée entre deux enregistrements est additionnée et qualifiée de « temps d'arrêt », jusqu'à ce qu'un prochain enregistrement atteigne une vitesse supérieure. Alors, si le « temps d'arrêt » est supérieur à une valeur seuil, les enregistrements

sont coupés ici. C'est-à-dire qu'on distingue deux traces différentes. Le seuil du temps d'arrêt proposé est 120 secondes.

La méthode de découpage se formalise comme suit

Si

$$\left(\sum_{i=k}^j In_i \geq S_{Ta} \right) \wedge (v_{es_{k-1}} \geq S_v) \wedge (v_{es_i} < S_v \forall i \in [k, j]) \wedge (v_{es_{j+1}} \geq S_v)$$

ou

$$(In_j \geq S_{Ta}) \wedge (v_{es_j} < S_v),$$

alors on découpe l'enregistrement au point j, il s'agit d'une trace différente à partir du point j+1.

Si $(In_j \geq S_{Ta}) \wedge (v_{es_j} \geq S_v)$, alors on découpe l'enregistrement aux points j-2 et j+1. Il s'agit de trois traces différentes : avant j-1, de j-1 à j+1, après j+1.

On définit le seuil de vitesse estimée pour le découpage :

$$S_v = \begin{cases} 1.1, & \text{si } In_i \leq In_{cf} \\ 1.4, & \text{si } In_{cf} \leq In_i \leq \frac{S_{Ta}}{2} \\ 0.00344In_i + 1.5, & \text{si } \frac{S_{Ta}}{2} < In_i \end{cases}$$

Contrôle des exceptions :

Bien que nous ayons filtré les points invalides, estimé et lissé la vitesse, il y a toujours des points exceptionnels où la vitesse est plus grande que le seuil que nous avons défini. En outre, l'arrêt n'est jamais sur un point, les gens bougent dans presque toutes les situations, attente du bus, tennis sur le terrain, sorti dans le jardin etc. Quand la vitesse n'est pas très élevée pendant un court laps de temps, on ignore cette période, et on considère qu'elle fait partie de la période où on reste sur le point d'arrêt. Mais quand la vitesse devient très élevée, comme les points invalides ont été éliminés, il n'est pas possible qu'il s'agisse d'un point sur un lieu d'arrêt.

Nous pouvons accepter quelques points exceptionnels, mais pas tous, donc combien d'exceptions peut-on tolérer ? Comment mesure-t-on les exceptions ? Quel facteur doit être considéré ?

Nous notons $N_{ex}(i)$ le nombre de points exceptionnels au segment i. On distingue trois cas pour calculer le nombre de points exceptionnels :

(a) Si l'intervalle de temps entre deux points successifs In_i est plus petit que l'intervalle de temps configuré In_{cf} , alors nous ajoutons 1 au N_ex .

(b) Si l'intervalle de temps entre deux points successifs In_i est plus grand que l'intervalle de temps configuré In_{cf} et plus petit que la moitié du seuil du temps d'arrêt S_{Ta} , alors nous ajoutons au N_ex une valeur $\sqrt{\frac{In_i v_{es}}{In_{cf} 1.4}}$.

(c) Si l'intervalle de temps entre deux points successifs In_i est plus grand que la moitié du seuil du temps d'arrêt S_{Ta} , alors nous considérons que N_ex est un nombre très grand.

L'expression mathématique pour calculer le nombre de points exceptionnels sera la suivante :

$$N_ex(i) = \begin{cases} N_ex(i-1) + 1, & \text{si } In_i \leq In_{cf} \\ N_ex(i-1) + \sqrt{\frac{In_i v_{es}}{In_{cf} 1.4}}, & \text{si } In_{cf} \leq In_i \leq \frac{S_{Ta}}{2} \\ \infty, & \text{si } In_i > \frac{S_{Ta}}{2} \end{cases}$$

Nous définissons un plafond de nombre de points exceptionnels :

$$\text{plafond_N_ex} = \begin{cases} \frac{8}{\sqrt{In_{cf}}}, & \text{si } In_{cf} < 16 \\ 2, & \text{si } 17 < In_{cf} \end{cases}$$

Si le plafond est dépassé, le segment ne peut plus être considéré comme une partie exceptionnelle dans un temps d'arrêt, mais comme une partie de trace.

Préciser les heures d'arrivée et de départ :

On a besoin de nettoyer les points sur les lieux aux extrémités de la trace afin de préciser les heures d'arrivée et de départ. Quand un enquêté atteint le lieu d'arrivée, il y a deux situations possible. Soit le récepteur GPS ne reçoit aucun signal sur ce lieu, l'heure du dernier enregistrement est l'heure d'arrivée. Soit le récepteur GPS reçoit quand même les signaux de satellites sur le lieu d'arrivée, même si c'est de façon intermittente. Dans la table brute d'une trace, il y a des enregistrements correspondant aux arrêts aux extrémités dont la vitesse doit être faible. Pour cette raison, l'heure d'arrivée doit être l'heure du premier enregistrement avec une faible vitesse. Après cet enregistrement, il faut que tous les suivants soient avec une vitesse faible.

Grâce au contrôle des points exceptionnels, quand le plafond de nombre de points exceptionnels est dépassé et le temps d'arrêt accumulé est supérieur à la valeur seuil de 120 secondes, le premier point exceptionnel sera le début d'une nouvelle trace. Sinon, il s'agit d'une partie de la même trace. L'heure enregistrée pour le premier point exceptionnel représentera raisonnablement l'heure de départ de la trace.

5.1.1.5 Estimation des données manquantes

On définit trois tranches pour la durée de perte du signal : courte, longue et très longue.

(a) Si dans la période de perte du signal, le statut (mouvement, arrêt) de l'enquêté varie peu ou que ce changement n'influence pas l'algorithme de découpage, on dit que la période est courte, par exemple quelques secondes perdues sur autoroute.

(b) Si dans la période où les signaux sont perdus, on ne peut plus être certain du statut (mouvement, arrêt, variation significative de la vitesse) de l'enquêté, on dit que la période est longue, par exemple une trentaine de minutes perdu dans le métro.

(c) S'il y a des déplacements ou activités totalement perdus, on dit que la période de perte du signal est très longue.

Cas 1 : Les signaux sont perdus pendant un temps court

Ce cas correspond à condition de $In_{cf} \leq In_i \leq \frac{S_{Ta}}{2}$, alors la vitesse estimée est la moyenne pondérée de la vitesse instantanée mesurée par le GPS et de la vitesse moyenne sur segment calculée sur les coordonnées de deux points consécutifs. Si cette vitesse est inférieure à 1.4 km/h, cette période sera interprétée comme un arrêt, sinon elle sera intégrée à la trace en mouvement.

Cas 2 : Si les signaux sont perdus pendant un temps long,

Nous pouvons distinguer 4 cas : le trou se produit pendant la trace ou une période d'activité, au début/ à la fin de la trace.

i) Dans le premier cas (le trou survient pendant la trace), la vitesse moyenne dans ce trou est élevée. Cette situation intervient dans le cas de mauvaises conditions d'utilisation du GPS comme par exemple lors du passage dans un tunnel, dans un bus/ train loin de la fenêtre. A l'aide de méthode d'estimation de la vitesse réelle et du découpage des enregistrements par trace, ce trou sera intégré à la trace ou séparé comme une trace complète, puis relié avec les traces précédente et suivantes pour reconstruire un déplacement.

ii) Dans le deuxième cas (le trou se produit pendant une période d'activité) si le début d'une trace est identique au point d'arrêt précédent ou à la fin de la trace précédente, la période sans données est une partie ou la totalité d'un point d'arrêt. Ce phénomène correspond plutôt à la perte du signal dans un bâtiment comme dans un bureau, au domicile, dans un centre commercial, etc. Pour vérifier ce cas, nous devons examiner si les points avant et après sont bien situés au même endroit.

iii) Dans le troisième cas (le trou se produit au début de la trace), quand le début d'une trace est assez éloigné mais pas trop du point d'arrêt précédent ou de la fin de la trace précédente et la vitesse instantanée au début de cette trace est très élevée. Ce phénomène dans la plupart des cas correspond à un temps assez long de démarrage du GPS pour reprendre les signaux après les avoir perdus, etc. Nous avons besoins de recalculer le moment où l'enquêté était immobile en rétropolant la vitesse au début de la trace et la distance éventuellement parcourue pour atteindre cette vitesse. Nous choisissons une distance seuil de 150 mètre sur laquelle on joue si cette trace commence au point d'arrêt précédent ou à la fin de la trace précédente.

La décision de prolongation de la trace au point d'arrêt précédent dépend de la distance entre le premier point enregistré sur la trace et le lieu d'arrêt, de la vitesse au début de la trace, et aussi du temps écoulé entre le premier point enregistré sur la trace et le dernier point sur le lieu d'arrêt. On prolonge le temps et la distance au début de la trace :

$$\text{Temps à compléter } T_c = \begin{cases} 0, & \text{si } \bar{v}_d \in [0; 1.8] \\ 1.67\bar{v}_d, & \text{si } \bar{v}_d \in [1.8; 18] \\ 1.25\bar{v}_d + 7.5, & \text{si } \bar{v}_d \in [18; 54] \\ 0.395\bar{v}_d + 53.68, & \text{si } \bar{v}_d \in [54; 130] \\ 105, & \text{si } \bar{v}_d \in [130; +\infty) \end{cases}$$

$$\text{Distance à compléter } D_c = \begin{cases} 0, & \text{si } \bar{v}_d \in [0; 1.8] \\ 0.321\bar{v}_d^2, & \text{si } \bar{v}_d \in [1.8; 18] \\ 0.24\bar{v}_d^2 - 4.33\bar{v}_d + 104, & \text{si } \bar{v}_d \in [18; 54] \\ 0.076\bar{v}_d^2 - 4.1\bar{v}_d + 570, & \text{si } \bar{v}_d \in [54; 130] \\ 1211.4, & \text{si } \bar{v}_d \in [130; +\infty) \end{cases}$$

avec \bar{v}_d : la vitesse au début de la trace (km/h)

La vitesse au début de la trace doit être la vitesse moyenne observée pendant un laps de temps assez court. On note ce temps au début de la trace comme T_{dt} .

$$T_{dt} = \begin{cases} 6, & \text{si } In_{cf} = 1 \\ 3In_{cf}, & \text{si } In_{cf} \geq 2 \end{cases}$$

Néanmoins, dans les applications, nous n'arrivons pas à prendre la valeur exactement à T_{dt} , à cause des intervalles de temps irréguliers.

1) Si les intervalles de temps du premier point au quatrième point sont tous une seconde, et l'intervalle du quatrième au cinquième point égale 10 secondes, nous ne prenons que les quatre premiers points pour calculer la vitesse moyenne.

2) Si l'intervalle de temps enregistré entre les deux premiers points est déjà supérieur à 6 secondes, nous sommes obligés de prendre la vitesse calculée sur ces deux points comme vitesse au début de la trace « \bar{v}_d ».

3) Si la durée total de la trace est de moins de 6 secondes, nous calculons la vitesse moyenne sur la trace comme T_{dt} .

Après avoir compensé en distance et en temps la partie éventuellement perdu au début de la trace, on doit fixer la valeur seuil de distance. Si la distance entre le premier point de la trace et le lieu d'arrêt est plus petite que la valeur seuil de distance plus la distance complétée, on prolonge la trace jusqu'au lieu d'arrêt précédent.

Si $D_{tl} - D_c < S_D$, alors nous prolongeons la trace jusqu'au lieu d'arrêt précédent.
avec

$$\text{le seuil de distance (unité: mètre) } S_D = \begin{cases} 150, & \text{si } T_{tl} < 0 \\ 1000, & \text{si } T_{tl} - T_c < 0 \\ 1000 - 4.1(T_{tl} - T_c), & \text{si } 0 < T_{tl} - T_c < 120 \\ \frac{360000}{(T_{tl} - T_c) + 600}, & \text{si } 120 < T_{tl} - T_c \leq 1800 \\ 150, & \text{si } T_{tl} - T_c > 1800 \end{cases}$$

D_{tl}, T_{tl} (unité: mètre, seconde): la distance et le temps entre le premier point de la trace et le dernier enregistrement sur lieu d'arrêt;

D_c, T_c : la distance et le temps à compléter au début de la trace;

iv) Dans le quatrième cas (trou situé à la fin de la trace). Après avoir pris le métro, le train, le bus ou la voiture, l'enquêté peut marcher à pied pendant un laps de temps. Sans aucun doute, il y a des cas où on a perdu les signaux à la fin de la trace à cause de batterie faible, d'une mauvaise utilisation du GPS, etc. En effet, il est plus fréquent que les signaux perdus en fin de trace contiennent une partie ou la totalité d'un point d'arrêt, quelquefois jusqu'au début de la trace suivante. En réalité, il est impossible de distinguer dans quelle partie de la trace se situent les signaux perdus. Nous allons donc prendre les coordonnées du lieu d'arrivée comme celles du lieu d'arrêt. Puis nous traitons les données manquantes dans le cas où le trou se produit au début de la trace.

Cas 3 : Les signaux sont perdus pendant un temps très long

Il y a éventuellement des déplacements ou activités qui ne sont pas enregistrés. Ce phénomène peut être entraîné par la défaillance humaine : oubli d'allumer/ recharger/ porter le GPS ou positionnement du récepteur dans de mauvaises conditions. L'estimation automatique de cette partie est impossible, on peut considérer chaque cas précis pour deviner l'origine et la destination des déplacements à partir de l'heure où les données manquantes ont lieu, de l'habitude de transport de l'enquêté. A partir de la destination et de l'origine, on peut estimer approximativement le temps de parcours et la distance de déplacement.

5.1.1.6 Imputation du mode de déplacement

Dans les enquêtes traditionnelles avec CAPI, les enquêtés informent les modes de transport pour chaque déplacement. Malheureusement, les informations de mode de transport ne sont pas fournies dans l'enquête à l'aide du récepteur GPS.

La construction d'algorithmes pour identifier les modes de déplacements obtenus par le récepteur GPS à partir de vitesse a été réalisée à l'IFSTTAR (Yuan, 2010). Dans cet algorithme « imputation des modes », afin de distinguer les modes, on a sélectionné 8 indices pour déterminer le mode de transport qui reflètent les différentes caractéristiques des divers modes de transport

- la vitesse moyenne sans arrêt sur la trace,
- le 95^{ème} percentile de la vitesse,
- le 95^{ème} percentile de l'accélération,
- la durée de la trace,
- la plus longue période de signaux perdus pendant la trace
- le ratio de signaux perdus (voir glossaire)

On a limité l'analyse à six modes de déplacement : la marche à pied, le vélo, la voiture/moto, l'autobus, le métro, le train, l'avion.

Table 5.1 : Les valeurs des indices appliquées pour déterminer le mode de transport dans la thèse de Yuan.

	Vitesse moyenne sans arrêt (km/h)	Vitesse médiane avec arrêt (km/h)	95% vitesse maximum (km/h)	95% accélération maximum (m/s ²)	Qualité des données	Durée de la trace (s)	Période signaux perdus (s)	Ratio de signaux perdus
Marche à pied	2--7	1,5--7	3--12	<4	>0,15	>30	<300	<0,5
En vélo	7--20	6--20	12--24	<4	>0,3	>30	<300	<0,5
Voiture/Moto	9--120	1,5--120	24--200	2--15	>0,3	>60		

Bus	8--80	6--80	15--100	<15	<0,3	>120	10-120	
Méto	7--50		7--80		<0,05	>240	>180	
En train	50--330		200-1500		<0,01	>1800	>1200	
En avion	200-800		50-400		<0,1	>1800	>600	

Source : Imputation de mode de transport (Yuan, 2010).

Ces indices sont variables et n'appartient pas totalement dans un fixe intervalle. Ces indices dépendent aux véhicules de transport, la santé, le temps, le GPS, la circulation de transport (en ville, en autoroute, la fréquence des feux, de priorités à droite forçant à ralentir, la présence de bandes cyclables, et la largeur des voies en cas d'embouteillage, ...). Par exemple, normalement, la vitesse moyenne de vélo est de 7 à 10km/h ; mais elle peut atteindre de 16 à 20km/h pour le cycliste non sportif ; de 20 à 25 km/h pour le sportif occasionnel, bonne forme sur route dégagée et dans de bonnes conditions. D'autre exemple, la vitesse de moyenne de voiture est de 12 à 20 km/h en ville, mais dans le cas de bouchons, voies étroites, la vitesse moyenne de voiture diminue de 1 à 3km/h qui correspond à la celle de Marche à pied. En plus, l'algorithme d'imputation pourrait ne marcher pas bien si l'une des données de ces indices est manquante (par exemple, quand on se déplace sous un bâtiment, dans un tuyau de métro, dans environnement avec beaucoup de bâtiments de grande hauteur, ...).

Suite au protocole de l'ENTD 2007-2008, les courtes questions complémentaires du CAPI-GPS nous fournir les informations de mode de transports de tous déplacements dans un jour choisir pour évaluer l'algorithme d'imputation de mode de transport. Nous notons que le nombre de déplacement avec les modes de bus, métro, train est trop petit, donc nous allons regrouper ces modes par un mode « Transport Commun ». Dans les déplacements quotidiens du questionnaire classiques et les déplacements déclarés par l'enquêté dans le questionnaire complémentaire CAPI-GPS, il n'y a pas de déplacement aérien, donc ça nous entraine de ne pas vérifier la précision de l'algorithme de l'imputation de mode aérien, donc nous allons négliger ce mode aérien. Nous construisons un algorithme d'imputation de 4 modes : Marche à pied, Voiture/Moto, Transport Commun (Bus, Métro, Train), Vélo, Autre en utilisant le modèle probabilité conditionnelle.

Nous nous intéressons à constituer un modèle pour décrire la relation entre la variable de réponse « MODE » et certaines des variables explicatives, comme vitesse maximum, vitesse minimum, distance linéaire, durée de trace GPS, ainsi que distance moyenne quotidienne de l'enquête réalisés la veille de l'interview, le temps quand l'enquêté se déroule avec vélo, ...

Nous prenons les probabilités :

$\theta_1 = \pi_1$, probabilité de « Autre »,

$\theta_2 = \pi_1 + \pi_2$, probabilité de « Autre » ou « Marche à pied »,

$\theta_3 = \pi_1 + \pi_2 + \pi_3$, probabilité de « Autre » ou « Marche à pied » ou « Voiture/Moto »,

$\theta_4 = \pi_1 + \pi_2 + \pi_3 + \pi_4$, probabilité de « Autre » ou « Marche à pied » ou « Voiture/Moto » ou « Transport commun »,

où

π_1 = probabilité de « Autre »,

π_2 = probabilité de « Marche à pied »,

π_3 = probabilité de « Voiture/Moto »,

π_4 = probabilité de « Transport commun »,

π_5 = probabilité de « Vélo »,

$$\text{logit}(\theta_1) = \log(\theta_1/(1 - \theta_1)) = \log(\pi_1/(\pi_2 + \pi_3 + \pi_4 + \pi_5)),$$

$$\text{logit}(\theta_2) = \log(\theta_2/(1 - \theta_2)) = \log((\pi_1 + \pi_2)/(\pi_3 + \pi_4 + \pi_5)),$$

$$\text{logit}(\theta_3) = \log(\theta_3/(1 - \theta_3)) = \log((\pi_1 + \pi_2 + \pi_3)/\pi_4 + \pi_5).$$

$$\text{logit}(\theta_4) = \log(\theta_4/(1 - \theta_4)) = \log((\pi_1 + \pi_2 + \pi_3 + \pi_4)/\pi_5).$$

Donc nous recalculons les probabilités à partir des probabilités cumulées comme suit

$$\pi_1 = \theta_1,$$

$$\pi_2 = \theta_2 - \theta_1,$$

$$\pi_3 = \theta_3 - \theta_2,$$

$$\pi_4 = \theta_4 - \theta_3,$$

$$\pi_5 = 1 - \theta_4.$$

En retenant les variables significatives, le tableau 5.2 montrent que les facteurs, qui jouent un rôle dans l'imputation de mode de transport de déplacement, sont : vitesse maximum, vitesse minimum, distance linéaire, durée de trace GPS, ainsi que distance moyenne quotidienne de l'enquête réalisés la veille de l'interview, le temps quand l'enquêté se déroule avec vélo. Toutefois la distance parcourue, le sexe, l'âge, l'état de santé, le goût pour la conduite (aimer vélo/ voiture/ Moto), l'équipement en automobile (possession de voiture/ moto/Vélo de ménage), avoir permis de conduite ne semblent pas avoir d'influence.

Tableau 5.2 : Analyse des effets de facteur dans l'algorithme d'imputation de mode

Facteur	DF	Wald Khi-2	Pr > Chi-2
Vitesse maximum	2	35.7459	<.0001
Distance linéaire	2	43.2646	<.0001
Durée	3	14.4783	0.0023
Quand l'enquêté se déroule avec vélo	3	12.0333	0.0073
Vitesse minimum	2	8.6202	0.0134

Distance moyenne quotidien de l'enquêté la veille de l'interview	4	12.6939	0.0129
--	---	---------	--------

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Tableau 5.3 : Analyse des odds ratios de facteur dans l'algorithme d'imputation de mode

Variabes	Modalités	Rapport des cotes (odds ratios)	Pr > Khi-2
Vitesse maximum (km/h)	1--7	19.548	<.0001
	7--20	0.588	0.0009
	>20	Référence	
Vitesse minimum (km/h)	1--7	0.340	0.0033
	7--20	1.857	0.0468
	>20	Référence	
Distance linéaire (km)	0--1	16.147	<.0001
	1--6	1.718	0.0018
	>6	Référence	
Durée (min)	0--5	0.212	0.0004
	5--15	0.575	0.4843
	5--30	0.528	0.8335
	>30	Référence	
Quand l'enquêté se déroule avec le vélo	Il y a moins d'une semaine	0.230	0.0010
	Entre une semaine et un mois	0.873	0.3916
	Entre un mois et plus de deux ans	0.947	0.0621
	Jamais	Référence	
Distance moyenne de déplacements quotidiens la veille de l'interview (km)	0	0.521	0.0056
	0--10	0.968	0.9793
	10--20	0.947	0.9568
	20--40	1.723	0.0074
	>40	Référence	

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

A l'aide des appariements entre les traces collectées par GPS et les déplacements quotidiens dans l'enquête classique d'une part, et les questions complémentaires du CAPI-GPS d'autre part, nous pouvons évaluer l'algorithme « imputation de mode ». Il est certain que la correspondance n'est pas parfaite : il y a toujours des cas qu'on ne peut pas traiter. Nous rencontrons également des difficultés pour les transports en communs tels que bus, métro ou train qui sont des modes peu fréquents. Il n'est d'ailleurs pas très facile de distinguer métro, train ou bus avec Voiture/Moto. Nous allons considérer le type de point d'intérêt (POI) pour mieux distinguer ces modes de transport. Par contre,

la distinction entre marche à pied et voiture/moto est satisfaisante. Les résultats de l'algorithme "imputation des modes" sera étudié en détail dans la partie 5.2.

Pour améliorer l'algorithme de l'imputation de mode de transport, lors d'une enquête légère CAPI-GPS en complément, on peut demander :

- Quand avez-vous fait du vélo pour la dernière fois ?
- Quand avez-vous fait de voiture/moto pour la dernière fois ?
- Quand avez-vous fait de transport commun (bus, métro, train, avion) pour la dernière fois ?
- Quelle moyenne de transport utilisez - vous fréquemment pour aller au travail/ à l'école/ au super marché ?

5.1.1.7 Repérage des points d'arrêt pour identifier les lieux visités

Par une méthode géométrique, on regroupe en un même lieu les points voisins trouvés pendant la période d'enquête. On identifie notamment quelques points remarquables tels que domicile, lieu de travail ou station de métro.

Combinaison des mêmes points d'arrêt

On commence au premier point d'arrêt qui est le lieu de départ de la première trace. Ce point est normalement le domicile parce que la visite au cours de laquelle on remet le GPS a lieu chez l'enquêté. Prenant ce point comme le centre que l'on appelle la semence, on dessine un cercle avec un rayon fixé à 45 mètres. S'il y a d'autres points inclus dans le cercle, on calcule le centre de gravité de ces points. Puis prenant ce centre de gravité comme centre, on dessine un autre cercle de même rayon. Si les points dans le nouveau cercle ont varié, on recalcule le centre de gravité, et on trace à nouveau le cercle... jusqu'à ce que les points dans le cercle ne changent plus, ni de nouveaux inclus, ni les anciens exclus. Comme cela, on trouve un cercle qui contient le maximum de points.

On considère le point visité (PV) qui est un ensemble de point d'arrêt (pa) dont la distance au centre du dernier cercle (cc_d) est inférieure à 45 mètres.

$$PV = \{pa \mid D(pa, cc_d) < 45\}$$

On adopte le centre du dernier cercle comme la position du point visité, et les points dans le dernier cercle sont considérés comme un lieu unique.

Le centre du dernier cercle est déterminé comme suit

$$PV_0 = \{pa \mid D(pa, cc_d) < 45\}; cc_i = \overline{pa}, pa \in PV_{i-1}; PV_i = \{pa \mid D(pa, cc_i) < 45\}$$

Si $PV_{i-1} = PV_i$ alors $cc_d = cc_i; i \geq 1$

Tableau 5.2 : Caractérisation des points d'arrêts en fonction de leur durée

Temps d'arrêt	Pas le même jour	> 3h	1h-3h	20m-1h	400s-20m	2m-400s	< 2m	Autre
Cat	1	2	3	4	5	6	7	8
	Passer la nuit au domicile ou à l'hôtel	Rester chez soi, travailler au bureau	Faire du shopping, travailler au bureau	Manger au restaurant, achat au supermarché	Attente de bus	Arrêt court : Attente de bus, achat rapide dans un magasin	Signal perdu > 300s et vitesse élevée	Déplacements manquants

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

5.1.2 Couche "Point d'Intérêt" (POI)

En considérant les points d'arrêt, on cherche le POI le plus proche des points visités. On vérifie qu'il est cohérent avec le POI précédent et le suivant, avec les modes de transport imputés pour les étapes correspondantes, ainsi qu'avec la durée de séjour en ce point. Le POI trouvé nous aide à identifier un point de correspondance (changement de mode au cours d'un déplacement) ou l'activité pratiquée, à la destination du déplacement qui indique son motif.

Une correction sur un POI peut conduire à reconsidérer le mode de transport qui est la sortie de la première couche.

- **Map-matching** : On rapproche les séquences de mouvement des arcs du réseau routier ou ferroviaire. Cet appariement aide à corriger le mode de transport.
- **Combinaison des segments en déplacement** : En fin nous relient les segments entre deux activités de manière à reconstituer un déplacement qui soit le plus conforme possible à la définition habituelle, c'est-à-dire notamment en rassemblant les segments séparés par un point de correspondance, car plusieurs modes peuvent être utilisés au cours d'un déplacement, qui est défini par l'activité à destination.

On choisit le point d'intérêt (POI) le plus proche de la position déterminée pour le point visité. La distance entre ces deux points doit être inférieure à 45 mètres.

$$POI \text{ choisi} = \text{Min}_{POI} \{D(POI, cc_d) < 45\}$$

5.2 Evaluation de l'algorithme DECOUTRACE par le court questionnaire complémentaire CAPI-GPS

A l'aide des appariements entre les traces GPS obtenues par le post-traitement Decoutrace (notées « traces GPS ») et les déplacements quotidiens décrits par l'enquêté dans les questions complémentaires du CAPI-GPS (notés « déplacements CAPI-GPS »), nous pouvons évaluer les résultats de l'imputation du mode par l'algorithme DECOUTRACE. Nous trouvons 841 couples de trace GPS et déplacement CAPI-GPS qui comportent des informations sur le mode de déplacement déclarés par l'enquêté dans CAPI-GPS et celui imputé par le post-traitement des traces GPS. Il est certain que la correspondance n'est pas parfaite : il y a toujours des cas qu'on ne peut pas traiter. Dans les traces CAPI-GPS, il n'y a pas de déplacement aérien ou par train, donc le rapprochement n'a pas pu être testé ; en outre, le petit nombre de déplacements utilisant les transports publics (Bus, Métro) déclarés par l'enquêté rend difficile le calibrage des paramètres correspondant à ces modes dans notre algorithme. Il n'est d'ailleurs pas très facile de distinguer métro ou bus avec Voiture/Moto. Par contre, la distinction entre marche à pied et voiture/moto est satisfaisante. En général, le taux d'imputation est de 90%. Ces taux ; qui sont 97% pour voiture/moto, 86% pour le vélo et 48% pour Marche à pied ; sont donc peu supérieurs à ceux que l'on trouvera au chapitre 6 par comparaison avec la description des déplacements quotidiens dans le questionnaire principal de l'ENTD 2007-2008.

Tableau 5.10 : Comparaison du mode entre déplacements CAPI-GPS et traces GPS

Mode de transport déclaré par l'enquêté dans CAPI-GPS	Mode de trace GPS imputé par le programme					
	Marche à pied	Vélo	Voiture/ Moto	Transport commun	Autre mode	Total
Marche à pied	43 (48%)	0	45	0	1	89
Vélo	0	6 (86%)	1	0	0	7
Voiture/ Moto	17	2	700 (97%)	1	1	721
Transport commun	0	0	14	2 (13%)	0	16
Autre mode	2	0	1	0	5 (63%)	8
Total	62	8	761	3	7	841 (90%)

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

CONCLUSION

Le post-traitement pour l'identification des traces et l'imputation des modes de transport à partir des données brutes GPS est un travail complexe. Il s'agit, en plusieurs étapes, de passer en revue les

déplacements et les informations complémentaires, d'examiner chacun et de les réarranger par rapport à la chronologie. Aucune étape n'est indépendante : le résultat de chaque étape peut modifier une autre. L'imputation des motifs de déplacement est trop incertaine pour avoir été tentée ; l'identification des « Points d'intérêt » (POI) et la durée de visite de ces points ne sont pas suffisantes pour identifier de manière satisfaisante les motifs/activités. Il faudrait expérimenter un court questionnaire permettant d'identifier les principaux lieux fréquentés pour les principaux types d'activité. En outre, chaque récepteur GPS a ses propres caractéristiques et le système de transport peut varier d'une ville à l'autre : certains variables explicatives et paramètres mentionnés dans cette thèse peuvent donc ne pas convenir à d'autres projets ou contextes. Ils doivent être ajustés. Le post-traitement des données GPS (sans les informations du questionnaire complémentaire CAPI-GPS) fournit une bonne image de la séquence des traces et une imputation satisfaisante des modes de déplacement pour la marche à pied et le mode « moto/voiture ». Mais la difficulté subsiste pour distinguer voiture et transport commun (bus, métro, train) parmi les modes routiers ; et le GPS ne fournit aucune information sur les conducteurs / passagers dans les véhicules, ce qui serait important pour l'analyse du taux d'occupation des véhicules (facteur majeur de la consommation de carburant par passager x km). Cela montre à nouveau que les enquêtes par GPS ne peuvent pas encore remplacer complètement les méthodes classiques : elles ne permettent bien sûr pas d'obtenir les caractéristiques socio-économiques des ménages et les caractéristiques techniques de leurs véhicules, les motifs et modes de déplacement, le nombre de personnes accompagnant l'enquêté, etc. Ces nouvelles formes d'enquête sont pour l'instant utilisées en complément des questionnaires classiques.

L'appariement des données collectées par GPS et par CAPI est l'outil de choisir les variables explicatives significatives dans modèle probabiliste conditionnelle de l'algorithme de post-traitement des données GPS et de l'évaluation de résultats statistique de l'enquête de transport. Les caractéristiques techniques de l'Enquête Nationale de Transport peuvent provoquer certains problèmes et compliquer la validation du volet GPS. Par exemple, il y a une différence de définition du jour d'enquête entre les deux types de méthodes, pour l'enquête classique, le jour commence de 4 heure jusqu'à 4 heure le lendemain, mais pour l'enquête GPS, c'est de 0h à minuit. En outre, le traitement des données GPS et l'appariement avec les déplacements quotidiens a été rendue plus difficile par le codage des lieux d'origine et de destination collectées par l'enquête classique au niveau communal, ce qui est moins précis que les coordonnées X et Y.

Le Système d'Information Géographique et le map-matching de Mappoint 2006 n'a pas fourni correctement les points d'intérêt pour le post-traitement Decoutrace des données brutes GPS. L'importance de la précision pour la cartographie du réseau routier/autoroutier et des réseaux de transport en commun ne sera jamais assez soulignée. Un bon réseau de transport en commun ou ferroviaire sont une nécessité absolue pour améliorer l'imputation des mode et motif lors du post-traitement des données GPS.

6 Validation des résultats par confrontation avec les données du questionnaire classique de l'ENTD 2007-2008

Un problème habituel dans les enquêtes sur la mobilité quotidienne conduites avec une méthode classique est la sous-estimation du nombre de déplacements et l'inexactitude de la déclaration de la durée et de la distance de déplacements. Plusieurs méthodes ont été développées pour dépasser ces limites. L'utilisation du récepteur GPS dans une enquête transport fournit une opportunité pour estimer les différences quant au nombre, à la durée, à la distance, aux heures de départ et d'arrivée des déplacements enregistrés par ces deux méthodes.

La collecte de données de mobilité par récepteur GPS a été réalisée pour la première fois en 1997 à Austin – Texas (Casas, Arce, 1999) et depuis lors, il y a eu plusieurs recherches sur l'utilisation de cette technologie pour les enquêtes sur la mobilité. En exploitant les données collectées par GPS, on peut montrer que le taux de déplacements manquants dans l'enquête classique peut être notable. A partir de données de 6 enquêtes de transport utilisant CATI et GPS aux Etats Unis entre 2001 et 2004, Wolf (2004) a montré que ce taux de déplacement manquant dans l'enquête classique peut aller de 11% à 81%. Toutes ces études considèrent qu'il n'y a pas de déplacements manquants dans les relevés GPS, ce qui signifie que ces taux devraient être plus élevés si l'on prend en compte les biais dans les deux sens.

Forrest et Pearson (2005), ont indiqué que dans l'enquête transport menée à Laredo – Texas, seulement 44 % des déplacements enregistrés par récepteur GPS ont pu être identifiés avec un déplacement déclaré par CATI. Ceci est un déficit significatif considérant que ces enquêtes sont utilisées pour la modélisation et la planification des transports.

Stopher (2007) a trouvé, en combinant le GPS et un rappel téléphonique, que le taux de déplacement non déclaré par cette méthode classique est seulement de 7,4 % ; mais on ne tient pas compte dans ce test des déplacements qui ont échappé au GPS à cause de multiples raisons, par exemples : oubli de porter ou d'allumer, non réception des signaux.

6.1 Panorama de l'appariement des traces GPS et déplacements quotidiens

A la deuxième visite, l'enquêté volontaire avec GPS répond aussi au questionnaire CAPI sur la description de sa mobilité quotidienne, c.à.d. les déplacements réalisés un jour de semaine (veille du passage de l'enquêteur) et au cours du dernier week-end. Grâce à ces sources de données, nous pouvons comparer les descriptions de la mobilité obtenues par deux méthodes (CAPI et GPS), sur les jours couverts par ces deux instruments d'observation.

La définition d'un jour entre l'enquête par CAPI et par GPS doit alors correspondre aux mêmes plages horaires. En effet, pour faciliter l'étude de la mobilité des personnes, les déplacements quotidiens sont décrits de 04h00 à 04h00 du lendemain. Le premier déplacement d'un jour est le premier qui se termine après 4h (soit commencé après 04h00, soit commencé avant 04h00 et terminé après 04h00). Le dernier déplacement doit finir avant 3h59 le lendemain, s'il finit après 04h00, ce déplacement est affecté au jour suivant. Par exemple : sortir de boîte de nuit à 3h 30 dimanche matin et arriver à 4h15 chez soi, correspond au premier déplacement du dimanche. Mais les données de traces GPS sont enregistrées en continu avec la chronologie habituelle, c.à.d. qu'un jour commence à 0h00 et se termine à minuit (24h00). Pour confronter ces deux sources de données, il faut veiller à ce que les tranches horaires soient bien les mêmes, quitte à décaler un jour enregistré par le récepteur GPS.

La comparaison entre déplacement quotidien déclaré par CAPI et la trace GPS a été effectuée automatiquement en s'appuyant sur leur chronologie et contrôlée manuellement. Pour trouver une/des trace(s) GPS correspondant à un déplacement quotidien, d'abord nous gardons seulement les deux traces GPS pour lesquelles soit l'heure de début est la plus proches de l'heure de départ du déplacement, soit l'heure de fin est la plus proche de l'heure d'arrivée du déplacement. Nous gardons aussi les traces GPS situées entre les traces trouvées ci-dessus. Pour chaque trace GPS, nous exécutons le même algorithme en échangeant les rôles des déplacements et des traces, afin de trouver un/des déplacements correspondant à la trace GPS considérée.

Nous disposons à la fois de la description des déplacements quotidiens déclarée par CAPI et des traces collectées par GPS pour 327 enquêtés volontaires sur 529 jours. Nous recensons au total 2294 déplacements quotidiens déclarés par CAPI et 2851 traces GPS. Parmi les 2294 déplacements quotidiens observés par CAPI, il y a 1288 déplacements quotidiens dont chacun apparié avec une seule trace GPS et 1006 déplacements quotidiens auxquels ne correspond aucune trace GPS. Parmi les 2851 traces collectées par GPS, il y a 1563 traces non appariées avec des déplacements quotidiens.

Tableau 6.1 : Appariement entre traces GPS et déplacements quotidiens décrits par CAPI

	Apparié	Non apparié	Total
Nombre de déplacements quotidiens	1288	1006	2294

Nombre de Traces GPS	1288	1563	2851
----------------------	------	------	------

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

6.2 Analyse des déplacements quotidiens/ traces GPS manquants

Après avoir apparié les traces GPS et les déplacements quotidiens, nous supposons que 1006 déplacements quotidiens non appariés sont des traces GPS manquantes et 1563 traces GPS non appariées correspondent à des déplacements quotidiens manquants. Le « Total déplacements quotidiens » est la somme des déplacements quotidiens déclarés par CAPI et des déplacements quotidiens manquants dans le CAPI, mais repérés par GPS. Les déplacements quotidiens manquants correspondent donc à la différence entre le « Total Déplacement quotidien » et le nombre total de déplacements quotidiens observés par CAPI. Similairement, le « Total trace GPS » est la somme des traces relevées par GPS et des traces manquantes par GPS mais recueillies dans le CAPI.

Le Tableau 6.2 résume l'analyse des déplacements quotidiens /traces GPS manquants qui ont été mis en évidence. Sur un « total » de 3857 déplacements quotidiens enregistré par les deux méthodes, 1536 ont été manquées par CAPI. Pour l'enquête classique CAPI, le taux de nombres des déplacements/traces de la sous-déclaration est donc de 41%. Sur un « total » de 3857 traces GPS enregistré par les deux méthodes, 1006 ont été manquées par GPS. Pour l'enquête à l'aide du GPS, le taux de nombres des déplacements/traces de sous-déclaration est seulement 26%.

Tableau 6.2 : Nombres des déplacements et traces manquants

Nombres des déplacements et traces	Déplacements/traces observés	Déplacement/trace manquant	Total Déplacement/trace	% Nombres des déplacement /trace manquant
Par CAPI	2294	1563	3857	41%
Par GPS	2851	1006	3857	26%

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Dans l'enquête classique CAPI, la distance linéaire de déplacement quotidien est la distance à vol d'oiseau entre la commune de départ et la commune d'arrivée. Dans l'enquête GPS, la distance linéaire de déplacement est la distance réelle entre les positions de départ et d'arrivée. Dans l'enquête classique CAPI, la distance linéaire de déplacement quotidien est la distance à vol d'oiseau entre la commune de départ et la commune d'arrivée. Dans l'enquête GPS, la distance linéaire de déplacement est la distance réelle entre les positions de départ et d'arrivée. Le Tableau

6.3 montre que les taux de sous-déclaration des distances linéaires de déplacements sont respectivement 19% et 31% pour l'enquête classique CAPI et pour l'enquête à l'aide du GPS. Suite à l'information de distance parcourue de déplacement du Tableau 6.4, le taux de sous-déclaration des longueurs de déplacements est 20% pour l'enquête classique CAPI et 34% pour l'enquête à l'aide du GPS. Dans le Tableau 6.5, pour la caractéristique de durée de déplacement, nous voyons que le taux de sous-déclaration est 31% pour l'enquête CAPI ; et 38% pour l'enquête GPS.

Nous remarquons que le taux de sous-déclaration de nombre de déplacement de l'enquête CAPI est plus grand que celui de l'enquête GPS ; tandis que les taux de sous-déclaration des durée, distance parcourue et distance linéaire de l'enquête CAPI sont plus petits que ceux de l'enquête GPS. Cela laisse supposer que les déplacements quotidiens manquants pour l'enquête CAPI qui sont imputés dans l'enquête GPS pourraient avoir des petits distances et durées.

Tableau 6.3 : Distance linéaire des déplacements et traces manquants

Distance Linéaire (km)	Déplacements/traces observés	Déplacement/trace manquant	Total Déplacement/trace	% Distance linéaires des déplacement /trace manquant
Par CAPI	13774.94	3257.27	17032.21	19%
Par GPS	12440.99	5720.19	18161.18	31%

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Tableau 6.4 : Distance des déplacements et traces manquants

Distance (km)	Déplacements/traces observés	Déplacement/trace manquant	Total Déplacement/trace	% Distance des déplacement /trace manquant
Par CAPI	18779.20	4666	23445.20	20%
Par GPS	16376.16	8290	24666.16	34%

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Tableau 6.5 : Durée des déplacements et traces manquants

Durée (minute)	Déplacements/traces Observés	Déplacement/trace Manquant	Total Déplacement/trace	% Durée des déplacement /trace Manquant
Par CAPI	24421	10820	35241	31%

Par GPS	21631	13197	34828	38%
---------	-------	-------	-------	-----

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

6.3 Identification des déplacements quotidiens non appariés avec des trace GPS

Nous avons dénombré précédemment 2294 déplacements quotidiens déclarés par CAPI dont 1006 déplacements ne sont pas appariés avec des trace GPS. Pour mieux comprendre les facteurs affectant l'appariement de déplacements quotidiens avec des traces GPS, nous estimons une régression logistique binaire de l'appariement sur l'échantillon de 2294 déplacements quotidiens observés par CAPI.

Le Tableau 6.6 montre que les facteurs qui jouent un rôle dans le non appariement des déplacements quotidiens avec les traces GPS sont : la longueur du déplacement (distance parcourue), le motif et le nombre de déplacements réalisés la veille de l'interview et aussi les caractéristiques de l'enquêté telles que son sexe, sa profession, son état de santé, le type du ménage auquel il appartient, la motorisation du ménage (nombre voiture dans le ménage), le mode de déplacement. Toutefois, l'âge de l'enquêté et sa zone de résidence ne semblent pas avoir d'influence.

Table 6.6 : Facteurs de l'appariement de déplacements quotidiens avec des traces GPS

Facteur	DF	Wald Khi-2	Pr > Chi-2
Motif de déplacement	4	32.30	<.0001
Distance de déplacement	3	42.55	<.0001
Nombre de déplacements quotidiens réalisés la veille de l'interview	3	20.94	0.0001
Profession de l'enquêté	7	28.11	0.0002
Type du ménage de l'enquêté	4	17.11	0.0018
Sexe de l'enquêté	1	7.40	0.0065
Etat de santé général de l'enquêté	3	10.69	0.0135
Motorisation (nombre voiture dans le ménage)	3	9.52	0.0231
Mode de déplacement	6	14.06	0.0290

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Table 6.7: Odds ratios sur l'appariement de déplacements quotidiens avec des traces GPS

Facteur de déplacement quotidien	Modalité	Odds ratios	Pr > Khi-2
Motif de déplacement	Aller au domicile	0.869	0.9268
	Achat	1.454	<.0001
	Loisir	0.681	0.0367
	Profession	0.595	0.0020
	Autre motif	Référence	
Distance de déplacement	moins de 2 km	2.191	0.0035
	de 2 à 5 km	2.141	0.0025
	de 5 à 10 km	1.643	0.8700
	égale ou plus de 10 km	Référence	
Nombre de déplacements quotidiens réalisés la veille de l'interview	1--2	0.545	<.0001
	3--4	0.955	0.2804
	5--6	1.139	0.0009
	>=7	Référence	
Profession	Agriculteurs exploitants	1.366	0.4679
	Artisans, commerçants	0.337	0.0416
	Cadres, professions intellectuelles supérieures	1.460	0.0757
	Professions intermédiaires	1.885	0.0001
	Employés	0.911	0.1314
	Ouvriers	1.351	0.2132
	Retraités	1.489	0.0387
	Autres non actifs	Référence	
Type de ménage	Personne seule	0.358	0.0041
	Famille monoparentale	0.372	0.1096
	Couple sans enfant	0.422	0.1202
	Couple avec enfant	0.530	0.5082
	Autres types de ménages	Référence	
Sexe	Homme	0.757	0.0065
	Femme	Référence	
Etat de santé	Très bon	2.292	0.8841
	Bon	2.961	0.3090
	Moyen	3.422	0.1449
	Mauvais	Référence	
Motorisation (nombre voiture dans le ménage)	0	0.592	0.0114
	1	1.141	0.0099
	2	0.995	0.2903
	>2	Référence	
Mode de déplacement	Autre	0.205	0.2720

	Bus	1.825	0.0559
	Métro	0.777	0.6225
	Train	0.304	0.2391
	Moto/ Voiture	0.541	0.5177
	Marche à pied	0.832	0.2698
	Vélo	Référence	

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

6.3.1 Effets des caractéristiques du déplacement

Distance de déplacement :

D'autres recherches sur les déplacements quotidiens oubliés montrent que la plupart d'entre eux étaient de courts déplacements (Stopher, 2006). Bien entendu, le taux de déplacements non appariés décroît en fonction de leur longueur (distance parcourue). En effet, ce taux est de 54% pour les déplacements de moins de 2 km et de 31% pour ceux qui dépassent 10 km.

Tableau 6.8: Odds ratio de distance pour les déplacements quotidiens non appariés

Facteur	Modalité	Odds ratio	Pr > ChiSq
Distance de déplacement	moins de 2 km	2.191	0.0035
	de 2 à 5 km	2.141	0.0025
	de 5 à 10 km	1.643	0.8700
	égale ou plus de 10 km	Référence	

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Tableau 6.9 : Longueur des déplacements quotidiens non appariés

	Modalité	Nombre de déplacements quotidiens non apparié (1)	Nombre total de déplacements quotidiens déclarés (2)	% de non appariés (1)/(2)
Distance	< 2 km	349	648	54%
	[2 ; 5[240	473	51%
	[5 ; 10[200	470	43%
	>= 10 km	217	703	31%
	Total	1006	2294	44%

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Nombre de déplacements quotidiens réalisés la veille de l'interview :

L'Odds ratios de déplacements quotidiens oubliés (non appariés) est proportionnel augmente par rapport au nombre de déplacements quotidiens réalisés la veille de l'interview. Le taux de déplacements quotidiens oubliés est plus fort pour les jours où l'enquêté a réalisé plus de 5 déplacements. Nous émettons l'hypothèse qu'un plus grand nombre de déplacements entrainerait une plus grande probabilité d'oubli de déclarer tous les déplacements. La base de cette théorie est qu'il y a une charge plus lourde sur les enquêtés et la probabilité d'oubli de déplacement est plus élevé.

Tableau 6.10 : Odds ratios du nombre de déplacement réalisé la veille de l'interview sur les déplacements quotidiens non appariés

Facteur de déplacement quotidien	Modalité	Odds ratios	Pr > Khi-2
Nombre de déplacements quotidiens réalisés la veille de l'interview	1--2	0.545	<.0001
	3--4	0.955	0.2804
	5--6	1.139	0.0009
	>=7	Référence	

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Tableau 6.11 : Nombre de déplacement réalisé la veille de l'interview pour les déplacements quotidiens non appariés

Caractéristiques du déplacement quotidiens sans trace	Modalité	Nombre de déplacements quotidiens non apparié (1)	Nombre total de déplacements quotidiens déclarés (2)	% de non appariés (1)/(2)
Nombre de déplacements quotidiens	1-2	83	292	28%
	3-4	266	637	42%
	5-6	325	661	49%
	>=7	332	704	47%
	Total		1006	2294

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Motif de déplacement :

Le motif est également un facteur influençant l'oubli de décrire des déplacements. Le taux de déplacements quotidiens non appariés est plus grand pour le motif achat (56%) et plus petit pour les

motifs professionnels (32%). Cela montre qu'on se souvient bien des activités qui ont eu lieu régulièrement telles que le domicile-travail.

Tableau 6.12 : Odds ratios du motif sur les déplacements quotidiens non appariés

Effect	Modalité	Odds ratio	Pr > ChiSq
Motif de déplacement	Aller au domicile	0.869	0.9268
	Achat	1.454	<.0001
	Loisir	0.681	0.0367
	Profession	0.595	0.0020
	Autre motif	Référence	

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Tableau 6.13 : Taux d'oublis selon le motif du déplacement

Caractéristiques de déplacements quotidiens sans trace	Modalité	Nombre de déplacements quotidiens non apparié (1)	Nombre total de déplacements quotidiens déclarés (2)	% de non appariés (1)/(2)
Motif de déplacement	Retour au domicile	370	883	42%
	Achats	214	379	56%
	Loisirs	98	255	38%
	Professionnels	81	253	32%
	Autre motif	243	524	46%
	Total	1006	2294	44%

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Pour mieux comprendre l'effet du motif de déplacement sur les oublis, nous le croisons avec l'heure de départ.

Tableau 6.14 : % de déplacements quotidiens non appariés

Heure de départ de déplacement quotidien déclaré	% de déplacement non apparié					
	Aller au domicile	Achats	Loisirs	Professionnel	Autre	Total
Avant 9h	53%	41%	43%	33%	53%	41%
de 9h à 11h30	45%	60%	35%	35%	52%	49%

de 11h30 à 14h	46%	47%	45%	27%	35%	41%
de 14h à 17h	40%	62%	36%	33%	46%	46%
de 17h à 21h	36%	53%	38%	27%	47%	41%
de 21h à 24h	48%	0%	38%	100%	44%	48%
Total	42%	56%	38%	32%	46%	44%

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Par rapport à l'heure de départ, nous trouvons que les déplacements pour motif achat oubliés tombent bien dans deux périodes de travail dans la journée, soit de 9h à 11h30 (60%), et 14h à 17h (62%). Nous constatons que les déplacements quotidiens liés au travail sont bien décrits en début de journée (avant 9h, 33% d'oublis) ou à la sortie des bureaux (17h00-23h59, 27%) ou à l'heure du déjeuner (11h30-14h, 27%).

Mode de déplacements quotidiens :

Le taux de déplacements quotidiens oubliés est plus fort pour les déplacements réalisés par le mode de bus. Ensuite, le taux de déplacements quotidiens diminue à l'ordre des modes de Vélo, Marche à pied, Métro, Moto/voiture, Train. L'enquête peut avoir haute probabilité de rappeler les déplacements de mode de Train. Les déplacements effectués par un mode non-motorisé ont été fréquemment manqués.

Tableau 6.15 : Odds ratios de mode de déplacements quotidiens non appariés

Facteur de déplacement quotidien	Modalité	Odds ratios	Pr > Khi-2
Mode de déplacement	Autre	Référence	
	Bus	8.907	0.0559
	Vélo	4.881	0.2310
	Marche à pied	4.061	0.2698
	Métro	3.790	0.6225
	Moto/ Voiture	2.643	0.5177
	Train	1.482	0.2391

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Tableau 6.16 : Taux d'oublis selon le mode de déplacements quotidiens non appariés

Caractéristiques de déplacements quotidiens sans trace	Modalité	Nombre de déplacements quotidiens non apparié (1)	Nombre total de déplacements quotidiens déclarés (2)	% de non appariés (1)/(2)

Mode de déplacement	Bus	9	14	64%
	Vélo	21	35	60%
	Marche à pied	203	368	55%
	Métro	11	26	42%
	Moto/ Voiture	757	1831	41%
	Autre	1	4	25%
	Train	3	15	20%
	Total	1005	2293	44%

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

6.3.2 Effets des caractéristiques de l'enquête

Nous constatons que les oublis sont plus fréquents chez les personnes des professions intermédiaires ou dans les couples avec enfant ou dans les ménages ayant plusieurs de moto/voiture ou chez les femmes, ainsi que chez ceux qui ont un mauvais état de santé.

Nous remarquons que plus d'enfant ou plus mauvais de santé ou plus de motorisation, plus de probabilité de l'oubli de déplacement quotidien.

Tableau 6.17: Odds ratios des caractéristiques de l'enquête sur les déplacements quotidiens non appariés

Facteur de déplacement quotidien	Modalité	Odds ratios	Pr > Khi-2
Profession	Agriculteurs exploitants	1.366	0.4679
	Artisans, commerçants	0.337	0.0416
	Cadres, professions intellectuelles supérieures	1.460	0.0757
	Professions intermédiaires	1.885	0.0001
	Employés	0.911	0.1314
	Ouvriers	1.351	0.2132
	Retraités	1.489	0.0387
	Autres non actifs	Référence	
Type de ménage	Personne seule	0.358	0.0041
	Famille monoparentale	0.372	0.1096
	Couple sans enfant	0.422	0.1202
	Couple avec enfant	0.530	0.5082
	Autres types de ménages	Référence	
Sexe	Homme	0.757	0.0065

	Femme	Référence	
Etat de santé	Très bon	2.292	0.8841
	Bon	2.961	0.3090
	Moyen	3.422	0.1449
	Mauvais	Référence	
Motorisation (nombre voiture dans le ménage)	0	0.518	0.0114
	1	Référence	
	2	0.872	0.2903
	>2	0.876	0.4589

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Tableau 6.18 : Oublis selon le genre du répondant

Caractéristiques de l'enquête	Modalité	Nombre de déplacements quotidiens non apparié (1)	Nombre total de déplacements quotidiens déclarés (2)	% de non appariés (1)/(2)
Genre	Homme	423	1017	42%
	Femme	583	1277	46%
	Total	1006	2294	44%

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

6.3.3 Raisons de l'oubli de décrire des déplacements quotidiens

Wolf et al. (2003) ont rapporté que les principales raisons de la sous-déclaration des déplacements quotidiens sont une durée d'enquête excessive et des déplacements oubliés, délibérément omis ou considérés comme sans importance par le répondant. Les déplacements effectués par un mode non-motorisé ont également été fréquemment manqués.

Tableau 6.19 : Raison des déplacements quotidiens sans trace GPS

Raison	Fréquence	%
Oubli chez l'enquête	441	44%
Démarrage à froid lors de la première sortie du matin (Le premier déplacement du matin sans trace GPS, mais à partir du second déplacement, le GPS marche bien)	110	11%
Achats	86	9%

Oubli au bureau	74	7%
Autre raison	295	29%
Total	1006	100%

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Dans le volet GPS de l'ENTD, en considérant le motif du déplacement précédent, nous avons un autre outil pour trouver la raison des déplacements quotidiens sans trace GPS. Le Tableau 6.15 montre que la moitié des déplacements quotidiens sans traces sont liés à une défaillance humaine, dont 44% correspondent à un oubli au domicile et 7 % à l'oubli au bureau, par exemple quand l'enquêté sort du bureau pour déjeuner ou travailler hors de l'endroit habituel, etc. 11% des déplacements quotidiens sans trace proviennent d'un démarrage à froid ou d'un redémarrage nécessitant un laps de temps pour capter les signaux des satellites, par exemple lors de la première sortie du matin. 9 % des déplacements quotidiens sans trace suivent un déplacement pour motif achat ; cela montre un mauvais positionnement du récepteur, par exemple dans un centre commercial ou dans un immeuble.

6.4 Identification des traces GPS sans déplacements quotidiens

Au tableau 6.1, nous avons constaté 2851 traces GPS dont 1563 trace GPS non appariées avec des déplacements quotidiens. Pour mieux comprendre les effets des caractéristiques des traces et aussi de l'enquêté sur les traces GPS non appariée avec un déplacement quotidien observé par CAPI, nous estimons une régression logistique binaire sur cet échantillon de 2851 traces GPS.

Le Tableau 6.20 montre que les facteurs qui jouent un rôle sur les traces GPS non appariées avec un déplacement quotidien sont : le nombre de traces GPS relevées la veille de l'interview, la vitesse, la durée et la distance linéaire de la trace GPS, la région de résidence de l'enquêté, la profession de l'enquêté, ainsi que la durée moyenne et le nombre moyen des déplacements quotidiens effectués par l'enquêté observés par CAPI.

Tableau 6.20 : Facteurs de l'appariement des traces GPS avec les déplacements quotidiens

Facteur	DF	Wald Khi-2	Pr > Chi-2
Nombre de traces GPS	3	166.7260	<.0001
Vitesse sur la trace GPS	4	51.2141	<.0001
Durée de la trace GPS	4	102.5806	<.0001
Nombre moyen de déplacements quotidiens observés par CAPI	4	60.1618	<.0001
Région de l'enquêté	11	27.7060	0.0036

Profession de l'enquêté	7	20.2003	0.0052
Distance linéaire de trace GPS	3	11.5147	0.0092

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Le tableau 6.21 fournit une analyse plus détaillée des traces GPS non apparié avec un déplacement quotidien observé par CAPI. Tout d'abord, nous avons analysé l'effet des caractéristiques du déplacement et de l'enquêté (odds-ratio) et p-value du Test Khi-2.

Tableau 6.21 : Odds ratios sur l'appariement de trace GPS avec déplacement quotidien

Facteur de trace GPS	Modalité	Odds ratios	Pr > Khi-2
Nombre de trace GPS la veille de l'interview	1--2	0.093	<.0001
	3--4	0.420	0.2014
	5--6	0.481	0.0106
	>=7	Référence	
Vitesse sur la trace (en km/h)	< 7	1.885	<.0001
	7--20	1.404	0.0054
	20--50	0.689	<.0001
	50--80	0.671	0.0037
	>=80	Référence	
Durée de trace GPS	moins de 5 minutes	11.918	<.0001
	de 5 à 15 minutes	6.221	<.0001
	de 15 à 30 minutes	3.733	0.5999
	de 30 à 60 minutes	1.911	0.0003
	>= 60 minutes	Référence	
Distance linéaire de trace GPS	moins de 2 km	1.629	0.0012
	de 2 à 5 km	1.026	0.2997
	de 5 à 10 km	1.023	0.3330
	égale ou plus de 10 km	Référence	
Région de résidence	11 - Ile de France	1.183	0.0694
	22 - Picardie	1.465	0.9963
	23 - Haute Normandie	1.599	0.5743
	26 - Bourgogne	1.078	0.1893
	31 - Nord Pas de Calais	1.643	0.6236
	41 - Lorraine	1.181	0.2896
	53 - Bretagne	1.235	0.2228
	54 - Poitou Charentes	1.454	0.9641
	72 - Aquitaine	2.462	0.0004

	73 - Midi Pyrénées	1.607	0.4775
	82 - Rhône Alpes	2.351	0.0093
	83 - Auvergne	Référence	
Profession	Agriculteurs exploitants	1.643	0.6236
	Artisans, commerçants	1.181	0.2896
	Cadres, professions intellectuelles supérieures	1.235	0.2228
	Professions intermédiaires	1.454	0.9641
	Employés	2.462	0.0004
	Ouvriers	1.607	0.4775
	Retraités	2.351	0.0093
	Autres non actifs	Référence	
Nombre moyen des déplacements quotidiens de l'enquête observés par CAPI	0	2.765	0.0040
	1--2	2.926	<.0001
	3--4	2.168	0.0892
	5--6	1.335	0.0003
	>=7	Référence	

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

6.4.1 Caractéristiques des traces GPS non appariées

Nombre de trace GPS la veille de l'interview

Nous notons que le taux de trace non appariée décroît avec la fréquence de mobilité. Le taux de trace GPS non appariés est plus haut pour les jours avec 1-2 traces GPS (faible mobilité) et plus grand pour les trace GPS qui ont lieu un jour de forte mobilité.

Tableau 6.22: Odds ratios de vitesse sur les traces GPS non appariées

Facteur de trace GPS	Modalité	Odds ratios	Pr > Khi-2
Nombre de trace GPS la veille de l'interview	1--2	0.093	<.0001
	3--4	0.420	0.2014
	5--6	0.481	0.0106
	>=7	Référence	

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Tableau 6.23 : Vitesse sur les traces GPS non appariées

Caractéristique de trace GPS	Modalité	Nombre de trace GPS non apparié (1)	Nombre total de trace GPS (2)	% de trace non appariés (1)/(2)
Nombre de déplacements quotidiens	1--2	41	266	15%
	3--4	189	476	40%
	5--6	196	414	47%
	>=7	1137	1695	67%
	Total	1563	2851	55%

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Vitesse sur les traces GPS

Nous remarquons que le taux de trace non appariée décroît avec la vitesse sur la trace. Ce taux est plus grand pour les vitesses inférieures à 7km/h qui correspondent à la "Marche à pied" ; il est plus petit pour les traces de vitesse supérieure à 80 km/h pour lesquelles le moyen de transport peut être "Train" ou "Moto/ Voiture". Les traces GPS, pour lesquelles le mode imputé est "Marche à pied", ne sont pas bien appariés avec les déplacements quotidiens. Les raisons seraient l'oubli des événements importants par l'enquêté ou une mauvaise imputation du mode pour la trace GPS.

Tableau 6.24: Odds ratios de vitesse sur les traces GPS non appariées

Facteur de trace GPS	Modalité	Odds ratios	Pr > Khi-2
Vitesse sur la trace (en km/h)	< 7	1.885	<.0001
	7--20	1.404	0.0054
	20--50	0.689	<.0001
	50--80	0.671	0.0037
	>=80	Référence	

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Tableau 6.25 : Vitesse sur les traces GPS non appariées

Caractéristique de trace GPS	Modalité	Nombre de trace GPS non apparié (1)	Nombre total de trace GPS (2)	% de trace non appariés (1)/(2)
------------------------------	----------	-------------------------------------	-------------------------------	---------------------------------

Vitesse sur la trace (en km/h)	< 7	576	745	77%
	7--20	280	427	66%
	20--50	259	671	39%
	50--80	67	247	27%
	>=80	381	761	50%
	Total	1563	2851	55%

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Durée de trace GPS

Nous remarquons que le taux de trace non appariée décroît avec la durée de la trace. Le GPS marche mieux pour des déplacements de plus de 30 minutes. Ce taux est le plus petit pour les traces de durée moins de 5 minutes. Cela peut être le résultat de l'habitude de l'oubli des courts déplacements quotidiens.

Tableau 6.26: Odds ratios de durée des traces GPS non appariées

Facteur de trace GPS	Modalité	Odds ratios	Pr > Khi-2
Durée de trace GPS	moins de 5 minutes	11.918	<.0001
	de 5 à 15 minutes	6.221	<.0001
	de 15 à 30 minutes	3.733	0.5999
	de 30 à 60 minutes	1.911	0.0003
	>= 60 minutes	Référence	

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Tableau 6.27 : Durée des traces GPS non appariées

Caractéristique de trace GPS	Modalité	Nombre de trace GPS non apparié (1)	Nombre total de trace GPS (2)	% de trace non appariés (1)/(2)
Durée de trace GPS	moins de 5 minutes	956	1275	75%
	de 5 à 15 minutes	398	851	47%
	de 15 à 30 minutes	148	445	33%
	de 30 à 60 minutes	48	203	24%
	>= 60 minutes	13	77	17%
	Total	1563	2851	55%

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Distance linéaire de trace GPS

Le taux de trace non appariée décroît avec la distance linéaire de la trace GPS. Le taux de trace GPS non appariée est plus élevé pour les traces GPS de distance linéaire moins de 2km.

Tableau 6.28 : Odds ratios de distance linéaire des traces GPS non appariées

Facteur de trace GPS	Modalité	Odds ratios	Pr > Khi-2
Distance linéaire de trace GPS	moins de 2 km	1.629	0.0012
	de 2 à 5 km	1.026	0.2997
	de 5 à 10 km	1.023	0.3330
	égale ou plus de 10 km	Référence	

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Tableau 6.29 : Traces GPS non appariées selon distance linéaire de trace

Caractéristique de trace GPS	Modalité	Nombre de trace GPS non apparié (1)	Nombre total de trace GPS (2)	% de trace non appariés (1)/(2)
Distance linéaire de trace GPS	moins de 2 km	1234	1761	70%
	de 2 à 5 km	155	423	37%
	de 5 à 10 km	95	296	32%
	égale ou plus de 10 km	79	371	21%
	Total	1563	2851	55%

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

6.4.2 Effets des caractéristiques de l'enquêté

Profession de l'enquêté

Suite à l'information de odds ratios et % de trace non appariés, nous remarquons qu'il y a un peu plus de traces appariées chez les retraités.

Tableau 6.30 : Odds ratios de profession de l'enquêté sur les traces GPS non appariées

Facteur	Modalité	Odds ratios	Pr > Khi-2
Profession	Agriculteurs exploitants	1.643	0.6236
	Artisans, commerçants	1.181	0.2896
	Cadres, professions intellectuelles supérieures	1.235	0.2228
	Professions intermédiaires	1.454	0.9641
	Employés	2.462	0.0004
	Ouvriers	1.607	0.4775
	Retraités	2.351	0.0093
	Autres non actifs	Référence	

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Tableau 6.31 : Traces GPS non appariées selon la profession de l'enquêté

Caractéristiques de l'enquêté	Modalité	Nombre de trace GPS non apparié (1)	Nombre total de trace GPS (2)	% de trace non appariés (1)/(2)
Profession	Agriculteurs exploitants	66	88	75%
	Artisans, commerçants	3	16	19%
	Cadres, professions intellectuelles supérieures	216	363	60%
	Professions intermédiaires	211	444	48%
	Employés	312	642	49%
	Ouvriers	230	397	58%
	Retraités	376	640	59%
	Autres non actifs	140	250	56%

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Région de résidence de l'enquêté

En comparant les odds ratio et le pourcentage de trace non appariés dans les 12 régions expérimentées l'enquête GPS, nous remarquons que les régions Aquitaine, Midi-Pyrénées et Rhône Alpes ont un taux de trace non appariée plus élevé. Il est possible que ces régions sont des milieux ruraux où le récepteur GPS marche bien et les traces GPS pourraient être enregistrées meilleur que les autres régions urbaines.

Tableau 6.32 : Odds ratios de la région de résidence sur les traces GPS non appariées

Facteur	Modalité	Odds ratios	Pr > Khi-2
Région de résidence	11 - Ile de France	1.183	0.0694
	22 - Picardie	1.465	0.9963
	23 - Haute Normandie	1.599	0.5743
	26 - Bourgogne	1.078	0.1893
	31 - Nord Pas de Calais	1.643	0.6236
	41 - Lorraine	1.181	0.2896
	53 - Bretagne	1.235	0.2228
	54 - Poitou Charentes	1.454	0.9641
	72 - Aquitaine	2.462	0.0004
	73 - Midi Pyrénées	1.607	0.4775
	82 - Rhône Alpes	2.351	0.0093
	83 - Auvergne	Référence	

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Tableau 6.33 : Traces GPS non appariées selon la région de résidence

Caractéristique de l'enquête	Modalité	Nombre de trace GPS non apparié (1)	Nombre total de trace GPS (2)	% de trace non appariés (1)/(2)
Région de résidence	11 – Ild de France	400	611	65%
	22 – Picardie	91	187	49%
	23 - Haute Normandie	131	242	54%
	26 - Bourgogne	37	96	39%
	31 - Nord Pas de Calais	47	92	51%
	41 - Lorraine	49	132	37%
	53 - Bretagne	137	304	45%
	54 - Poitou Charentes	112	237	47%
	72 - Aquitaine	168	278	60%
	73 - Midi Pyrénées	248	408	61%
	82 - Rhône Alpes	118	186	63%
	83 - Auvergne	25	78	32%
	Total	1563	2851	55%

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Mobilité de l'enquêté

Nous notons que le taux de trace non appariée décroît avec la fréquence de mobilité. Le taux de trace GPS non appariés est plus haut pour les jours avec 1-2 traces GPS (faible mobilité) et plus grand pour les trace GPS qui ont lieu un jour de forte mobilité.

Tableau 6.34 : Odds ratios de la fréquence des déplacements de l'enquêté sur les traces GPS non appariées

Facteur de trace GPS	Modalité	Odds ratios	Pr > Khi-2
Nombre moyen des déplacements quotidiens de l'enquêté observés par CAPI	0	2.765	0.0040
	1--2	2.926	<.0001
	3--4	2.168	0.0892
	5--6	1.335	0.0003
	>=7	Référence	

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Tableau 6.35 : Traces GPS non appariées selon la fréquence de déplacement

Caractéristiques de l'enquêté	Modalité	Nombre de trace GPS non apparié (1)	Nombre total de trace GPS (2)	% de trace non appariés (1)/(2)
Nombre de déplacements réalisés la veille	0	138	257	54%
	1-2	295	499	59%
	3-4	577	964	60%
	5-6	292	595	49%
	>=7	261	536	49%
	Total	1563	2851	55%

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

6.5 Identification des traces GPS appariés avec des déplacements

Nous venons d'étudier les caractéristiques de deux types de déplacements :

- déplacements quotidiens déclarés par CAPI non appariés avec les traces GPS ;
- déplacements collectés par GPS (traces GPS) non décrits par CAPI ;

Nous continuons par l'analyse de l'ensemble complémentaire : les déplacements appariés, c.à.d. ceux qui ont été enregistrés par GPS et noté aussi par CAPI. Cette comparaison nous permettra de mieux comprendre la spécificité des estimations de mobilité par ces deux méthodes d'enquête.

Comme nous l'avons mentionné dans le Panorama de l'appariement des traces GPS et des déplacements quotidiens, il y a 1288 cas qu'un déplacement quotidien est apparié avec une trace GPS.

Le GPS peut enregistrer des informations précises sur les heures de départ et d'arrivée, la durée et la distance de déplacement, ce que les enquêtes classiques fournissent difficilement. Pour évaluer ces deux méthodes d'enquête, nous allons comparer les descriptions des déplacements sur la base des horaires (heures de départ, d'arrivée), des durées, et des distances.

6.5.1 Comparaison des heures de départ et d'arrivée

La différence des heures de départ entre les traces GPS et les déplacements quotidiens est présenté dans le tableau 6.36. Il y a 17 % des déplacements quotidiens pour lesquels l'heure de départ déclarée diffère de moins de deux minutes de l'heure de départ "réelle" enregistrée par le GPS. Il y a environ 61% de déplacements quotidiens dont le décalage avec l'heure de départ donnée par le GPS est inférieur à 15 minutes, 74 % à moins de 30 minutes, 86% à moins d'une heure. Cela montre bien l'imprécision des déclarations par rapport au temps. Cette imprécision peut résulter de l'arrondi de l'heure de départ par la plupart des enquêtés.

Tableau 6.36 : Différence absolue sur l'heure de départ

Différence absolue sur l'heure de départ	Nombre de déplacement quotidien	%
De 0 à 2 minutes	221	17%
De 0 à 5 minutes	462	36%
De 0 à 10 minutes	681	53%
De 0 à 15 minutes	790	61%
De 0 à 30 minutes	959	74%
De 0 à 1 heure	1111	86%
Total	1288	100%

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Le Tableau 6.37 montre que 63 % des déplacements quotidiens ont une heure de départ déclarée antérieure à celle de la trace GPS. En effet, un démarrage à froid ou un redémarrage nécessitent un laps de temps pour capter les signaux des satellites, ce qui fait très souvent perdre le début du déplacement.

Tableau 6.37 : Différence sur l'heure de départ

Différence sur l'heure de départ	Nombre de déplacement quotidien	%	% cumulé
Plus de 1 heure (-)	142	11,0%	11,0%
De 30 à 1 heure (-)	107	8,3%	19,3%
De 15 à 30 minutes (-)	89	6,9%	26,2%
De 10 à 15 minutes (-)	73	5,7%	31,9%
De 5 à 10 minutes (-)	140	10,9%	42,8%
De 2 à 5 minutes (-)	141	10,9%	53,7%
De 0 à 2 minutes (-)	123	9,5%	63,3%
0	2	0,2%	63,4%
De 0 à 2 minutes (+)	96	7,5%	70,9%
De 2 à 5 minutes (+)	100	7,8%	78,6%
De 5 à 10 minutes (+)	79	6,1%	84,8%
De 10 à 15 minutes (+)	36	2,8%	87,6%
De 15 à 30 minutes (+)	80	6,2%	93,8%
De 30 à 1 heure (+)	45	3,5%	97,3%
Plus de 1 heure (+)	35	2,7%	100,0%
Total	1288	100,0%	100,0%

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Le Tableau 6.38 illustre la différence absolue de l'heure d'arrivée entre les traces GPS et les déplacements quotidiens. La déclaration est plus précise que pour l'heure de départ. Il y a 77 % de déplacement quotidien dont le décalage de l'heure d'arrivée est inférieur à 30 minutes avec la trace GPS ; et seulement 9% à plus d'une heure.

Tableau 6.38 : Différence absolue sur l'heure de l'arrivée

Différence absolue sur l'heure d'arrivée	Nombre de déplacement quotidien	%
De 0 à 2 minutes	268	21%
De 0 à 5 minutes	541	42%
De 0 à 10 minutes	727	56%
De 0 à 15 minutes	839	65%
De 0 à 30 minutes	995	77%
De 0 à 1 heure	1169	91%
Total	1288	100%

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Le Tableau 6.39 montre que, contrairement à l'heure de départ, il n'y a pas de biais sur l'heure d'arrivée.

Table 6.39 : Différence sur l'heure d'arrivée

	Différence sur l'heure de départ	Nombre de déplacement quotidien	%	% cumulé
Heure d'arrivée du déplacement antérieure à celle de la trace GPS	Plus de 1 heure (-)	83	6,4%	6,4%
	De 30 à 1 heure (-)	124	9,6%	16,1%
	De 15 à 30 minutes (-)	86	6,7%	22,7%
	De 10 à 15 minutes (-)	62	4,8%	27,6%
	De 5 à 10 minutes (-)	104	8,1%	35,6%
	De 2 à 5 minutes (-)	119	9,2%	44,9%
	De 0 à 2 minutes (-)	137	10,6%	55,5%
Heure d'arrivée du déplacement postérieure à celle de la trace GPS	0	2	0,2%	55,7%
	De 0 à 2 minutes (+)	129	10,0%	65,7%
	De 2 à 5 minutes (+)	154	12,0%	77,6%
	De 5 à 10 minutes (+)	82	6,4%	84,0%
	De 10 à 15 minutes (+)	50	3,9%	87,9%
	De 15 à 30 minutes (+)	70	5,4%	93,3%
	De 30 à 1 heure (+)	50	3,9%	97,2%
	Plus de 1 heure (+)	36	2,8%	100,0%
	Total	1288	100,0%	100,0%

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Suite à la comparaison des différences pour les heures de départ et d'arrivée, nous constatons que l'heure de départ est biaisée et moins précise que celle de l'arrivée, pour laquelle les sur-estimations et sous-estimations semblent s'équilibrer. Cela peut s'expliquer par le fait que le GPS a besoin d'un certain laps de temps pour recevoir les signaux après le début du déplacement. Une fois que le GPS fonctionne, il va bien enregistrer les données jusqu'à la fin du déplacement.

6.5.2 Comparaison des durées entre déplacement quotidien et trace GPS

Nous trouvons une proportion très élevée de traces GPS qui durent moins de 5 minutes (27% par GPS contre 7% par CAPI). Si nous regroupons en 3 tranches telles que moins de 5 minutes, de 5 à

15 minutes, plus de 15 minutes, alors pour les traces GPS, ce taux est stable à environ 32%, tandis que pour les déplacements quotidiens, ce taux est respectivement de 7%, 34% et 54%.

Tableau 6.40 : Fréquences des déplacements quotidiens et traces GPS par tranches de durée

Durée de déplacement (minutes)	Nombre de déplacements quotidiens appariés avec des traces GPS	%	Nombre de traces GPS appariés avec des déplacements quotidiens	%
[0 -5[91	7%	343	27%
[5 - 10[231	18%	271	21%
[10 - 15[274	21%	173	13%
[15-30[395	31%	290	23%
[30 ou plus [297	23%	211	16%
Total	1288	100%	1288	100%

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Le tableau 6.41 montre que, parmi 1589 déplacement comparé, il y en a 81% pour lesquels le décalage des durées de déplacements est moins de 30 minutes.

Table 6.41 : Différence absolue des durées entre déplacements quotidiens et traces GPS

Ecart absolue des durées entre déplacements quotidiens et traces GPS (minutes)	Fréquence	%
[0-2[267	21%
[0- 5[578	45%
[0-10[873	68%
[0-15[1021	79%
[0-30[1168	91%
[0-60[1288	100%
Total	1288	100%

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Parmi les 1288 déplacements traités, il y a seulement 8 cas d'estimation précise de la durée de déplacement.

La plupart des durées déclarées par CAPI sont sur-estimées, et seulement 30% sous-estimées (tableau 6.42). Ceci peut résulter de l'arrondi de l'heure de départ, de l'heure d'arrivée par la plupart des enquêtés.

Tableau 6.42 : Différence absolue des durées entre déplacements quotidiens et traces GPS

	Différence absolue des durées de déplacement quotidien et trace GPS (minutes)						
	[0,01-2[[2-5[[5-10[[10-15[[15-30[>=30	Total
Sous-estimation de durée de déplacement	102	75	81	37	50	45	390
	7,9%	5,8%	6,3%	2,9%	3,9%	3,5%	30,3%
Sur estimation de durée de déplacement	157	236	214	111	97	75	890
	12,2%	18,3%	16,6%	8,6%	7,5%	5,8%	69,1%
Total	259	311	295	148	147	120	1280
	20,1%	24,1%	22,9%	11,5%	11,4%	9,3%	99,4% (*)

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

(*) plus 0.6% d'estimations précises de durée de déplacement.

6.5.3 Comparaison des distances parcourues entre déplacement quotidien et trace GPS

Parmi les 1288 déplacements traités, il y a seulement 17 cas d'estimation précise de la distance parcourue ; la déclaration par CAPI la sur-estime dans environs les trois quarts des cas (tableau 6.43).

Table 6.43 : Différence absolue des distances parcourues entre déplacements quotidiens et traces GPS

	Différence absolue des distances parcourues de déplacement quotidien et trace GPS (km)						
	[0,01-0,5[[0,5-1[[1-2[[2-5[[5-10[[10 ou +[Total
Sous-estimation de la distance parcourue de déplacement	76	50	58	102	40	67	393
	5,9%	3,9%	4,5%	7,9%	3,1%	5,2%	30,5%
Sur estimation de la distance parcourue de déplacement	119	102	153	229	131	144	878
	9,2%	7,9%	11,9%	17,8%	10,2%	11,2%	68,2%
Total	195	152	211	331	171	211	1271
	15,1%	11,8%	16,4%	25,7%	13,3%	16,4%	98,7% (**)

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

(**)1.3% d'estimations précises.

6.5.4 Comparaison des distances linéaires entre déplacement quotidien et trace GPS

Parmi les 1288 déplacements traités, il y a seulement 13 cas d'estimation précise de la distance linéaire. En général, les nombres des cas de sous-estimation estimation de la distance linéaire de déplacement par CAPI est égale à ceux de sur-estimation. Mais pour plus de détail, si la différence absolue des distances linéaires entre déplacement quotidien et trace GPS est moins que 1km, alors la déclaration par CAPI la sous-estime dans environs un tiers des cas. Dans d'autres cas où la différence absolue des distances linéaires entre déplacement quotidien et trace GPS est plus que 1km, la déclaration par CAPI la sur-estime dans environs les 40% des cas (Tableau 6.44).

Table 6.44 : Différence absolue des distances parcourues entre déplacements quotidiens et traces GPS

	Différence absolue des distances linéaires de déplacement quotidien et trace GPS (km)						
	[0,01-0,5[[0,5-1[[1-2[[2-5[[5-10[[10 ou +[Total
Sous-estimation de la distance linéaire de déplacement	239	105	101	101	39	47	632
	18,6%	8,2%	7,8%	7,8%	3,0%	3,6%	49,1%
Sur estimation de la distance linéaire de déplacement	71	85	126	184	79	98	643
	5,5%	6,6%	9,8%	14,3%	6,1%	7,6%	49,9%
Total	310	190	227	285	118	145	1275
	24,1%	14,8%	17,6%	22,1%	9,2%	11,3%	99,0%

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

(**)1% d'estimations précises.

6.5.5 Comparaison des modes de déplacement imputés sur les traces GPS avec la description des déplacements quotidiens

A l'aide de l'appariement entre les traces GPS et les déplacements quotidiens décrits par l'enquêté, nous pouvons évaluer les résultats de l'imputation du mode de transport par l'algorithme DECOUTRACE. Il est certain que la correspondance n'est pas parfaite : il y a toujours des cas qu'on ne peut pas traiter. Dans les déplacements quotidiens du questionnaire classiques et les déplacements déclarés par l'enquêté dans le questionnaire complémentaire CAPI-GPS, nous rencontrons des difficultés pour les transports en communs tels que bus, métro ou train qui sont des

modes peu fréquents. Il n'est d'ailleurs pas très facile de distinguer les transport communs (métro, train ou bus) avec Voiture/Moto. Par contre, la distinction entre marche à pied et voiture/moto est satisfaisante (le taux d'imputation de Marche à pied est 61%, pour Voiture/Moto est 97%). En général, le taux d'imputation est de 90%.

Tableau 6.45 : Comparaison du mode entre déplacements quotidiens et traces GPS

Mode de déplacement quotidien déclaré par l'enquête	Mode de déplacement imputé par le programme					
	Marche à pied	Vélo	Moto /Voiture	Transport commune	Autre moyen	Total
Marche à pied	100 (61%)	2	63	0	0	165
Vélo	1	3 (21%)	10	0	0	14
Moto /Voiture	23	0	1049 (97%)	1	2	1075
Transport commune	6	0	23	2 (6%)	0	31
Autre moyen	0	0	3	0	0	3
Total	130	5	1148	3	2	1288 (90%)

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

7 Conclusion générale

7.1 Les principaux résultats

Aujourd'hui, les évolutions des modes de vie et les avancées des technologies font que la mobilité des personnes est de plus en plus variée dans l'espace et dans le temps. Ce changement peut s'expliquer notamment par la fragmentation spatiale des activités, par l'augmentation jusqu'à présent du nombre de personnes actives dans la population et aussi par la croissance de l'équipement automobile des ménages (Roux, 2012).

D'autre part, pour bien dimensionner les infrastructures à une époque où la circulation globale tend à plafonner, nous avons besoin d'estimations plus précises sur la mobilité, vers une meilleure compréhension des comportements individuels, non seulement au cours d'une journée mais encore tout au long de l'année. Ainsi, les méthodes d'enquêtes sur la mobilité s'améliorent, de la première méthode simple papier-crayon (PAPI), puis à l'aide des dispositifs assistés par ordinateur (CAPI, CATI...), jusqu'à l'utilisation des nouvelles technologies (GPS, GSM...).

En tous cas, les données recueillies dans le cadre des enquêtes de transport doivent être très précises. Or la capacité des enquêtés à rapporter avec exactitude des informations sur les déplacements qu'ils ont effectués durant un laps de temps relativement long est remise en cause dans de nombreux travaux (Stopher et al, 2007). Ceci s'explique notamment par la durée de l'étude, les défauts de mémoire de l'enquêté et la sélectivité de ses déclarations. Les enquêtes classiques, bien que de bonne qualité, sont d'un coût très élevé que les commanditaires d'études ne sont pas toujours prêts à payer. Ainsi, de plus en plus d'enquêtes reposent sur des protocoles complexes associant plusieurs modes ou méthodologies pour augmenter le taux de réponse global, sans renoncer à la qualité (Couper et al, 2000).

Il semble maintenant clair que les données recueillies en utilisant les technologies mobiles (principalement mais pas exclusivement GPS) fournissent des informations précieuses pour l'analyste des transports. Elles apportent une haute résolution sur le cadre non seulement spatial, mais aussi temporel de la mobilité.

Le suivi GPS permet par ailleurs d'obtenir des informations que les méthodes classiques ne peuvent pas fournir : description des trajets très courts et des parcours terminaux, estimation des vitesses réelles, choix de l'itinéraire, etc. L'apport de l'enquête GPS est essentiellement la précision, qui devrait permettre de « détecter » des déplacements, notamment courts, que les enquêtés passent

parfois sous silence mais qui peuvent au total représenter un temps et des distances non négligeables. Il n'est pas question bien sûr de remplacer les enquêtes classiques, puisqu'on veut suivre des données sur des séries longues de plusieurs dizaines d'années (la première Enquête Nationale Transport en France date de 1966). Le souhait pour les futures ENT-D serait que, comme pour le recensement, elles soient désormais conduites de manière partielle chaque année, plutôt qu'une fois chaque dix ou douze ans. Pour le volet GPS, cela permettrait aussi de bénéficier des nouveaux progrès techniques. Cependant, la collecte par GPS présente aussi des limites, telles que son faible taux d'acceptation, la difficulté d'en déduire des informations sur le mode et le motif de déplacement, ainsi que sur le statut de conducteur ou passager (donc le taux d'occupation des véhicules), ...

La première enquête de mobilité urbaine reposant essentiellement sur le GPS (avec validation sur Internet) a été mise en œuvre aux États-Unis à Cincinnati (Ohio) en 2009 pendant 12 mois avec la participation d'environ 5000 ménages. Son résultat est positif avec une précision élevée pour les imputations des modes et des motifs de déplacement (Stopher et al. 2012). Toutefois, sur la base d'un test négatif, les Britanniques ont renoncé à utiliser des traceurs GPS pour leur National Travel Survey (Department of Transport of UK, 2012), il est donc prématuré de pronostiquer la substitution complète des enquêtes classiques de mobilité par des enquêtes basées sur les nouvelles technologies (GPS, GSM, etc.). Les raisons principales pour lesquelles cela n'a pas été tenté à ce jour sont que ces types de récepteur sont dotés d'une autonomie encore insuffisante et restent donc à perfectionner, et que les algorithmes de traitement et d'analyse des données doivent être améliorés (Bricka et al 2012) (Lu et al 2012).

A l'occasion du volet GPS de l'ENT-D 2007-2008, l'IFSTTAR a développé des logiciels de collecte et traitement automatique de ces données. D'une part, au niveau du PC portable de l'enquêteur INSEE, un logiciel (développé en Visual Basic) est complètement indépendant du logiciel INSEE permettant l'enquête classique. Les données GPS sont téléchargées par liaison sans fil Bluetooth pendant l'entretien, le logiciel les utilise pour générer automatiquement les questions complémentaires concernant le GPS qui seront posées à l'enquêté (l'objectif étant que ces questions ne durent que quelques minutes). Elles apportent des précisions sur le motif, le mode et le nombre de personnes accompagnant le répondant, pour les déplacements d'une des journées comportant des enregistrements GPS, plus des questions éventuelles concernant le bon fonctionnement du récepteur et les jours sans enregistrement. D'autre part, un outil d'analyse des positions enregistrées ; il s'agit d'une base de données associées à une cartographie (développée sous Delphi pour Microsoft Mappoint). Ceci permet des traitements automatiques visant à découper les enregistrements en trace, l'imputation des modes de déplacement, la visualisation des trajectoires suivies par les enquêtés pour pouvoir analyser visuellement des déplacements particuliers.

Dans le cadre de ce mémoire, nous nous sommes efforcés de comprendre les enjeux des enquêtes classiques et des enquêtes par GPS. D'abord, en étudiant l'acceptabilité et les biais des enquêtes de mobilité à l'aide de GPS, nous avons confirmé l'hypothèse de Stopher (2010) selon laquelle les enquêtés qui répondent à une enquête de transport à l'aide de récepteur GPS sont plus consciencieux et décrivent mieux leur mobilité. Nous trouvons que les personnes qui acceptent d'être suivies par GPS ont une mobilité supérieure à la moyenne de la population (+0.32 déplacement/jour ; +3.5 kilomètres/jour et + 6 minutes/jour). Ceci montre que plus une personne se déplace, plus la probabilité qu'elle accepte une enquête GPS est importante. Le profil des personnes qui acceptent l'enquête GPS est un homme jeune, exerçant une profession intellectuelle supérieure, vivant dans un ménage de structure complexe à haut revenu (plus de 2400 € par mois), équipé en ordinateur et multi-motorisé, habitant dans une zone rurale ou péri-urbaine d'une aire urbaine de 100000 à 10000000 habitants.

L'une des raisons principales du faible taux de réponse GPS était le nombre insuffisant du nombre de dispositifs GPS ou le retard dans leur mise à disposition aux enquêteurs. Nous pouvons augmenter le taux de remise des GPS aux enquêtés par l'accroissement de taux de disponibilité de GPS. Pour ce faire, nous aurions notamment besoin d'optimiser les durées entre les deux visites au domicile de l'enquêté pour baisser le taux d'immobilisation forcée des GPS.

Le traitement des traces GPS sans information supplémentaire donne une bonne image de la succession des déplacements, qui dépend de l'intervalle entre les enregistrements (10s dans l'ENTD). Sur la base des réponses à quelques questions complémentaires, des règles probabilistes peuvent être définies à partir du modèle logit, dans le cas où l'algorithme de traitement des traces est incertain pour déterminer le mode ; dans bien des cas, la probabilité est quasi-nulle (jamais d'utilisation de vélo, ...). L'imputation de mode de transport par méthode probabiliste conditionnelle satisfait du mode de déplacement (la marche à pied et le mode « moto/voiture »). Mais la difficulté subsiste pour distinguer vélo / voiture / bus parmi les modes routiers ; et le GPS ne fournit aucune information sur les conducteurs / passagers dans les véhicules, ce qui serait important pour l'analyse du taux d'occupation des véhicules (facteur majeur de la consommation de carburant par passager x km). Cela montre à nouveau que les enquêtes par GPS ne peuvent pas encore remplacer complètement les méthodes classiques : elles ne permettent bien sûr pas d'obtenir les caractéristiques socio-économiques des ménages et les caractéristiques techniques de leurs véhicules, les motifs et modes de déplacement, le nombre de personnes accompagnant l'enquêté, etc. Ces nouvelles formes d'enquête sont pour l'instant utilisées en complément des questionnaires classiques.

Les rôles du temps, de la position et de la vitesse diffèrent entre les enquêtes classiques et par GPS. Dans les enquêtes classiques, les enquêtés donnent souvent des informations erronées sur la durée, souvent déduite des heures de départ et d'arrivée. La distance n'est pas fournie par

l'enquêté, mais est calculée à partir de l'information géocodée des points de départ et d'arrivée des déplacements ; elle n'est donc pas précise. La vitesse « réelle » n'est pas cernée par les enquêtes classiques. En revanche, le GPS indique exactement les heures, les positions de départ et d'arrivée, la durée et la distance. L'enquête par GPS enregistre la vitesse ponctuelle de déplacement et calcule la vitesse moyenne sur un segment à partir de la distance et de la durée de déplacement. La vitesse moyenne estimée en pondérant la vitesse ponctuelle et la vitesse moyenne sur un segment est un élément important pour l'imputation du mode de déplacement. De plus, cette vitesse peut déterminer la raison pour laquelle le GPS manque des déplacements (dans le cas où la vitesse est élevée et la distance est petite) dans la partie de filtrage des données aberrantes.

Pour la comparaison entre interrogation classique sur la mobilité quotidienne et GPS sur les mêmes jours décrits, les données obtenues par le post-traitement GPS et les déplacements déclarés par CAPI sont assemblées automatiquement dans un tableau en se basant sur leurs chronologies. Ceci diffère de la confrontation entièrement manuelle des déplacements observés par GPS et par carnet de déplacements (Stopher, 2011).

Les raisons des déplacements non appariés avec des traces GPS sont identifiées. Premièrement, ces déplacements quotidiens non appariés peuvent être de courte distance, moins de 2 km (voire, moins de 5 km), plutôt les jours de forte mobilité (5 déplacements ou plus). Deuxièmement, l'enquêté peut avoir oublié de porter l'appareil pendant ce déplacement. Pour déterminer quelle en est la raison, en regardant les motifs du déplacement précédent, nous trouvons que la moitié de ces déplacements se produisent après que l'enquêté ait quitté son domicile/ son lieu de travail. Par conséquent, la raison pour laquelle le GPS n'a pas enregistré ces déplacements est très probablement que la personne a oublié l'appareil chez elle ou ne l'a pas emporté en sortant du bureau. Troisièmement, les appareils GPS peuvent manquer quelques déplacements à cause de problèmes techniques du récepteur GPS (démarrage à froid). Quatrièmement, les enquêtés peuvent sur/ sous-évaluer leurs déplacements en raison d'erreurs de mémoire. D'une part, si les données GPS sont enregistrées sans erreurs et toutes les traces GPS ce jours ont été connectées en séquence, tandis que l'enquêté déclare plus de déplacements que les GPS pour ce jour, on peut supposer que l'enquêté a sur-déclaré ses déplacements quotidiens. D'autre part, l'enquêté se souvient bien des activités qui ont lieu régulièrement telles que le travail, les retours au domicile mais a tendance à mal décrire des activités qui diffèrent des habitudes, par exemple un achat pendant les heures de travail, le loisir en début de matinée ou à l'heure du déjeuner, etc. Nous avons examiné les défauts d'appariement de déplacement quotidien en fonction du profil démographique de l'enquêté : les hommes des professions intermédiaires en couple sans enfant ont un taux de déplacements sans trace GPS plus élevé. Nous notons que l'état de santé influe sur la qualité de la déclaration des déplacements : les individus de moyen/mauvais état de santé ont un taux plus grand que ceux qui sont en bonne/très bonne santé.

Quant à l'appariement des traces GPS avec un déplacement déclaré par CAPI, plus la vitesse GPS est élevée, plus forte est la probabilité d'appariement. Les traces GPS de vitesse inférieure à 7 km/h (souvent "à pied") trouvent plus difficilement un déplacement correspondant. Au contraire, les traces à plus de 80km/h (probablement en moto/voiture) s'apparient plus facilement. Les déplacements de longue durée sont plus souvent enregistrés par GPS que les plus courts. Plus l'enquêté se déplace, plus la probabilité de la collecte de ses traces GPS est importante. Les retraités ont plus de traces non appariées en comparant avec d'autres professions. Les régions Aquitaine, Midi-Pyrénées et Rhône Alpes ont un taux de trace non appariée plus élevé. Il est possible que ces régions sont des milieux ruraux où le récepteur GPS marche bien et les traces GPS pourraient être enregistrées meilleur que les autres régions urbaines.

Les raisons possibles de non appariement des traces sont :

- L'enquêté a des problèmes de mémoire ;
- Les enquêtés peuvent hésiter à signaler des déplacements en raison de la sensibilité des lieux ou motifs ;
- Il peut s'agir de problèmes de post-traitement de données collectées par GPS ou de problèmes techniques (données manquantes à cause d'un démarrage à froid ou d'un mauvais positionnement du GPS, etc.) ; le post-traitement peut mal imputer le mode de transport, ou fournir quelques traces GPS qui pourraient ne pas correspondre à un déplacement, dans les cas où la vitesse et la distance de trace sont presque nulles lorsque le GPS est au repos, mais continue à enregistrer la position. Ils sont généralement éliminés lors du post-traitement, mais quelques-uns peuvent échapper à la détection et rester dans les données ;
- le répondant a mal compris la définition d'un « déplacement » (c.à.d. une séquence de trajets continue entre deux activités, chaque trajet correspondant à un mode de transport); les trajets qui ont pour but un transfert entre modes de transport peuvent être omis par l'enquêté ; le répondant peut déclarer seulement un seul « déplacement » vers le travail, tandis que le GPS a enregistré une séquence de trajets qui s'est déroulée durant le déplacement vers le travail, tels que marcher à pied jusqu'à une station de train, prendre un train jusqu'à la station suivante, prendre un tramway en correspondance, etc. Les déplacements effectués par un mode non-motorisé ont également été fréquemment manqués, puisque la marche à pied n'est identifiée que si c'est le mode exclusif du déplacement.

Environ 74% / 75% de déplacement quotidien ont un décalage de l'heure de départ / heure d'arrivée avec la trace GPS à moins de 30 minutes, 86% / 91% à moins d'une heure. La déclaration de l'heure d'arrivée paraît plus précise que celle de l'heure de départ. Ceci indique l'imprécision de la collecte des données par rapport aux heures de départ ou d'arrivée. Cette imprécision est le résultat des arrondis des heures déclarées par la plupart des enquêtés. 63% des déplacements quotidiens ont une heure de départ antérieure à celle de la trace GPS. La raison possible est le laps de temps (en cas de démarrage à froid) nécessaire pour repérer les satellites et obtenir les premières

localisations. Si ce biais est réglé, le GPS va enregistrer normalement la trajectoire jusqu'à la fin du déplacement. Nous remarquons aussi que plus long est le déplacement, plus la probabilité de déclaration de ce déplacement par l'enquêté est importante. Nous avons constaté un niveau élevé de sur-déclaration de la distance parcourue et de la durée du déplacement. Le pourcentage de déplacement sur-estimant la distance parcourue et la durée du déplacement est deux fois plus élevé que celui des sous-estimations.

Cette étude démontre que l'enquête à l'aide de GPS peut être utilisée avec succès pour compléter les enquêtes de transports classiques. En synergie avec cette enquête, le GPS peut constituer un outil de comparaison et une méthode pour produire des facteurs de correction sur quelques variables importantes.

7.2 Recommandations

Cette dernière partie présente les recommandations pour le choix du récepteur GPS, la collecte de données GPS, la formation des enquêteurs, le post-traitement de données GPS et l'harmonisation de l'organisation des différentes équipes afin d'améliorer l'enquête par GPS.

Le récepteur GPS

Il est préférable de choisir un récepteur GPS de petite taille, le moins encombrant possible et susceptible d'être porté facilement dans différentes situations. Les nouvelles générations de récepteurs GPS avec une autonomie plus grande vont permettre de continuer à collecter les données même si l'enquêté a oublié de charger l'appareil pendant la nuit. Cette amélioration nous aidera à convaincre les enquêtés qui ont peur d'oublier de charger l'appareil GPS d'accepter de participer aux enquêtes à l'aide de récepteurs GPS.

La formation de l'enquêteur

En plus de suivre les procédures normales de l'Enquête Nationale Transport, les enquêteurs devront expliquer clairement et s'assurer que les enquêtés comprennent exactement ce que des données GPS sont, comment le GPS fonctionne, comment il sera utilisé et qui aura accès à aux données de l'enquête. Pour ce faire, les enquêteurs seront informés sur toutes les questions évoquées ci-dessus et on devra leur fournir des documents qui peuvent les aider à expliquer tout ce qui concerne l'enquête GPS.

Il est probable que les enquêteurs peuvent rencontrer quelques soupçons sur le but de la collecte des données GPS et les implications de la confidentialité des enquêtes. Les enquêteurs ont besoin

de suivre les formations sur ces questions. Il est recommandé que les enquêteurs portent eux-mêmes les appareils GPS pendant quelques jours où les séances de formation ont lieu afin qu'ils se familiarisent avec le processus et les questions que les enquêtés peuvent rencontrer.

La qualité de collecte des données GPS a diminué en raison du départ de responsables initiaux de l'enquête et de la dégradation du travail des enquêteurs. Ces problèmes sont susceptibles de s'aggraver au fur et à mesure que l'enquête se prolonge. Il faut aussi prendre des précautions dans la gestion du réseau des enquêteurs, surtout avoir une équipe qui soit spécialisée dans le suivi de ces données, de la collecte à l'analyse et qui soit suffisamment motivée.

Qualité de l'interview

Les enquêteurs doivent expliquer exactement l'enquête de transport avec GPS et s'assurer que les enquêtés comprennent exactement ce qu'on leur demande de faire. L'enquêteur doit pouvoir clarifier les doutes que l'enquêté peut avoir.

Les enquêteurs devront faire des efforts pour convaincre les enquêtés de participer à l'enquête de transport avec GPS en soulignant l'aspect expérimental d'une nouvelle méthode d'enregistrement et de post-traitement des données GPS, permettant de faciliter la collecte de données, réduire la charge de l'enquêté en portant seulement quelques jours le GPS au lieu de remplir complètement un carnet de déplacements. Les enquêteurs devront également mettre en avant la simplicité du dispositif GPS, la précision des données. Les enquêteurs doivent faire comprendre aux enquêtés que le récepteur GPS est seulement un outil de suivi passif, et que l'enquêté ne doit pas se sentir surveillé, s'il n'y a pas de transmission des données en temps réel. Les enquêteurs devront également souligner le caractère anonyme de la collecte de données et la confidentialité des données obtenues qui seront utilisées uniquement pour la recherche par une équipe de recherche. L'enquêteur devra également rassurer l'enquêté quant aux risques de perte ou d'endommagement de l'appareil : il ne devra bien sûr pas prendre en charge lui-même les frais associés.

Pour ce faire, les enquêteurs seront formés sur toutes les questions évoquées ci-dessus et recevront des documents qui les aideront à expliquer le dispositif et qui leur permettront de s'assurer que les enquêtés n'ont pas de doutes et comprennent clairement ce qu'on leur demande de faire.

L'enquêteur peut faire référence aux instructions, montrer le GPS afin d'expliquer à l'enquêté : comment et quand charger, transporter, allumer, éteindre l'appareil GPS. Les informations des notices et brochures sont très utiles pour rappeler l'utilisation du GPS et l'objectif de l'expérimentation de nouvelles technologies.

Appui et motivations

Il sera également important que les enquêtés sentent un appui tout au long de période de l'enquête par la possibilité de contacter l'enquêteur ou quelqu'un de l'INSEE ou l'IFSTTAR par l'email, téléphone ou web, en particulier lorsqu'ils ont des questions sur l'appareil GPS ou sur l'enquête par GPS. Une copie des notices et brochures comportant des conseils sur l'utilisation de l'appareil GPS sera laissée à chaque enquêté volontaire pour le volet GPS.

Tout comme le ménage reçoit un cadeau s'il remplit complètement le carnet de déplacements, l'enquêté volontaire pour le volet GPS devrait recevoir également un autre cadeau s'il rend le GPS à l'enquêté et répond au questionnaire complémentaire CAPI-GPS lors de la deuxième visite.

Contrôle de l'utilisation du GPS – Rappel des données manquantes à cause de l'oubli du GPS

Savoir si l'enquêté a pensé à prendre le récepteur GPS et à le charger chaque jour pendant la période d'enquête est très important. Nous aurions besoin d'un outil qui pourrait nous aider à l'interprétation des informations de l'utilisation de GPS et à l'identification des causes potentielles de déplacements manquants.

On pourrait demander à l'enquêté de remplir un petit document papier très simple sur l'utilisation du GPS : il pourrait y noter pour chaque jour au cours de la période d'enquête s'il a pris l'appareil et s'il a chargé le GPS ou bien s'il ne s'est pas déplacé du tout ce jour-là. Pour combler les données manquantes provenant de l'oubli de prendre / charger le GPS, nous pourrions demander à l'enquêter de noter ces déplacements perdus dans un petit carnet ou par le biais d'un site web dédié au volet GPS.

Disponibilité de l'appareil GPS

En nous remémorant le processus de collecte des données GPS, nous remarquons que la raison principale du faible taux de réponse GPS était le manque de dispositifs GPS ou le retard pour les rendre à nouveau disponibles. Nous pouvons augmenter le taux de remise des GPS par l'enquêté en augmentant le taux de disponibilité des GPS. Pour ce faire, nous avons besoin de calculer au mieux durées entre deux visites de l'enquêté pour baisser le taux d'immobilisation forcée des GPS. Nous entendons par "immobilisation forcée" les cas où des GPS restent dans une Direction Régionale de l'INSEE ou à l'INRETS (maintenant IFSTTAR), et ne sont proposés à aucun enquêteur GPS : soit parce qu'ils servent de "réserve", soit parce qu'aucun enquêteur GPS n'est disponible durant une certaine période dans une Direction Régionale.

Post-traitement des données GPS

L'appariement des données collectées par GPS et par CAPI est la source pour choisir les variables significatives dans modèle probabiliste conditionnelle de l'algorithme de post-traitement des données

GPS et pour l'évaluation des résultats statistiques de l'enquête transport. Les caractéristiques techniques de l'Enquête Nationale Transport peuvent poser certains problèmes et rendre complexe une comparaison avec l'enquête GPS. Il y a par exemple des différences de définition dans le jour d'enquête entre deux types de méthodes, pour l'enquête classique, le jour commence de 4 heures jusqu'à 4 heures de lendemain, mais pour l'enquête GPS, c'est pareil de définition de jour d'habitude. En outre, le traitement des données GPS et l'appariement avec déplacement déclaré par CAPI a été rendu plus difficile par le niveau de données d'origine et de destination collectées par l'enquête classique qui sont moins précises au niveau de code de commune. Le codage de ménage, enquêté entre deux enquêtes cause la difficulté de l'appariement des données. Nous avons besoin d'une harmonisation au niveau des définitions, du codage, de la précision relatives de données, etc.

Le Système d'Information Géographique et la carte map-matching de Mappoint 2006 n'a pas fourni bien les points d'intérêt dans le post-traitement Decoutrace de données brutes GPS. Le développement de nouvelles technologies nous aide d'améliorer les précisions de cartographie.

Questionnaire complémentaire CAPI-GPS

Le court questionnaire CAPI-GPS devrait permettre d'obtenir davantage d'informations sur les habitudes de conduite, la fréquence d'utilisation de Vélo/ Voiture/Moto/ Transport commun (bus, métro, train, aérien) ; les lieux fréquemment visités par l'enquêté, tels que le travail, l'étude, le supermarché, la crèche / école de l'enfant, le logement des parents ou amis de l'enquêté, ..., sous réserve des règles de confidentialité.

Il est possible que ces adresses seraient incomplètes, par exemple, l'enquêté pourrait informer qu'il est allé au supermarché Carrefour Belle Epine, mais il ne serait pas susceptible de connaître ses noms de rue ou code postal. Nous avons provisoirement un programme pour filtrer les adresses de lieux que l'enquêté fournit. Si l'adresse était encore incomplète à ce programme, nous avons besoin de consultation sur l'Internet par exemple Google Map ou annuaire téléphonique pour s'assurer que les adresses étaient aussi complétées que possible.

Enquête permanente

L'Enquête Nationale Transport est effectuée en général environ tous les dix ans en France. Sept ans ont déjà passé depuis sa dernière réalisation en 2007-08. Il faut prévoir beaucoup de temps dans la préparation avant de pouvoir lancer une nouvelle enquête au niveau national. S'appuyer sur les nouvelles technologies comme le GPS est très prometteur pour collecter régulièrement des données de mobilité. Il nous donne l'opportunité d'obtenir des données plus fréquentes avec certains avantages que les enquêtes classiques ne fournissent pas facilement. Nous pouvons penser qu'une opération permanente dont le budget serait prévu chaque année, à condition de contrôler que les résultats correspondent bien à ce qu'on attend, serait moins lourde et moins difficile à monter qu'une

opération exceptionnelle comme c'est actuellement le cas pour l'Enquête Nationale Transports en France. Et pour des questions d'encadrement, de montage et de suivi de l'enquête, des équipes et des enquêteurs, il y aurait sûrement beaucoup d'avantage à opérer en continu.

7.3 Prospective

Cette étude démontre que la technologie GPS est très prometteuse pour améliorer la collecte de données de transport. L'enquête à l'aide de GPS peut être utilisée avec succès pour compléter les enquêtes de transport classique. En synergie avec cette enquête, le GPS peut constituer un outil de comparaison et une méthode pour produire des facteurs de correction sur quelques variables importantes. Cependant, il y a un certain nombre de questions qui doivent être résolues avant de réaliser une enquête sur la mobilité à grande échelle.

Il existe trois principaux domaines où la poursuite des travaux est recommandée :

- Notre études peuvent imputer le moyen de transport, mais pas encore de mode de transport à cause de limites de protocole de l'Enquête de Transport 2007-08. Si dans l'immeuble de l'enquêté, il y a de supermarché, crèche, centre de loisirs, cabinet médicale, avec les limites de technologie de récepteur GPS et la cartographie de la pauvreté de points d'intérêt au moment de l'année 2007-08, il est difficile de distinguer que l'enquêté retourne au domicile, fait de l'achat, de loisirs ou de visite médicale, etc.
- Côté des avantages que l'enquête traditionnelle ne peut pas fournir, le GPS a aussi des inconvénients. Pour les améliorer, d'une part, nous devons réaliser régulièrement les tests avec les nouvelles générations du GPS pour choisir un GPS pertinent relativement avec l'enquête de transport personnel. Nous pouvons combiner les GPS et d'autres technologies différentes pour obtenir les informations supplémentaires comme utilisation d'internet Web, du téléphone, de l'application de Smartphone, etc. D'autre part, nous pouvons développer le traitement de données GPS pour imputer des modes et motifs de transport.
- Nous avons besoin d'échanger et de communiquer des sources de données GPS, ainsi que les méthodes de collecte et de traitement des données GPS entre différents organismes au niveau régional ou national pour faire progresser les connaissances dans le domaine des enquêtes sur la mobilité des personnes à l'aide de récepteurs GPS.

8 Annexes

8.1 Annexe 1 : Notice GPS





Objet de la notice GPS : expliquer le contenu du « pack GPS », fonctionnement du récepteur, quelques conseils d'utilisation, consignes et remarques pour l'enquêté ; ainsi que les matériaux et documents fournis aux enquêteur, le principe de la chaîne GPS pour l'enquêté.

Contenu du « pack GPS » fournis aux enquêtés

Le « pack GPS » remis aux enquêtés, est un petit sac contenant l'ensemble du matériel : Le récepteur GPS

- Un petit étui (qui permet de fixer le récepteur à la ceinture ou à un sac à main, etc.)
- Un câble allume-cigare
- Un adaptateur secteur
- Une notice GPS pour l'enquêté (recto-verso)

Figure 8.1 : Contenu du pack GPS

	
1. Récepteur GPS	2. Etui
	
3. Câble allume-cigare	4. Adaptateur secteur
5. La présente Notice GPS pour l'enquêté	

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Fonctionnement du récepteur GPS

Le récepteur GPS comprend uniquement un interrupteur : arrêt-marche.

Quand l'appareil est allumé : un voyant vert s'allume.

Ce voyant vert clignote quand il enregistre les données, donc quand le récepteur « voit » suffisamment de satellites.

Lorsque la batterie commence à être faible, un voyant orange clignote.

Lorsqu'elle est complètement « à plat », plus aucun voyant ne s'allume.

Quand on recharge la batterie (soit par le cordon allume-cigare, soit par l'adaptateur secteur) : le voyant orange s'allume, puis s'éteint quand elle est rechargée.

Quelques conseils d'utilisation

Si on utilise le sac, il faut placer le GPS dans la poche avant (avec fermeture éclair)

Il est conseillé de vérifier le clignotement du voyant vert, avant de placer le GPS dans son sac, son étui, ou dans le véhicule.

Dans les transports en commun, il est préférable, dans la mesure du possible de se placer à proximité d'une fenêtre pour améliorer la réception.

Si vous oubliez d'allumer le récepteur lors d'un trajet aller, vous pouvez malgré tout l'allumer pour le retour : l'information recueillie nous aidera.

Consignes importantes

(À reprendre sur le bordereau d'accompagnement)

Recharger l'appareil tous les soirs à l'aide de l'adaptateur fourni.

Ne pas laisser l'appareil en fonctionnement dans une voiture en stationnement (risque de vol et de saturation de la mémoire).

Ne pas le laisser non plus en fonctionnement dans un lieu ouvert (terrasse) ou proche d'une fenêtre (saturation de la mémoire)

En cas d'utilisation en voiture, il n'est pas souhaitable que les voyants soient visibles du conducteur ni qu'il surveille aucunement le GPS.

Autres remarques

(Si repris sur le bordereau, en caractères de taille illisible)

Cet appareil se contente d'enregistrer les emplacements où vous vous trouvez.

Il ne communique aucun renseignement durant l'enquête (ne permet pas de déterminer où vous êtes)

Les données sont utilisées seulement lors du retour de l'enquêteur pour le second entretien. En cas de nécessité, vous pourrez demander à ce que les traces ne soient pas exploitées.

L'appareil n'a aucun cadran permettant de savoir où l'on est, il est donc différent des GPS de loisir. L'appareil ne contient pas de service de "guidage", il est donc différent d'un modèle utilisé comme navigateur dans une automobile.

Le GPS RBT-3000 ou RBT1100 est robuste, mais pas spécialement étanches, il doit normalement être utilisé dans une poche, un sac ou une voiture.

En cas de besoin momentané de confidentialité, l'utilisateur peut glisser le commutateur sur la position 'O' : la diode verte s'éteint.

L'appareil se recharge à l'aide du cordon fourni sur le secteur 220 volts (éventuellement 12 volts d'allume cigare).

L'autonomie électrique de l'appareil est d'une dizaine d'heures pour le GPS RTB-3000, seize heures pour le GPS RBT-1100, nous vous invitons à l'éteindre si possible, par exemple pendant votre journée de bureau.



Certains GPS RBT-3000 sont équipés d'un dispositif d'économie d'énergie qui les mettent en sommeil en cas d'immobilité (plus de 10 minutes) ou de mouvement de très longue durée (plus de 2 heures)

L'appareil n'est pas particulièrement rayonnant, mais il ne dispose pas de permission particulière, la réglementation vous oblige à l'éteindre lors d'un trajet aérien.

L'appareil est capable de mémoriser environ 30000 points, il est réglé pour enregistrer un point toutes les 10 secondes, il conservera donc plus de 40 heures de relevé.

Matériel et documents fournis aux enquêteurs

Figure 8.2 : Matériel et documents fournis aux enquêteurs

	
<p>1. Carte mémoire SD: doit être insérée dans son logement du début à la fin des enquêtes</p>	<p>2. Adaptateur USB – Bluetooth: doit être inséré avant chaque déchargement sur le poste</p>
<p>3. Mode d'emploi GPS pour l'enquêteur</p>	<p>4. Memento GPS pour l'enquêteur</p>

Source : INSEE- SOeS-INRETS, Enquête Nationale Transports et Déplacements 2007-08

Principe de la chaîne GPS pour l'enquêteur

Lors de la première visite, l'enquêteur remet le « Pack GPS » à l'enquêté volontaire, remplit la « Fiche de suivi pack GPS », et fait signer la « Fiche de remise du pack GPS » à l'enquêté.

Lors de la seconde visite, il récupère le « Pack GPS » et complète les deux fiches.

Il prépare le déchargement des données enregistrées par le GPS dans la période entre les deux visites en exécutant les étapes suivantes sur son ordinateur portable :

- Connecter la clé Bluetooth dans un port USB du portable
- Placer la carte mémoire
- Placer le GPS à proximité de l'ordinateur
- Allumer le GPS
- Sur l'ordinateur, faire : 'Démarrer', puis 'Invite de commandes'
- Dans la fenêtre de commandes, frapper la ligne suivante (terminée par la frappe de la touche <Enter>) :
 - START MENUGPS
 - Decouverte du GPS, Couplage, Transfert, Effacement
- MenuGPS assure successivement :
 - Recherche de GPS
 - Couplage du GPS
 - Connexion au GPS.
 - Lancement du programme de déchargement (ENTGPS)
- Dans la fenêtre ' Moniteur ' qui apparaît, entrez la FICHE_ADRESSE. Le programme de déchargement est lancé, nous pouvons maintenant passer au CAPI de la deuxième visite.
- Lorsque vous aurez fini le CAPI de la deuxième visite, le programme ANATRACES nous permet de retrouver la fenêtre de Questionnement complémentaire GPS.

A la fin du questionnaire, la fenêtre se fermera automatiquement.

8.2 Annexe 2 : Imputation des modes de traces GPS à partir des données complémentaires CAPI-GPS

Informations sur le modèle	
Table	WORK.COMPARE_CAPITRACE
Variable de réponse	log_mode
Nombre de niveaux de réponse	5
Modèle	logit cumulé
Technique d'optimisation	Score de Fisher

Nb d'observations lues	841
Nb d'observations utilisées	841

Profil de réponse		
Valeur ordonnée	log_mode	Fréquence
1	Autre Mode	8
2	Marche à pied	89
3	Voiture/Moto	721
4	Transport commun (Bus, Métro, Train)	16
5	Vélo	7

Estimations par l'analyse du maximum de vraisemblance						
Paramètre		DDL	Estimation	Erreur type	Wald Khi-2	Pr > ChiSq
Intercept	Autre mode	1	-5.5717	0.4401	160.2704	<.0001
	Marche à pied	1	-2.1624	0.2244	92.8710	<.0001
	Voiture/Moto	1	4.8969	0.3010	264.6439	<.0001
	Transport commun	1	6.1325	0.4369	197.0571	<.0001
Vitesse maximum	1--7	1	2.1589	0.3626	35.4461	<.0001
	7--20	1	-1.3450	0.4037	11.1018	0.0009
Vitesse minimum	1--7	1	-0.9261	0.3156	8.6138	0.0033

	7--20	1	0.7726	0.3886	3.9522	0.0468
Distance linéaire	0--1	1	1.6741	0.2555	42.9461	<.0001
	1--6	1	-0.5666	0.1814	9.7548	0.0018
Durée	0--5	1	-0.8659	0.2447	12.5245	0.0004
	5--15	1	0.1331	0.1903	0.4892	0.4843
	5--30	1	0.0471	0.2241	0.0442	0.8335
Quand l'enquête se déroule avec le vélo	Il y a moins d'une semaine	1	-1.0549	0.3199	10.8739	0.0010
	Entre une semaine et un mois	1	0.2792	0.3259	0.7340	0.3916
	Entre un mois et plus de deux ans	1	0.3609	0.1934	3.4803	0.0621
Distance moyenne de déplacements quotidiens la veille de l'interview	0	1	-0.6137	0.2215	7.6796	0.0056
	0--10	1	0.00680	0.2621	0.0007	0.9793
	10--20	1	-0.0153	0.2833	0.0029	0.9568
	20--40	1	0.5832	0.2178	7.1708	0.0074

Variables	Modalités	Rapport des cotes (odds ratios)	95% Intervalle de confiance de Wald	
Vitesse maximum (km/h)	1--7	19.548	6.456	59.184
	7--20	0.588	0.167	2.075
	>20	Référence		
Vitesse minimum (km/h)	1--7	0.340	0.133	0.864
	7--20	1.857	0.554	6.231
	>20	Référence		
Distance linéaire (km)	0--1	16.147	6.436	40.510
	1--6	1.718	0.863	3.418
	>6	Référence		
Durée (min)	0--5	0.212	0.087	0.514
	5--15	0.575	0.266	1.245
	5--30	0.528	0.245	1.140
	>30	Référence		
Quand l'enquête se déroule avec	Il y a moins d'une semaine	0.230	0.094	0.562

le vélo	Entre une semaine et un mois	0.873	0.349	2.183
	Entre un mois et plus de deux ans	0.947	0.540	1.662
	Jamais	Référence		
Distance moyenne de déplacements quotidiens la veille de l'interview (km)	0	0.521	0.268	1.013
	0--10	0.968	0.448	2.093
	10--20	0.947	0.417	2.149
	20--40	1.723	0.887	3.346
	>40	Référence		

8.3 Annexe 3 : Identification des déplacements quotidiens par CAPI non trace GPS

Informations sur le modèle	
Table	WORK.DEP
Variable de réponse	log_depapparie
Nombre de niveaux de réponse	2
Modèle	logit cumulé
Technique d'optimisation	Score de Fisher

Nb d'observations lues	2294
Nb d'observations utilisées	2294

Response Profile		
Variable: log_depapparie	Modalité de déplacements quotidiens	Fréquence
0	Déplacement quotidien non apparié avec trace GPS	1006
1	Déplacement apparié avec trace GPS	1288
	Total	2294

Probability modeled is log_depapparie=0.

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	216.1747	34	<.0001
Score	204.9374	34	<.0001
Wald	187.3911	34	<.0001

Facteur	DF	Wald Khi-2	Pr > Chi-2
Motif de déplacement	4	32.30	<.0001
Distance de déplacement	3	42.55	<.0001
Nombre de déplacements quotidiens réalisés la veille de l'interview	3	20.94	0.0001
Profession de l'enquêté	7	28.11	0.0002
Type du ménage de l'enquêté	4	17.11	0.0018

Sexe de l'enquêté	1	7.40	0.0065
Etat de santé général de l'enquêté	3	10.69	0.0135
Motorisation (nombre voiture dans le ménage)	3	9.52	0.0231
Mode de déplacement	6	14.06	0.0290

Analysis of Maximum Likelihood Estimates						
Paramètre	Modalité	DF	Estimate	Standard Error	Wald Chi-2	Pr > Khi-2
Intercept		1	-0.7147	0.3815	3.5104	0.0610
Motif de déplacement quotidien	Aller au domicile	1	-0.00686	0.0747	0.0084	0.9268
	Achat	1	0.5082	0.0987	26.4917	<.0001
	Loisir	1	-0.2500	0.1196	4.3653	0.0367
	Profession	1	-0.3852	0.1249	9.5074	0.0020
Distance de déplacement quotidien	moins de 2 km	1	0.2738	0.0936	8.5509	0.0035
	de 2 à 5 km	1	0.2507	0.0830	9.1205	0.0025
	de 5 à 10 km	1	-0.0140	0.0857	0.0268	0.8700
Nombre de déplacements quotidiens réalisés dans la veille de l'interview	1-2	1	-0.4768	0.1104	18.6642	<.0001
	3-4	1	0.0846	0.0784	1.1652	0.2804
	5-6	1	0.2612	0.0785	11.0551	0.0009
Profession	Agriculteurs exploitants	1	0.2068	0.2849	0.5270	0.4679
	Cadres, professions intellectuelles supérieures	1	-1.1933	0.5856	4.1520	0.0416
	Professions intermédiaires	1	0.2731	0.1538	3.1549	0.0757
	Employés	1	0.5288	0.1392	14.4322	0.0001
	Ouvriers	1	-0.1987	0.1317	2.2763	0.1314
	Retraités	1	0.1955	0.1571	1.5493	0.2132
	Autres sans actifs	1	0.2928	0.1417	4.2727	0.0387
Type du ménage	Personne seule	1	-0.3238	0.1129	8.2280	0.0041
	Famille monoparentale	1	-0.2861	0.1788	2.5605	0.1096
	Couple sans enfant	1	-0.1600	0.1030	2.4142	0.1202

	Autre (ménage complexe)	1	0.0674	0.1018	0.4378	0.5082
Sexe	Femme	1	-0.1389	0.0511	7.4007	0.0065
Etat de santé	Très Bon	1	2.292	0.239	21.947	2.292
	Bon	1	2.961	0.309	28.333	2.961
	Moyen	1	3.422	0.354	33.088	3.422
Motorisation (nombre voiture dans le ménage)	0	1	0.518	0.337	0.797	0.518
	1	1	0.872	0.695	1.093	0.872
	2	1	0.876	0.613	1.253	0.876
Mode de déplacement	Bus	1	8.907	0.661	120.044	8.907
	Métro	1	3.790	0.322	44.643	3.790
	Train	1	1.482	0.102	21.453	1.482
	Moto/ Voiture	1	2.643	0.261	26.793	2.643
	Marche à pied	1	4.061	0.399	41.318	4.061
	Vélo	1	4.881	0.435	54.793	4.881

Odds Ratio Estimates				
Effect	Modalité	Point Estimate	95% Wald Confidence Limits	
Motif de déplacement	Aller au domicile	0.869	0.690	1.093
	Achat	1.454	1.100	1.922
	Loisir	0.681	0.490	0.947
	Profession	0.595	0.423	0.837
	Autre motif	Référence		
Distance de déplacement	moins de 2 km	2.191	1.647	2.914
	de 2 à 5 km	2.141	1.651	2.777
	de 5 à 10 km	1.643	1.269	2.127
	égale ou plus de 10 km	Référence		
Nombre de déplacements quotidiens réalisés la veille de l'interview	1-2	0.545	0.394	0.753
	3-4	0.955	0.752	1.212
	5-6	1.139	0.899	1.443
	>=7	Référence		
Profession	Agriculteurs exploitants	1.366	0.687	2.717
	Artisans, commerçants	0.337	0.089	1.277

	Cadres, professions intellectuelles supérieures	1.460	0.953	2.236
	Professions intermédiaires	1.885	1.258	2.825
	Employés	0.911	0.626	1.326
	Ouvriers	1.351	0.872	2.092
	Retraités	1.489	0.994	2.230
	Autres non actifs	Référence		
Type de ménage	Personne seule	0.358	0.206	0.622
	Famille monoparentale	0.372	0.194	0.715
	Couple sans enfant	0.422	0.248	0.718
	Couple avec enfant	0.530	0.314	0.895
	Autres types de ménages	Référence		
Sexe	Homme	0.757	0.620	0.925
	Femme	Référence		
Etat de santé	Très bon	2.292	0.239	21.947
	Bon	2.961	0.309	28.333
	Moyen	3.422	0.354	33.088
	Mauvais	Référence		
Motorisation (nombre voiture dans le ménage)	0	0.518	0.337	0.797
	1	0.872	0.695	1.093
	2	0.876	0.613	1.253
	>2	Référence		
Mode de déplacement	Autre	Référence		
	Bus	8.907	0.661	120.044
	Métro	3.790	0.322	44.643
	Train	1.482	0.102	21.453
	Moto/ Voiture	2.643	0.261	26.793
	Marche à pied	4.061	0.399	41.318
	Vélo	4.881	0.435	54.793

8.4 Annexe 4 : Identification des traces GPS non appariées avec déplacement quotidien par CAPI

Informations sur le modèle	
Table	WORK.DEP
Variable de réponse	log_traceapparie
Nombre de niveaux de réponse	2
Modèle	logit cumulé
Technique d'optimisation	Score de Fisher

Nb d'observations lues	2851
Nb d'observations utilisées	2851

Response Profile		
Variable:	Modalité de déplacements quotidiens	Fréquence
log_depapparie		
0	Trace GPS non appariée avec déplacement quotidien	1563
1	Trace GPS appariée avec déplacement quotidien	1288
	Total	2851

Probability modeled is log_traceapparie=0.

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	937.8913	39	<.0001
Score	827.2972	39	<.0001
Wald	625.0406	39	<.0001

Facteur	DF	Wald Khi-2	Pr > Chi-2
Nombre de traces GPS	3	166.7260	<.0001
Vitesse sur la trace GPS	4	51.2141	<.0001
Durée de la trace GPS	4	102.5806	<.0001

Nombre moyen de déplacements quotidiens observés par CAPI	4	60.1618	<.0001
Région de l'enquête	11	27.7060	0.0036
Profession de l'enquête	7	20.2003	0.0052
Distance linéaire de trace GPS	3	11.5147	0.0092

Analysis of Maximum Likelihood Estimates						
Paramètre	Modalité	DF	Point Estimate	Standard Error	Wald Chi-2	Pr > Chi-2
Intercept		1	-1.2894	0.1376	87.7653	<.0001
Nombre de traces GPS dans la veille de l'interview	1-2	1	-1.3808	0.1444	91.4099	<.0001
	3-4	1	0.1260	0.0986	1.6323	0.2014
	5-6	1	0.2619	0.1025	6.5259	0.0106
Vitesse de trace GPS	< 7	1	0.5936	0.1036	32.7992	<.0001
	7-20	1	0.2991	0.1076	7.7253	0.0054
	20-50	1	-0.4130	0.0888	21.6447	<.0001
	50-80	1	-0.4391	0.1513	8.4260	0.0037
Durée de la trace GPS	moins de 5 minutes	1	1.2238	0.1227	99.5076	<.0001
	de 5 à 15 minutes	1	0.5738	0.1120	26.2443	<.0001
	de 15 à 30 minutes	1	0.0631	0.1203	0.2751	0.5999
	de 30 à 60 minutes	1	-0.6066	0.1674	13.1372	0.0003
Nombre moyen de déplacements quotidiens observés par CAPI	0	1	0.3863	0.1342	8.2924	0.0040
	1--2	1	0.4429	0.1060	17.4567	<.0001
	3--4	1	0.1431	0.0842	2.8885	0.0892
	5--6	1	-0.3416	0.0935	13.3595	0.0003
Région de résidence de l'enquête	11 - Ile de France	1	-0.2130	0.1173	3.2968	0.0694
	22 - Picardie	1	0.000813	0.1743	0.0000	0.9963
	23 - Haute Normandie	1	0.0883	0.1572	0.3155	0.5743
	26 - Bourgogne	1	-0.3056	0.2328	1.7230	0.1893
	31 - Nord Pas de Calais	1	0.1156	0.2357	0.2408	0.6236
	53 - Bretagne	1	-0.2150	0.2030	1.1215	0.2896
	54 - Poitou Charentes	1	-0.1700	0.1394	1.4864	0.2228
	72 - Aquitaine	1	-0.00701	0.1558	0.0020	0.9641
	73 - Midi Pyrénées	1	0.5199	0.1479	12.3541	0.0004
	82 - Rhône Alpes	1	0.0935	0.1316	0.5047	0.4775
83 - Auvergne	1	0.4736	0.1821	6.7667	0.0093	

Profession de l'enquêté	Agriculteurs exploitants	1	0.2635	0.2794	0.8888	0.3458
	Cadres, professions intellectuelles supérieures	1	-1.2149	0.6143	3.9117	0.0479
	Professions intermédiaires	1	0.0590	0.1559	0.1431	0.7052
	Employés	1	0.1798	0.1464	1.5083	0.2194
	Ouvriers	1	-0.1904	0.1321	2.0775	0.1495
	Retraités	1	0.3945	0.1577	6.2553	0.0124
	Autres non actifs	1	0.2677	0.1323	4.0957	0.0430
Distance linéaire	moins de 2 km	1	0.3536	0.1094	10.4474	0.0012
	de 2 à 5 km	1	-0.1080	0.1042	1.0756	0.2997
	de 5 à 10 km	1	-0.1114	0.1150	0.9372	0.3330

Odds Ratio Estimates				
Facteur de trace GPS	Modalité	Point Estimate	95% Wald Confidence Limits	
Nombre de trace GPS la veille de l'interview	1--2	0.093	0.063	0.137
	3--4	0.420	0.326	0.541
	5--6	0.481	0.370	0.626
	>=7	Référence		
Vitesse sur la trace (en km/h)	< 7	1.885	1.440	2.469
	7--20	1.404	1.053	1.872
	20--50	0.689	0.539	0.882
	50--80	0.671	0.459	0.983
	>=80	Référence		
Durée de trace GPS	moins de 5 minutes	11.918	5.985	23.733
	de 5 à 15 minutes	6.221	3.148	12.295
	de 15 à 30 minutes	3.733	1.900	7.334
	de 30 à 60 minutes	1.911	0.923	3.957
	>= 60 minutes	Référence		
Nombre moyen des déplacements quotidiens de l'enquêté observés par CAPI	0	2.765	1.868	4.092
	1--2	2.926	2.102	4.075
	3--4	2.168	1.644	2.859
	5--6	1.335	0.993	1.796

	>=7	Référence		
Région de résidence	11 - Ile de France	1.183	0.651	2.151
	22 - Picardie	1.465	0.759	2.829
	23 - Haute Normandie	1.599	0.842	3.037
	26 - Bourgogne	1.078	0.515	2.256
	31 - Nord Pas de Calais	1.643	0.784	3.446
	41 - Lorraine	1.181	0.589	2.366
	53 - Bretagne	1.235	0.666	2.291
	54 - Poitou Charentes	1.454	0.769	2.747
	72 - Aquitaine	2.462	1.310	4.627
	73 - Midi Pyrénées	1.607	0.873	2.960
	82 - Rhône Alpes	2.351	1.200	4.604
	83 - Auvergne	Référence		
Profession	Agriculteurs exploitants	1.023	0.527	1.985
	Artisans, commerçants	0.233	0.057	0.950
	Cadres, professions intellectuelles supérieures	0.834	0.554	1.254
	Professions intermédiaires	0.941	0.636	1.392
	Employés	0.650	0.448	0.942
	Ouvriers	1.166	0.769	1.768
	Retraités	1.027	0.712	1.482
	Autres non actifs	Référence		
Distance linéaire de trace GPS	moins de 2 km	1.629	1.066	2.488
	de 2 à 5 km	1.026	0.681	1.547
	de 5 à 10 km	1.023	0.682	1.535

	égale ou plus de 10 km	Référence		
--	------------------------	-----------	--	--

9 Bibliographie

- Abdel-Aty, M., Kitamura, R., Jovanis, P., Reddy, P. and Vaughn, K. (1995) 'New Approach to Route Choice Data Collection: Multiphase, Computer-Aided Telephone Interview Panel Surveys Using Geographic Information Systems Data Base', *Transportation Research Record*, pp.159–169.
- Ben-Akiva, H. and Lerman, S. (1985) *Discrete Choice Analysis: Theory and Application to Travel Demand*, MIT Press.
- Ampt, E.-S., Wundke, J. and Stopher, P. (2006) *Households on the Move: Experiences of a New Approach to Voluntary Travel Behaviour Change*, Sydney, Institute of Transport and Logistics Studies.
- Armoogum, J. (2000) *Correction de La Non-réponse et de Certaines Erreurs de Mesure Dans Une Enquête Par Sondage : Application à L'enquête Transports et Communications 1993-94*, Thèse de doctorat, Université Libre de Bruxelles.
- Armoogum, J., Castaigne, M., Hubert, J.-P. and Madre, J.-L. (2005) 'Immobilité et Mobilité Observées à Travers Les Enquêtes Ménages de Transport Ou D'emploi Du Temps'. In *Journées de méthodologie statistique*, Paris, France.
- Armoogum, J., Hubert, J.-P., Bonnel, P. and Madre, J.-L. (2007) *Préparer la prochaine enquête nationale transport avec un regard international*, Rapport sur convention.
- Armoogum J., Hubert J.-P., François D., Roumier B., Roux S., B. Robin M., (2009) *Enquête nationale transport et déplacements 2007-2008 (ENTD 2007-2008)*, Rapport technique du SOeS
- Armoogum, J. and Madre, J.-L. (1997a) 'Du Redressement Des Non-réponses Totales Aux Contrôles Sur La Cohérence Des Réponses', *Recherche, transports et sécurité*, pp67–77.
- Armoogum, J. and Madre, J.-L. (1997b) 'Item Non-response in Travel Surveys', Grainau.
- Armoogum, J. and Madre, J.-L. (1997c) 'Accuracy of Data and Memory Effects in Home Based Surveys on Travel Behaviour', *Transportation Research Board*.
- Armoogum, J., Madre, J.-L., Gascon, M.-O. and François, D. (2010) 'Les Enquête Nationales et Locales Sur La Mobilité : Sources et Méthodes', *La revue CGDD*, 49–64.
- Armoogum, J. & al. (2012) *COST Action TU 0804 Survey HARmonisation with New Technologies Improvement (SHANTI), Final report*.
- Auld, J., Williams, C., Mohammadian, A. and Nelson, P. (2009) 'An Automated GPSbased Prompted Recall Survey with Learning Algorithms', *Transportation Letters: The International Journal of Transportation Research*, 1, 59–79.
- Auriat, N. (1996) : *Les défaillances de la mémoire humaine. Aspect cognitifs des enquêtes rétrospectives*, presses universitaires de France, INED.
- Axhausen, K.-W. (2003) 'Definitions and Measurement Problems', In *Capturing Long Distance Travel*, (2003).
- Bohte, W. and Maat, K. (2009) 'Deriving and Validating Trip Purposes and Travel Modes for Multi-day GPS-based Travel Surveys: A Large-scale Application in the Netherlands', *Transportation Research Part C: Emerging Technologies*, 17, 285–297.
- Bricka, S. and Bhat, C. (2006) 'Comparative Analysis of Global Positioning System-Based and Travel Survey-Based Data', *Transportation Research Record: Journal of the Transportation Research Board*, 1972, 9–20.

- Bricka, S. G., Sen, S., Paleti, R. and Bhat, C. R. (2012) 'An Analysis of the Factors Influencing Differences in Survey-reported and GPS-recorded Trips', *Transportation Research Part C: Emerging Technologies*, 21, 67–88.
- Brog, W. and Meyburg, A. H. (1980) 'Nonresponse Problem in Travel Survey: An Empirical Investigation', *Transportation Research Record*.
- Brog, W. and Meyburg, A. H. (1981) 'Consideration of Nonresponse Effects in Large-scale Mobility Surveys', *Transportation Research Record*.
- Bricka, S. and C. Bhat (2006). *Comparative Analysis of Global Positioning System-Based and Travel Survey-Based Data*, *Transportation Research Record No. 1972*, pp. 9-20.
- Casas, J. and Arce, C. (1999) 'Trip Reporting in Household Travel Diaries', paper presented at the 78th Annual Meeting of the Transportation Research Board, Washington D.C, January 1999.
- CERTU (2009) *Calcul à Posteriori Des Distance Dans Les Enquêtes Ménages Déplacements*, Guide méthodologique.
- CERTU (2011) 'La Mobilité Dans Les Villes Moyennes : Exploitation Des Enquêtes Déplacements Villes Moyennes 2000-2010'.
- CETE Nord-Picardie and INRETS (2006) *Le Diagnostic Environnemental de La Mobilité 2006*.
- Chalasanani, V. S., Denstadli, J. M., Engebretsen, O. and Axhausen, K. W. (2005) 'Precision of Geocoded Locations and Network Distance Estimates - TRB Transportation Research Information Services', *Journal of Transportation and Statistics*, 8, 1–15.
- Chen, C., Gong, H., Lawson, C. and Bialostozky, E. (2010) 'Evaluating the Feasibility of a Passive Travel Survey Collection in a Complex Urban Environment: Lessons Learned from the New York City Case Study', *Transportation Research Part A: Policy and Practice*, 44, 830–840.
- Collet, R. (2007) *Dynamique de la Motorisation, et Usage de l'Automobile En France*, Paris, Thèse de doctorat de l'université Paris 1 Panthéon Sorbonne.
- Dejoux, V. (2010) *Situation de Handicap Lors Des Déplacements : Caractéristiques Individuelles, Pratiques de Mobilité, Environnement Physique et Social*, Thèse de doctorat de l'Université Paris 1 Panthéon-Sorbonne.
- Department of Transport of UK (2012) *National Travel Survey 2011 GPS Pilot: Summary Analysis*, DfT
- Doherty, S. T., Papinski, D. and Lee-Gosselin, H. (2006) 'An Internet-Based Prompted Recall Diary with Automate GPS-Trip Detection: System Design', *TRB Annual Meeting (2006)*.
- Doherty, S.-T., Noël, N., Lee-Gosselin, H., Sirois, C. and Ueno, M. (2001) 'Moving Beyond Observed Outcomes: Integrating Global Positioning Systems and Interactive Computer-based Travel Behavior Surveys', *Transportation Research Circular E-C026 (March 2001)*.
- Draijer, G., Kalfs, N. and Perdok, J. (2000) 'Global Positioning System as Data Collection Method for Travel Research', *Transportation Research Record*, 1719, 147–153.
- Ettema, D.F., H.J.P. Timmermans and L. van Veghel., (1996) *Effects of Data Collection Methods in Travel and Activity Research*. European Institute of Retailing and Services Studies, Eindhoven, University of Technology, Eindhoven.
- Flavigny, P.-O. and Marchal, P. (2008) 'Explication Du CAPI - GPS de l'Enquête Nationale Sur Les Transports et Les Déplacements 2007-08', Note interne de l'Inrets.
- Forrest, T. and Pearson, D. (2005) 'Comparison of Trip Determination Methods in Household Travel Surveys Enhanced by a Global Positioning System', *Transportation Research Record*, 1917, 63–71.

- Godard, X., Diaz-Olvera, L. and Dieng, A. (2011) *Guide méthodologique d'enquête-ménages de mobilité dans les villes en développement*, MAE.
- Gourieroux, C. (1989) *Econométrie Des Variables Qualitatives*, Edition Economica, Paris.
- Greaves, S.-P. (1997) 'Applications of GIS Technology in Recent Household Travels Surveys', Proceedings of the Geographic Information System for Transportation (GIS-T) Symposium (March 1997).
- Hato, E. (2006) 'Development of MoALs (Mobile Activity Loggers Supported by Gpsphones) for Travel Behavior Analysis - Publications Index', TRB 85th Annual Meeting Compendium of Papers CD-ROM.
- Hubert, J.-P. (2009) 'Dans Les Grandes Agglomérations, La Mobilité Quotidienne Des Habitants Diminue, et Elle Augmente Ailleurs', Insee première.
- IAURIF (2012) *Faisabilité D'une Enquête Globale Transports (EGT) Intégrale Par Association D'un GPS, D'un SIG et D'un Système Expert En Île-de-France*, Rapport PREDIT, Paris.
- Itsubo, S. and Hato, E. (2006) 'Effectiveness of Household Travel Survey Using GPSEquipped Cell Phones and Web Diary: Comparative Study with Paper-Based Travel Survey', TRB 85th Annual Meeting Compendium of Papers CD-ROM.
- Kochan, B., Bellemans, T., Janssens, D. and Wets, G. (2006) 'Dynamic Activity-Travel Diary Data Collection Using a GPS-Enabled Personal Digital Assistant'. In AATT, Chicago, Illinois, USA.
- De Leeuw, E. (1992) 'Data Quality in Mail, Telephone and Face to Face Surveys', TT Publikaties (1992).
- De Leeuw, E. and Van de Joewen, J. (1992) 'Data Quality and Mode of Data Collection: Methodology and Explanatory Model'. In *Qualite de l'Informations dans les enquêtes*, Paris, pp. p11–31.
- Lollivier, S., Marpsat, M., Verger D. (2000) *L'économétrie et l'étude des comportements. Présentation et mise en œuvre de modèles de régression qualitatifs. Les modèles univariés à résidus logistiques ou normaux (LOGIT, PROBIT)*, Document de travail de l'INSEE série « Méthodologie Statistique » N°0001.
- Lu, Y., Zhu, S. and Zhang, L. (2012) 'A Machine Learning Approach to Trip Purpose Imputation in GPS-Based Travel Surveys', In the Proceedings of 91th Annual Conference Transportation Research Board (2012).
- Madre, J.-L. and Bussière, Y. (2002) 'Comment Développer L'approche Démographique ?' In *Démographie et transport : Villes du Nord et Villes du Sud*, Yves Bussière et Jean-Loup Madre.
- Marchal, P., Roux, S., Yuan, S., Hubert, J.-P., Armoogum, J., Madre, J.-L. and Lee-Gosselin, M. (2008) 'A Study of Non-Response in the GPS Sub-sample of the French National Travel Survey 2007-2008', Annecy, France.
- Moritz, G. and Brog, W. (1999) 'Redesign of the Dutch Travel Survey: Response and Improvement'. In TRB Conference on Personal Travel Survey: The long and short of it, Washington D.C, p. 13p.
- Murakami, E. and Wagner, D.-P. (1999) 'Can Using Global Positioning System (GPS) Improve Trip Reporting?', *Transportation Research Part C: Emerging Technologies*, 7, 149–165.
- Ohmori, N., Nakazato, M., Sasaki, K., Nishii, K. and Harata, N. (2006) 'Activity Diary Surveys Using GPS Mobile Phones and PDA', TRB 85th Annual Meeting Compendium of Papers CD-ROM (2006).
- Ortúzar, J. D. D., Armoogum, J., Madre, J. and Potier, F. (2011) 'Continuous Mobility Surveys: The State of Practice', *Transport Reviews*, 31, 293–312.
- Polak, J., Plaxton, J., Jackson, P. and Wofinden, D. (1999) 'Using the Internet to Improve Long-Distance Travel Data Collection', *Transportation Research Record: Journal of the Transportation Research Board*, 1660, 148–155.

- Potier, F. and Terrier, C. (2007) *Atlas des mobilités touristiques en France métropolitaine*, Paris, Autrement.
- Quételard, B. (2010) 'Se Rendre Au Travail Ou Faire Ses Courses Motive Toujours Un Déplacement Quotidien Sur Deux. Le Recours à La Voiture Se Stabilise', *La revue CGDD*, p25–47.
- Rietveld, P., Zwart, B., Van Wee, B. and Hoom, V.-D. (1999) 'On the Relationship Between Travel Time and Travel Distance of Commuters. Reported Versus Network Travel Data in the Netherlands', *The Annals of Region Science* 33 (1999).
- Roux, S. (2012) *Transition de La Motorisation En France Au XXème Siècle*, Thèse de doctorat, Université de Paris 1 Panthéon Sorbonne.
- Sammer, G. (1995) 'Problems and Solutions in Urban Travel Survey'. In *Les Enquêtes de Déplacement Urbains : Mesurer le Présent, Stimuler le Future*, Lyon, Bonnel P, et all., pp. pp145–159.
- Scheiner, J. (2010) 'Interrelations Between Travel Mode Choice and Trip Distance: Trends in Germany 1976-2002', *Journal of Transport Geography*, 18, 75–84.
- Schönfelder, S., Rich, J., Nielsen, O.-A., Würtz, C.-J. and Axhausen, K.-W. (2007) 'Road Pricing and Its Consequences for Individual Travel Patterns', *Mobilities*, 2, 75–98.
- Schuessler, N. and Axhausen, K.-W. (2009) 'Processing Raw Data from Global Positioning Systems Without Additional Information', *Transportation Research Record: Journal of the Transportation Research Board*, 2105, 28–36.
- Shen, L. and Stopher, P. R. (2013) 'A Process for Trip Purpose Imputation from Global Positioning System Data', *Transportation Research Part C: Emerging Technologies*, 36, 261–267.
- Solère, R. and Vulliet, D. (2008) 'Mobilités et Transports - Les Recueils de Données Sur La Mobilité Urbaine'.
- Stopher, P., (1992) *Use of an activity-based diary to collect household travel data*. *Transportation*, 19(2), 159-176.
- Stopher, P., Clifford, E., Zhang, J. and FitzGerald, C. (2008) *Deducing Mode and Purpose from GPS Data*, Working Paper of the Austrian Key Centre in Transport and Logistics, Sydney, N.S.W.
- Stopher, P., FitzGerald, C. and Xu, M. (2007) 'Assessing the Accuracy of the Sydney Household Travel Survey with GPS', *Transportation*, 34, 723–741.
- Stopher PR, Prasad C and Zhang J (2010) *Comparing GPS and Prompted Recall Data Records*, Working Paper ITLS-WP-10-24
- Stopher, P., Madre, J.-L., Armoogum, J., Bonnel, P. and Zhang, J. (2011) 'Harmonisation in National Travel Surveys'. In ISCTSC, Chili.
- Stopher, P. (2012) *Final Report for GPS-based Household Interview Survey for the Cincinnati, Ohio Region: Fiscal Year 2009*, Abt SRBI.
- Varshavsky, A., Chen, M., Froehlich, J., Haehnel, D., Hightower, J., Lamarca, A., Potter, F., Sohn, T., Tang, K. and Smith, I. (2006) 'Are GSM Phones THE Solution for Localization', In 7th IEEE Workshop on Mobile Computing System and Application, 20–28.
- Wagner, D. (1997) 'Lexington Area Travel Data Collection Test', Final Report for OHIM, OTA, and FHWA.
- Wagner, D., Malcosky, J., Kindateder, J. and Nguyen, J. (1998) 'Determining Heavy Duty Truck Activity Using GPS Technology', paper presented at the Presented at the 8th CRC On Road Vehicle Emission, San Diego, April 1998.
- Walmsley, D.-J. and Belloff, H. (1989) 'Urban Living: The Individual in the City', *Journal of Environmental Psychology*, 9, 174–176.

- Witlox, F. (2007) '*Evaluating the Reliability of Reported Distance Data in Urban Travel Behaviour Analysis*', Journal of Transport Geography, 15, 172–183.
- Wolf, J. (2000) '*Using GPS Data Loggers to Replace Travel Diaries in the Collection of Travel Data*', Dissertation, Georgia Institute of Technology, School of Civil and Environmental, 58–65.
- Wolf, J., Bricka, S., Ashby, T. and Gorugantua, C. (2004) '*Advances in the Applications of GPS to Household Travel Survey*', paper presented at the Transportation Research Board National Household Transportation Survey Conference, Washington D.C, 2004.
- Wolf, J., Guensler, R. and Bachman, W. (2001) '*Elimination of the Travel Diary: Experiment to Derive Trip Purpose from Global Positioning System Travel Data*', Transportation Research Record: Journal of the Transportation Research Board, 1768, 125–134.
- Yuan, S. (2010) '*Méthodes D'analyse des Données GPS Dans Les Enquêtes Sur La Mobilité Des Personnes - Les Données Manquantes et Leur Estimation*', Paris, Thèse de doctorat de l'Université Paris 1 Panthéon-Sorbonne.
- Zmud, J. & Arce, C., (2000) '*Item non-response in travel surveys: causes and solutions*'. Dans International Conference on Transport Survey Quality and Innovation. Grainau.
- Zmud, J. and Wolf, J. (2003) '*Identifying the Correlates of Trip Misreporting –Results from the California Statewide Household Travel Survey GPS Study*', Lucerne.