Cerema

La mesure des comportements de mobilité en France : une collaboration de longue date entre ingénieurs et statisticiens, sous le regard du chercheur

Barbara CHRISTIAN, Cerema

Un peu d'histoire : les outils du *traffic engineer*

XX^{ème} siècle : apparition et développement très rapide de la voiture d'abord aux Etats-Unis, puis en Europe, surtout après 1945

⇗ de la congestion dans les villes
⇗ accidents mortels dûs à la voiture

⇨ Construire des infrastructures pour la voiture « au bon endroit »

Mise au point de la « **matrice technique** » de l'ingénieur :

- Dispositif de quantification de la mobilité = estimer une photo du présent
- Dispositif de modélisation = estimer une photo du futur
- Dispositif d'évaluation socio-économique = analyse coût/avantage

⇨ L'objet « **déplacement** » lie les trois éléments de la matrice

⇨ Premières enquêtes ménage-déplacement (EMD) dès 1960 à Rennes, Nancy, Strasbourg par les ingénieurs du ministère (RST)

⇨ Le besoin local de ces enquêtes précède la mesure nationale avec un fort soutient de l'Etat encore aujourd'hui

Construction du
périphérique à Paris

1960-73

Mazamet ville morte
1973

Les enquêtes nationales, réalisées par les statisticiens

L'Enquête Nationale Transport

- Première ENT en 1966-67 réalisée par l'Insee, puis avec une fréquence tous les dix ans environ, la dernière étant en cours.
- L'Insee et les chercheurs du ministère des Transports (Ifsttar) ont œuvré souvent ensemble pour élaborer les différentes éditions de ces ENT.
- Basée sur les mêmes concepts de mesure du déplacement que les EMD, elle estime aussi la mobilité les we ou en période de vacances, et s'intéresse aux longs déplacements, contrairement aux EMD qui visent la mobilité quotidienne.

Le panel « Parcs autos »

- Enquête annuelle menée par TNS-Sofres (Kantar) ; existe depuis 1983.
- Elle renouvelle son échantillon par tiers ce qui permet de suivre le comportement des ménages enquêtés dans le temps et de mieux mesurer les changements à l'œuvre, ce que ne fait pas explicitement l'ENT.
- Par contre, le questionnement est plutôt centré sur l'usage de la voiture par le ménage, au contraire des EMD et de l'ENT qui captent tous les modes de déplacement.

Les grands principes de l'enquête ménage déplacement : l'échantillonnage

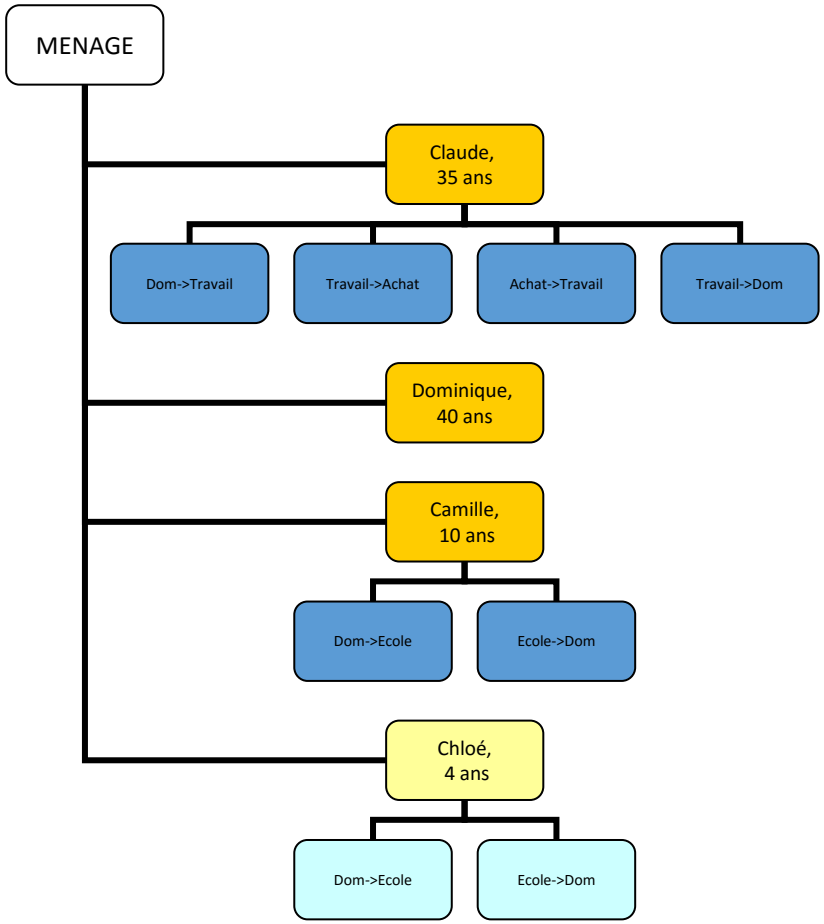
L'EMD alimente les besoins du modèle :

- Estimer **la matrice des flux** de déplacement de zone à zone avec une **précision à peu près équivalente** sur le périmètre du modèle (calibrage)
- Recueillir les **déplacements** effectués en les reliant aux **individus** qui les réalisent et aux **ménages** auxquelles ils appartiennent afin de modéliser les comportements (lois de comportement)

Plan de sondage :

- Base de sondage : liste de logements issue de source fiscale anonyme
- Echantillon de logements **stratifié géographiquement** selon h-strates de taille équivalente en population, et homogène en terme de comportement de mobilité (N_h variant de 1 à 3 en moyenne)
- Dans chaque strate, **sondage aléatoire simple** de logements selon un taux de sondage de 1 à 2% avec l'objectif d'atteindre 70 ménages et 160 individus minimum par strate => n_h / N_h peu dispersés

Les grands principes de l'enquête ménage déplacement : le protocole



Mode de collecte : Fàf ou Tél

Champ :

- Individu > 5 ans

Tous en Fàf ; 1 ou 2 au Tél selon taille

- Tous les déplacements effectués la veille de l'enquête quel que soit le motif ou le mode
- Du lundi au vendredi, hors vacances
- Entre octobre et mars

⇨ **Estimation de la mobilité des résidents « un jour moyen de semaine » au moment de la plus forte utilisation des réseaux**

Les grands principes de l'enquête ménage déplacement : définition du déplacement

Mouvement d'une personne effectué sur la voie publique pour…

un certain MOTIF *(pourquoi?)* entre une ORIGINE *(de ?)*

et une DESTINATION *(vers ?)*

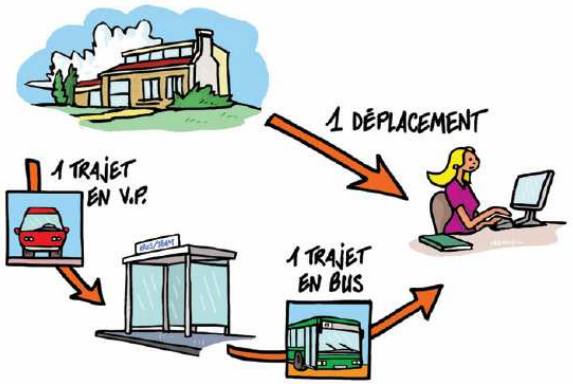
selon une heure de DÉPART *(quand ?)* et une heure d'ARRIVÉE

à l'aide d'un ou de plusieurs MODES de transport *(comment ?)*

=> la description des **individus**, des **ménages**, du **parc de véhicules**, des **conditions de stationnement** et **d'occupation de la voiture** s'y ajoutent

Les grands principes de l'enquête ménage déplacement : la hiérarchisation des modes

- Dans une EMD, un déplacement peut être effectué avec plusieurs modes de transport. L'EMD les distingue si ils sont mécanisés et décompose alors le déplacement en **trajets** pour prendre en compte l'intermodalité.
- Il faut alors déterminer le mode principal du déplacement : les modes de déplacement sont hiérarchisés en fonction des transports en commun, sans critère, ni de temps passé, ni de distance effectuée
- La marche est recensée comme mode si le déplacement est intégralement effectué à pied



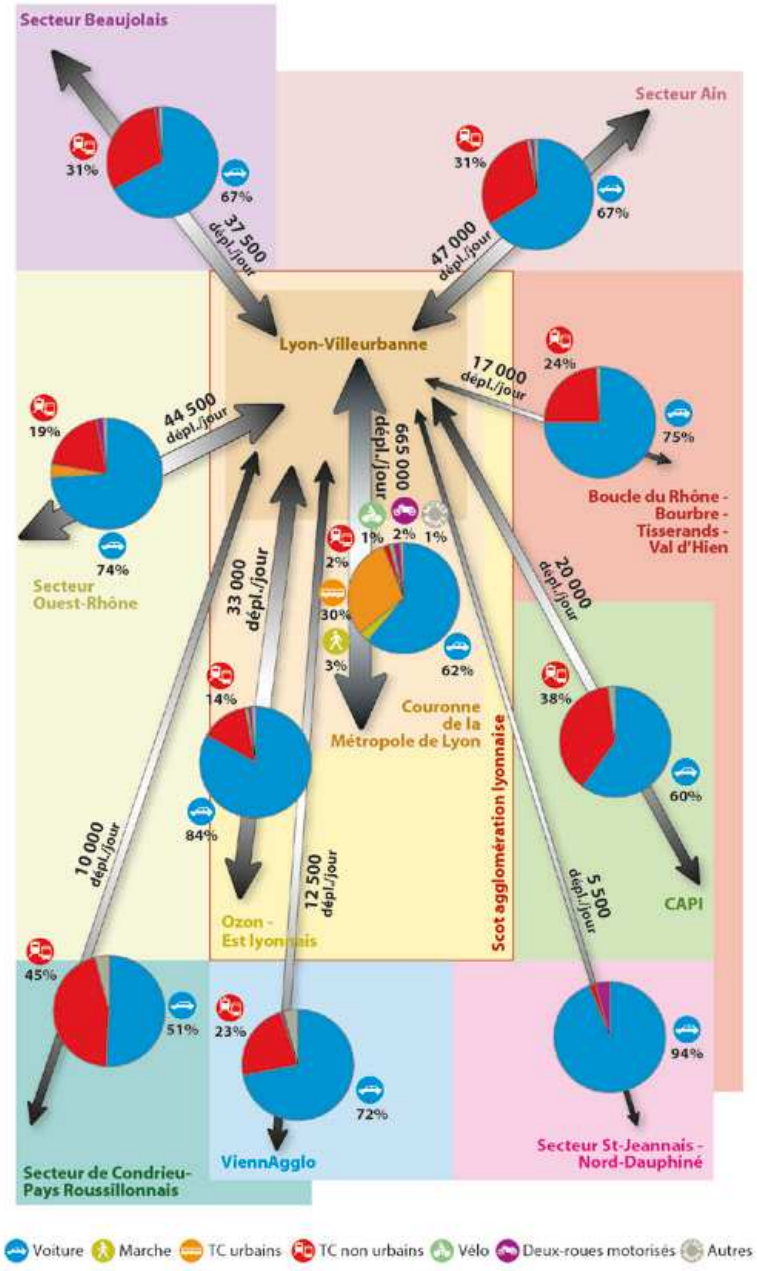
La non-réponse totale est non-ignorable

La non-réponse dépend du mode de collecte :

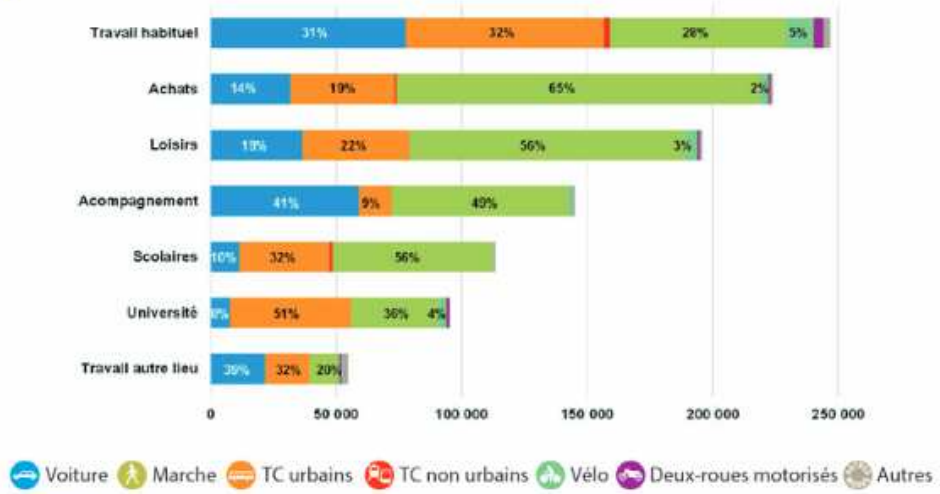
	Face à face	**Téléphone**
Taux de réponse	50-60%	20%
Déformation de l'échantillon	Moins de très grands ménages	Moins de 18-24 ans Moins de personnes non motorisées Plus de + 65 ans Moins de petits ménage

- La collecte par téléphone souffre d'un biais de constitution de la base de sondage du fait de l'enrichissement non homogène en n° de téléphone fixes
- Un effet de mode de collecte a été identifié sur le nombre moyen de déplacements/personne qui a mis en évidence une déclaration plus importante de petits déplacements, notamment à pied au téléphone
- La non-réponse est non-ignorable :
  - Ceux qui se déplacent peu ou pas sont moins motivés à répondre
  - Ceux qui se déplacent longtemps et loin restent difficilement joignables
  - Ceux qui ont des mobilités complexes (mulitmodalité) sont plus motivés à répondre

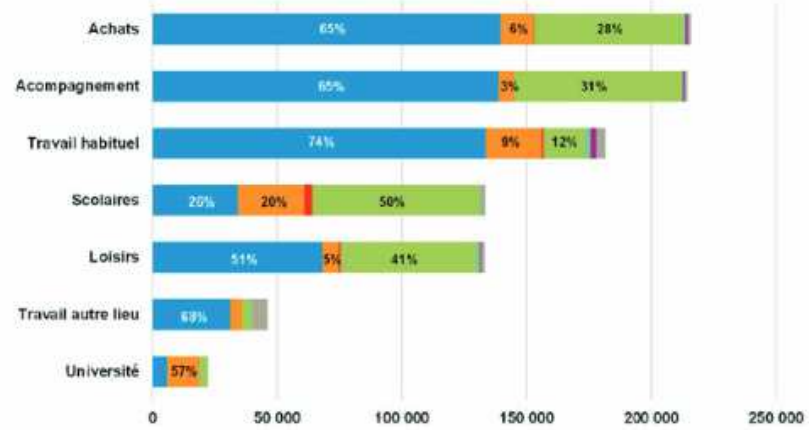
Monsols
Belleville
Montmerle-sur-Saône
Villefranche-sur-Saône
Trévoux
Anse
Tarare
Chalamont
Ambérieu-en-Bugey
Villars-les-Dombes
Meximieux
Montluel
Miribel
L'Arbresle
St-Laurent-de-Chamousset
Vaugneray
Lyon
Genas
Pont-de-Cheruy
Crémieu
Morestel
Thurins
Brignais
St-Symphorien-sur-Coise
Mornant
St-Symphorien-d'Ozon
Heyrieux
l'Isle d'Abeau
Bourgoin-Jallieu
La Tour-du-Pin
Les Abrets
St-Victor-de-Cessieu
St-Jean-de-Bournay
Vienne
Condrieu
Le Péage-de-Roussillon
Avril 2016
Urba Lyon


Secteur Beaujolais
31%
67%
37 500 dépl./jour
Secteur Ain
31%
67%
47 000 dépl./jour
Lyon-Villeurbanne
17 000 dépl./jour
24%
75%
Boucle du Rhône - Bourbre - Tisserands - Val d'Hien
44 500 dépl./jour
19%
74%
Secteur Ouest-Rhône
665 000 dépl./jour
33 000 dépl./jour
1%
2%
1%
2%
30%
3%
62%
Couronne de la Métropole de Lyon
20 000 dépl./jour
14%
84%
38%
60%
CAPI
10 000 dépl./jour
12 500 dépl./jour
Ozon - Est lyonnais
Scot agglomération lyonnaise
5 500 dépl./jour
45%
51%
23%
72%
94%
ViennAgglo
Secteur St-Jeannais - Nord-Dauphiné
Secteur de Condrieu-Pays Roussillonnais
Voiture
Marche
TC urbains
TC non urbains
Vélo
Deux-roues motorisés
Autres

Répartition des déplacements des habitants du Scot lyonnais dans et à destination de Lyon-Villeurbanne selon le mode principal et le motif



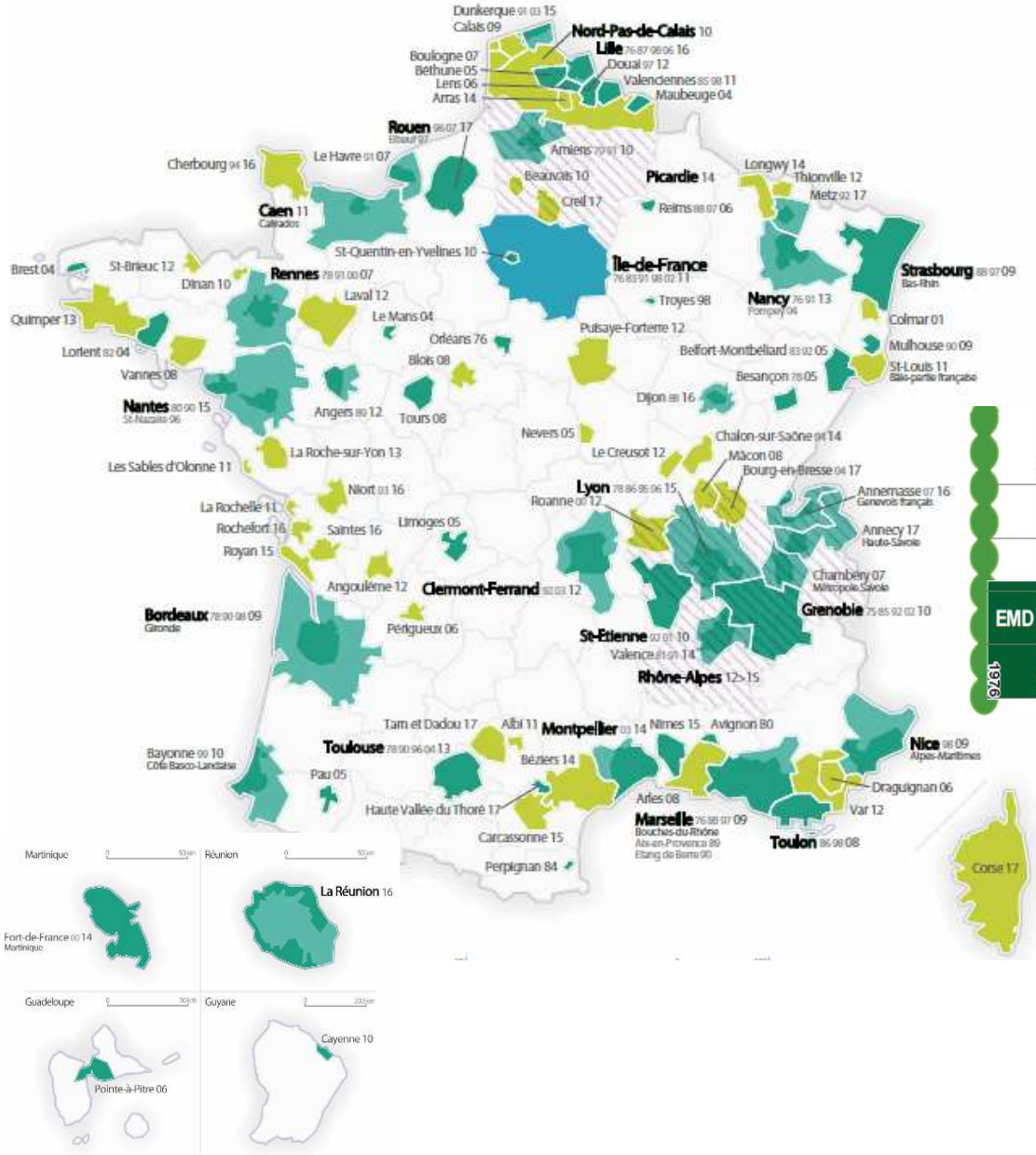
Répartition des déplacements des habitants du Scot lyonnais dans et à destination de la Couronne de la Métropole de Lyon selon le mode principal et le motif



Outils développés avec les EMD

Outre l'implémentation des modèles de prévision de la demande de transport, l'usage des EMD s'est diversifié avec le temps et d'autre outils ont été développés :

- Analyse de **bassins de déplacement** selon un motif particulier (algorithme d'agrégation successives selon critère de distance) pour mesurer l'autonomie d'un territoire
- Mesure de **co-présence** de personnes, ou de véhicules, sur un territoire un jour moyen de semaine (*Pulsations urbaines*)
- Elaboration du **Diagnostic Energie Emissions des Mobilités** grâce à la description du (*Deem*)
- Typologie de **programme d'activité** (*ADN de la mobilité*)


Dunkerque 91 03 15
Calais 09
Nord-Pas-de-Calais 10
Lille 76 87 98 06 16
Boulogne 07
Béthune 05
Douai 97 12
Lens 06
Valenciennes 85 98 11
Arras 14
Maubeuge 04
Rouen 96 07 17
Elbeuf 07
Cherbourg 94 16
Le Havre 91 07
Amiens 79 91 10
Longwy 14
Thionville 12
Beauvais 10
Picardie 14
Metz 02 17
Caen 11
Calvados
Creil 17
Reims 88 07 06
St-Quentin-en-Yvelines 10
Île-de-France
76 83 91 98 02 11
Strasbourg 88 97 09
Bas-Rhin
Brest 04
St-Brieuc 12
Dinan 10
Rennes 78 91 00 07
Laval 12
Troyes 98
Nancy 76 91 13
Pompey 04
Quimper 13
Le Mans 04
Colmar 01
Pulsaye-Forterre 12
Orléans 76
Belfort-Montbéliard 83 02 05
Lorient 82 04
Blois 08
Mulhouse 90 09
Vannes 08
Besançon 78 05
St-Louis 11
Bâle-partie française
Dijon 88 16
Nantes 80 90 15
St-Nazaire 96
Angers 89 12
Tours 08
Nevers 05
Chalon-sur-Saône 94 14
Le Creusot 12
La Roche-sur-Yon 13
Mâcon 08
Les Sables d'Olonne 11
Bourg-en-Bresse 04 17
Lyon 78 86 95 06 15
Niort 03 16
Roanne 07 12
Annemasse 07 16
Genevois français
La Rochelle 11
Rochefort 16
Saintes 16
Limoges 05
Annecy 17
Haute-Savoie
Royan 15
Chambéry 07
Métropole Savoie
Angoulême 12
Clermont-Ferrand 80 03 12
Grenoble 75 85 92 02 10
Bordeaux 78 90 08 09
Gironde
Périgueux 06
St-Etienne 03 01 10
Valence 81 01 14
Rhône-Alpes 12>15
Tarn et Dadou 17
Albi 11
Montpellier 03 14
Nîmes 15
Avignon 80
Nice 98 09
Alpes-Maritimes
Bayonne 99 10
Côte Basco-Landaise
Toulouse 78 90 96 04 13
Béziers 14
Draguignan 06
Pau 05
Arles 08
Haute Vallée du Thoré 17
Var 12
Marseille 76 88 97 09
Bouches-du-Rhône
Aix-en-Provence 89
Etang de Berre 90
Toulon 86 98 08
Carcassonne 15
Perpignan 84
Corse 17
Martinique
Réunion
La Réunion 16
Fort-de-France 00 14
Martinique
Guadeloupe
Guyane
Cayenne 10
Pointe-à-Pitre 06
Nombre d'enquêtes selon les méthodes par période
EMD
EDVM
EDGT
autres
1976
1979
1982
1985
1988
1991
1994
1997
2000
2003
2006
2009
2012
2015
0
5
10
15
20

Pourquoi rénover le dispositif d'enquête EMD ?

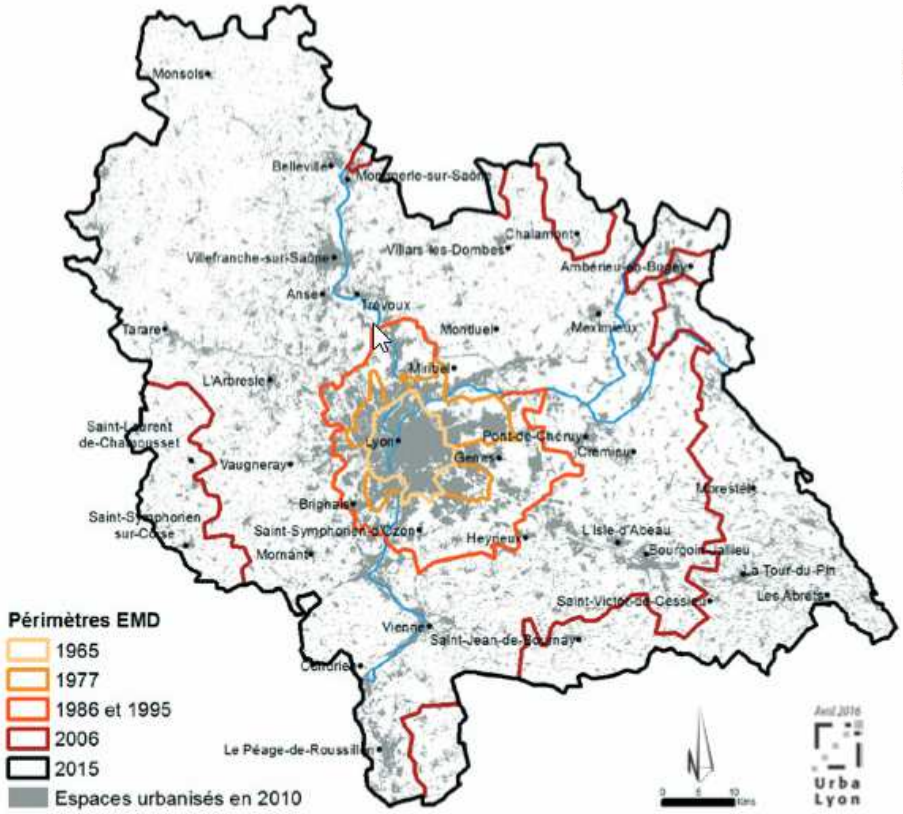
Les EMD ont déjà rendu de grands services depuis 40 ans du fait de leur forte standardisation mais **elles sont chères** car les périmètres se sont étendus.

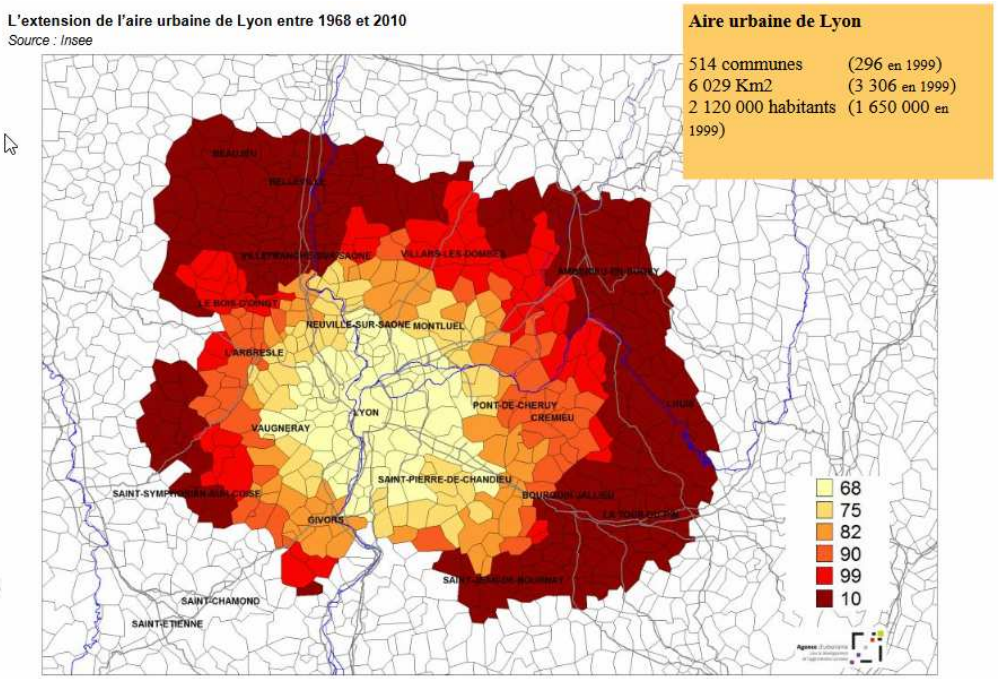
Si les besoins de connaissance des collectivités se sont **diversifiés**, ils restent **différents** d'une collectivité à l'autre => essayer de **moduler le contenu** de l'enquête sans perdre la **comparabilité dans le temps et l'espace :**

- Besoin de connaissance au-delà du « jour moyen de semaine »
- Besoin de mieux connaître certaines sous-populations
- Besoin de données plus fréquentes que tous les 10 ans

Tout en préservant les objectifs d'alimentation de modèles pour les collectivités qui en disposent.

Exemple d'extension de périmètre : Lyon





Le nouveau dispositif des EMC²

Un dispositif modulable préservant les indicateurs historiques des EMD :

- Définition d'un **« cœur » d'enquête**, centré sur les déplacements et sa définition historique
- Proposition de **modules optionnels** bâtis sur des enquêtes de seconde phase **adossées à l'enquête-coeur** sur des échantillons plus petits :
  - Enquête sur les déplacements le week-end (Tél)
  - Enquête pour mesurer l'opinion des personnes sur les conditions de mobilité (transport, stationnement etc.) et la vision qu'ils ont des modes de transport (internet)
  - Enquête pour approfondir une thématique ou mieux connaître une sous-population (internet)
  - Enquête de type « GPS » pour mieux connaître les itinéraires, les temps parcours, la variabilité des comportements dans la semaine (expérimentale pour l'instant)

  ⇨En plus d'être optionnelles, par construction, leur redressement peut s'appuyer sur les marges estimées par le cœur d'enquête qui mesure la mobilité semaine

Le nouveau dispositif des EMC²

La baisse du coût de l'enquête est réalisée en introduisant **plus de téléphone** y compris sur les zones historiquement interrogées en face-à-face en :

- améliorant la base de sondage : accès à des bases fiscales plus riches
- améliorant le redressement des enquêtes : mieux prendre en compte la base de sondage initiale dans le redressement et introduire un calage sur marge qui permet d'utiliser plus de variables simultanément selon le mode de collecte

Et internet ?

Ce passage à internet n'est pas qu'un changement de mode de collecte,

- il fait passer l'enquête d'un mode administré par un enquêteur à un mode auto-administré, d'un recueil fastidieux qui peut être sujet à des **oublis**, voire des **refus « mous »** ou **« faux immobiles »** déjà démontrés dans quelques expérimentations
- Ile de France Mobilité expérimente actuellement un test de grand ampleur avec une application très ergonomique dont on espère qu'elle pourrait surmonter les défauts ci dessus

Le nouveau dispositif des EMC²

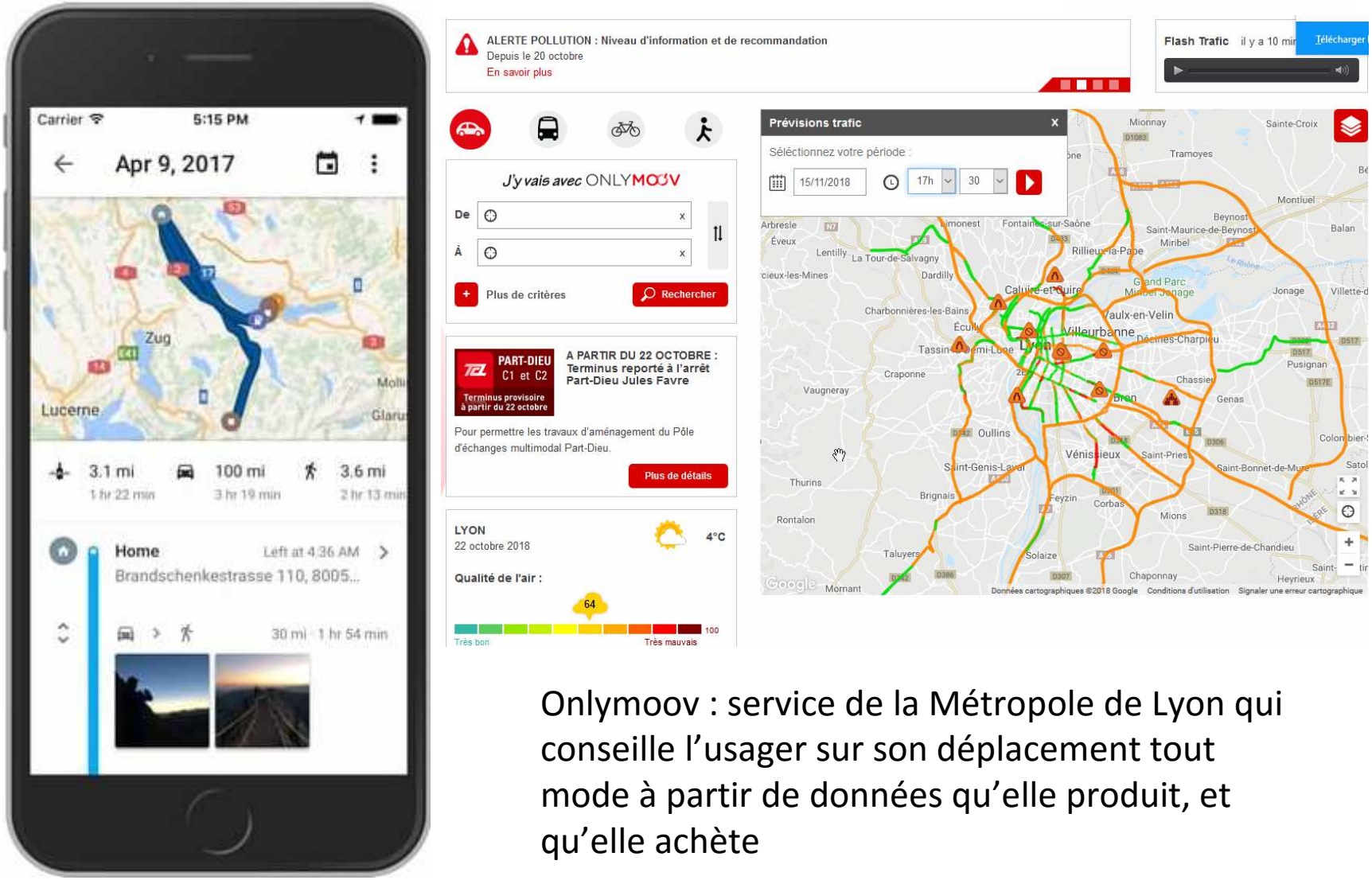
Des informations plus fréquemment qu'une fois tous les 10 ans ?

La mise en place de documents de **planification d'urbanisme à moyen terme** en vue de maitriser l'augmentation de l'usage de la voiture nécessite que les collectivités suivent des indicateurs intermédiaires d'atteinte des objectifs.

Plutôt que de réaliser une enquête tous les ans pour estimer la variation des ces indicateurs, le Cerema préconise le **suivi de sources basées sur la fréquentation** de lieu ou de mode de transport (billettique, comptages routiers, cycliste) car les comportements changent lentement. Mais ces indicateurs ne donnent pas une vision d'ensemble du système de transport et des ses interactions.

- L'idéal : suivre un **panel renouvelé** régulièrement qui permettrait de mesurer les changements de comportements et d'estimer les indicateurs en coupe -> nécessite une infrastructure statistique de suivi importante peu présente dans les collectivités (Ile de France Mobilité essaie de se lancer)
- Pour l'instant le Cerema propose de **réaliser une photo à mi-chemin** afin d'estimer quelques grands indicateurs de l'enquête-cœur initiale

Les nouvelles sources à côté des EMC²



Onlymoov : service de la Métropole de Lyon qui conseille l'usager sur son déplacement tout mode à partir de données qu'elle produit, et qu'elle achète

Les données massives et les EMC²

Les données massives peuvent elles répondre à tous ces objectifs à la fois ?

Ces données sont « **très populaires** » dans le domaine de la mobilité puisque chacun peut expérimenter tous les jours différents outils qui « tracent » ses déplacements, enregistrent les lieux visités sur un smartphone pour aider par exemple à choisir mieux son itinéraire, son mode de transport, son heure de départ, trouver un vélo en libre service, une place de stationnement etc. , le tout en « temps réel ».

Les opérateurs de ces services constituent des **bases de données gigantesques** avec de données de différents types :

- agrégées ou désagrégées,
- publiques (ex : trafic, géographie, équipement) ou privées. Dans les sources privées on peut distinguer celles d'autres entreprises de celles récupérées auprès des utilisateurs. En effet, pour bénéficier du service de ces applications, les utilisateurs, ou leurs véhicules, fournissent des informations
- récoltées de manière plus ou moins « passive »
- qui viennent alimenter les bases de connaissance multiples des opérateurs.

Les données massives et les EMC²

Les avantages / inconvénients de ces sources par rapport aux EMC²

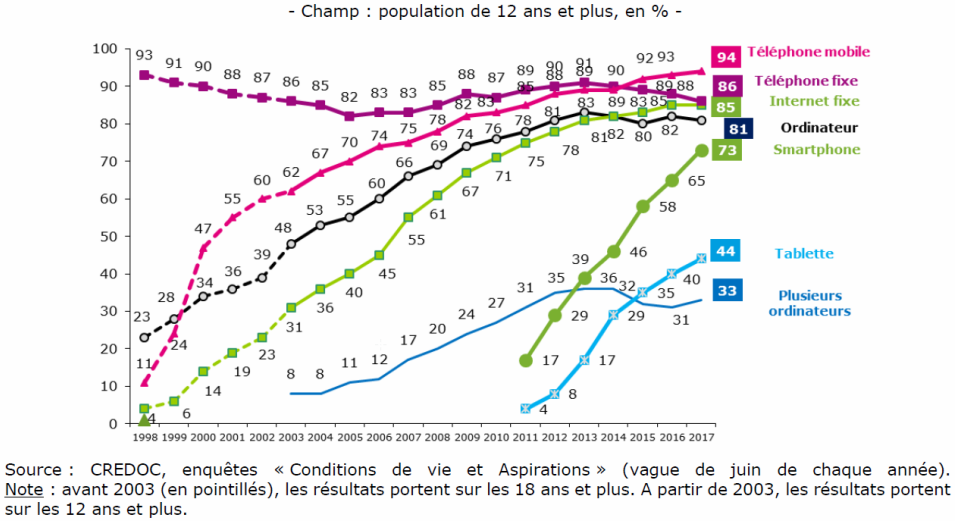
Cette analyse peut être faite de manière générale. Très souvent, comme pour les sources administratives, la donnée est un sous produit d'une activité principale sauf pour les quelques opérateurs dont c'est la finalité ultime

	Données massives	EMC²
Exhaustivité	Oui sur leur champ, qui peut varier dans le temps	Echantillon / géographie
Rapidité d'obtention	Oui, souvent mesure en continu => Problème de stockage	Délai d'un an en moyenne pour la production Réalisée tous les 10 ans
Coût/Droits de propriété	La donnée individuelle privée n'est pas vendue, on achète de la donnée reconstruite; EMC² hors SSP	Données publiques depuis la Loi Lemaire
Transparence et stabilité des traitements dans le temps	- Le processus de collecte peut varier dans le temps (techno, services) - Les données sont souvent « reconstruites » : a minima redressées, pour exploiter les multiples corrélations => algorithme d'apprentissage « boite noire »	Données dont le processus de collecte et le redressement sont explicités et connus
Biais de collecte	Inconnu surtout si l'action de l'utilisateur est importante pour l'acquisition de données	Biais mieux connu (pas forcément corrigé)

Les données massives et les EMC²

Les données des opérateurs de téléphonie mobile : les principes

- Utilise l'activité des téléphones portables (sms, appels, données, changement de zone) qui génèrent des événements localisés (précisions différentes selon la zone urbaine ou rurale)≠ traces
- Peut mesurer une présence en un lieu et des flux de déplacements entre zone
- Echantillons très grands dont on peut penser qu'ils sont représentatifs au vu du taux de détention d'un téléphone portable dans la population
- Recueil en continu permettant un suivi temporel fin et sur de longues périodes



Source : CREDOC, enquêtes « Conditions de vie et Aspirations » (vague de juin de chaque année). Note : avant 2003 (en pointillés), les résultats portent sur les 18 ans et plus. A partir de 2003, les résultats portent sur les 12 ans et plus.

Les données massives et les EMC²

Les données des opérateurs de téléphonie mobile : les difficultés

➢Il peut être difficile de reconstituer le zonage d'étude / au zonage antenne

➢Echantillonnage spatio-temporel irrégulier embêtant pour identifier exhaustivement les déplacements effectués et où

➢Lors du traitement des données, on doit traiter l'effet ping-pong (effet frontière de LA) et à partir de quand on décide qu'il y a déplacement (temps d'arrêt)

➢On récupère toujours de la donnée agrégée (confidentialité) reconstituée (redressement et identification de sous populations)

➢Redressement en population générale délicat :

- Plusieurs opérateurs sur le marché dont on ne connaît pas exactement la part de marché qui peut être différente selon le territoire
- On ne connaît pas la part de téléphone éteint, le taux de double équipement
- On estime une population présente dans laquelle il faut identifier les résidents

➢ Identification des modes difficile, parfois impossible en urbain

➢Détermination des motifs très limitée (domicile / travail)

➢Les technologies de détection et les algorithmes de traitement évoluent rapidement dans le temps, quid de la comparabilité à moyen terme ?

Les données massives et les EMC²

Evidemment, la téléphonie mobile ne représente pas tout le champ de ce qu'on entend par « données massives » car elle apparie peu de sources concernant les individus.

Les opérateurs fournissant de multiples services aux individus qui révèlent leur préférences, leurs contacts, leurs activités et les lieux qui s'y rattachent ont d'autant plus de chances d'estimer le programme d'activité d'un individu et par la même ses déplacements. Mais les algorithmes de traitements seraient ils efficaces pour estimer les comportements de mobilité sur une zone réduite du territoire sans biais ?

Par ailleurs ces traitements restent coûteux, il n'est pas sûr qu'ils soient adaptés à un territoire et vendus pour un prix beaucoup moins cher qu'une enquête à une collectivité. Les expériences montrent que lorsqu'il faut acheter plusieurs lots de données produits par des opérateurs privés pour alimenter des algorithmes performant, la facture monte vite…

Les données numériques et les EMC²

Les données obtenues par GPS : les principes

- Utilise un boitier GPS ou le GPS d'un smartphone pour enregistrer les traces d'un individu => nécessite une participation active minimale (il faut activer le GPS ou penser à prendre son boitier)
- Permet de recueillir les itinéraires/distances/durées de manière plus ou moins précise selon les capacités de l'appareil, d'interroger l'enquêté en cours de déplacement
- Sur une période de plusieurs jours
- Meilleur si on couple avec d'autres informations (accélération, wifi)
- Peut être couplé à une enquête « classique » ou pas (application à télécharger librement à partir d'un site « communiquant »)

Beaucoup d'expériences depuis 2010 mais peu de passage « en production »

- Souvent sous forme de sous-échantillons d'enquête principale (Entd 2007-08)
- Avec ou sans confrontation des données d'entretien
- Sur une période de un jour à quelques semaines

Les données numériques et les EMC²

Les données obtenues par GPS : les difficultés

➢Le recrutement reste une opération difficile :

- En France, malgré un bon taux de pénétration du smartphone, le taux d'acceptation explicite est de un tiers environ (confidentialité, problème de batterie etc.)
- Ceux qui acceptent ne représentent pas la population (plus diplômé, plus aisé, plus mobile, plus urbain, aime « jouer » avec les applications)
- Les applications doivent s'adapter à tous les systèmes d'exploitation/navigateur quels que soient la version et ne pas consommer trop ce qui nécessite un développement de bon niveau
- Il faut anonymiser presque en temps réel les traces des enquêtés

➢Le traitement des traces : une opération d'ingénierie de traitement du signal

- Adapter des algorithmes de filtrages/lissage des traces pour réduire « les bruits »…

⇨ Pour pouvoir mieux identifier les arrêts = trajets (surestimation de 10% environ)

⇨ Pour identifier correctement les itinéraires (85% de « bons chemins »)

⇨ Pour identifier correctement les modes (65% de macro modes correctement identifiés) mais le taux diffère selon les zones urbaines/rurales

- La détermination des motifs reste imprécise en dehors du domicile/travail
- Les budgets distance et temps sont mieux estimés
- Le traitement est nettement meilleur si on peut lister les lieux fréquemment visités avec l'enquêté, lui poser quelques questions pendant le déplacement

L'avenir des EMC² ?

A ce jour, les données issues d'appareils mobiles produites plus ou moins passivement enrichissent la connaissance qu'on a de la mobilité des individus / recueil « classique » EMC²:

- des itinéraires, distances, vitesses, durée avec le GPS, surtout si on récolte auprès de l'enquêté des informations => couplage EMC² intéressant
- Une vision de la variabilité dans le temps et l'espace des flux de déplacements, sur des grandes zones, avec la téléphonie mobile, et prise en compte de la population non-résidente qui fréquente également les infrastructures de transport de la zone => sensible quand la zone est touristique et/ou professionnellement attractive.

Ces données sont donc plutôt **complémentaires aux EMC²** qui continuent d'estimer explicitement les liens ménages-individus-(im)mobilité-véhicule, en maîtrisant tant que possible la représentation de toute la population, y compris des personnes qui bougent peu.

Améliorer encore la connaissance

Outils

Les enquêtes « cordon », un autre élément historique de la matrice technique

Si les données données « mobiles » restent complémentaires des EMC², elles sont en passe de remplacer les enquêtes « cordon », de plus en plus difficile à réaliser sur la route. Outre le fait qu'on ne mesure plus un flux de véhicules, les matrices en volume restent difficiles à reconstituer, surtout en milieu dense

Mieux prendre en compte les changements de comportement

Les comportements de mobilité peuvent changer conjoncturellement (prix du carburant) ou plus durablement. La France ne s'est doté d'aucun outil pour évaluer ces effets. La mise en place d'un panel maintenu au niveau national au lieu d'une ENT tous les 10 ans serait fort utile.

Thèmes

Mieux connaître l'immobilité subie ou choisie

Le vieillissement de la population et le renchérissement du coût des carburants laissent des personnes sans solution de mobilité. Par ailleurs, les nouveaux modes d'organisation pourraient « éviter » de se déplacer.