

ÉTUDE TECHNIQUE

# Modélisation de la relation entre le comportement au volant et la consommation de carburant

De nombreuses entreprises investissent dans des services d'accompagnement destinés aux conducteurs professionnels qui visent à inciter ces derniers à adopter des habitudes d'écoconduite afin de réduire leur consommation de carburant. Afin que ces services soient efficaces, les entreprises doivent avoir une perception claire de l'effet qu'a le comportement au volant sur la consommation de carburant.

À l'heure actuelle, les véhicules exploités par Scania Group, chef de file mondial dans le secteur du transport et client de CGI, sont munis de dispositifs de surveillance qui produisent des données sur le conducteur et sur le véhicule. Ces dispositifs nous permettent de mettre en relation les données sur la consommation de carburant et les données recueillies à l'aide de systèmes d'information électronique et d'autres sources (p. ex. les conditions météorologiques).

Afin de modéliser la relation entre le comportement au volant et la consommation de carburant, nous avons effectué une analyse prédictive des données de Scania. Celles-ci ont été recueillies lors de plus de trois millions de déplacements effectués dans sept pays de l'Union européenne. Dans la présente étude, nous décrivons les méthodes et les modèles utilisés pour l'évaluation de l'efficacité des services d'accompagnement axés sur l'écoconduite dans différents parcs de véhicules et pays.

Nous présentons également certaines relations statistiques inattendues qui ont été révélées par l'analyse. En outre, nous proposons une estimation de l'efficacité des services d'accompagnement des conducteurs, qui fournit un aperçu réaliste de la réduction de la consommation de carburant pouvant être obtenue grâce à ces services d'accompagnement axés sur l'écodconduite.

## TABLE DES MATIÈRES

|  |           |
|--|-----------|
| <b>PROGRAMME D'ÉCOCONDUITE DE SCANIA .....</b>   | <b>3</b>  |
| <b>METTRE À PROFIT L'ABONDANCE DE DONNÉES DE SCANIA<br/>SUR SES CONDUCTEURS .....</b>                | <b>4</b>  |
| <b>GROUPES.....</b>  | <b>5</b>  |
| <b>MODÈLE DE CONSOMMATION DE CARBURANT.....</b>  | <b>6</b>  |
| <b>UTILISATION DES MODÈLES DANS LE CADRE DES SERVICES-<br/>CONSEILS EN GESTION DU CARBURANT.....</b> | <b>8</b>  |
| <b>DÉFIS DURANT LE PROCESSUS .....</b>   | <b>9</b>  |
| <b>FILTRAGE.....</b>   | <b>9</b>  |
| <b>PARADOXE DE SIMPSON.....</b>  | <b>10</b> |
| <b>COMMENTAIRES DES UTILISATEURS .....</b>   | <b>10</b> |
| <b>CONCLUSION.....</b>   | <b>10</b> |
| <b>RÉFÉRENCES .....</b>  | <b>11</b> |

## Programme d'écoconduite de Scania

Scania a mis en place un programme d'accompagnement axé sur les conducteurs. Ce programme utilise un portail et des jeux pour encadrer les conducteurs et les motiver à réduire leur consommation de carburant à long terme, et non seulement après la formation initiale.

Le portail d'accompagnement destiné aux conducteurs est intégré aux systèmes de gestion de parcs de véhicules exploités par les propriétaires de parcs et est utilisé pour former les conducteurs à la pratique de l'écoconduite.

Les déplacements sont communiqués et comparés à des valeurs de référence dans le portail. Une fonction de surveillance permet aux accompagnateurs de voir la progression de leurs conducteurs. Les services d'accompagnement sont offerts par téléphone ou par d'autres moyens.

Les conducteurs participent également à un jeu qui leur fournit des commentaires quotidiens sur leur comportement au volant. Le jeu n'est pas directement axé sur l'économie de carburant. Il a plutôt pour fonction d'aider les participants à devenir de meilleurs conducteurs en contrôlant mieux leurs véhicules. Ce jeu les incite à perfectionner six techniques de conduite liées au carburant : plus de roulage, moins de freinages brusques, moins de ralentis, moins d'accélération brusques, moins de régime élevé, et plus de régulation de la vitesse. Toutes ces techniques sont implicitement reliées à la consommation de carburant.

Le portail permet aux conducteurs de voir leurs résultats et de mettre d'autres conducteurs au défi de faire mieux, le tout dans une ambiance conviviale. Les conducteurs qui se classent parmi les premiers sont ensuite reconnus par leurs collègues et peuvent publier leurs accomplissements dans les médias sociaux. Une fois tous les deux mois environ, les meilleurs conducteurs reçoivent également des récompenses financières, et le jeu est réinitialisé pour permettre à de nouveaux conducteurs de participer.

L'objectif principal du jeu est de renforcer les techniques d'écoconduite nouvellement apprises et de faciliter leur intégration au comportement habituel au volant.

Les données recueillies sur ces déplacements étaient d'abord été utilisées uniquement à des fins d'analyse comparative et d'évaluation au sein du portail. Cependant, au fil du temps, cette base de données est devenue une ressource précieuse dans l'évaluation de la relation entre le comportement au volant et la consommation de carburant. Puisque ce jeu s'adresse aux conducteurs de plusieurs pays qui utilisent différentes marques de camions et différents schémas de déplacements logistiques, les données permettent d'effectuer une analyse beaucoup plus approfondie.

**En utilisant des modèles statistiques et des techniques de regroupement, CGI a analysé les données abondantes de Scania sur ses conducteurs pour étudier la relation entre la consommation de carburant et le comportement au volant.**

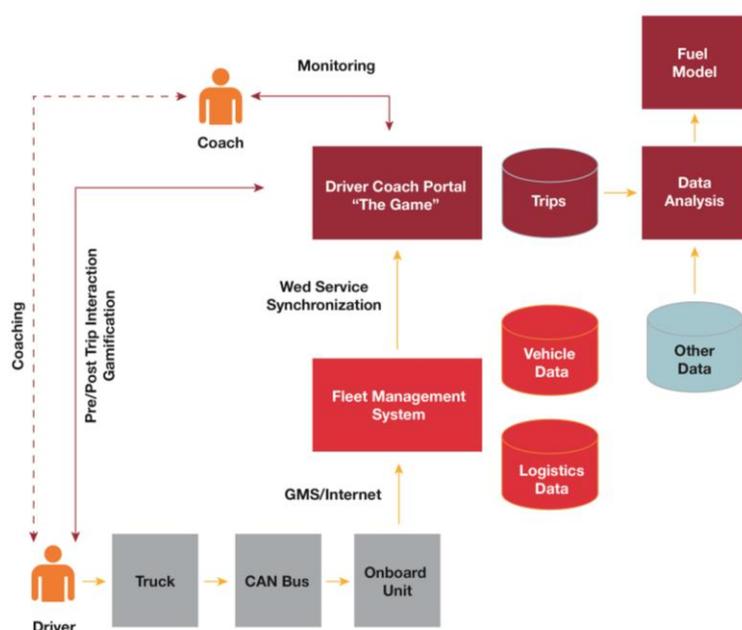


Figure 1 – Composantes du service d'accompagnement axé sur l'écoconduite

## Mettre à profit l'abondance de données de Scania sur ses conducteurs

En utilisant des modèles statistiques et des techniques de regroupement, CGI a analysé les données abondantes de Scania sur ses conducteurs pour étudier la relation entre la consommation de carburant et le comportement au volant.

Nous avons adopté une approche Agile pour l'analyse des données sur les déplacements, et avons effectué plusieurs itérations afin de mieux préciser les définitions des données, les champs de données et les modèles qui en résultent. L'analyse avait pour but de fournir un modèle permettant de relier le comportement au volant, tel que mesuré par l'entremise du portail, à la consommation réelle du carburant.

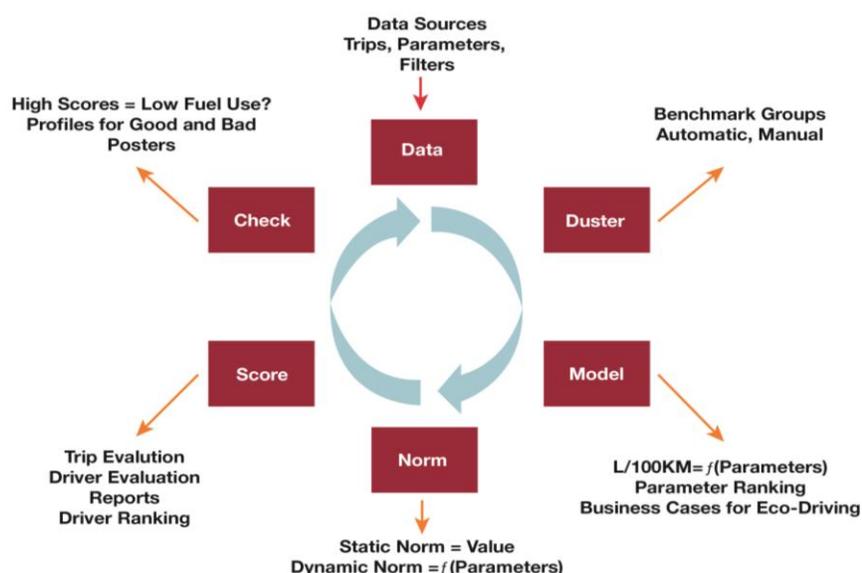


Figure 2 – Cycle d'analyse des données et résultats des différentes étapes

L'approche adoptée comportait plusieurs étapes.

- **Données** – La première étape consistait à filtrer les données et à les combiner avec d'autres sources. Nous avons combiné les données sur les déplacements avec les données météorologiques, par exemple la température et la vitesse du vent moyennes du jour. Les renseignements concernant le véhicule et le moteur étaient reliés aux déplacements au moyen des numéros de série des châssis.
- **Groupe** – Pour modéliser correctement la relation entre la consommation de carburant et le comportement au volant, nous avons comparé chaque déplacement à un groupe de référence en utilisant plusieurs techniques de regroupement manuel et automatique. Nous avons finalement choisi l'approche qui consiste en une grille bidimensionnelle représentant la *distance de déplacement quotidienne* versus la *vitesse moyenne du déplacement*.
- **Modèle** – Pour chaque groupe, nous avons construit un modèle reliant la consommation de carburant aux paramètres de déplacement. Plusieurs techniques de modélisation ont été utilisées, notamment les réseaux de neurones<sup>1</sup>, CHAID<sup>2</sup> et

<sup>1</sup> [http://en.wikipedia.org/wiki/Competitive\\_learning](http://en.wikipedia.org/wiki/Competitive_learning) (en anglais)

<sup>2</sup> <http://fr.wikipedia.org/wiki/CHAID>

plusieurs types de modèles de régression linéaire<sup>3</sup>.

Ces modèles nous permettent de classer les paramètres en fonction de leurs incidences sur la consommation de carburant. En outre, grâce aux services d'accompagnement des conducteurs et au jeu, nous pouvons établir une projection des changements dans la consommation de carburant. Cela nous permet d'utiliser le modèle en tant que calculateur d'analyse de rentabilité pour établir un lien entre les effets de l'accompagnement et une économie de carburant possible. Les modèles linéaires se sont révélés légèrement moins précis que les réseaux de neurones, mais plus faciles à utiliser dans la pratique.

- **Norme** – Grâce aux modèles et aux rangs centiles, nous pouvons établir des normes pour les paramètres des déplacements dans un groupe donné. Nous pouvons évaluer les comportements au volant non seulement en relation à la norme, mais aussi en relation à l'effet escompté sur la consommation de carburant. Nous avons testé des normes statiques et des normes qui sont une fonction de la vitesse moyenne, par exemple.
- **Note** – En utilisant les normes, nous pouvons calculer une note par déplacement. Ces notes peuvent être utilisées pour les évaluations des conducteurs, car elles reflètent le rendement du déplacement individuel dans le contexte du groupe.
- **Vérification** – Il est possible de tester les modèles et leurs algorithmes de classement et de notation avec différents ensembles de données pour vérifier si les déplacements ayant obtenu des notes plus élevées consomment réellement moins de carburant. À cette étape, l'ensemble du processus est passé en revue, ce qui donne généralement lieu à de nouvelles idées quant aux sources et aux champs de données à utiliser. Puis le cycle recommence avec une liste d'éléments à définir de façon plus précise ou des effets inattendus qui nécessitent une étude plus approfondie. Les commentaires des utilisateurs du modèle sont très importants, et il faut souvent apporter des changements pour rendre les résultats plus pratiques.

## GROUPES

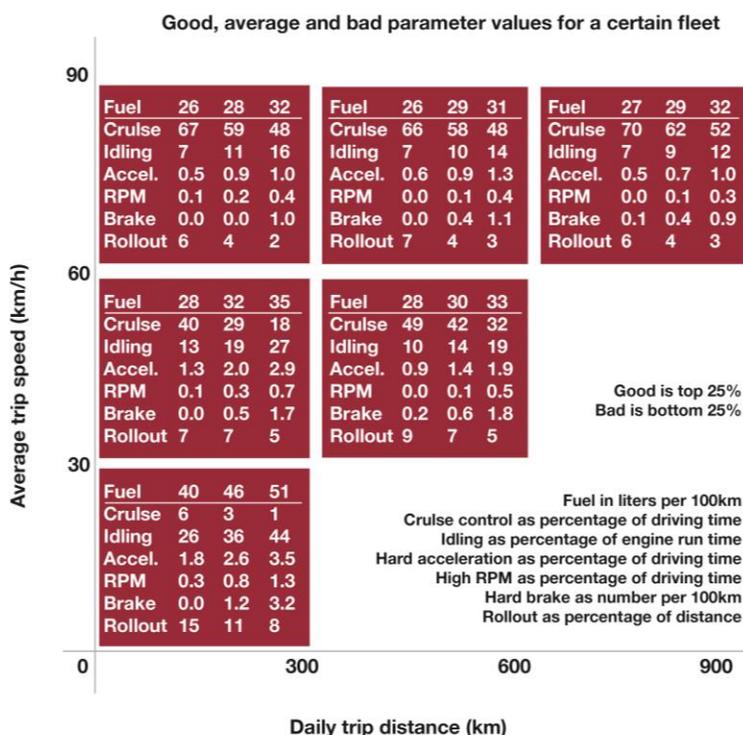
En utilisant la distance de déplacement quotidienne et la vitesse de déplacement moyenne pour répartir les données sur une grille, nous obtenons des résultats comparables à d'autres techniques de regroupement automatique. Ce format s'est révélé facile à comprendre tout en offrant une base statistique suffisamment significative (voir la figure 3). Cette grille permet également d'indiquer les zones pertinentes pour plusieurs schémas logistiques : les longs parcours (groupes du haut à droite), les courtes distances (groupes du haut à gauche), et la distribution (groupes du bas à gauche).

Cette grille peut également servir à présenter des résultats d'analyses. Nous l'avons utilisée pour fournir une plage de notes « efficace, moyennement efficace, inefficace » pour les différents paramètres et la consommation de carburant associée pour certains parcs, pays ou types et marques de véhicules.

Un paramètre ayant obtenu la note « efficace » permettra de réduire la consommation de carburant, et ce, en fonction d'un modèle linéaire par groupe. Un conducteur « efficace » se classera dans les 25 centiles supérieurs au chapitre de la consommation de carburant.

***Dans chaque modèle, nous pouvons voir les différents ensembles de paramètres et leur contribution globale aux variations dans la consommation de carburant.***

<sup>3</sup> [http://fr.wikipedia.org/wiki/Régression\\_linéaire](http://fr.wikipedia.org/wiki/Régression_linéaire)



**Figure 3 – Un tableau normatif par groupe utilisé pour évaluer le rendement d'un déplacement pour un parc de véhicules donné**

Nous avons développé ces tableaux pour plusieurs pays européens, et ils sont maintenant utilisés durant l'accompagnement afin de déterminer des objectifs d'accompagnement réalistes. Nous avons découvert que ces tableaux peuvent varier considérablement d'un pays à l'autre en raison des différences dans les conditions routières, les types de terrain et l'âge moyen du matériel roulant.

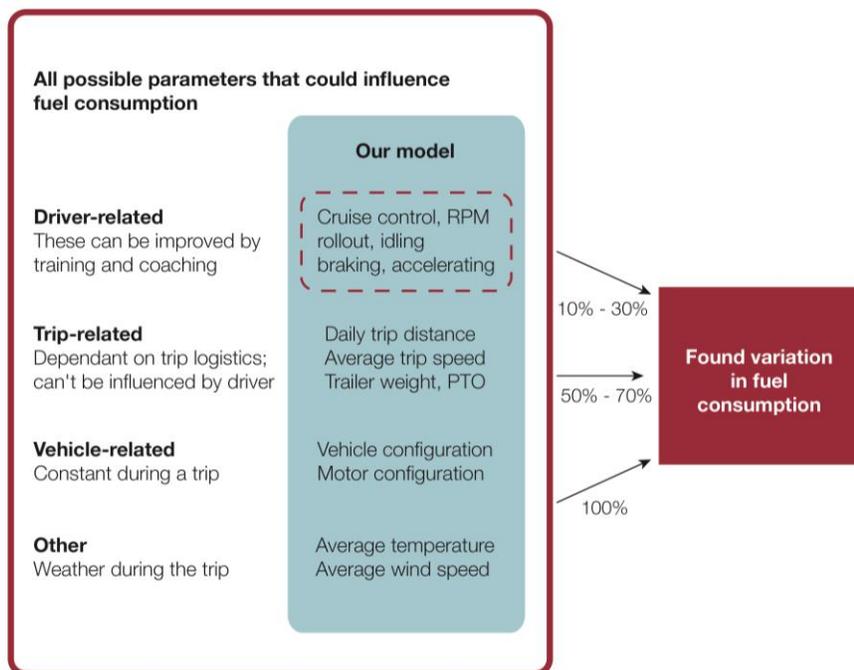
### MODÈLE DE CONSOMMATION DE CARBURANT

En utilisant les plaques d'immatriculation ou les numéros de série des châssis, nous pouvons relier les renseignements sur les déplacements à tous les types de données de configuration des véhicules et des moteurs. Nous pouvons également relier les déplacements à des renseignements additionnels provenant de bases de données logistiques et utiliser un horodateur pour les relier à l'information sur les conditions météorologiques et routières. De cette manière, nous pouvons juger du comportement au volant durant le déplacement dans un contexte plus étendu de paramètres qui influencent également la consommation de carburant, mais sur lesquels le conducteur n'a aucun contrôle.

Le modèle qui en résulte, fondé sur les déplacements en tenant compte de ces données supplémentaires, est un sous-ensemble de tous les facteurs qui influencent la consommation de carburant (voir la figure 4). La partie de ce modèle sur laquelle la formation et l'accompagnement ont une incidence est limitée.

Dans chaque modèle, nous pouvons voir les différents ensembles de paramètres et leur contribution globale aux variations dans la consommation de carburant. Dans l'ensemble, seulement 10 % à 30 % de la variation dans la consommation de carburant est attribuable à des facteurs liés au conducteur. Mais ce sont les variations au sein même de cette contribution qui s'avèrent les plus intéressantes. Les modèles correspondant aux groupes de long parcours dans les pays nordiques montrent des variations de plus de 70 %, ce qui indique que le conducteur a une grande influence sur la consommation de carburant.

Cependant, nous constatons également que, dans un modèle comparable aux Pays-Bas, les variations se limitent à 10 %, ce qui indique que l'accompagnement d'un conducteur aura moins d'incidence puisque la contribution des facteurs liés au conducteur est bien moindre.



**Figure 4 – Pourcentage de la variation dans la consommation de carburant attribué à différents ensembles de paramètres du modèle**

L'étude de cette incidence nous amène à rechercher une manière d'utiliser les données sur les déplacements pour calculer l'incidence prévue de la formation et de l'accompagnement pour différents pays, en fonction des différences observées dans la contribution des facteurs liés aux conducteurs à la consommation de carburant.

| Country | Estimated effect of coaching (L/100KM) |      |       |
|---------|--|------|-------|
|         | Min                                    | Mean | Max   |
| SWE     | 3,32                                   | 6,82 | 14,58 |
| DK      | 2,11                                   | 4,87 | 9,59  |
| IT      | 2,18                                   | 3,58 | 7,26  |
| UK      | 2,16                                   | 3,44 | 4,86  |
| FR      | 1,46                                   | 2,84 | 7,08  |
| NL      | 1,63                                   | 2,81 | 4,13  |
| PL      | 1,31                                   | 2,40 | 4,91  |

Figure 5 – Économies de carburant attendues à la suite de la formation et de l'accompagnement

Pour ce faire, nous avons formulé une estimation de l'efficacité de l'accompagnement. L'estimation de l'efficacité de l'accompagnement fournit une indication des avantages possibles de la formation et de l'accompagnement dans le but d'aider les conducteurs qui se classent dans les 25 derniers centiles pour leur consommation de carburant à se classer parmi les conducteurs des 25 premiers centiles. Cette migration moyenne de la population de conducteurs peut être exprimée en litres économisés par 100 km par pays. Le tableau ci-dessus illustre que le même niveau de formation et d'accompagnement pourrait avoir une incidence plus importante dans les pays où les contributions des facteurs liés aux conducteurs sont beaucoup plus élevées.

## Utilisation des modèles dans le cadre des services-conseils en gestion du carburant

Il est intéressant d'obtenir une image plus précise de la manière dont plusieurs paramètres de conduite changent pour différents types de conducteurs et de déplacements. En utilisant les données disponibles, il est possible de créer des cartes des points chauds. La figure ci-dessous présente un aperçu de la consommation moyenne de carburant pour les conducteurs de Scania en fonction de la distance parcourue et de la vitesse moyenne.

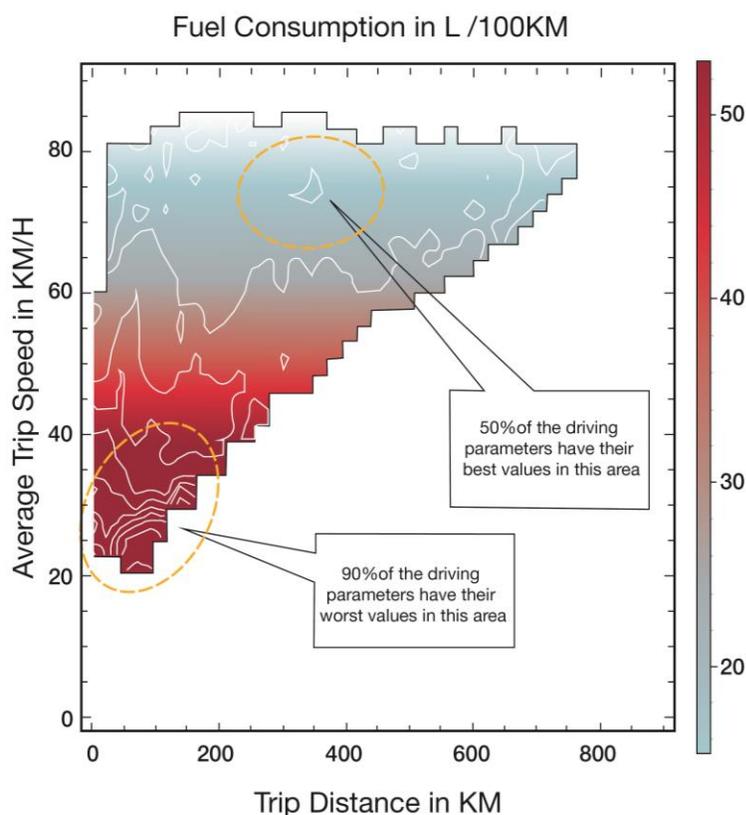


Figure 6 – Consommation de carburant en fonction de la distance parcourue et de la vitesse moyenne

Ce diagramme a été élaboré afin d'aider le propriétaire d'un parc de véhicules à déterminer les types de déplacements qui consomment le moins de carburant. Le propriétaire avait envisagé de concentrer ses services d'accompagnement sur les longs parcours (coin supérieur droit). Toutefois, l'analyse des données a révélé que ses déplacements les plus efficaces s'effectuaient sur une distance moyenne. Il pourrait cependant réduire davantage sa consommation de carburant en se concentrant plutôt sur la distribution (coin inférieur gauche). À l'aide de ce modèle, nous avons été en mesure de calculer la réduction dont pourrait bénéficier le propriétaire afin d'appuyer l'analyse de rentabilité élaborée en vue de son investissement dans les services d'accompagnement.

Grâce à ces techniques de regroupement et de modélisation, nous pouvons désormais répondre à un ensemble de questions sur la gestion du carburant. Il suffit d'analyser les données du propriétaire du parc de véhicules ou de les comparer au groupe de référence approprié.

Nous pouvons maintenant répondre aux questions ci-dessous grâce à des méthodes normalisées d'analyse.

- À quel point dois-je améliorer mes paramètres de conduite afin de réduire ma consommation de carburant d'au moins 5 % durant la première année de formation et d'accompagnement? Cette cible est-elle réaliste?
- Les services d'accompagnement des conducteurs visant à réduire la consommation de carburant seront-ils plus efficaces pour mes déplacements de courte distance ou pour mes déplacements internationaux sur de longues distances?
- Quelle est la différence entre la consommation de carburant de mes conducteurs « efficaces » et celle de mes conducteurs « inefficaces »?
- Parmi l'ensemble des paramètres de conduite, lequel a la plus grande influence sur la consommation de carburant? Mes services d'accompagnement devraient-ils d'abord se concentrer sur ce paramètre?
- À quel point les conditions météorologiques agissent-elles sur ma consommation de carburant?

## Défis durant le processus

Nous nous sommes heurtés à plusieurs difficultés dans notre processus d'analyse. Celles-ci sont susceptibles d'intéresser d'autres chercheurs.

### FILTRAGE

Nous avons utilisé un grand éventail de filtres pour extraire les données et les formater en vue de l'analyse. Il s'agissait entre autres de filtres simples – par exemple, rejeter les déplacements avec consommation de carburant négative, et de filtres complexes – comme s'assurer qu'il s'agissait d'un déplacement générateur de revenus plutôt que d'une visite au garage pour l'entretien courant.

Nous avons découvert que plus de 40 % des déplacements étaient rejetés durant le filtrage. Il s'agit d'un pourcentage plus élevé que prévu, et nous cherchons à déterminer comment cela pourrait créer des artefacts dans les autres données. Il est clair que le filtrage est une étape très importante dans ce processus. La configuration des filtres a plus d'incidence sur l'efficacité prédictive et la portée du résultat final global que le choix d'une stratégie de regroupement ou d'un format de modèle d'analyse. Cependant, le nombre total des déplacements n'a pas beaucoup d'importance. Nous avons souvent constaté qu'il était possible d'atteindre 80 % de l'efficacité prédictive finale d'un modèle en utilisant seulement un échantillon aléatoire de 20 % des données.

***Bien qu'il soit facile de communiquer les résultats numériques, il est plus difficile d'évaluer la portée de ces résultats et leur incidence sur l'utilisation possible des données numériques.***

## PARADOXE DE SIMPSON

Nous avons remarqué plusieurs exemples du paradoxe de Simpson (1953) durant nos études. Ce phénomène se produit lorsqu'une tendance linéaire apparaît dans différents groupes, mais est soudainement inversée lorsque ces groupes sont combinés en une seule analyse. Pendant un temps, il semblait qu'un roulage accru augmentait la consommation de carburant, en réalité, plutôt que de la réduire. Cependant, en examinant les données de plus près, il est apparu que les données étaient séparées en plusieurs groupes selon le poids du véhicule. Pour chaque groupe individuel, un roulage accru réduisait la consommation de carburant. Mais lorsque les groupes ont été combinés en une seule analyse sans l'inclusion de facteur de poids, la tendance inverse est apparue. Ce phénomène nous a incités à effectuer une inspection visuelle beaucoup plus approfondie des cas afin de pouvoir repérer ces effets.

## COMMENTAIRES DES UTILISATEURS

Les progiciels d'analyse de données actuels comportent un grand nombre d'options de visualisation pour les grands ensembles de données et les résultats prédictifs. Ces options sont certainement utiles pour les experts en statistique. Mais lorsqu'il s'agit de communiquer de l'information obtenue à partir de données massives à un vaste auditoire, la plupart de ces formats sont trop complexes. Dans notre approche itérative, nous nous sommes limités à des modèles plus simples et plus pratiques, même s'il fallait pour ce faire recourir à des modèles ayant une efficacité prédictive réduite.

Bien qu'il soit facile de communiquer les résultats numériques, il est plus difficile d'évaluer la portée de ces résultats et leur incidence sur l'utilisation possible des données numériques. Nous aimerions favoriser une utilisation étendue de l'information et en faciliter la compréhension en ne nous attardant pas outre mesure sur des aspects complexes comme la signification statistique. Cependant, il nous faut tenir compte de la signification pour nous assurer que les données sont utilisées correctement, même si le fait d'y accorder trop d'importance limite généralement l'accessibilité de l'information.

## Conclusion

Il apparaît clairement que ce type d'analyse de grands ensembles de données sur les déplacements peut fournir de précieux renseignements sur la relation entre le comportement au volant et la consommation de carburant. Les techniques d'analyse des données permettent également de combiner différents ensembles de données afin de comprendre le contexte entourant les facteurs liés aux conducteurs. Les modèles qui en résultent nous ont fourni une image claire de cette relation, que l'on peut directement traduire en objectifs d'accompagnement réalistes pour les conducteurs et en attentes réalistes au chapitre des économies de carburant pour les différents propriétaires de parcs de véhicules.

## Remerciements

Cette étude a été menée par CGI dans le cadre du projet sur les méthodes d'évaluation du rendement de Scania. Ce projet avait pour but de développer une méthode d'évaluation du comportement au volant durant les déplacements afin de guider les services d'accompagnement visant à réduire la consommation de carburant.

### À PROPOS DE CGI

Fondée en 1976, CGI est un fournisseur de services informatiques et de gestion des processus d'affaires fournissant aux entreprises des services-conseils en management, des services d'intégration de systèmes et d'impartition de haute qualité.

Grâce à ses 68 000 membres présents dans 40 pays, CGI a un bilan inégalé de projets réalisés selon les échéances et budgets prévus. Nos équipes s'arriment aux stratégies d'affaires des clients afin d'obtenir des résultats probants sur toute la ligne.

Pour plus de renseignements sur les sujets présentés dans cette étude ou sur CGI, visitez le site [www.cgi.com](http://www.cgi.com) ou écrivez-nous à [info@cgi.com](mailto:info@cgi.com).

## Références

1. Lapré, L. (2013). « Scania eco-driving development », Workshop SIS39, 9<sup>th</sup> ITS European Congress, Dublin. ERTICO (ITS Europa).
2. Stam, M. (2012). « The use of gamification to drive behavioral change – lessons from the Strategic Platform for ITS (SPITS) and Scania », Workshop TS086, 19<sup>th</sup> ITS World Congress, Vienne, ERTICO (ITS World).
3. Constantinescu, Z. (2010). « Driving style analysis using data mining techniques », *Int. J. of Computers, Communications & Control*, Vol. V (2010), N° 5, pp. 654-663.
4. Simpson, E.H. (1951). « The Interpretation of Interaction in Contingency Tables », *Journal of the Royal Statistical Society*, Ser. B, 13, 238-241.