

## Étude sur l'utilisation des véhicules au Canada : collecte de données électronique

Émile Allie<sup>1</sup>

### Résumé

L'Étude sur l'utilisation des véhicules au Canada est menée par Transports Canada, en partenariat avec Environnement Canada, Ressources naturelles Canada et les registraires provinciaux. L'étude se divise en deux composantes : les véhicules légers, comme les voitures, minifourgonnettes, VUS et camions dont le poids nominal brut du véhicule (PNBV) est inférieur à 4,5 tonnes métriques; et la composante des véhicules moyens et lourds, c'est-à-dire les camions dont le PNBV est de 4,5 tonnes métriques et plus. L'étude est la première qui recueille des données sur l'activité directement dans les véhicules, au moyen de méthodes de collecte électronique exclusivement. Cela permet d'obtenir plus de renseignements opportuns et fiables.

Mots clés : collecte électronique, véhicule, mégadonnées, Canada.

### 1. Introduction

#### 1.1 Objectifs

L'Étude sur l'utilisation des véhicules au Canada (EUV) se divise en deux composantes : les véhicules légers (EUV-L) et les véhicules moyens et lourds (EUV-ML). La composante des véhicules légers vise à mesurer diverses quantités liées aux véhicules légers, comme les kilomètres parcourus par véhicule, les kilomètres parcourus par passager, la consommation de carburant, la vitesse, le ratio de consommation de carburant, etc. La composante des véhicules moyens et lourds porte davantage sur le tonnage des cargaisons, le tonnage-kilomètre, la configuration des véhicules, le type de cargaison, etc. Les estimations sont fournies aux niveaux national et provincial pour les quatre derniers trimestres, afin d'appuyer l'analyse et l'élaboration de politiques.

#### 1.2 Enquête sur les véhicules au Canada

L'Enquête sur les véhicules au Canada (EVC) a été menée par Statistique Canada, en vertu d'un contrat avec Transports Canada et Ressources naturelles Canada, entre 1999 et 2009. L'EVC a été affectée par de faibles taux de réponse des participants tout au long de son existence, attribuables pour une large part à la méthode fastidieuse de collecte au moyen d'un journal papier. La qualité des estimations a aussi été amoindrie par des erreurs importantes dans la façon dont le parc de véhicules routiers était classé.

#### 1.3 Étude sur l'utilisation des véhicules au Canada

Transports Canada a décidé de procéder à une EVC révisée, au moyen d'une méthode de collecte des données améliorée, en collaboration avec Ressources naturelles Canada, Environnement Canada et les registraires provinciaux. À l'heure actuelle, l'étude englobe l'Île-du-Prince-Édouard, le Québec, l'Ontario, le Manitoba et la Saskatchewan. Il s'agit de plus des deux tiers du parc automobile au Canada. Par suite des changements récents apportés à la *Loi sur la sécurité automobile*, il est permis de s'attendre à couvrir toutes les provinces et tous les territoires à partir du deuxième trimestre de 2015.

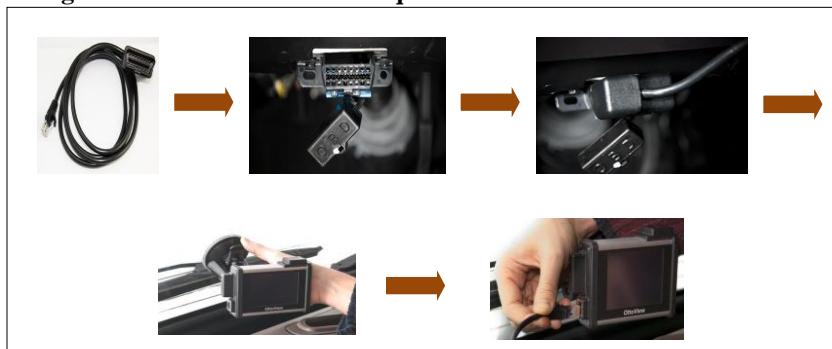
---

<sup>1</sup>Emile Allie, Transports Canada, 339, rue Sparks, Ottawa, Canada, K1A 0N5, emile.allie@tc.gc.ca.

L'enquête est menée chaque trimestre et comprend deux composantes : les véhicules légers, qui comprend les voitures, minifourgonnettes, VUS et camions dont le poids nominal brut du véhicule (PNBV) est inférieur à 4,5 tonnes métriques, et les véhicules moyens et lourds (PNBV) de 4,5 tonnes métriques et plus. La section 2 abordera la base de sondage et la section 3, le plan d'échantillonnage.

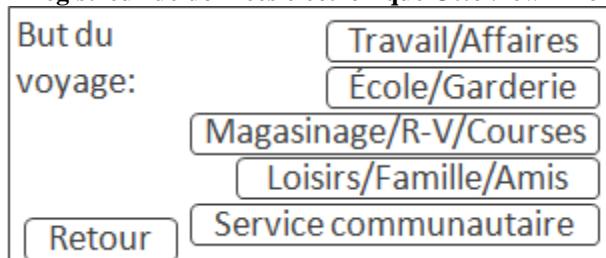
L'étude utilise un enregistreur de données électronique pour réduire le fardeau de réponse, augmenter l'exactitude et la quantité de données recueillies, et accélérer le traitement des données. La figure 1.3.1 illustre les enregistreurs de données électroniques OttoView utilisés pour l'étude et la facilité d'installation au moyen du port de données ODB-II/J1979. À notre connaissance, l'EUVC est la première étude à recueillir des données sur l'activité directement dans les véhicules au moyen de méthodes de collecte électronique exclusivement.

**Figure 1.3.1**  
**Enregistreur de données électronique OttoView**



Le fardeau de réponse est réduit grâce à l'enregistreur de données électronique OttoView, et par l'entremise d'une petite séquence de questions concernant le numéro d'identification du conducteur, le nombre de passagers selon le groupe d'âge, le but du voyage et l'achat de carburant. Les réponses à toutes ces questions sont fournies au moyen de l'écran de l'enregistreur de données électronique OttoView (figure 1.3.2). Toutes les données qui restent concernant le voyage sont recueillies à partir du port de données ODB-II/J1979, sans l'intervention du conducteur.

**Figure 1.3.2**  
**Enregistreur de données électronique OttoView – Écran du but du voyage, EUVC-L**



L'enregistreur de données électronique recueille des données plus précises parce que le conducteur n'a pas à dépendre des données du journal ou des données qu'il recueille ou dont il se souvient au début et à la fin d'un voyage, comme la distance parcourue, le temps de conduite et le type de route. Tous ces renseignements sont recueillis automatiquement.

L'enregistreur de données électronique fournit plus de renseignements parce que ceux-ci sont recueillis **chaque seconde**, dès que la voiture est mise en marche. En outre, il est relié à l'ordinateur de la voiture, ce qui fait qu'il recueille des données impossibles à recueillir autrement, comme la vitesse instantanée et la consommation de carburant, les tours/minutes, la température de l'air d'admission, la température de refroidissement, le moteur qui tourne au ralenti, le moment du jour (à la seconde près), etc. Cela pourrait être utile pour analyser les comportements de conduite, les économies de carburant et les émissions. L'enregistreur de données électronique utilise aussi la technologie GPS et consigne des coordonnées spatiales qui pourraient être utilisées pour l'analyse de la congestion

routière, la sécurité routière et la planification des infrastructures. Étant donné que l'étude recueille des données sur les voyages toutes les secondes, plus de 90 millions d'enregistrements sont produits chaque trimestre. La section 6 porte sur l'ensemble des données disponibles.

Le temps de traitement des données est significativement réduit, parce que toutes les données sont recueillies électroniquement, ce qui permet d'éviter l'étape du transfert des données d'un journal sur papier à une base de données électronique. Habituellement, des rapports statistiques sont disponibles 60 jours après la fin d'un trimestre. La section 4 aborde le traitement des données, et la section 5, la procédure de pondération.

## **2. Base de sondage de l'Étude sur l'utilisation des véhicules au Canada**

### **2.1 Base de sondage**

La base de sondage comprendra les fichiers d'enregistrement de véhicules automobiles fournis par chaque secteur de compétence avant le début d'un trimestre, afin de réduire les changements d'adresse et de maximiser l'exactitude des données sur le parc automobile. Les données sont limitées à ce qui est requis pour sélectionner les véhicules dans la base de sondage. Les véhicules dans la base de sondage sont par la suite divisés en deux groupes : les véhicules légers, et les véhicules moyens et lourds. Cette division est fondée sur la capacité de chargement du véhicule. Toutes les voitures et tous les camions dont le PNVB est inférieur à 4,5 tonnes métriques constituent la base de sondage de l'EUVCL. Les camions moyens et légers (PNVB de 4,5 tonnes métriques et plus) constituent la base de sondage de l'EUVCL.

### **2.2 Population hors du champ de l'enquête**

La base de sondage de l'EUVCL exclut certains véhicules (hors du champ de l'enquête), comme les motocyclettes, les remorques, les grues, les autobus, les ambulances, les camions d'incendie, l'équipement agricole, les caravanes motorisées, les véhicules de police, etc. Afin d'éliminer les véhicules hors du champ de l'enquête et de répartir les véhicules entre l'EUVCL ou l'EUVCL-ML, nous utilisons le numéro d'identification du véhicule (NIV), un décodeur de NIV Polk et d'autres renseignements dans le fichier d'enregistrements (type de véhicule, année du modèle et évaluation du poids brut du véhicule). Le NIV est un code alphanumérique à 17 caractères pour tous les véhicules dont l'année du modèle est 1981 et après (on utilise une norme différente pour les véhicules avant 1981). Les 11 premiers caractères définissent la marque, le modèle et d'autres caractéristiques du véhicule, tandis que les six derniers chiffres identifient de façon unique chaque véhicule.

## **3. Échantillonnage**

### **3.1 Plan d'échantillonnage**

Le plan d'échantillonnage est illustré à la figure 3.1.1. Dans chaque secteur de compétence, selon l'année du modèle des véhicules, ces derniers sont classés comme ayant trois ans ou moins, quatre à huit ans et neuf ans ou plus.

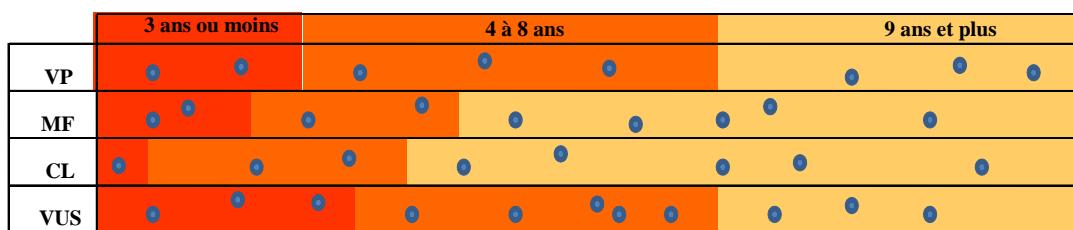
Dans l'EUVCL, il y a une distinction entre les voitures, les minifourgonnettes, les camions légers (moins de 4,5 tonnes métriques) et les VUS. Dans l'EUVCL-ML, les véhicules sont répartis entre les camions moyens légers, dont le PNVB se situe entre 4,5 et 10 tonnes métriques, les camions moyens lourds, dont le PNVB se situe entre 10 et 15 tonnes métriques, et les camions lourds, dont le PNVB est de 15 tonnes métriques et plus. La discussion qui suit sera axée uniquement sur l'EUVCL.

### **3.2 Sélection de l'échantillon de l'EUVCL**

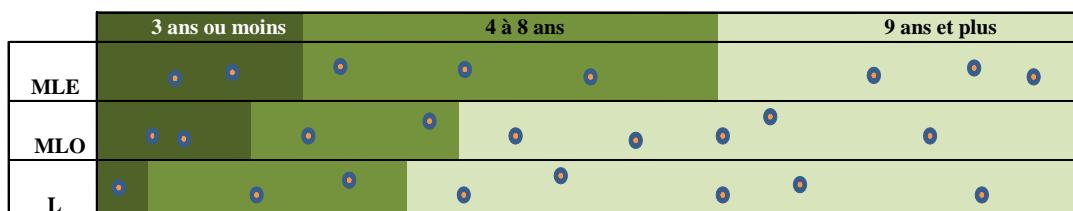
Chaque trimestre, 6 000 véhicules sont sélectionnés. La répartition entre les quatre provinces participantes est la suivante : 2 000 chacun pour le Québec et l'Ontario et 1 000 chacun pour le Manitoba et la Saskatchewan. Récemment, l'Île-du-Prince-Édouard a accepté de participer à l'étude, et une répartition préliminaire de 350 a été faite dans cette province, jusqu'à ce qu'il soit possible d'évaluer le taux de réponse. Par suite des changements récents apportés à la *Loi sur la sécurité automobile*, il est permis de s'attendre à inclure toutes les provinces et tous les territoires à partir du deuxième trimestre de 2015.

L'échantillon de chaque province est stratifié selon le type de véhicule et l'âge (groupe), et l'affectation à une strate est proportionnelle à la racine carrée de la population de la strate. Au niveau du Canada, l'échantillon définitif estimé après non-réponse (environ 1 000 par trimestre) justifie un niveau de confiance global de 95 % et un intervalle de confiance de 3 %. Cela ne sera pas le cas pour chaque groupe d'une province, mais la qualité des données des groupes comportant une très petite population atteint un niveau satisfaisant sur une base annuelle pour la plupart des groupes. L'affectation à la strate est proportionnelle à la racine carrée de la taille de la population de la strate.

**Figure 3.1.1**  
**Plan d'échantillonnage de l'Étude sur l'utilisation des véhicules au Canada**  
**Composante des véhicules légers**



**Composante des véhicules moyens et lourds**



VP : Voiture personnelle

MF : Minifourgonnette

CL : Camion léger

VUS : Véhicule utilitaire sport

MLE : PNBV de 4,5 à 10 tonnes

MLO : PNBV de 10 à 15 tonnes

L : PNBV de 15 tonnes et plus

## 4. Traitement des données de l'Étude sur l'utilisation des véhicules au Canada

### 4.1 Étape 1 : Sélection de l'échantillon

Le processus de collecte des données commence par une entente avec une province et le service d'enregistrement des véhicules. Un mois avant la fin d'un trimestre, chaque province participante envoie à Transports Canada un fichier comprenant tous les véhicules actifs. Le fichier comprend uniquement les données limitées requises pour

stratifier l'échantillon. Lorsque l'échantillon est sélectionné, la liste de véhicules choisis est envoyée à la province pour obtenir des données de contact sur le propriétaire du véhicule.

## 4.2 Étape 2 : Collecte des données

Un tiers est recruté pour gérer toutes les communications avec les propriétaires des véhicules sélectionnés. Chaque échantillon est réparti de façon aléatoire en 13 lots pour couvrir les activités des véhicules au cours du trimestre. Le propriétaire d'un véhicule sélectionné reçoit une lettre officielle de Transports Canada qui l'invite à participer à l'enquête. Il est invité à répondre par l'entremise du site Web de Transports Canada, par la poste ou au moyen d'un numéro sans frais et à fournir certains renseignements de base concernant le véhicule et les conducteurs (voir l'annexe C). S'il est impossible de communiquer avec le propriétaire, une procédure d'appariement de numéros de téléphone est appliquée.

Lorsque les propriétaires acceptent de participer, l'expert-conseil envoie l'enregistreur et les câbles, une trousse d'information et la date de début et de fin de la collecte des données. Il faut environ 60 jours (y compris les 21 jours de l'enquête) pour un cycle complet, du premier contact avec un propriétaire de véhicule jusqu'à l'enregistreur.

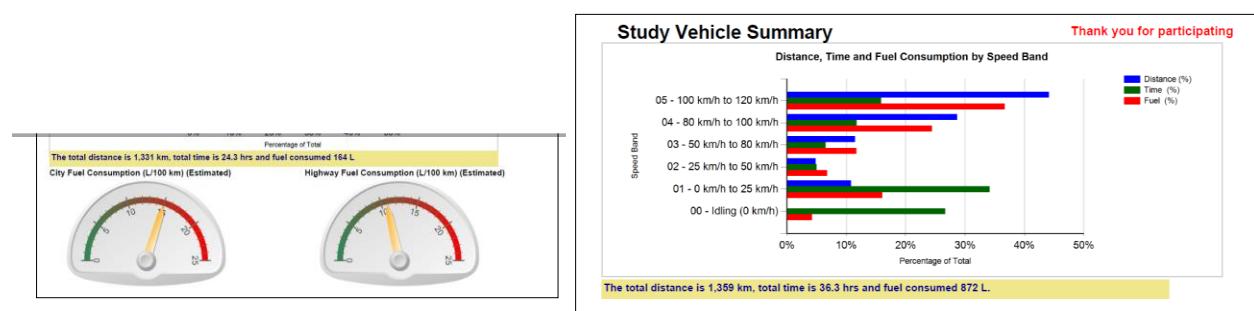
Des mesures incitatives sont utilisées auprès des participants afin qu'ils renvoient l'enregistreur OttoView. L'un d'eux est un rapport de leur utilisation du véhicule (figure 4) et de celle des autres conducteurs. Des mesures incitatives pécuniaires sont aussi utilisées, comme un tirage pour les premiers participants (500 \$), c'est-à-dire pour ceux qui renvoient l'enregistreur au plus tard sept jours après la fin de la période d'enquête; et un tirage mensuel (1 000 \$) pour ceux qui répondent à l'enquête et la renvoient le même mois.

Figure 4 : Extrait du rapport de l'utilisateur

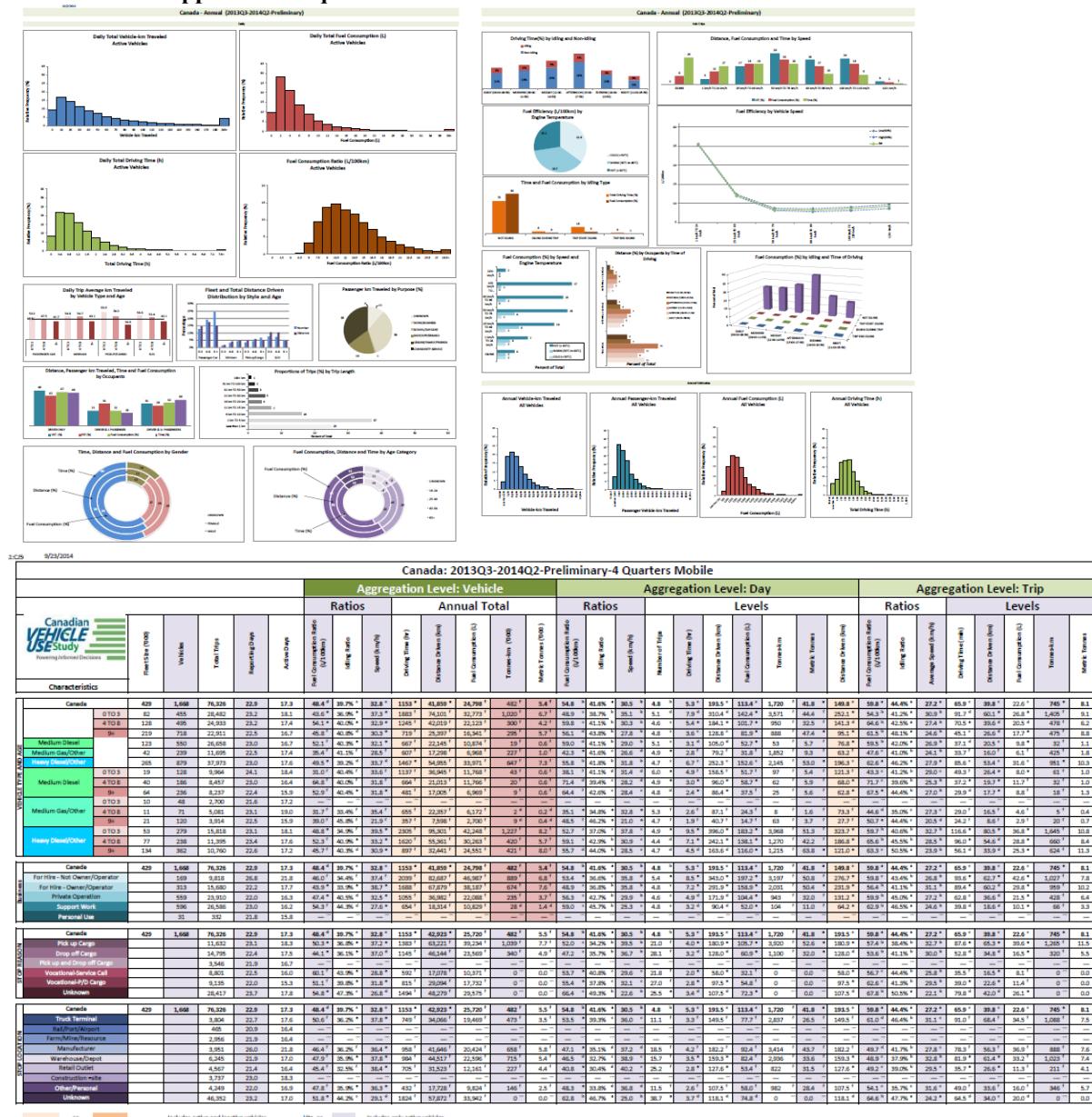
## 4.3 Étape 3 : Production de statistiques

L'ensemble de l'épuration et du traitement des données se fait à Transports Canada. Le fichier mis à la disposition de l'analyste ne comprend pas de renseignements personnels concernant le propriétaire du véhicule. Les données privées sont chiffrées et accessibles uniquement à un nombre limité de personnes. La figure 4.3.1 illustre certains des renseignements disponibles dans le rapport du participant. La figure 4.3.2 est un exemple des renseignements partagés avec nos partenaires. Les données recueillies servent à produire des tableaux sommaires mis en commun avec nos partenaires. Une version réduite est disponible sur le site Web de Transports Canada. Les données pour un voyage en particulier sont fondées sur un échantillon aléatoire de 10 % du fichier des données propres à chaque voyage (environ 90 millions d'enregistrements).

**Figure 4.3.1**  
Échantillon de rapport du participant



**Figure 4.3.2**  
**Échantillon de rapport statistique de l'EUVCL-L**



## 5. Procédure de pondération

Étant donné que l'échantillon est stratifié, toutes les observations sélectionnées comportent un poids égal à l'inverse de la probabilité de sélection dans la strate avec rajustement pour tenir compte de la non-réponse. Des rajustements supplémentaires des poids sont effectués avant de les utiliser pour la production de statistiques :

- Une valeur limite de 21 a été attribuée aux poids, ce qui signifie qu'un véhicule ne pourrait pas représenter plus de 21 000 véhicules pendant un trimestre.
- Pour les observations comportant une période inférieure à 21 jours, les poids sont réduits en proportion de la période de réponse pour une période uniformisée de 21 jours, par exemple, si un répondant recueille l'information uniquement pendant cinq jours, avant d'abandonner l'étude, le poids associé, w, est réduit à  $w^*5/21$ , parce qu'il est impossible de savoir quel est le comportement de conduite après les cinq jours.

- De même, lorsque la période d'observation est supérieure à 21 jours, il ne faut pas que les observations associées au véhicule pèsent trop sur les résultats. Si la période d'observation est de 30 jours, le poids associé est réduit à  $w*21/30$ , compte tenu de la période d'observation uniformisée de 21 jours.

## 6. Somme considérable d'information (mégadonnées)

Chaque trimestre, l'EUVc produit une somme importante d'information. Cinq ensembles de données sont disponibles pour l'analyse. Par exemple, le fichier des données propres à chaque voyage décrit ci-dessus comprend environ 90 millions d'enregistrements. En couplant ces fichiers, plus de 200 variables peuvent être disponibles pour l'analyse.

- Fichiers des **registraires** des provinces, qui fournissent le NIV.
- Fichier du **décodeur**, qui associe le NIV et un ensemble de caractéristiques du véhicule, comme la marque, le modèle, la date de production, la cylindrée du moteur, le nombre de cylindres, le type de transmission, etc.
- Données de **contact**, qui déterminent les caractéristiques des conducteurs, comme la catégorie d'âge et le sexe. Ces données sont recueillies lorsque le participant convient de participer à l'étude. Tous les autres renseignements sont recueillis au niveau du véhicule.
- Données sur le **voyage**, qui proviennent des réponses aux questions posées au début d'un voyage, par exemple, le but et la catégorie d'âge des passagers, le numéro d'identification du conducteur, le numéro d'identification du voyage et le numéro d'identification du véhicule.
- Les données **propres à chaque voyage**, qui fournissent un ensemble de plus de 65 variables enregistrées chaque seconde, comme la distance parcourue, la consommation de carburant, la position géographique, les tours/minutes, la température de l'air d'admission, la température de refroidissement, etc.

## 7. Conclusion

L'EUVc recueille plus de renseignements, de façon plus opportune et précise, grâce à l'utilisation d'un enregistreur de données électronique. La somme de renseignements est 1 000 fois plus élevée qu'au moyen d'un journal papier. Transports Canada s'attend à publier certains documents de recherche dans un proche avenir.