

Entreprise de parrainage

Partenaires

L'assurance automobile comportementale (Pay How you Drive):

Impact sur la gestion du risque automobile

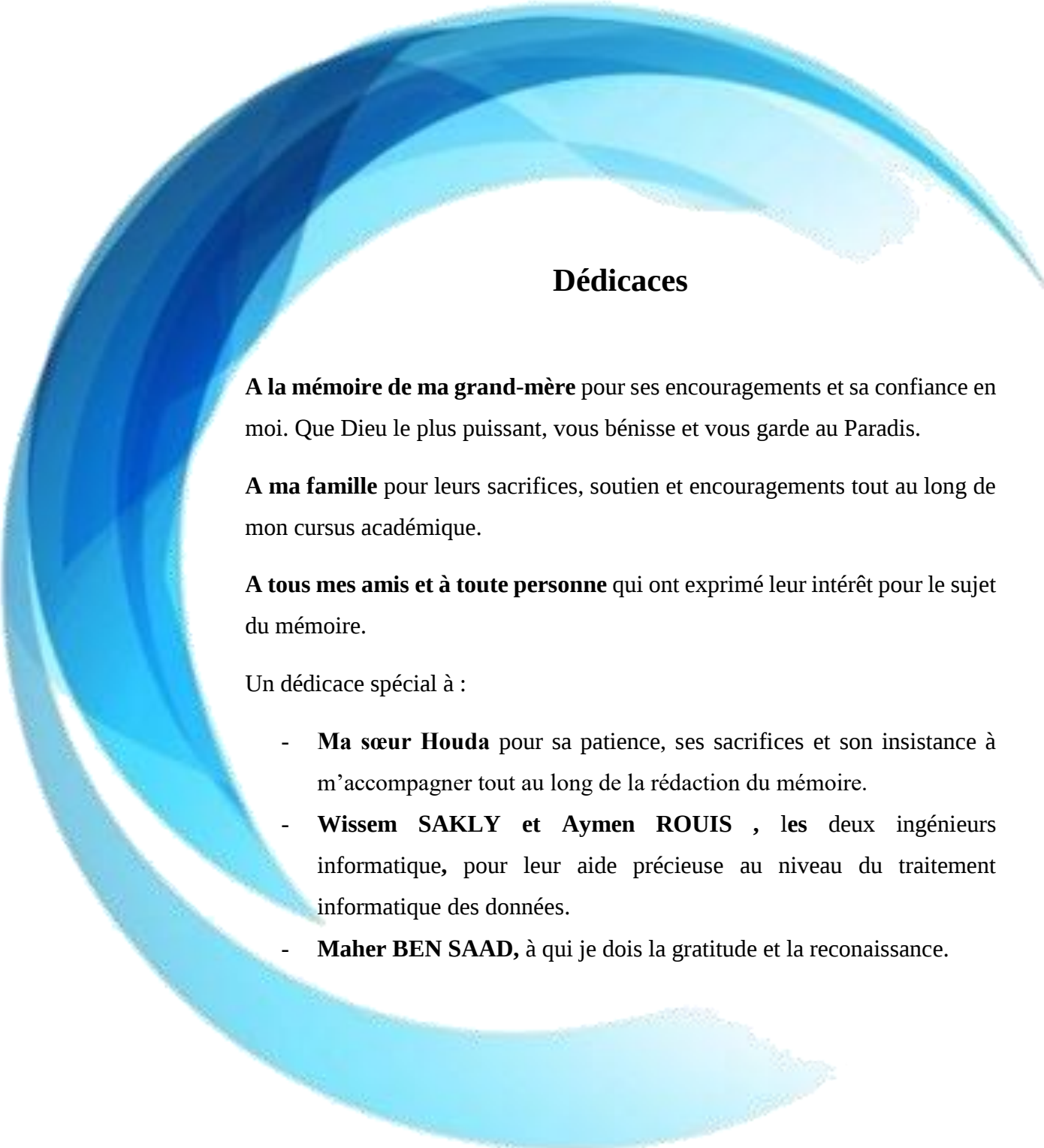
*Analyse théorique avec application sur des données télématiques du produit
«Sayartech » commercialisé par Ooredoo*

Présenté et soutenu par :

Amal ROUIS

Encadrée par :

Farouk KRIAA

A large, stylized, blue curved graphic that resembles a thick, multi-layered arc or a partial circle, framing the central text. It has various shades of blue, from light to dark, creating a sense of depth and movement.

Dédicaces

A la mémoire de ma grand-mère pour ses encouragements et sa confiance en moi. Que Dieu le plus puissant, vous bénisse et vous garde au Paradis.

A ma famille pour leurs sacrifices, soutien et encouragements tout au long de mon cursus académique.

A tous mes amis et à toute personne qui ont exprimé leur intérêt pour le sujet du mémoire.

Un dédicace spécial à :

- **Ma sœur Houda** pour sa patience, ses sacrifices et son insistance à m'accompagner tout au long de la rédaction du mémoire.
- **Wissem SAKLY et Aymen ROUIS** , les deux ingénieurs informatique, pour leur aide précieuse au niveau du traitement informatique des données.
- **Maher BEN SAAD**, à qui je dois la gratitude et la reconnaissance.

Remerciements

Je tiens, tout d'abord, à remercier, **M. Farouk KRIAA**, mon encadrant du mémoire, pour son sérieux, sa compétence et pour son sens du devoir.

Je remercie **la direction générale de COMAR assurances** pour l'intérêt qu'elle porte au sujet du mémoire et pour l'opportunité de le traiter.

J'exprime la gratitude à **M. Mustapha BEN GHACHEM**, Directeur adjoint VAS & New Business Opportunities (NBO) chez Ooredoo Tunisie, d'avoir accepté de me fournir les données des clients du produit Sayartech malgré le souci de confidentialité.

Je remercie également **Mme. Nouha HAMMAR**, Chef de projet NBO chez Ooredoo Tunisie, pour sa précieuse collaboration.

Je remercie particulièrement **M. Faysal JILANI**, Fondateur de la start-up Telematik et concepteur du produit Sayartech, et toute son équipe, pour leur patience, leur courtoisie et le temps consacré afin de me fournir la base des BIG DATA bien traitée malgré les conditions difficiles du travail pendant le mois de Ramadan .

Je remercie **M. Hatem MANAI**, Directeur technique, **M. Maher MARZOUKI**, Chef service Automobile et son équipe, de m'avoir acceptée de passer un stage de deux mois au service automobile à la COMAR, et de me fournir de l'information nécessaire pour le mémoire.

Sans oublier, **M. Sami HIDRI**, Directeur de la formation à la COMAR et **Mme. Imen LOUSAIEF**, sa collaboratrice qui étaient toujours présents tout au long de mon parcours académique à l'IFID pour me conseiller et pour assurer la continuité et le bon déroulement de la formation.

Un remerciement spécial à **M. Bechir ELLOUMI**, Directeur central « Production » et **M. Badr BEN SLIMANE**, Directeur « Prestations » à la COMAR pour leurs conseils et leur intérêt manifesté pour le sujet du mémoire.

Je remercie infiniment **M. Racem KTATA**, Directeur des risques chez GAT Assurances, pour son intérêt pour le sujet et pour son avis professionnel.

J'adresse un remerciement tout particulier à **M. Mohamed Achraf MAIZA**, Ex-Actuarial Data Scientist à Axa Global Direct France, pour son éclaircissement précieux à propos de la méthodologie à adopter dans le travail.

Je remercie enfin, l'administration et tout le personnel de l'IFID pour leurs efforts déployés pour instaurer les meilleures conditions de travail.

Résumé

Le déséquilibre tarifaire dont souffre la branche automobile en Tunisie trouve ses raisons dans la variété des risques individuels de conduite souvent mal évalués. Avec le développement de la télématique automobile, l'approche de l'assurance comportementale Pay How You Drive (PHYD) présente une solution inédite pour l'optimisation de la gestion du risque automobile qu'est la tarification individuelle sur la base du comportement au volant.

L'objectif de l'étude est d'identifier les nouveaux facteurs associés au risque individuel de conduite et de prédire le niveau du risque de chaque conducteur en se basant sur les variables comportementales caractérisant le profil de conduite.

Pour l'application et le développement de la méthodologie, ont été utilisées les données télématiques de 48 conducteurs abonnés à l'unique produit PHYD « Sayartech » offert en Tunisie par Telematik et Ooredoo et surveillées pendant une période variant de 3 à 14 mois.

Une analyse bivariée a été conduite pour identifier les facteurs significatifs de risque. Les résultats indiquent que les événements d'incidents critiques (freinage brusque, accélération et virage rapides), la conduite en excès de vitesse ($> 120\text{km/h}$) et la conduite en nuit sont des facteurs fortement associés au risque de quasi-accident (substitut du risque d'accident). Un modèle de régression linéaire multiple a été développé pour prédire le niveau du risque individuel de conduite. Les résultats montrent que le modèle présente un fort pouvoir prédictif avec un R^2 de 71%.

Cette étude conclut que le risque individuel de quasi-accidents est associé aux événements d'incidents critiques (accélération et freinage brusques), à la conduite en nuit et à la conduite en excès de vitesse. De plus, les variables comportementales sont des excellents prédicteurs pour le profilage des conducteurs risqués avant même la réalisation des sinistres.

Mots-clés

Voiture connectée ; télématique ; Pay How You Drive ; incidents critiques ; quasi-accidents ; risque individuel de conduite ; comportement au volant ; Big Data ; UBI ; régression linéaire multiple ; évaluation du risque

Sommaire

Introduction générale.....	9
Chapitre Premier : L'assurance automobile comportementale : une opportunité pour les assureurs	12
Introduction	12
Section 1 Assurance automobile conventionnelle.....	13
1.1. Fonctionnement de l'assurance automobile conventionnelle	13
1.2. Risques automobiles	18
1.3. Problèmes soulevés de l'assurance automobile conventionnelle.....	21
Section 2 Nouvelles formules d'assurance automobile : Assurance automobile comportementale	23
2.1. Comportement au volant.....	23
2.2. Caractérisation de profils de conduite.....	25
2.3. Technique de la télématique automobile	27
2.4. Variables de conception d'un modèle d'assurance comportementale	31
2.5. Principe de l'assurance comportementale.....	36
Section 3 Apports de l'assurance comportementale	38
3.1. Expériences à l'étranger.....	38
3.2. Bénéfices tridimensionnels	43
3.3. Défis de l'assurance comportementale	47
Deuxième chapitre : Analyse du comportement au volant par la télématique automobile du produit « Sayartech » d'Ooredoo	51
Introduction	51
Section 1 Cadre général de l'étude	52
1. Entreprise de parrainage.....	52
2. Les partenaires de l'étude	53
3. Description du produit Sayartech	54
4. Fonctionnement du produit Sayartech	56
5. Processus de sélection de risque.....	61
Section 2 Données télématiques	64
1. Description de l'échantillon	64
2. Définition du phénomène d'accident, quasi-accident et incidents critiques...	65
3. Traitement de données	67
4. Variables prédictives	68
5. Exemple d'analyse du comportement de deux conducteurs	75

Section 3 Méthodologie adoptée.....	79
1. Etude descriptive.....	79
2. Etude analytique	79
Section 4 Résultats.....	85
1. Etude descriptive.....	85
2. Etude analytique	86
Section 5 Discussion et conclusions.....	99
Conclusion générale	103
Annexes	105
Bibliographie.....	127

Liste des abréviations

Acc	Nombre d'accélération fortes et excessives
AIC	Akaike information criterion
CRM	Customer Relationship management
DD	Day Driving
Decc	Nombre de freins forts et excessifs
GPRS	General Packet Radio Service
GPS	Global Positioning System
GSM	Global System for Mobile Communications
HTTPS	HyperText Transfer Protocol Secure
M2M	Machine To Machine
MOD	Main d'œuvre directe
ND	Night Driving
OBD	On-Board Diagnostics
OMS/ WHO	Organisme Mondiale de Santé
PAYD	Pay As You Drive
PHYD	Pay How You Drive
QA	Quasi-accident
RC	Responsabilité Civile
RDBMS	Relational Database Management System
SSL	Secure Sockets Layer
TIC	Technologies de l'information et de la communication
UBI	Usage-Based Insurance
VMT/VTM	Véhicule Terrestre à Moteur

Liste des tableaux

Tableau 1 Critères de tarification a priori [11].....	15
Tableau 2 Formules conventionnelles d'assurance automobile [17].....	17
Tableau 3 Principaux aléas liés aux véhicules et leurs différents motifs[3]	20
Tableau 4 Comparaison des catégories de la technologie de la télématique [49].....	28
Tableau 5 Un ensemble de Données brutes collectées d'un dispositif télématique [45].....	30
Tableau 6 Liste non exhaustive des données communément extraites [45]	30
Tableau 7 Variables prédictives du risque d'accident selon le modèle UBI	32
Tableau 8 Exemples de produits d'assurance télématique [44]	41
Tableau 9 les évènements détectés par le dispositif télématique	58
Tableau 10 Nombre d'évènements moyen d'un conducteur par unité de temps ou par km	66
Tableau 11 Taux d'accident par 1000 habitants par pays en 1991	66
Tableau 12 Les victimes de la route par 100 000 habitants par année par pays	67
Tableau 13 Liste des données télématiques agrégées	71
Tableau 14 Liste finale des variables du modèle UBI à étudier	75
Tableau 15 Statistiques descriptives de l'échantillon	85
Tableau 16 Répartition de la populaion selon la durée de conduite sous une vitesse supérieure à 150 km/h	86
Tableau 17 Répartition de la population selon la durée de conduite sous une vitesse entre 120 et 150 km/h.....	86
Tableau 18 Variation du degré d'association de la variable «Speed1 » avec QA en présence d'autres variables.....	89
Tableau 19 L'interdépendance significative des variables explicatives mesurée par le coefficient de corrélation.....	90
Tableau 20 Evaluation des paramètres du modèle quadratique (VI = distance parcourue)	Error! Bookmark not defined.
Tableau 21 Evaluation des paramètres du modèle initial	93
Tableau 22 Evaluation des paramètres du modèle final.....	96
Tableau 23 Analyse de sensibilité du modèle par le changement de la variable dépendante .	99

Liste des figures

Figure 1 Continuum des évènements de circulation [19].....	19
Figure 2 Les deux voies d'accident [35].....	24
Figure 3 Caractérisation des comportements au volant [36].....	26
Figure 4 Aperçu de l'architecture technique du système télématique [45]	29
Figure 5 Les différents termes des nouvelles formules de l'assurance automobile [45]	36
Figure 6 Nombre de programmes actifs d'assurance automobile connectée par pays [75]	39
Figure 7 Impact estimatif de la télématique automobile sur la chaîne de valeur de l'assurance[84].....	47
Figure 8 Le CA par branche d'activité en 2017	53
Figure 9 Captures d'écran de l'application mobile de Sayartech	55
Figure 10 Interface client /agent prototypée	61
Figure 11 Nouveau processus d'évaluation du risque automobile	62
Figure 12 Outputs du prototype d'assurance automobile comportementale	63
Figure 13 Causes des accidents de la circulation en Tunisie en 2018 (les cinq premiers mois) [98]	69

Figure 14 Nombre d'accidents observés en 2018 (5 premiers mois) en fonction du jour de semaine (à gauche), du moment de la journée (à droite) et selon le type de la route (en bas)	70
Figure 15 les quatre les plus fortes relations linéaires avec QA	88
Figure 16 les deux les plus faibles relations avec QA.....	88
Figure 17 Modèles linéaire et quadratique associant le nombre des QA avec la distance parcourue / 1000km.....	92
Figure 18 Qualité d'ajustement du modèle	95
Figure 19 Distribution de valeurs résiduelles.....	Error! Bookmark not defined.

Introduction générale

La télématique automobile permet la collecte d'informations sur le conducteur via un dispositif électronique. Utilisant la technologie de GPS, ce dispositif enregistre la distance des trajets et toute autre information liée au comportement du conducteur. L'utilité de cette technique dans le domaine de transport et de la sécurité routière est largement étudiée par plusieurs chercheurs à savoir Toledo *et al* (2008), Klauer *et al* (2009) et Guo *et al* (2013). Cette technique est susceptible d'apporter des changements radicaux dans l'assurance automobile dans le futur proche. Avec l'essor de la télématique automobile et du développement du BIG DATA, l'industrie de l'assurance automobile semble en être bénéficiaire avec le développement rapide des produits Pay-As-You-Drive et Pay-How-You-Drive -fruit de l'approche de l'assurance comportementale-. Les bénéfices potentiels de cette approche sont essentiellement l'amélioration de la précision statistique de la tarification du risque et une opportunité pour les meilleurs conducteurs de tirer profit de leur contrat d'assurance.

L'évaluation conventionnelle du risque automobile dans l'assurance est basée sur les modèles de fréquence et de coût. Ces derniers permettent de prédire le nombre de sinistres futurs et leurs coûts en se référant à l'historique de sinistralité. Les données liées au véhicule et au conducteur incluant ces modèles sont généralement collectées le jour de l'émission du contrat. Par contre, les données comportementales notamment le kilométrage ne peuvent pas être mesurées objectivement par la simple déclaration de l'assuré. C'est alors que la collecte des données comportementales objectives devient possible avec la technologie de la télématique et les outils BIG Data. Dans le monde, les experts du métier et les chercheurs ont toujours essayé avec l'essor des Big Data de montrer les opportunités qu'offre l'analyse du comportement au volant à l'assureur automobile ; l'amélioration de la tarification du risque automobile est le centre d'intérêt des travaux de Liu *et al* (2017), Ayuso *et al* (2016), Paefgen *et al* (2014, 2013), Jun *et al* (2011), de l'étude de Progressive Insurance (2005) et de l'étude de Ptolemus consulting¹ (2012).

A notre connaissance, la présente étude sera la première tentative en Tunisie de mettre en évidence l'utilité de l'approche comportementale dans l'amélioration de la modélisation du risque individuel de conduite. Cela implique l'utilisation des données télématiques d'un portefeuille d'abonnés tunisiens dans un programme UBI offert par une start-up tunisienne de services télématiques. Par risque individuel de conduite, on entend le risque d'être impliqué dans un accident sans pour autant

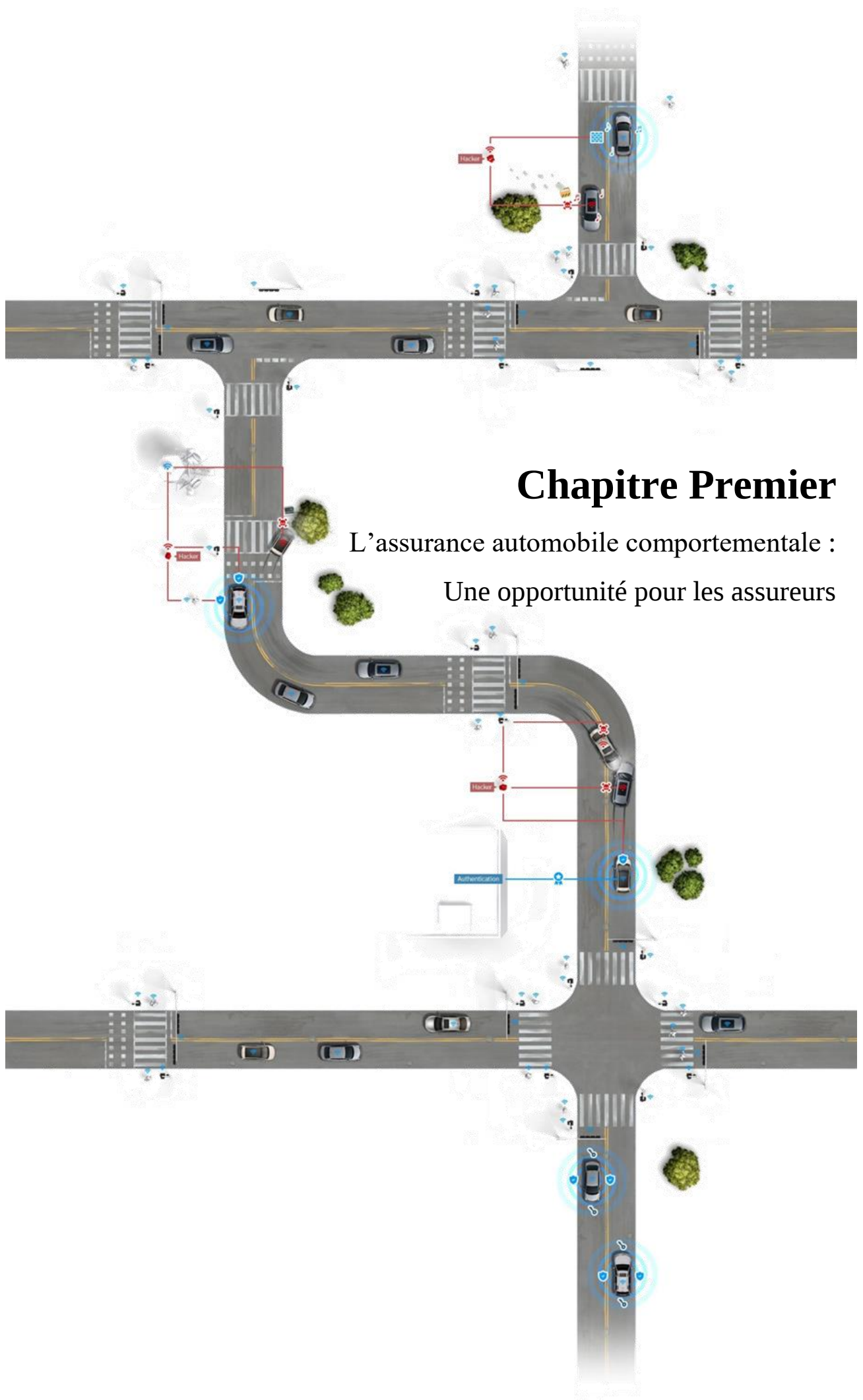
¹ Cabinet de conseil européen spécialisé en services télématiques et assimilés

qu'il y ait collision. Il s'agit des situations de quasi-accidents au lieu des accidents sur lesquelles l'étude se focalise.

L'approche de l'assurance comportementale suscite le recours aux techniques d'analyse du comportement humain au volant appelées techniques de « Béhaviorisme » afin de juger objectivement la qualité de conduite des assurés et de mieux prédire le risque individuel de conduite. Elles se basent sur l'observation du comportement individuel en interaction avec son environnement qu'est le trafic routier. L'observation est effectuée en temps réel à l'aide des dispositifs télématiques pour analyser et profiler le comportement de l'assuré. Ce profil, par la suite, est traduit en paramètres d'estimation du niveau de risque individuel de conduite.

A cet effet, l'objectif du mémoire est d'identifier les nouveaux facteurs associés au risque individuel de conduite et de prédire le niveau du risque de chaque conducteur en se basant sur les variables comportementales caractérisant le profil de conduite.

Ce mémoire servira alors à montrer l'intérêt de l'approche de l'assurance comportementale par une mesure qualitative et quantitative de son impact sur la gestion du risque automobile ; Dans le premier chapitre, une comparaison de cette approche avec la logique conventionnelle de l'assurance automobile est proposée, soutenue par l'exposition des expériences réussies à l'étranger et des résultats des études scientifiques traitant ce sujet. Le deuxième chapitre dévoile le fonctionnement de l'unique produit PHYD en Tunisie dans une première partie. Ensuite, il expose une tentative de mise en place d'un nouveau processus de sélection de risque et, enfin, une dernière partie servira à l'étude expérimentale qui se penche sur la méthode de profilage des comportements au volant, par le moyen de l'information fournie par la télématique automobile de la start-up Telematik.



Chapitre Premier

L'assurance automobile comportementale :

Une opportunité pour les assureurs

Chapitre Premier : L'assurance automobile comportementale : une opportunité pour les assureurs

Introduction

L'évolution industrielle et technologique a considérablement développé les aléas dans le monde et toute personne cherche à protéger son patrimoine et/ou son intégrité physique à travers l'assurance [1]. L'assureur, à travers le calcul de probabilités, se permet de « couvrir la plupart des risques de l'existence »[2] y compris le risque automobile. Le coût de la couverture représente pour l'assuré la prime payée en amont et pour l'assureur la compensation en aval en cas de survenance de l'événement couvert. Néanmoins, avec l'évolution de l'offre d'assurance, le client qui paye souvent une prime statique et inadaptée à son profil de risque est à la recherche des tarifs les plus avantageux. Cette logique classique de l'assurance est face à un dilemme pour la branche automobile: Le parc automobile mondial passe de 2 millions [3] en 1930 à 1,8 milliards² en 2018, le risque automobile lui-même évolue avec l'évolution des technologies dans l'industrie automobile. Les besoins des clients, par conséquent, requièrent des réponses innovantes et efficaces. La concurrence acharnée exige aux assureurs d'adopter une stratégie d'innovation. Cette dernière vise d'une part, à mieux connaître les clients et d'estimer leur niveau de risque pour leur proposer une prime personnalisée et adaptée à leurs profils de risque. D'autre part, elle vise à renforcer la culture de l'assurance à travers la proposition des services adaptés à leurs besoins.

L'objectif de ce chapitre consiste à montrer l'importance du comportement du conducteur dans l'évaluation du risque automobile. Plusieurs chercheurs en sécurité routière ont révélé l'importance de comprendre le comportement des conducteurs afin de maîtriser le risque automobile notamment Nääätänen et Summala, (1976), Fuller (1984), Ajzen (1988) et Evans (1991). La première forme de l'assurance comportementale était initialement proposée par Vickrey (1968) et l'introduction de la technologie de la télématique automobile a beaucoup développé les recherches en assurance automobile comportementale à savoir les travaux de Jun *et al* (2007, 2011), Paefgen *et al* (2013, 2014), Ayuso *et al* (2016), Weidner *et al* (2016) et Soleymanian *et al* (2017). En pratique, Le premier produit d'assurance automobile comportementale dans le monde est proposé par Progressive aux Etats-Unis (1998) suivi par Unipol en Italie (2005) –la première en Europe– pour arriver aux quatre coins du monde au bout de dix ans.

² Estimation en temps réel selon <http://carfree.fr/parc-automobile-mondial-temps-reel.html>

Ainsi, ce chapitre vise à présenter un panorama de l'assurance automobile, de ses problèmes et de son évolution dans le monde à l'ère du numérique. A cet effet, un zoom sur les nouvelles formules de l'assurance automobile sera exposé avec l'introduction de la télématique automobile. Enfin, il détaille les résultats des études antérieures tant au niveau national qu'au niveau international traitant la problématique du thème de l'assurance comportementale.

Section 1 Assurance automobile conventionnelle

Cette première section sert à donner un aperçu sur l'assurance automobile conventionnelle. En premier lieu, elle expose le fonctionnement de l'assurance automobile et ses principes de tarification, ensuite, le risque automobile et enfin les problèmes soulevés de l'activité.

1.1. Fonctionnement de l'assurance automobile conventionnelle

Afin de comprendre la logique de l'assurance automobile, il convient de présenter les bases théoriques de l'assurance avant de passer aux principes de tarification du risque automobile.

a. Bases théoriques

Le risque automobile est la probabilité que le véhicule sur lequel porte la couverture soit impliqué dans un accident de circulation. Cette probabilité peut être calculée à partir des études statistiques sur la fréquence et le coût des accidents tout en prenant compte des grands principes d'assurance suivants [4] :

- *La mutualité* : La solidarité entre les personnes assurées contre la survenance d'un même événement. En cas d'aggravation des risques ou de fraude des assurés, l'ensemble de la mutualité subit le sort.
- *La loi de grand nombre* : Plus une population homogène d'assurés est grande, plus la fréquence des accidents est approximativement égale à la probabilité d'accident. Lorsque les dommages subis par les assurés sont distribués d'une manière identique et homogène, l'indemnité moyenne par assuré –qui est aléatoire– devient quasi-constante et prévisible par l'assureur. Donc, plus le nombre d'assurés est grand, plus le coût moyen par assuré est estimé avec précision [5]. Pour assurer l'équilibre budgétaire de l'assureur, la prime (actuarielle) doit être égale à l'espérance de perte. Par contre, l'augmentation du nombre d'assurés amplifie l'hétérogénéité de la population ce qui augmente le risque moyen et donc les primes payées par les assurés. Rothschild et Stiglitz (1978) ont montré qu'en présence d'une population hétérogène et d'une asymétrie d'information, une classification des risques en sous-populations homogènes s'avère primordiale pour offrir des primes adaptées aux différents niveaux des risques [6]. Pour le faire, l'assureur doit être en mesure de disposer

d'un minimum d'informations nécessaires sur ses assurés. Arrow (1963) montre effectivement qu'une tarification imparfaite peut produire « une sélection adverse défavorable des risques » ; Les contrats non assez discriminants sont susceptibles d'attirer principalement les mauvais risques présents sur le marché [7].

De ce qui précède, on retient que l'assureur doit tenir compte de deux problèmes majeurs à savoir l'anti sélection et l'aléa moral :

- *Aléa moral ou moral hasard* est le fait qu'une personne physique et/ou morale assurée contre un risque peut se comporter de manière plus risquée que si elle était totalement exposée au risque [8].
- *Anti-sélection ou sélection adverse* initialement introduite par Akerlof (1970) se définit, en assurance, comme l'asymétrie d'information entre assureur et assuré qui amène à mal estimer la probabilité de risque. L'assureur alors n'est plus en mesure de différencier les primes en fonction des risques. Les primes reflètent le coût moyen des sinistres des individus dans la même mutualité. Il s'agit d'une situation de déséquilibre qui bénéficie aux « mauvais risques » et empêchent les « bons risques » à s'assurer [9].

Plusieurs travaux de recherche ont montré que la classification statistique des risques permet d'améliorer leur sélection à la souscription et que cela contribue à réduire l'effet de l'anti sélection. Hoy (1982) montre que la classification conduit à la constitution de groupes de moindres risques qui peuvent obtenir de meilleures conditions d'assurance, sans pour autant pénaliser les assurés des autres groupes [10].

Compte tenu des problèmes d'asymétrie d'information précités, la logique de l'assurance automobile conventionnelle peut se résumer comme suit :



L'assureur collecte les informations déclarées par l'assuré détaillées ci-après sur lesquelles il se base pour estimer sa classe de risque. La prime actuarielle compatible avec le niveau du risque de l'assuré sera révisée en cours du contrat souscrit par les observations des sinistres encourus a posteriori.

b. Principe de tarification du risque automobile

D'une façon générale, deux types de tarification existent pour le risque automobile selon la nature des critères utilisés ; la *tarification a priori* qui permet de calculer la prime de base, qui sera au fur et à mesure modulé de l'historique des implications dans les accidents via la *tarification a posteriori*.

Tarification a priori

Les principaux critères de tarification du risque automobile sont exposés dans le tableau suivant:

Tableau 1 Critères de tarification a priori [11]

Critères	Sous-critères	Exemples
Liés au véhicule	Catégorie de véhicule ³	1 ^{ère} catégorie : véhicule de tourisme et commerciaux à 4 roues 2 ^{ème} catégorie : camions et véhicules industriels 3 ^{ème} catégorie : cyclomoteurs
	Zone de circulation	
	Usage du véhicule	Privé, professionnel, année de 1 ^{ère} mise en circulation, catégorie socio-professionnel
Liés au conducteur	Qualité du conducteur [3]	Exclusif, principal, secondaire, occasionnel, novice ⁴
	Age	
	Sexe	
	Expérience de conduite	Ancienneté de permis
	Antécédents (Bonus-Malus) ⁵	Nombre de sinistres ou infraction, nombre d'années sans sinistre
Packages de garanties		

³ La tarification des catégories :

- 2^{ème} catégorie est fonction du tonnage et de type de matière transportée par le véhicule.
- La 1^{ère} et la 3^{ème} catégorie sont selon le groupe et la classe. Le groupe est fonction du type et de la puissance du véhicule et caractérise le risque RC tandis que la classe

⁴ **Conducteur novice** : Tout individu ayant récemment eu son permis de conduite indépendamment de son âge

⁵ Informations collectées du **relevé d'informations** qui est un document obligatoirement délivré par l'assureur à l'assuré en cas de résiliation de contrat et retraçant le passé de sinistres de l'assuré

Néanmoins ces critères ne sont pas universels ; leur adoption varie d'un pays à un autre. Dans l'annexe n°1, les différents critères de tarification adoptés par un échantillon de pays sont présentés. A partir de ces critères, l'assureur peut classer les conducteurs en sous-populations plus homogènes, et calculer un coefficient de tarification multiplié par la valeur du point (unité de valeur tarifaire d'assurance) ce qui donne la prime de base de l'assuré [12].

Cependant, malgré la multitude de critères utilisés pour la tarification, il est possible d'observer une certaine variabilité au sein même d'une seule classe de risque. D'où l'utilité de la tarification a posteriori.

Tarification a posteriori

Le tarif calculé a priori repose sur des données acquises mais anciennes. Pour affiner son calcul, il faut prendre en compte des données du système Bonus-Malus ⁶ [11].

La prime de base est alors réduite en l'absence de sinistre, ou majorée en fonction du nombre de sinistres enregistrés. Le système a trois objectifs principaux tels que proposés par Henriet et Rocher (1986) [13] :

- Responsabiliser le conducteur, puisque la prime augmente en cas de sinistre,
- Affiner davantage la sélection des risques
- Réduire les coûts de gestion, en incitant les assurés à ne pas déclarer les petits sinistres,

Ce mécanisme se justifie par l'amélioration de la classification des risques d'une part et de la réduction de l'aléa moral d'autre part. Lambert (1983) a montré que l'effet de l'aléa-moral s'atténue avec l'augmentation de la durée du contrat d'assurance [14]. Cooper et Hayes (1987) également, généralisant le modèle de Rothschild et Stiglitz (1976) à un cadre multi-périodique, ont montré qu'il est possible de réduire l'effet de l'anti-sélection sur le bien-être des assurés à bas risque, en introduisant une tarification prenant compte de l'expérience passée [15]. Par conséquent, le système Bonus-Malus permet d'atténuer les effets de la sélection adverse.

Dans le monde, il existe deux approches du système [12] :

- Le changement de classe de risque, où l'assuré peut passer d'une catégorie de risque à une autre, et donc d'une prime de base à une autre, supérieure ou inférieure, en fonction de son taux de sinistralité,

⁶ **Le système Bonus-Malus** permet de réviser le montant de la prime selon la survenance ou non de sinistre impliquant la responsabilité de l'assuré, en fonction de paramètres fixés à l'avance.

- La multiplication de la prime de base, où le taux de sinistralité est converti en un coefficient de réduction/majoration multiplié, par la suite, à la prime de base.

Chaque pays a ses propres méthodes de calcul des bonus et des malus attribués aux assurés. L'annexe 2 regroupe les différents modèles du système adoptés par les différents pays cités.

Bien que le système Bonus-Malus aboutisse à séparer les mauvais risques des bons, il ne conduit pas à une segmentation plus fine. Henriet et Rocher (1986) ont montré qu'en fait, la prime doit dépendre, non seulement, des sinistres passés, mais aussi de leur répartition dans le temps. Les sinistres récents doivent être plus pénalisants que les anciens ; Les deux auteurs supposent déjà, que le comportement de l'individu peut modifier son niveau de risque [13].

L'assureur, donc, outre les renseignements demandés lors de la souscription du contrat d'assurance, doit suivre le comportement de son assuré au niveau de sa sinistralité.

Le comportement du conducteur est expliqué par deux hypothèses ainsi admises par Vasechko et Grun-Rehomme (2012): l'existence des différents degrés d'aversion au risque et l'existence des différents niveaux de risque [16].

A cet effet et face à la situation d'aversion au risque, l'assureur propose différentes formules de contrats d'assurances automobiles détaillées ci-après afin de cibler les différents segments de marché ayant chacun un degré d'aversion au risque automobile et plus particulièrement les bas risques.

c. Formules du contrat d'assurance automobile conventionnelle

En général, trois formules d'assurance automobile tirées des approches initiées par Rothschild & Stiglitz se distinguent en fonction des différentes garanties ainsi exposées dans le tableau 2:

Tableau 2 Formules conventionnelles d'assurance automobile [17]

Responsabilité civile (Au tiers)	Responsabilité civile étendue	Tous Risques
Garantie de base RC	Garantie de base RC	Garantie de base RC
Défense et recours	Défense et recours	Défense et recours
Protection juridique	Protection juridique	Protection juridique
Assistance	Assistance	Assistance
	Incendie + Vol	Incendie + Vol
	Bris de glace	Bris de glace
	Evénements climatiques	

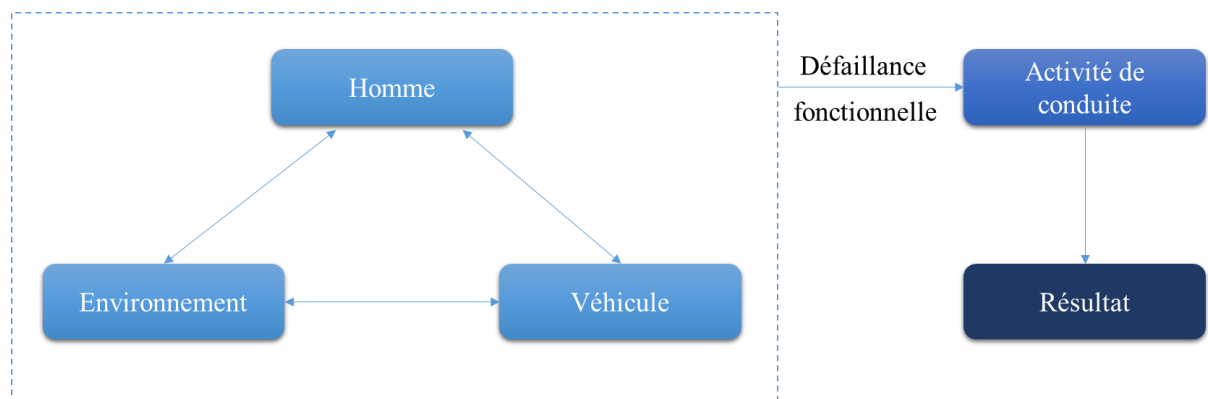
	Collision (tiers identifié)	Evénements climatiques Collision (avec ou sans tiers)
--	-----------------------------	--

En effet, le caractère obligatoire universel appliqué à tout conducteur et lié à la responsabilité civile de l'assuré a pour avantage de réduire l'effet de la sélection adverse dans le sens où en l'absence d'obligation, seuls les hauts risques s'assureraient [16].

1.2. Risques automobiles

Comme cité ci-dessus, le risque automobile est la probabilité qu'un véhicule soit impliqué dans un accident de circulation causant des dommages corporels et/ou matériels à autrui ou subis par le conducteur ou par son véhicule.

Van Elslande *et al* (1997) ont étudié les scénarios type de production de l'erreur humaine dans l'accident de circulation et ont déduit que les accidents sont le résultat de l'interaction entre les différents éléments du système suivant [18] :



- L'Homme : l'expérience de conduite, les motifs (distraction, inattention, fatigue...), les traits de personnalité, les habitudes...
- Le véhicule : l'ensemble de paramètres techniques comme les amortisseurs, les pneumatiques, le freinage, la direction...
- L'environnement : tout facteur autre qu'humain et technique comme la visibilité, la lisibilité d'une signalisation, la qualité des infrastructures...

Cet ensemble de facteurs aboutit à la production de la défaillance fonctionnelle qui amène à une activité de conduite particulière au conducteur.

- Activité de conduite : elle diffère d'un conducteur à un autre selon sa perception de la situation, sa connaissance des règles de sécurité, sa capacité de réaction et d'exécution d'une manœuvre d'urgence.

Cette activité de conduite avec le degré d'ajustement de la conduite à la situation aboutissent à la réalisation ou non d'accident qu'est le résultat.

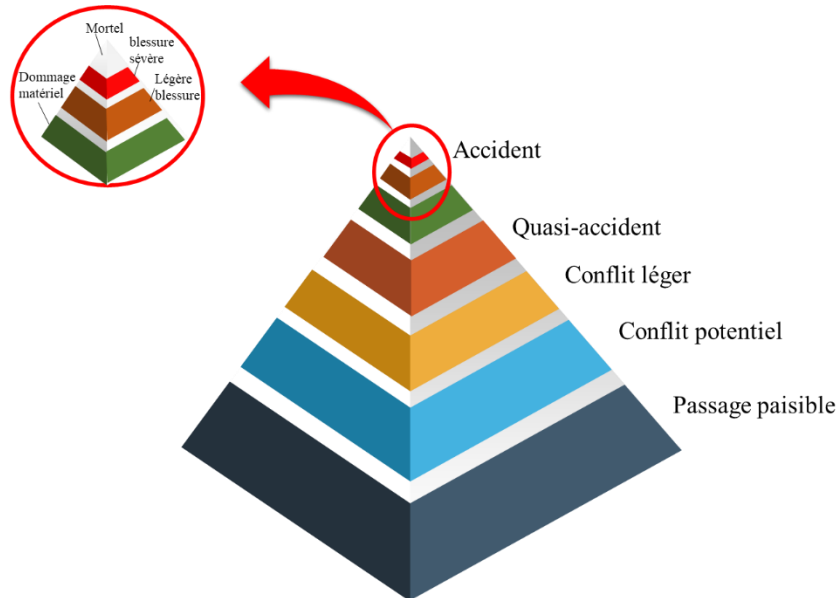


Figure 1 Continuum des événements de circulation [19]

Le graphique 1 présente la série d'évènements de circulation qui précèdent l'accident telle que décrite par Hydén (1987) [20]:

- Le passage paisible réfère aux conditions du déplacement à savoir l'itinéraire choisi, le motif du déplacement etc. où il n'y a aucun conflit⁷.
- Le conflit potentiel est le comportement à proximité du lieu du futur accident : c'est la situation de conduite normale et stable;
- Le conflit léger est le moment où une rupture survient (la défaillance fonctionnelle);
- Le quasi-accident est la situation où le conducteur essaye de revenir à la normale, en engageant une manœuvre d'urgence;
- L'accident est la nature et la configuration du choc (blessure/mort et/ou dommage matériel).

De ce fait, l'assurance automobile est assimilée à l'offre d'un triptyque [12]:

- Une protection juridique qui fournit au conducteur assuré la couverture des frais et conseils juridiques en cas d'engagement de sa responsabilité civile ou encore en cas de recours ;

⁷ Un conflit est la situation observable de rapprochement entre deux véhicules avec un risque de collision en l'absence de mesure d'évitement.

- Une indemnisation financière qui couvre les frais engagés en cas de sinistre résultant des dommages corporels causés au conducteur et/ou matériels causés au véhicule assuré ;
- Une assistance à l'assuré en cas d'accident, de panne, de vol ou de détérioration de son véhicule...

Ces trois éléments constituent le contrat d'assurance automobile souvent conçu selon les besoins de l'assuré afin de couvrir deux types de risque : les risques liés au véhicule et les risques liés au conducteur.

a. Risques liés au véhicule terrestre à moteur ⁸

Le tableau 3 ci-dessous retrace les principaux aléas frappant les véhicules et leurs différents motifs :

Tableau 3 Principaux aléas liés aux véhicules et leurs différents motifs[3]

Type d'aléa	Motif de survenance
Accident de circulation	Soudain, fortuit résultant d'une collision
Incendie	Technique, émeutes et mouvements populaires, vengeance, fraude à l'assurance ...
Vol	Joyriding ⁹ , crime organisé, fraude à l'assurance...

b. Risques liés au conducteur

Mis à part les facteurs inhérents au véhicule causant les accidents de circulation, les facteurs humains reflétant son comportement au volant comptent parmi les principaux facteurs de survenance de ces accidents. Evans (1996) a démontré que les facteurs inhérents au comportement du conducteur contribuent à la survenance de 95% des cas d'accidents [21].

Parmi les facteurs à risques relevés par l'OMS, on retrouve [22]:

- L'excès de vitesse ;
- La conduite en état d'ébriété ou sous l'influence de substances psychoactives ;
- Le non-respect ou l'absence de dispositions de sécurité (casque, ceinture de sécurité siège-auto pour enfants...) ;
- La distraction au volant en raison de l'usage du téléphone portable ;
- Une infrastructure routière dangereuse ;

⁸ VMT : un véhicule doté d'un dispositif, servant à transformer une énergie quelconque en une énergie mécanique (Robert), lui permettant de circuler sur le sol sans être lié à une voie ferrée. Les VTM sont les automobiles et deux-roues, les engins de chantier, les machines agricoles, les chariots élévateurs, les remorques et semi-remorques attelées à un VTM

⁹ Joyriding : vol pour l'amusement effectué par les jeunes

- Le non-respect du code de la route

Il est clair que d'ores et déjà, le comportement du conducteur est l'un des facteurs les plus importants de survenance d'un accident. Ceci affirme que la négligence des paramètres comportementaux dans l'évaluation du risque automobile conduit à un cercle vicieux de problèmes que l'assureur automobile essaie de traiter au quotidien.

1.3. Problèmes soulevés de l'assurance automobile conventionnelle

L'un des principaux problèmes de l'assureur auto est la *sous-tarification*. Une récente étude de Verisk Insurance Solutions¹⁰ a montré qu'au minimum 29 milliards \$ de primes ne sont pas collectées annuellement par les assureurs américains suite à l'omission ou la fausse déclaration des informations par les assurés lors de la souscription. Entre profitabilité et fidélisation des clients dans un marché très concurrentiel, l'assureur, face à ce dilemme, se trouve obligé de faire payer les bons assurés les conséquences de l'anti-sélection. Les principales sources de ce phénomène, selon la même étude, est la non reconnaissance des conducteurs (Changement du profil de risque de l'assuré ou de nouveaux conducteurs du même véhicule assuré...), la sous-estimation du kilométrage, les infractions de conduite et la non déclaration de sinistres, la non reconnaissance du changement d'usage du véhicule [23].

Un autre problème majeur est *la difficulté à évaluer les profils à risque* notamment les jeunes, les novices et les âgés. Selon le rapport d'Ingenie¹¹ sur les conducteurs jeunes britanniques âgés entre 17 et 25 ans, 1 sur cinq conducteurs a fait un accident durant les six premiers mois de conduite et 20% des accidents graves impliquent la même tranche d'âge de conducteur [24]. Afin de limiter les conséquences de l'anti-sélection, ce profil de risque subit une surprime « jeune conducteur » atteignant 6 fois la prime d'assurance. (La prime moyenne pour un jeune conducteur est de £1,997 contre £372 la prime moyenne du marché auto britannique) [25].

La prime extrêmement élevée généralement payée par les parents, pousse ces derniers à inclure leurs enfants dans leur police d'assurance. On trouve parfois même, que le véhicule assuré au nom des parents n'est utilisé que par les enfants. Selon Verisk, les parents ne font cette inclusion qu'après en moyenne 13,8 mois de conduite coûtant à l'assureur plus que \$1 000 de perte de prime par cas [23].

Un troisième problème de l'assurance automobile est *le vol des véhicules* notamment modernes généralement équipés des dispositifs sophistiqués de sécurité quasi-inviolables, ce qui conduit à une

¹⁰ Verisk Insurance solutions : cabinet américain de gestion de risques des assurances

¹¹ Ingenie : Assurance automobile connectée pour les jeunes conducteurs au Royaume Uni

augmentation de la violence avec le « carjacking¹² ». Ce phénomène amène à faire supporter l'assureur un coût supplémentaire d'éventuels dommages corporels ou psychologiques en plus du coût élevé de remboursement du véhicule volé.

Un autre problème concerne *le taux d'accident*. En effet, l'exposition au risque d'accident est plus forte avec l'augmentation du parc automobile mondial et du nombre de kilomètres parcourus [26]. En Chine, le nombre des voitures à usage privé augmente de 25% chaque année [27]. Aux Etats-Unis, 5 trillions de km sont parcourus en 2015 soit une augmentation de 3,5% par rapport à 2014 (l'évolution annuelle la plus élevée depuis 25ans) [27] suite à l'amélioration des conditions économiques et à la baisse des prix de pétrole.

La fréquence de sinistres n'est pas le seul souci de l'assureur automobile, leur *coût* également est une contrainte pour la performance de la branche qui est souvent déficitaire. Les sinistres corporels deviennent plus coûteux suite à deux facteurs ; l'inflation médicale dépassant le taux d'inflation dans plusieurs pays développés [26] d'une part, et le montant élevé des frais de gestion des dossiers dont la cadence de règlement peut atteindre 15 ans d'autre part. Les sinistres matériels sont plus coûteux vu l'augmentation du coût de réparation (MOD et pièces de rechange). Cela est dû principalement à l'inflation et aux nouvelles technologies incluses dans les véhicules modernes.

Enfin, *la fraude à l'assurance* est le souci majeur de l'assureur. Les coûts estimés de cette pratique en Europe est de 10% du montant total des sinistres. Ce pourcentage est d'autant plus élevé que le contrôle dans les assurances est faible. En Espagne, le coût de la fraude à l'assurance automobile atteint 22% du coût de sinistres [28]. Aux Etats Unis, 13 à 17% du montant de sinistres automobiles corporels est de source frauduleuse [29]. Au Royaume-Uni, 350 sinistres frauduleux ont été détectés chaque jour en 2014 soit une valeur estimée de £1,3 milliards. La fraude à l'assurance automobile représente 52% des cas [30].

Il est à remarquer que d'ores et déjà certains de ces problèmes pourraient être solutionnés grâce à une meilleure connaissance des comportements au volant, et à la disposition d'informations fiables, objectives et en temps-réel, sur les conducteurs et leurs véhicules. L'évolution technologique et l'innovation en matière d'assurance sont derrière l'apparition des nouvelles formules d'assurance automobile exposées en détail dans la deuxième section.

¹² Carjacking : piraterie routière ou encore vol de véhicule avec menaces ou violence sur le conducteur.

Section 2 Nouvelles formules d'assurance automobile : Assurance automobile comportementale

Dans cette deuxième section, sont exposées les différentes études de recherche effectuées sur l'introduction du comportement du conducteur dans la logique d'évaluation du risque automobile. Ensuite, sont présentées les différentes formules d'assurance automobile comportementale et plus précisément l'assurance PHYD, centre d'intérêt du présent mémoire.

En effet, l'assurance comportementale est connue sous plusieurs noms : Assurance connectée, assurance télématique ou les termes anglo-saxons Usage-Based Insurance (UBI), Pay How You Drive (PHYD) et Pay As You Drive (PAYD). La tarification du risque automobile en utilisant cette formule est basée sur de nouvelles informations méconnues auparavant par l'assureur tel que le comportement de conduite, le kilométrage et le moment et les lieux de circulation ...

2.1. Comportement au volant

Le profil de conduite est composé de deux éléments séparés ; styles de conduite et compétences de conduite [31] ou encore performance et comportement de conduite [32]. Les compétences de conduite comprennent les aptitudes cognitives et motrices¹³ susceptibles de s'améliorer avec l'apprentissage et l'expérience [31]. En plus de l'apprentissage, les compétences de conduite sont affectées par les capacités cognitives du conducteur mais ces dernières peuvent se dégrader comme dans le cas de la conduite sous un état alcoolique. Les styles de conduite concernent les habitudes individuelles– la manière de conduite choisie par un individu –. Ils se construisent après des années d'expérience mais le conducteur ne devient pas nécessairement prudent avec l'expérience [31]. La pratique et l'exposition plus élevée aux différentes situations du trafic améliorent les compétences certes, mais augmentent le contrôle subjectif de la conduite¹⁴ et la conduite habituelle avec de faibles marges de sécurité¹⁵ [33]. En effet, les aptitudes à conduire en sécurité (marges de sécurité, anticipation, balayage visuel) peuvent se détériorer en l'absence des feedbacks d'un mentor [34]. Mais, l'implication dans un accident n'a pas de lien direct avec la prise des leçons de conduite défensive car les conducteurs prennent généralement leurs propres marges d'erreur et les styles de conduite sont considérés comme l'ensemble des croyances, des traits de personnalité, des attitudes et des motifs [33].

¹³ Les connaissances en code de la route et en conduite et les réflexes

¹⁴ Contrôle subjectif de conduite est l'évaluation de ses propres comportements (sécuritaires et insécuritaires) de conduite

¹⁵ Marge de sécurité est l'ensemble des mesures de sécurité prises face à une situation de circulation afin d'éviter l'accident

Le graphique 2 ci-dessous ainsi établi par Lajunen et Özkan (2012) montre la relation entre la performance et le comportement de conduite d'une part et la probabilité d'erreurs, de prise de marges de sécurité et finalement de la probabilité d'accident d'autre part.

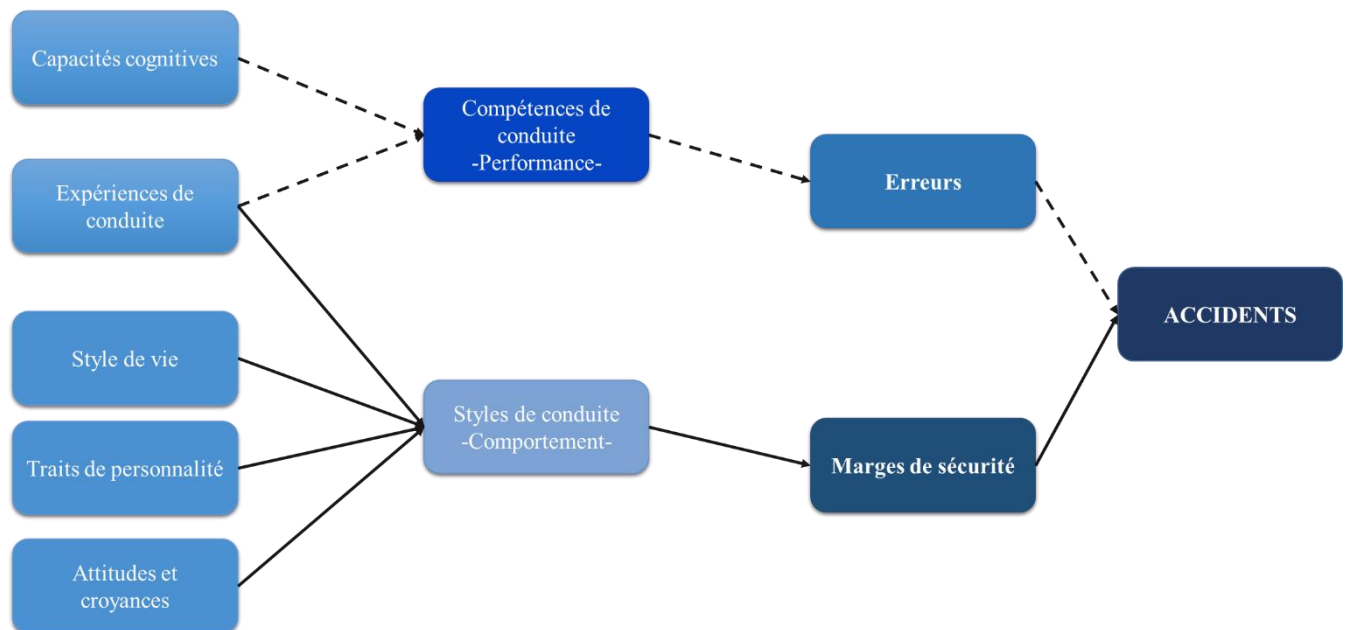


Figure 2 Les deux voies d'accident [35]

Selon ce modèle, les erreurs et les marges de sécurité sont considérées comme résultats de la performance et du comportement. La voie de la performance (flèches pointillées) montre que l'expérience en conduite¹⁶ est liée aux compétences et donc elle détermine la probabilité d'erreurs. La voie du comportement montre que l'expérience avec les traits de personnalité, les attitudes, les croyances et le style de vie sont liés au comportement de conduite ce qui influe sur la taille de la marge de sécurité. Plus le comportement du conducteur est risqué, plus faible est la marge de sécurité prise par le conducteur. Les erreurs de conduite fréquentes avec les faibles marges de sécurité intensifient le risque d'accident. En effet, les capacités maxima de performance d'un conducteur ne sont pas nécessairement en mesure de prédire l'implication dans un accident. De même, les mauvaises attitudes ne peuvent pas être la seule cause de perte de contrôle du véhicule. Les conducteurs peuvent déterminer les mesures de contrôle des différentes situations du trafic. Par contre, si la conduite est assez risquée, les capacités du conducteur deviennent insuffisantes pour contrôler la situation. Par conséquent, les contre-mesures efficaces doivent inclure au même temps, les compétences et les éléments du style de conduite [35].

¹⁶ Expérience de conduite désigne la pratique et l'exposition aux différentes situations de circulation

Ce qui intéresse les assureurs automobiles est la sécurité des conducteurs, le taux d'accident, les sinistres potentiels et les comportements au volant [36]. Par conséquent, la question dont la littérature cherche la réponse est comment un conducteur est jugé « prudent » et comment la performance de sa conduite est évaluée pour prévenir les sinistres futurs et améliorer la sécurité routière [37].

D'après une étude de Svenston (1981) sur la conduite des étudiants américains, l'excès de confiance de ces jeunes conducteurs a amené à proposer plusieurs méthodes pour encourager le comportement au volant « prudent » [38]. Parmi ces méthodes, on trouve le contrôle parental à travers les dispositifs électroniques embarqués dans les véhicules (Road Safety International, 2008) [39]. Ces dispositifs présentés en détail plus loin sont utilisés par plusieurs chercheurs et praticiens notamment Lotan et Toledo (2005) pour identifier les différentes manœuvres de conduite et classer les conducteurs en prudents, non prudents ou dangereux [37]. Plusieurs compagnies d'assurance ont demandé aux clients d'installer volontairement les dispositifs embarqués dans leurs véhicules pour enregistrer leur conduite quotidienne [40]. D'après Jensen et Wagner (2011), la conduite quotidienne est une bonne méthode pour mesurer et classer la performance au volant. Le fait de savoir que les conducteurs sont surveillés, leur comportement se diffère. Ainsi, les comportements anormaux des conducteurs peuvent se réduire. Lehmann (1986) affirme que l'analyse objective des données collectées des dispositifs embarqués détaillées plus loin semble une solution prometteuse d'amélioration de la sécurité [41].

2.2. Caractérisation de profils de conduite

La littérature économique a largement étudié la corrélation du comportement du conducteur avec les accidents de circulation par une panoplie de méthodes : enquêtes, questionnaires d'auto-évaluation et modélisation statistique. En présence des dispositifs télématiques embarqués dans les véhicules, la modélisation statistique s'avère la meilleure méthode pour mesurer le comportement du conducteur et son impact sur le risque automobile. Jensen et Wagner (2011) ont proposé la caractérisation de comportements des conducteurs illustrée dans le graphique 3 ci-dessous. Six classes de conducteurs ont été ressorties allant du timide à l'agressif. Un conducteur neutre et/ou conservateur est considéré ayant une conduite normale (zone cible de sécurité). Alors, les deux classes extrêmes de conducteurs (timide et agressif) peuvent refléter un comportement dangereux. De plus, il s'est avéré selon les deux auteurs que la conduite non prudente peut se produire même chez les conducteurs attentifs et confiants. Par exemple, un conducteur confiant peut avoir une conduite non prudente du fait de son comportement comme le talonnage, l'excès de vitesse ou le

changement rapide de voie. Par analogie, la conduite attentive peut être classée non prudente en raison d'une vitesse de conduite inférieure à celle du trafic (par souci de maintenir la limite de vitesse légale) ou de l'excès de balayage avant le changement de voie. Le comportement général des conducteurs est normalement distribué avec des pourcentages basés sur le nombre des écarts-type par rapport à la moyenne. La zone cible varie entre $\pm\sigma_1 = 0.67$ par rapport à la moyenne, μ . Par contre, les zones non prudentes varient entre $\pm\sigma_1$ et $\pm\sigma_2 = 1.64$. Un conducteur dépassant $|\sigma_2|$ sera classé dangereux [36].

Figure 3 Caractérisation des comportements au volant [36]

% de population	5%	20%	50%		20%	5%
Niveau de sécurité	Dangereux	Non-prudent	Zone cible		Non-prudent	Dangereux
Comportement	• Très lent • Non confiant • Déranger le flux de circulation • Imprévisible	• Respecter souvent les limites de vitesse indépendamment du flux de circulation • Excès de Balayage • Non instinctif	• Adapter la vitesse en fonction de la circulation • Prévisible • Confiant • Technique de balayage correcte		• Vitesse habituelle • Talonnage • Insuffisance de balayage	• Excès de vitesse • Excès de confiance • Déranger le flux de circulation • Imprévisible
Classe du conducteur	Timide	Attentif	Conservateur	Neutre	Confiant	Agressif
Statistiques	$-\sigma_2$		$-\sigma_1$	μ	$+\sigma_1$	$+\sigma_2$

Weidner *et al* (2016) ont montré qu'un style de conduite spécifique à chaque assuré peut émaner d'un profil de conduite basé sur la vitesse, l'accélération et le freinage. De plus, ils ont mis en évidence qu'un même véhicule peut être conduit de deux styles différents au minimum [42]. Comme est précisé plus haut, le style de conduite est un ensemble de traits de personnalité, de attitudes et de styles de vie. C'est-à-dire, le contexte de la situation du trafic affecte le style de conduite d'un même conducteur (par exemple, un conducteur devient plus nerveux durant les heures de pointe donc plus accidentable ou un jeune conducteur en présence de ses parents devient plus prudent). Biecheler et Moget (1985) ont introduit le concept du comportement de base d'un conducteur qui désigne le style habituel de conduite en appliquant préférentiellement son propre système de règles de conduite. Selon les mêmes auteurs, le degré de dangerosité de la conduite diffère selon le type de comportement de base qu'adopte un conducteur [43]. Toute déviation au comportement de base conduit à l'observation d'un comportement occasionnel qui dépend de la situation. De ce fait, le profil de conducteur peut être représenté en différents profils de conduite (agressif, confiant, attentif, neutre...) où le profil dominant est le comportement de base.

L'objectif final de l'assureur à travers le modèle de l'UBI est d'affiner la tarification sur la base du comportement de chaque assuré. Sans aller jusqu'à l'individualisation, l'assureur arrivera à attirer

les bons conducteurs et à encourager les mauvais conducteurs à améliorer le style de conduite et donc à mieux maîtriser le risque automobile. Cela devient possible avec la technique de la télématique automobile expliquée dans la sous-section suivante.

2.3. Technique de la télématique automobile

a. Définition de la télématique

Le terme « télématique » provient de la combinaison de télécommunication et informatique et la fusion des deux ont contribué à la naissance de la télématique. La télématique automobile est alors l'utilisation de la technologie d'information et de communication dans les véhicules[44]. Elle est basée sur la communication M2M (machine-to-machine) et représente l'échange de données entre les dispositifs à distance utilisant un réseau de communication avec câble et / ou sans fil pour la télémétrie et le contrôle à distance. La disponibilité du traitement informatique puissant et de la connexion réseau dans les véhicules ont conduit à l'évolution des applications et de services disponibles aux conducteurs [45] à savoir:

- Navigation GPS en temps réel
- Assistance routière,
- Suivi des véhicules,
- Gestion de flotte automobile¹⁷,
- Assurance automobile (UBI).

b. Moyens de fonctionnement de la télématique

Il existe actuellement quatre catégories de technologie UBI offerts sur le marché[46]:

- **Dongle OBD¹⁸** : C'est le premier système embarqué dans la voiture connectée. L'assuré installe lui-même l'appareil dans le port diagnostic situé dans l'habitacle du véhicule [46].
- **Black Box** : Un boîtier connu sous le nom « Black box » monté à bord du véhicule d'une façon permanente et installé par un professionnel. Il s'agit de la technologie la plus sécurisée [46]. Cet appareil contient un *GPS* pour mesurer la distance parcourue, la vitesse dépassée selon le type de routes fréquentées et le moment de la journée, un *accéléromètre* pour enregistrer la célérité et le freinage¹⁹, un *gyroscope* pour mesurer la qualité de prise de virage

¹⁷ Flotte automobile : un parc automobile rassemblant plusieurs véhicules d'une entreprise

¹⁸ L'OBD ou encore On-Board Diagnostics est le système informatique du véhicule. Il informe le conducteur sur l'état du véhicule via les alertes affichées dans le tableau de bord après contrôle des différents composants à travers leurs capteurs spécifiques.

¹⁹ Célérité : grande variation de vitesse ou d'accélération

et une *carte Sim* [47] pour stocker la data enregistrée et l'envoyer à l'assureur ou au fournisseur du service télématique²⁰ pour l'analyser et calculer le score de conduite [48].

- **Embedded ou intégré** : En 2013, 11 constructeurs automobiles offrent l'équipement de la télématique intégrée fournissant les services UBI. Avant, la télématique intégrée n'offrait que certains services comme la navigation et le diagnostic à distance.
- **Smartphone** : Il s'agit de l'outil le plus récent de la télématique. Les smartphones, appareils autonomes ou liés au système du véhicule, transmettent les informations vers et depuis le véhicule connecté [46]. Il devient alors, un capteur en lui-même.

Le tableau 4 ci-après présente la comparaison entre les différentes catégories de la technologie de la télématique offertes sur le marché.

Tableau 4 Comparaison des catégories de la technologie de la télématique [49]

	Véhicule (Embedded)	Boîtier lié (Black box/ Dongle)	Smartphone
Capteurs	Intégrés dès la construction dans le véhicule et pilotés par l'ordinateur à bord	Dédiés si installation dans le boîtier ou réutilisation des données émises par le port OBD	Intégrés dans le smartphone
Autonomie	Batterie du véhicule	Alimentation par le véhicule (OBD assure l'alimentation électrique). Des solutions sur batterie existent aussi.	Batterie du téléphone
Transmission des données	Stockage pour partage uniquement via le port OBD (diagnostic interprétable par un équipement dédié) ou une connexion data GSM.	Temps réel ou début/ fin d'usage du véhicule, via une connexion GSM autonome ou le smartphone du conducteur.	Temps réel et/ou déchargement asynchrone du smartphone, empruntant la connexion data GSM
Traitement et analyse	Temps réel pour l'interprétation par l'ordinateur de bord, asynchrone pour le gestionnaire de flotte.	Variable et multiusage : conducteur (via smartphone) pour action immédiate+ gestionnaire	Temps réel, souvent exploitation partielle via une application sur smartphone
Limites	- Combinaison de capteurs spécifique à chaque	- Coût du boîtier assez élevé souvent imputé sur	- Absence d'identification

²⁰ Intermédiaire entre l'assuré et l'assureur

	véhicule : toute solution industrielle de récupération des données requiert des accords bilatéraux avec tous les constructeurs - Exploitation et transmission à la main du constructeur.	la prime de risque à la souscription ; l'installation ou la maintenance peuvent nécessiter l'intervention professionnelle.	propriétaire du smartphone (conducteur ou passager) - Pas de reconnaissance du véhicule - Capteurs peu précis et partagés
--	---	---	---

c. Fonctionnement de l'assurance télématique

Un aperçu sur l'architecture « de bout-en-bout » du système télématique est exposé dans le graphique 4. Un listing des données du modèle UBI et le processus de tarification seront représentés en détail par la suite [45].

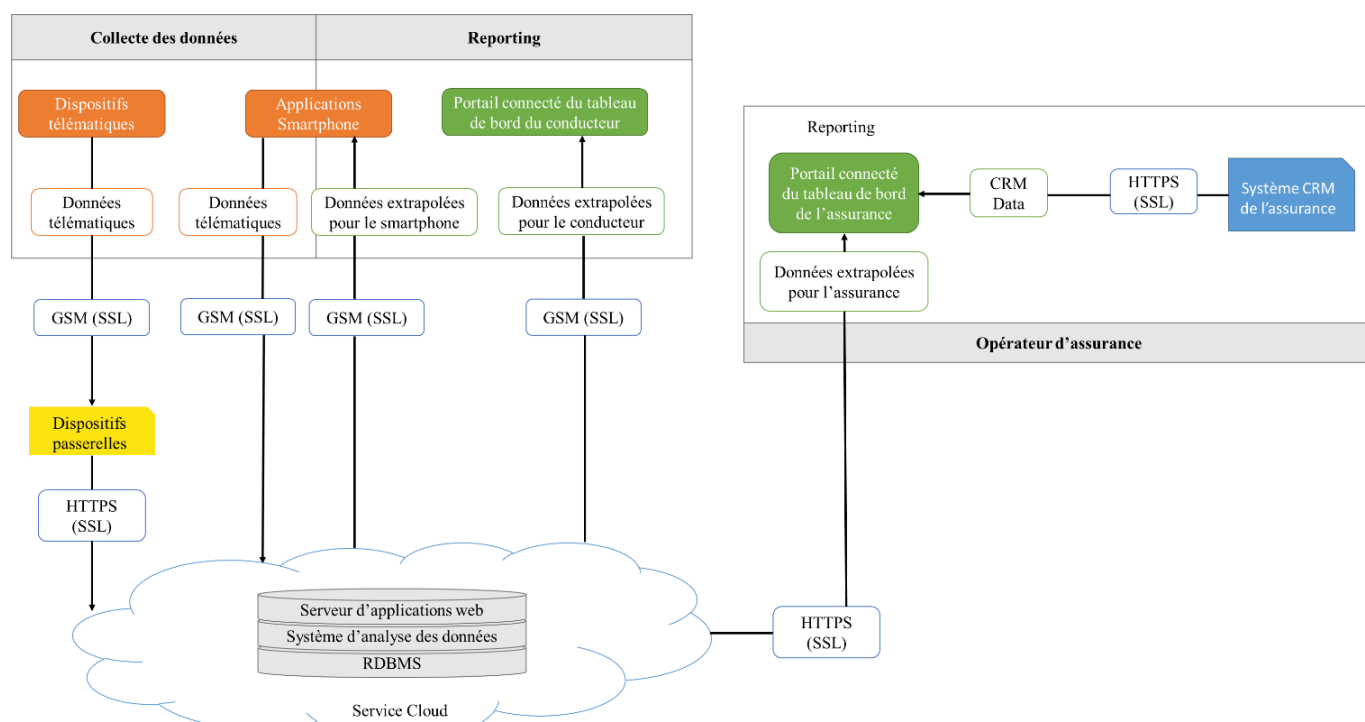


Figure 4 Aperçu de l'architecture technique du système télématique [45]

Les données sont enregistrées par un dispositif télématique dédié et sont utilisées par :

- L'assureur ou le fournisseur du service télématique : pour calculer le score de conduite et offrir la prime de risque correspondante au profil de risque ;
- Le conducteur : pour suivre son score de conduite et recevoir le coaching et les feedbacks de son assureur [50]

La majorité des données sont extraites de la base de données brute acquise par le dispositif télématique. Le tableau 5 expose un exemple d'un ensemble de données brutes collectées et traitées pour le besoin d'une compagnie d'assurance dans l'un des projets implantés à l'Europe centrale et orientale.

Tableau 5 Un ensemble de Données brutes collectées d'un dispositif télématique [45]

Données	Données brutes	Description
Emplacement du véhicule (Point GPS)	Latitude	Latitude globale du véhicule
	Longitude	Longitude globale du véhicule
	Horodatage	Temps en millisecondes lors de l'enregistrement du point GPS
	Vitesse	Vitesse du véhicule basée sur les deux derniers points GPS
Forces excessives exercées au véhicule	Force Axe -X	Force supérieure à $0.2G^{21}$ exercée perpendiculairement à la longueur du véhicule
	Force Axe -Y	Force supérieure à $0.2G$ exercée parallèlement à la longueur du véhicule

Une fois les données brutes collectées, leur extraction diffère d'une compagnie d'assurance à une autre selon ses besoins. Une liste non exhaustive des données communément extraites est retracée dans le tableau 6.

Tableau 6 Liste non exhaustive des données communément extraites [45]

Données extraites	Description
Liste de tous les trajets réalisés	Liste de tous les trajets parcourus par le véhicule avec leurs itinéraires
Durée du trajet	Durée de chaque trajet individuel
Distance du trajet	Distance de chaque trajet individuel
Début du trajet	Début de chaque trajet individuel
Fin du trajet	Fin de chaque trajet individuel
Vitesse moyenne	Vitesse moyenne de chaque trajet individuel
Vitesse maximale	Vitesse maximale de chaque trajet individuel
Style de conduite ²²	Score combiné basé sur deux niveaux d'évènements de freinage, d'accélération et de virage
Itinéraire routier	Itinéraire détaillé de chaque trajet individuel
Durée totale	Durée totale de tous les trajets enregistrés
Distance totale	Distance totale de tous les trajets enregistrés

²¹ $0.2\text{ g} = 2\text{ m/s}^2$; m/s^2 est l'unité d'accélération

²² Les données de la force-g (y compris les données extraites d'elles) ne sont disponibles que si le dispositif télématique est équipé d'un accéléromètre à trois axes.

Moment de la journée	Table de stockage d'heures quotidiennes de conduite
Longs trajets	Pourcentage de trajets sans pause dont la durée dépasse le seuil prédéfini par l'assureur
Temps d'inactivité	Durée totale passée en circulation sans mouvement (embouteillage, congestion routière)

Rappelant que l'objectif de l'assurance télématique est de procéder à une tarification dynamique et adaptée à chaque profil de risque évalué selon le comportement au volant (style de conduite). Pour une évaluation fiable, l'assureur doit veiller à la précision des données du GPS. Toutefois, la qualité de ces données est sujette aux circonstances environnementales²³ qui peuvent biaiser l'évaluation. Pour cela, l'assureur doit tenir compte de la contextualisation des données à travers les algorithmes avancés des données GPS afin d'éliminer tout biais possible [45].

Outre la veille à la qualité des données collectées, l'assureur vise à renforcer le système CRM afin de fidéliser ses clients à travers l'intégration des données de la télématique avec les données existantes dans le système CRM. Un minimum de données est requis pour construire un modèle de tarification performant [45] :

- Données liées à l'assuré : nom et prénom, e-mail, âge
- Données liées au véhicule : marque, modèle, année, Immatriculation, usage, puissance fiscale...
- Données liées à la police d'assurance : type de police, date d'expiration

La prochaine sous-section servira à exposer les résultats de la littérature sur la qualité de prédiction du risque automobile par les variables exposées plus haut.

2.4. Variables de conception d'un modèle d'assurance comportementale

Comme cité auparavant, les données télématiques nouvellement introduites dans le schéma de tarification du risque automobile transforment toute la logique d'estimation du risque à travers le « profilage » des conducteurs. Il est alors utile de dégager les variables prédictives d'un comportement de conduite tels que la littérature expose, fortement corrélées au taux d'accident. Plusieurs chercheurs jusqu'à présent ont montré que le kilométrage parcouru est l'un des facteurs les plus prédictifs du risque d'accident [51]. Jun *et al* (2011) ont proposé que les facteurs de conduite comme la vitesse et l'accélération puissent avoir une relation directe avec les accidents de circulation [52]. Paefgen *et al* (2013) ont combiné les différents moments de conduite (jour/nuit,

²³ Effets climatiques

jour de la semaine), le type de la route (autoroute, urbaine, extra-urbaine), la vitesse moyenne et le kilométrage comme des facteurs d'impact sur l'assurance automobile[53].

Le tableau suivant résume les variables de conduite utilisées pour la modélisation d'assurance automobile comportementale ainsi ressorties par la littérature [54] :

Tableau 7 Variables prédictives du risque d'accident selon le modèle UBI

Catégories	Variables	Référence
Informations basiques	Age [55], [56]	Rhodes et Pivik (2011) Yannis et al. (2005)
	Sexe [55]	Rhodes et Pivik (2011)
	Salaire [56]	Yannis et al. (2005)
	Niveau d'éducation [56]	Yannis et al. (2005)
	Années de conduite (expérience) [56]	Yannis et al. (2005)
	Infractions routières [52]	Jun et al. (2011)
	Statut familial [56]	Yannis et al. (2005)
	Statut physique [55]	Rhodes et Pivik (2011)
Informations sur le véhicule	Durée de circulation par véhicule [57]	Beirao et Cabral (2007)
	Valeur à neuf [57]	Beirao et Cabral (2007)
	Usage (privé/utilitaire) [57]	Beirao et Cabral (2007)
Informations géospatiales	Type de route [58], [53]	Jun et al. (2007) Paefgen et al. (2013)
	Localisation d'un accident à proximité [58], [53]	Jun et al. (2007) Paefgen et al. (2013)
	Type d'objet de collision (véhicule, moto, piéton...) [59]	Zheng et al. (2014)
	Type de collision ²⁴ [59]	Zheng et al. (2014)
	Facteurs déclenchants ²⁵ [59]	Zheng et al. (2014)
Informations télématiques	Distance parcourue [58]	Jun et al. (2007)
	Moment de la journée [58], [53]	Jun et al. (2007) Paefgen et al. (2013)
	Vitesse moyenne [58], [53]	Jun et al. (2007)

²⁴ Collision par l'arrière, collision frontale, collision latérale...

²⁵ Feux de circulation, réduction de voie, changement de voie, évitement de collision...

		Paefgen et al. (2013)
	Excès de vitesse [58], [53]	Jun et al. (2007) Paefgen et al. (2013)
	Accélération rapide [52]	Jun et al. (2011)
	Freinage brutal [59]	Zheng et al. (2014)
	Virage brusque [60]	Quddus et al. (2002)
Variables liés au conducteur	Appétence au risque [61]	Guo et Fang (2013)
	Agressivité [61]	Guo et Fang (2013)
	Capacité de prédiction de collision [61]	Guo et Fang (2013)
	Technologie de conduite assistée [61]	Guo et Fang (2013)
	Prudence [61]	Guo et Fang (2013)
	Sensibilisation à la sécurité [61]	Guo et Fang (2013)
	Plaisir de conduire [61]	Guo et Fang (2013)
	Distracted au volant [62]	Gupta et al. (2016)

Il existe 5 grandes catégories de variables directement ou indirectement liées au risque d'accident. Les informations basiques et celles liées au véhicule sont essentielles pour la prédiction du risque d'accident. Les informations géo spatiales sont largement discutées dans la littérature mais les variables comportementales collectées d'un des dispositifs télématiques cités plus haut sont au centre d'intérêt de ce présent mémoire. Elles peuvent refléter les choix opérationnels de conduite du véhicule prise par le conducteur en temps réel. Ces choix sont directement liés à la probabilité d'être impliqué dans un accident [54]. La variable *Distance parcourue* est considérée comme l'un des facteurs de risque les plus importants dans l'évaluation du risque automobile et son influence est largement discutée par plusieurs chercheurs ; certains considèrent que la relation est linéaire. Par contre, d'autres ont montré la non linéarité de la relation, notamment Boucher *et al* (2013). L'étude de Progressive Insurance en 2005 a trouvé une relation linéaire entre la distance parcourue et le nombre de sinistres[63] . Au même temps, les travaux de Paefgen *et al* (2014) ont mis en évidence la non-linéarité de la relation entre la distance et le risque d'accident en introduisant l'association entre la variable distance comme une mesure de « l'étendu » de l'exposition au risque et plusieurs autres variables mesurant « le degré » de l'exposition telles que le moment de la journée, le jour de semaine, le type de route, la vitesse moyenne. Litman (2005) a trouvé une corrélation positive entre la distance parcourue et le risque d'accident dans le sens où plus le véhicule est utilisé, plus la probabilité d'être impliqué dans un accident est forte [64].

Les variables [distance parcourue, moments de circulation, vitesse moyenne] jouent un rôle très important dans la prédiction du niveau du risque du conducteur certes, mais elles deviennent plus prédictives en présence d'autres facteurs comportementaux (freinage brutal, accélération rapide...) [65]. En effet, Jun *et al* (2011) ont observé et déduit que les conducteurs sinistrés ont la tendance à rouler à une très grande vitesse sans respect des signes de limite de vitesse. L'accélération rapide, le freinage brutal et les virages brusques fréquents peuvent augmenter la fréquence d'accident et l'intensité du dommage [60]. D'ailleurs, Klauer *et al* (2009) ont montré que le freinage excessif et l'excès de vitesse sont des comportements au volant assez risqués qu'ils sont susceptibles d'augmenter le risque d'implication dans un accident plus que la conduite normale. Ils ont également mis en évidence que dans des conditions normales de conduite (absence de situation de conflit), la fréquence du freinage, virage, et accélération sévères est beaucoup plus importante chez les conducteurs risqués que chez les conducteurs modérés ou prudents ce qui fait augmenter leur risque d'être impliqué dans un accident ou un quasi-accident²⁶. L'excès de vitesse a un impact néfaste sur la sécurité du conducteur suite à la réduction de son champ visuel d'une part et de la stabilité du véhicule d'autre part. Jokschi *et al* (1993) tiennent à montrer que l'excès de vitesse est le facteur le plus important de la gravité des accidents. Davis *et al* (2006) ont montré que le risque d'accident a clairement tendance à augmenter avec l'augmentation de la vitesse.

Ayuso *et al* (2017) ont étudié la performance du modèle d'assurance comportementale en comparant les résultats du modèle conventionnel d'estimation de sinistres futurs aux résultats du modèle incluant les variables comportementales du conducteur. Il est montré que l'estimation s'améliore en présence des variables télématiques. La distance totale parcourue, l'excès de vitesse et la conduite dans les zones urbaines ont une forte corrélation positive avec le nombre de sinistres estimés. Pour aller plus loin, la conduite durant la nuit fait augmenter le risque d'accident après un certain nombre de kilométrage [66]. Néanmoins, plus la durée du trajet parcouru augmente, plus le conducteur gagne en expérience et donc plus faible est la probabilité de survenance d'accident [67]. Langford *et al* (2013) ont montré que les conducteurs avec une distance parcourue élevée enregistrent moins de sinistres comparés aux conducteurs roulant peu [68]. En effet, cela est expliqué par plusieurs facteurs comportementaux. Par exemple, les conducteurs roulant peu, ont tendance à conduire –par manque de pratiques de conduite– dans les zones urbaines encombrées où le risque automobile est plus élevé [69]. Concernant l'effet des variables classiques, il s'est avéré qu'en introduisant les nouvelles variables télématiques, l'âge devient une variable plus significative ayant un effet non linéaire sur le nombre de sinistres estimé ; un modèle combiné montre qu'il existe une certaine

²⁶ Une notion à expliciter dans le deuxième chapitre

hétérogénéité au sein même d'un groupe d'âge. Le nombre de sinistres est négativement corrélé à l'expérience de conduite contrairement à l'âge du véhicule. La puissance du véhicule a un effet positif autant sur le modèle conventionnel que sur le modèle UBI. Par contre, le sexe et l'éducation deviennent moins prédictifs dans le modèle UBI [66],[70].

Les variables liées au conducteur sont des facteurs étroitement liés à la personnalité du conducteur. Guo et Fang (2013) ont trouvé que les traits de personnalité (nervosité, extraversion, agréabilité, conscience, appétence au risque) ont une relation significative avec le risque d'accident [61]. Klauer *et al* (2009) ont montré que la distraction est parmi les facteurs de risque les plus importants [71].

L'augmentation de nombre de variables explicatives du risque d'accident aura pour effet une segmentation très fine des assurés et une meilleure connaissance du risque. Comme montré auparavant, la classification des profils de conduite en fonction des variables télématiques a été étudiée dans la littérature. Par contre, la mesure de la performance de conduite généralement estimée par un score n'a pas fait l'objet d'assez d'études de recherche.

Liu *et al* (2017) ont procédé à la modélisation du score de comportement de conduite pour montrer son effet sur le niveau du risque automobile. Les variables du score de conduite choisies sont les suivantes [72] :

- Distance mensuelle totale
- Heure de pointe (jour de semaine)
- Durée de conduite à la nuit
- Durée de conduite dans les weekends
- Taux d'excès de vitesse
- Nombre de coups d'accélération rapides
- Nombre de freins secs
- Nombre de virages brusques
- Nombre d'infractions

Chaque variable est pondérée par un coefficient en utilisant la méthode EW-AHP (Entropy weight analytic hierarchy process). La mise en application de ce modèle construit sur un portefeuille d'assurés UBI a montré que plus élevé est le score de conduite, moins l'assuré est impliqué dans un accident. Par conséquent, le modèle de scoring du comportement au volant permet effectivement de refléter le niveau du risque de l'assuré ce qui permet à l'assureur de mieux anticiper le risque des nouveaux assurés à travers la classification des profils de conduite[72].

2.5. Principe de l'assurance comportementale

Avant la découverte de l'utilité de la télématique dans l'assurance automobile, seules les caractéristiques « conventionnelles » du conducteur et l'historique des sinistres pourraient servir à l'évaluation du risque automobile. C'est à partir de l'introduction des nouvelles informations via la télématique détaillées plus haut qu'un « profilage » des conducteurs est devenu une bonne approche pour mieux anticiper les phénomènes de survenance des accidents. Dorénavant, l'assureur se base sur les « profils de risque » pour estimer la prime de risque automobile. Le profil de risque est déterminé en fonction des paramètres fixés par l'assureur ce qui donne lieu à la naissance de plusieurs formes innovatrices de l'assurance automobile retracées dans le graphique 5. L'axe horizontal montre l'ensemble d'informations disponibles sur le véhicule et le style de conduite alors que l'axe vertical montre l'ensemble d'informations disponibles sur le conducteur [45] .

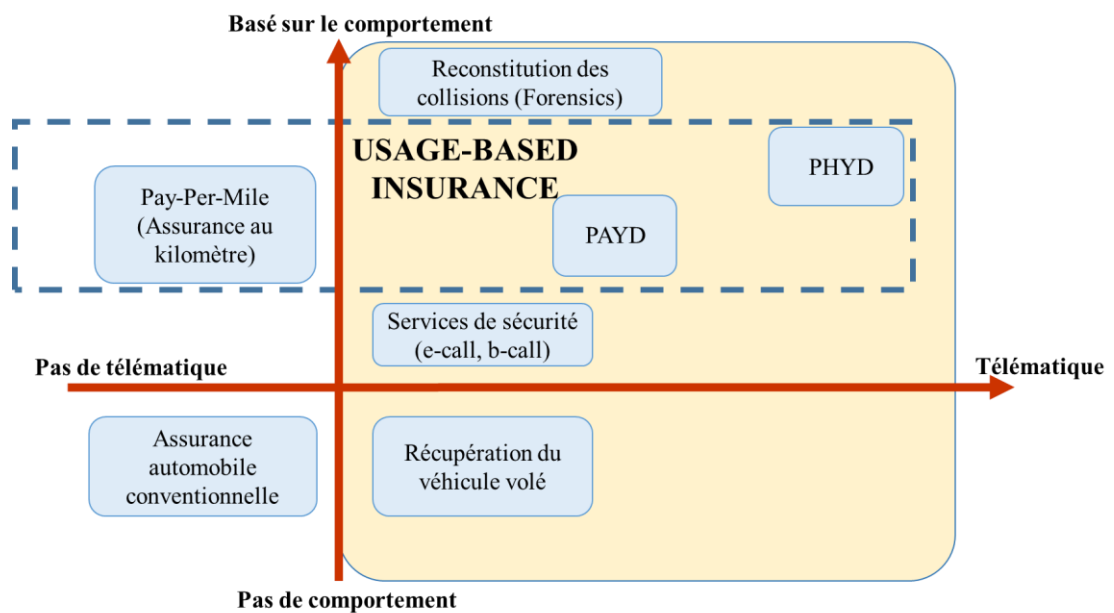


Figure 5 Les différents termes des nouvelles formules de l'assurance automobile [45]

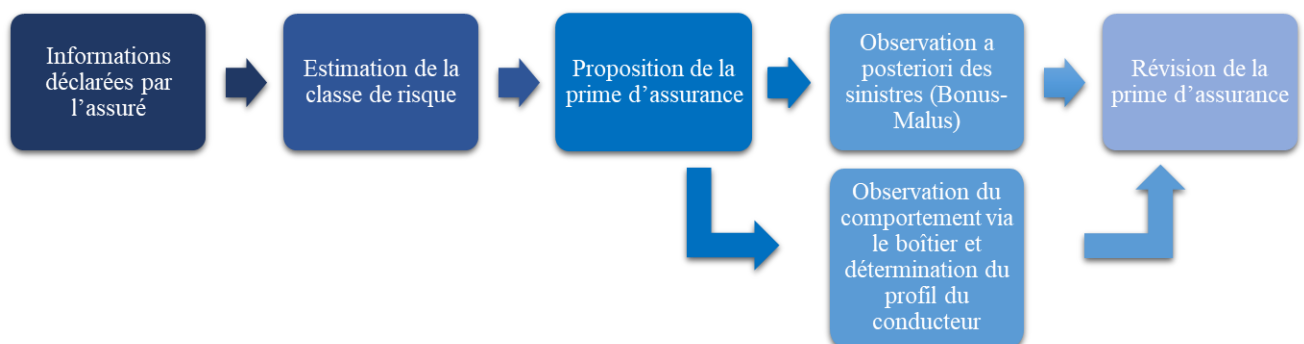
Les différents modèles de tarification de l'assurance comportementale sont comme suit:

- **Pay-Per-Mile (Assurance au kilomètre)** : la prime est basée sur un kilométrage forfaitaire annuel. Dépassant le chiffre fixé, la garantie cesse à moins de souscrire une tranche supplémentaire. La déclaration du kilométrage parcouru –affiché par le compteur de bord acquis par l'assuré– reste à sa volonté. Elle se fait d'abord à la souscription, puis à chaque échéance annuelle [3].

- *Pay-As-You-Drive (PAYD)*: la prime est proportionnée soit au temps de conduite, soit à la distance parcourue. Ces paramètres sont enregistrés au moyen d'un équipement de géolocalisation (GPS). D'autres paramètres peuvent être considérés par l'assureur dans la tarification tels que la vitesse et les lieux de circulation [3]. Toutes les données collectées sont transférées à l'assureur via le boîtier noir (Black Box) [44].
- *Pay-How-You-Drive (PHYD)*: La prime dépend de la qualité de conduite évaluée par la technologie GPS. Le concept de base est de construire un modèle de tarification basé sur le kilométrage, les lieux et la période²⁷ de circulation, et la qualité de conduite (excès de vitesse, accélération et freinage brusque, virage, usage de téléphone au volant). Toutes ses informations sont utiles à l'assureur pour évaluer le risque du conducteur. Ce dernier est mesuré par un score de conduite²⁸ qui sert à réviser la prime régulièrement selon des échéances prédéterminées ; Un conducteur prudent (score élevé) bénéficie plus de ristournes de prime qu'un conducteur moins prudent. PHYD est la plus mature des trois formules, offrant aux assureurs et aux clients de l'information plus détaillée. Le client n'a pas à payer uniquement une prime en fonction du nombre total de kilomètres couverts mais il subit aussi une prime plus élevée en présence des comportements susceptibles de faire augmenter l'exposition au risque. [44]

L'apparition de la télématique automobile dans l'industrie d'assurance est susceptible de bouleverser la logique conventionnelle de l'assurance automobile ainsi schématisée dans la première section. Il existe deux modes d'observation du comportement du conducteur : soit en amont pour les nouveaux souscripteurs soit en cours du contrat d'assurance [12].

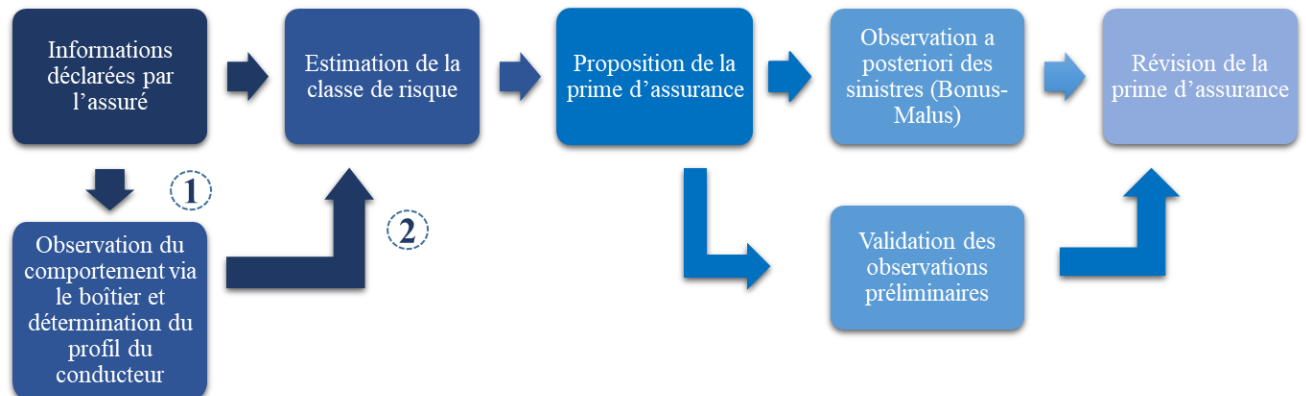
❖ La correction de la tarification a posteriori grâce aux informations collectées par le boîtier :



²⁷ Jour ou nuit/ Weekend ou au cours de la semaine

²⁸ Score de conduite : combinaison des paramètres du modèle PHYD pondérés par des coefficients

- ❖ La sélection fine des assurés en amont par l'instauration d'une période d'observation préalable à la signature du contrat définitif :



Le potentiel client s'inscrit dans le programme d'assurance comportementale pour être observé via la télématique durant une période donnée. Au bout de la période, l'assureur collecte toute l'information nécessaire sur le conducteur pour dégager « un profil de risque ». Selon le profil, l'assureur offre une réduction sur la prime de risque conventionnelle. Le taux de réduction offert peut être révisé à la hausse si les observations préliminaires s'améliorent mais aucune pénalité n'est appliquée dans le cas contraire. Néanmoins, le lien (1) entre les informations déclarées par l'assuré et le comportement et le lien (2) entre le comportement et l'estimation du risque ne sont possibles à établir que si tout le modèle de tarification est reconstruit par l'assureur dans le but d'introduire les nouveaux paramètres « comportementaux » permettant d'affiner la segmentation des assurés en amont.

Section 3 Apports de l'assurance comportementale

L'assurance connectée est susceptible de transformer tout le Business model de la branche. Il s'agit de la *disruption* digitale modifiant la manière de concevoir la sécurité routière mais également l'assurance automobile. Dans cette dernière section, les apports de l'assurance comportementale seront exposés dans un ordre conceptuel, scientifique et pratique après la présentation des différentes expériences en assurance connectée à l'étranger. Un bref aperçu sur les défis auxquels fait face l'assurance connectée sera présenté à la fin de la section.

3.1. Expériences à l'étranger

Depuis les années 1930, l'assurance automobile évoluait avec les profondes modifications de la fiscalité automobile. Pour plus d'équité entre les assurés, les assureurs se basaient sur la distance parcourue à la place de la puissance du véhicule pour la tarification. Une brève apparition de cette

formule en France en 1934 était reprise en 1949 et les expériences en ce sens se multipliaient à la fin des années 1980 partout dans le monde. Par contre, que des tentatives de mettre en place une tarification automobile basée sur la consommation d'essence étaient observées aux Etats-Unis dans la décennie 1990. En 1998, la première formule d'assurance à base de distance parcourue connue sous le nom « Pay As You Drive » est apparue avec la compagnie d'assurance américaine *Progressive* [3]. Depuis 2010, la disruption digitale a métamorphosé l'assurance automobile avec l'apparition des objets connectés et mobiles rendant l'information facilement accessible. Le développement de la formule « Pay As You drive » et l'apparition de la formule « Pay How You Drive » transforment la tarification automobile statique en une tarification dynamique à base de comportement.

Le marché de l'assurance automobile représente 42% du marché global de l'assurance non-vie en 2015 soit \$700 milliards [73]. Il est estimé à atteindre \$819 milliards en 2018 [74].

Actuellement, il existe plus que 30 millions de polices d'assurance à base d'usage actives dans le monde selon Ptolemus consulting. Le cabinet estime qu'elles dépassent 90 millions de polices à l'horizon 2020 dans le monde [75]. Les principaux assureurs actifs dans le marché d'assurance auto connectée sont les Etats-Unis, l'Italie et le Royaume Uni comme le montre le graphique ci-dessous [75].

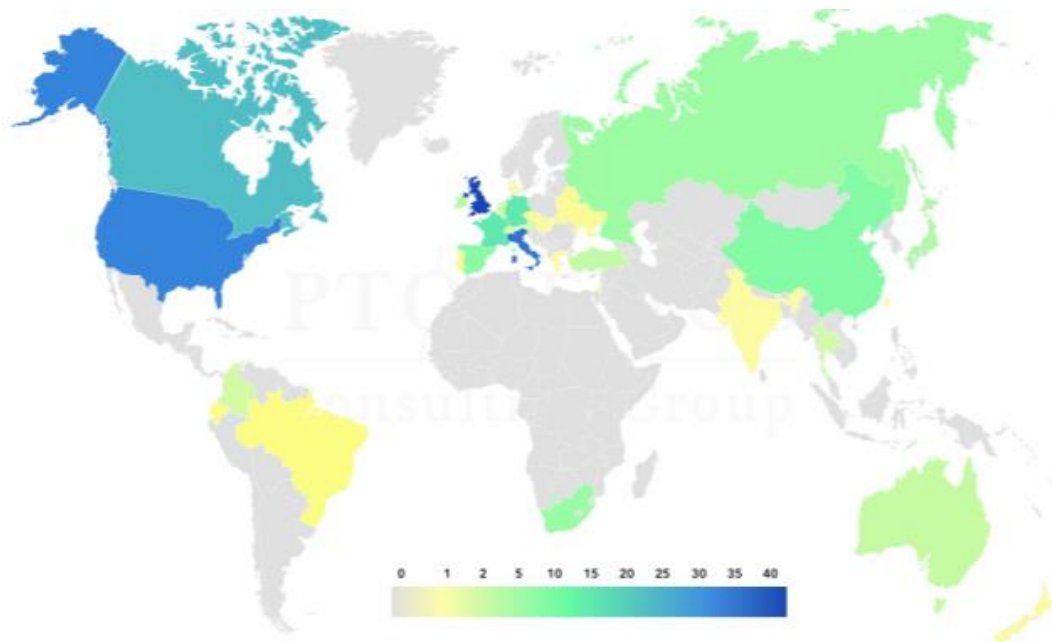


Figure 6 Nombre de programmes actifs d'assurance automobile connectée par pays [75]

Selon Swiss Re, il est estimé que 260 millions de véhicules connectés circuleraient dans le monde à l'horizon 2020. Dans la même étude, 2/3 des véhicules vendus dans le monde seront équipés de toute forme de système de connectivité [73].

Deloitte estime qu'à l'horizon 2020, le potentiel de part de marché global de l'assurance connectée est de 17% soit \$18 milliards avec deux pays très en avance ; L'Italie suivie par le Royaume-Uni. Ces deux pays ont un potentiel d'avoir une part de 27% et 23% respectivement de leur marché [76]. Il est à noter qu'en Italie, 4,5 millions de polices d'assurance connectée sont déjà souscrites contre 450 000 au Royaume-Uni. Malgré la faible pénétration du marché de l'assurance automobile connectée dans le monde, une croissance rapide est estimée pour plusieurs raisons expliquées par la suite. D'après une étude de SMA²⁹, les assureurs automobiles estiment une pénétration de l'assurance connectée de 36% du marché à l'horizon 2020. De plus, SAS Institute prévoit que les assureurs recevront plus que 25% de leurs chiffres d'affaires de l'assurance connectée [77] .

Les raisons d'essor du marché d'assurance automobile connectée sont résumées comme suit [78] :

- **Aux Etats-Unis** : Primes d'assurance Auto chères surtout pour les jeunes, réussite des campagnes publicitaires et de l'installation des OBD dans les véhicules.
- **Au Royaume-Uni** : Primes d'assurance Auto chères surtout pour les jeunes, nombre élevé de sinistres frauduleux très coûteux.
- **En Italie** : Primes d'assurance Auto chères, nombre élevé de sinistres frauduleux très coûteux, imposition réglementaire aux assureurs d'offrir l'assurance télématique avec des ristournes.
- **En Allemagne** : La télématique automobile est dominée par les assistants routiers offrant de l'assurance en tant que service additionnel.
- **L'Union Européenne** : Réglementation en vigueur sur la tarification unisexe pour l'assurance automobile, Nouvelles directives en 2018 du système e-call³⁰ et de règlement général sur la protection des données personnelles.
- **En Chine** : Plus de flexibilité réglementaire dans la tarification du risque automobile
- **Au Japon et en Chine**: les constructeurs automobiles développent la télématique intégrée dans les nouveaux modèles des véhicules japonais et chinois.
- **En Brésil** : Mise en place du Mandat « Contran 245 » obligeant les nouveaux véhicules d'intégrer la télématique pour lutter contre le vol.

²⁹ Strategy Meets Action

³⁰ E-Call : Service d'appel d'urgence automatique déclenché en cas d'accident de circulation par un système télématique obligatoirement installé dans tout véhicule circulant dans l'Europe.

- **En Afrique de Sud** : Implantation d'un marché de services de gestion de flottes et de récupération des véhicules volés suite à l'amplification du phénomène de vol des véhicules.

Le tableau 8 expose des exemples des principaux produits d'assurance connectée dans le monde.

Tableau 8 Exemples de produits d'assurance télématique [44]

Spécificités du produit	Axa	AVIVA	Progressive	StateFarm	Generali	AA
Nom du produit	Drive Save	Aviva Drive	Snapshot	OnStar/In Drive/ SYNC	Clear Box	Drive Safe
Pays	Irlande	Royaume-Uni	Etats-Unis	Etats-Unis	Italie	Royaume-Uni
Echelle de score de conduite	1-100	0-10 (10 le plus prudent)	n.a.	n.a.	n.a.	0-100
Paramètres de conduite	Position géographique, vitesse, accélération forcée et freinage brusque	Accélération, freinage et virage à vitesse élevée	Freinage, kilométrage, conduite entre minuit et 4h du matin	Distance parcourue, style de conduite	Distance parcourue	Moment de journée, vitesse, virage à vitesse élevée, accélération, freinage
Technologie	Smartphone	Smartphone	OBD-II Dongle	Black Box	Black Box	Black Box
Ristourne maximum	Jusqu'à 30%	10% si la prime entre £200-£400 20% si la prime > £400	n.a.	Jusqu'à 50%	Jusqu'à 18% (garantie RC) jusqu'à 15% (Tous Risques)	n.a.
Accès en temps réel	oui	non	non	oui	oui	oui
Intervalle de révision de prime	4 mois	n.a.	6 mois	6 mois	n.a.	3 mois
Pénalités en cas de faible score	non	non	non	non	n.a.	oui
Signal GPS requis	oui	oui	non	oui	oui	oui

Cependant et d'après les raisons précitées, l'innovation de l'assurance connectée dépasse la tarification individuelle pour aller jusqu'offrir des services à valeur ajoutée n'ayant pas de relation

directe avec le métier d'assurance tel que l'assistance. Une offre assez diversifiée permet au mieux de répondre aux besoins de ses clients et de les assister en temps réel.

Une panoplie de service peut être divisée en 3 catégories [76]:

- **Services en relation avec l'automobile:** assistance routière, notifications en cas de vol...
- **Services non liés à l'automobile:** offres promotionnelles géolocalisées...
- **Services liés au comportement de conduite:** informations sur les trajets et les dépenses en carburant, comparaison de la conduite avec les amis ...

Une enquête dirigée par Towers Watson aux Etats-Unis en 2013 a montré que 63% des participants sont prêts à payer un abonnement pour les services complémentaires proposés par l'assurance connectée. 80% de ces derniers ont le même niveau d'intérêt aux services suivants [79]:

- **Alerte en cas de vol et détection des véhicules volés :** Le système GPS dans le dispositif télématique permet de détecter l'emplacement du véhicule en cas de vol augmentant les chances de sa récupération.
- **Appel d'urgence automatique :** grâce au dispositif télématique, le choc sur le véhicule est détecté ce qui déclenche un appel automatique d'urgence fournissant la position GPS exacte pour être aidé ou dépanné plus rapidement.
- **Alertes d'état :** L'anticipation des pannes éventuelles à travers des alertes d'état et des rappels des échéances d'entretien du véhicule.

D'autres services offerts par l'assurance télématique notamment [79] :

- Conseils d'économie de carburant par l'identification des routes les plus efficaces
- Navigation GPS pour le choix du meilleur itinéraire routier
- Feed-back de conduite en temps réel
- Conseils de conduite prudente
- Carnet de bord dans lequel l'assuré trouve les enregistrements de ses déplacements en voiture
- Les notifications de géo-repérage: le géo-repérage (ou geo-fencing en anglais) est une technique de repérage par l'utilisation du GPS. Dès que le conducteur pénètre dans une zone donnée, une notification est envoyée immédiatement au smartphone connecté. Cette technique est utile pour les parents des jeunes conducteurs ou les enfants adultes des conducteurs âgés pour s'assurer que le conducteur reste dans les périmètres d'une zone particulière. Une fois la zone est dépassée, le véhicule déclenche une alerte dans le tableau de bord et une notification est envoyée au smartphone de la personne concernée. Le rôle de

l'assureur est d'aider cette dernière à mieux comprendre les circonstances de ces alertes via d'autres informations sur les conditions de circulation et l'entretien des routes [80].

- Gestion de flottes : la technique de géo-repérage est utile également pour les managers des flottes de véhicules à usage professionnel et de services de livraison. Elle permet d'identifier les routes les plus efficaces (qui accélèrent la conduite et réduisent le coût de carburant), de recevoir les alertes pour les livraisons tardives (avec la contextualisation de la situation) et de prévenir une activité non autorisée par le conducteur [80].

3.2. Bénéfices tridimensionnels

Les bénéfices tirés de l'assurance connectée peuvent être classés en fonction de l'impact :

- *Impact sur les revenus du marché* : Dispositif de la télématique automobile, Chiffre d'affaires des polices d'assurance PHYD/PAYD
- *Impact sur la qualité de service* : Polices d'assurance sur mesure répondant aux besoins spécifiques des conducteurs
- *Impact sur les coûts* : Réduction des coûts d'assurance au profit des conducteurs, réduction des coûts d'asymétrie d'information au profit des assureurs.

a. Bénéfices aux assureurs

L'assurance connectée basée sur le comportement du conducteur pour la tarification permet à l'assureur de mieux contrôler l'exposition des assurés au risque automobile. La promesse de réduction de prime pour les conducteurs prudents incite d'une manière très puissante les assurés à changer leur comportement au volant. Soleymanian *et al* (2017) ont mis en évidence l'effet de l'incitation financière offerte par une police UBI d'une compagnie d'assurance américaine sur le changement de comportement au volant. Ils ont montré que le score de conduite s'améliore pour tous les groupes d'âge³¹ durant les six mois du programme UBI. Il est également montré que le nombre de freins secs quotidien a diminué et particulièrement de 20% pour les jeunes conducteurs (âge < 35 ans) au bout de la même période [77].

Ces programmes d'assurance comportementale profitent aux assureurs pour plusieurs avantages compétitifs. Premièrement, les assureurs peuvent identifier les conducteurs à risque faible et améliorer le taux de rétention des risques préférés. Soleymanian *et al* (2017) ont montré que le taux de renouvellement du contrat d'assurance par les assurés abonnés au programme UBI³² est de 9%

³¹ Classification des assurés UBI étudiés en 4 groupes d'âge : (>35 ans) ; (35-50) ; (51-65) ; (>65ans)

³² Programme d'assurance UBI : Monitoring du comportement du conducteur par l'assureur pour une période préfixée (Flexibilité d'entrée et de sortie du programme pendant cette période). Un score de conduite quotidien est calculé et la ristourne permanente est déterminée en fonction du score et de la durée d'abonnement au programme.

plus élevé que celui des assurés conventionnels. Il est prouvé dans la même étude que 63% des assurés UBI restent abonnés au programme pour toute la période de six mois. En effet, le désabonnement prématuré réduit le niveau de la ristourne permanente reçue par l'assuré au terme du programme [77]. Par ailleurs, l'assureur italien Unipol a déclaré que la rétention des clients s'est améliorée de 10% entre 2009 et 2011 (Après l'introduction du produit d'assurance connectée Unibox) [81]. Deuxièmement, de nouveaux clients sont susceptibles de participer au programme UBI pour bénéficier d'une prime moins chère. Cela pourrait aider l'assureur à attirer les jeunes conducteurs souvent plus risqués mais probablement plus aptes à modifier leur comportement au volant en contrepartie d'une ristourne [77]. Une étude d'Ingenie, l'assureur télématique britannique ciblant les jeunes conducteurs (17-24ans) en 2014, a montré les résultats suivants :

- Plus que 90% de ses assurés suivent les feedbacks sur leur comportement au volant avec un rythme mensuel moyen de 14 fois.
- Les nouveaux clients réduisent leur risque d'accident de 50% durant la première année avec Ingenie.
- 91% des conducteurs reçoivent une ristourne pour l'absence de sinistres à la fin de la première année.
- 70% des conducteurs reçoivent une ristourne pour leur bonne conduite [24].

Troisièmement, les assureurs profitent de la réduction de coûts de gestion de sinistres grâce à la technologie avancée de détection de choc leur fournissant un premier avis de perte (FNOL³³). Cet avis – générateur des économies de charges totales de sinistres – permet aux assureurs de répondre efficacement aux demandes de secours de leurs assurés ; Avec de l'information précise et disponible en temps réel sur les accidents combinée avec les données géospatiales sur les routes, les assureurs peuvent régler les sinistres, détecter les fraudes et offrir de l'assistance immédiate (réponses aux urgences, voiture de remplacement, dépannage...) d'une manière efficace. Cela sous-entend que les partenaires de service d'assistance auront l'accès à l'information précise sur l'accident en temps réel ce qui rationalise le traitement de demandes leur est délégué par les assureurs. Une gestion des sinistres plus rapide implique une meilleure satisfaction du client, une réduction de coût de réparation et de dépannage, moins de risque de litiges et une réduction de coûts de sinistres [82]. Allianz Italie a déclaré que l'introduction de l'assurance télématique a amélioré son ratio combiné de 15% et son ratio de vente croisée de 83% grâce aux services complémentaires [83].

³³ First Notice Of Loss

Selon un rapport de Cisco sur les voitures connectées, ces dernières permettent la prévention de 80% des accidents déclarés et de réduire de 20% à 30% des coûts générés des infractions de circulation [84]. Zurich Insurance a déclaré qu'après essai de l'assurance télématique sur un groupe de véhicule d'une flotte, leur taux d'accident annuel a baissé de 66,4% pour les sinistres dommages et 72% pour les sinistres RC [83]. Unipol a déclaré qu'après deux ans de participation dans le programme UBI, les assurés participants, comparés aux assurés automobile conventionnels, ont enregistré un taux d'accident réduit de 20% à 30% et les vols de véhicules ont diminué de 30% à 50% [81]. Unipol a affirmé que le changement de comportement au volant (meilleure conduite) a contribué à la baisse du taux d'accident.

Selon le rapport de cisco précité, les sinistres frauduleux peuvent être détectés par les assureurs à travers l'analyse des données collectées des dispositifs de la télématique en temps réel et de les réduire jusqu'à 80% [84]. En effet, la technologie d'octo telematics (partenaire d'Unipol) a aidé, selon Unipol à détecter les sinistres frauduleux et à réduire de 30% à 50% les coûts d'investigations y liées [81] .

Les premiers assureurs adoptant l'assurance connectée ont un avantage compétitif grâce à la base de données comportementale collectée pour la tarification. Cette base de donnée leur permet une meilleure classification de profils de risque et de réduire le risque d'anti sélection. De plus, selon une étude de LexisNexis sur l'assurance télématique en 2014, les clients abonnés à un programme de surveillance d'une compagnie sont moins susceptible de changer d'assureur qui devrait le surveiller une deuxième fois afin de recevoir la ristourne [85]. D'ailleurs, Allianz Italie a déclaré que son *lapse ratio*³⁴ s'est amélioré de 43% en 2012 après l'introduction de la formule UBI [83].

b. Bénéfices aux clients

L'assurance télématique offre plusieurs avantages aux clients :

Primo, ils bénéficient d'une réduction du coût de l'assurance. Cette réduction a plusieurs sources à savoir la participation au programme UBI, l'amélioration du comportement au volant ou la réduction volontaire du kilométrage effectué [77]. Un jeune conducteur britannique de 17 ans paye en moyenne selon le modèle d'assurance automobile conventionnelle £2 204 contre £827 la moyenne du marché³⁵ [86]. Si ce conducteur souscrit une police UBI, il paye en moyenne £1 443. Selon le modèle d'assurance télématique à Ingenie, le conducteur peut bénéficier d'une ristourne trimestrielle jusqu'à 21%[87] . Le leader de l'assurance télématique britannique « Insure-The-Box »

³⁴ Lapse ratio: le pourcentage des polices d'assurance non renouvelées

³⁵ Polices Tous Risques, Willis Towers Watson / Confused Index, 2017

assure – en cas de bonne conduite – une ristourne annuelle moyenne de £601 lors de renouvellement du contrat UBI.

Secondo, la participation au programme UBI est sans risque pour le client. Le principe est le suivant :

- L'assureur propose la participation au programme UBI pendant une période donnée pour surveiller le comportement au volant des participants ;
- Les assurés reçoivent une réduction initiale sur leur taux de prime régulier (selon la tarification de l'assurance conventionnelle) pour leur participation au programme ;
- Au terme de la période de surveillance, ils reçoivent un taux de ristourne permanent correspondant à leur performance de conduite ;
- En cas de mauvaise performance, aucune pénalité n'est appliquée.

Tertio, tous les assurés seront traités sur le même pied d'égalité qu'ils soient de conducteurs à risque fort ou à risque faible. Selon une étude de Brookings Institute, 63,5% des ménages pourraient économiser \$496 annuellement sur leur contrat d'assurance automobile avec la formule UBI (en moyenne, 28% de réduction sur primes) [88]. Cette économie provient à la base des conducteurs avec un kilométrage élevé qui, en absence de cette formule, payent une prime disproportionnellement moins élevée. Par conséquent, les conducteurs roulant peu –généralement à bas salaire – profitent d'une prime plus équitable. L'absence de kilométrage élevé chez certains conducteurs ne les empêche pas de bénéficier d'une économie sur leur prime suite au changement de leur comportement au volant [77].

c. Bénéfices à la société

Il existe une reconnaissance croissante de la part des assureurs, des clients, des régulateurs et des écologistes que l'assurance télématique offre beaucoup d'avantages sociaux et écologiques (Les raisons d'essor de l'assurance télématique cités plus haut dans la section 2). La police UBI fournit une incitation pour les conducteurs à mieux conduire, par conséquent à réduire les accidents et leur intensité et à sauver des vies. Plusieurs autres avantages sont offerts, notamment l'amélioration du système de transport routier, la réduction des congestions routières, la réduction de la consommation de carburant et d'émission du CO₂, la réduction des coûts de récupération de véhicules volés et le développement des services d'assistance [46]. Selon le rapport d'Ingenie cité auparavant, 90% des véhicules volés des assurés Ingenie ont été récupérés grâce à la télématique embarquée. En moyenne, ces véhicules ont été localisés par Ingenie et récupérés par la police dans les trois heures

qui suivent leur vol [24]. Le graphique 7 en dessous résume l'impact estimé de la télématique sur la chaîne de valeur de l'assurance aux Etats-Unis selon Cisco [84] .

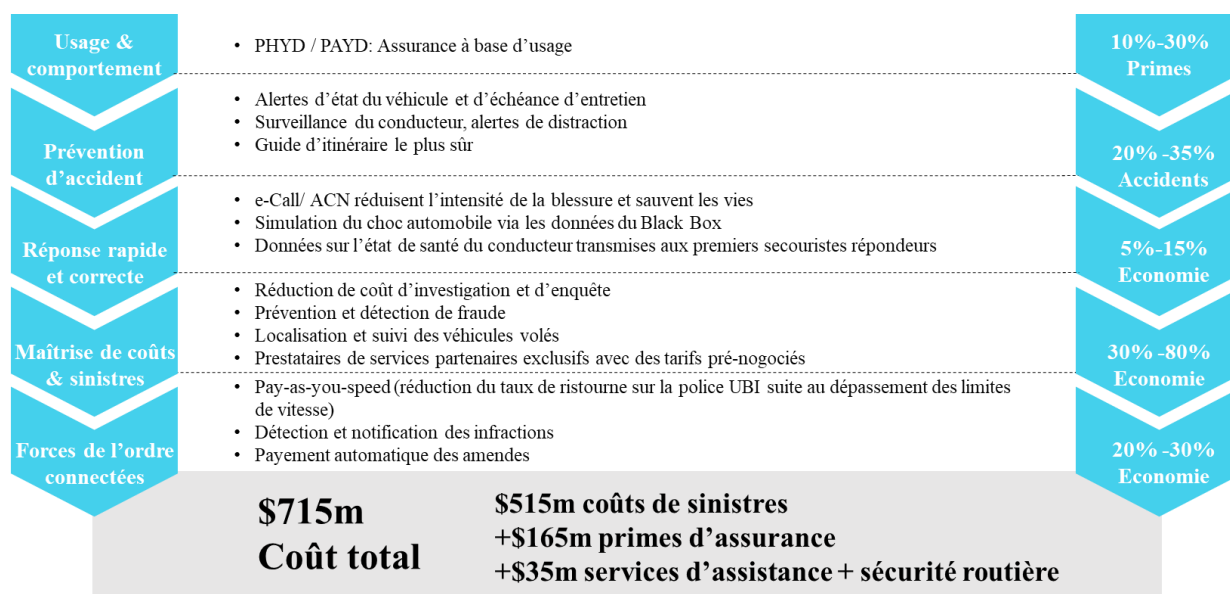


Figure 7 Impact estimatif de la télématique automobile sur la chaîne de valeur de l'assurance[84]

3.3. Défis de l'assurance comportementale

Comme toute innovation, cette nouvelle formule d'assurance trouve certaines difficultés pour se développer comme il se doit à savoir les difficultés techniques et de confidentialité.

a. Difficultés techniques

La question primordiale que l'assureur doit se poser avant d'offrir cette nouvelle formule est quel type de données à stocker. Selon NAIC, l'ensemble de données représentent entre 5Mo et 15 Mo annuellement par assuré. Un assureur avec 100 000 véhicules assurés peut collecter jusqu'à 1 téraoctet de données par année. Les assureurs sont déjà sous la pression du stockage des données massives et le coût de stockage des données - particulièrement sensibles et personnelles - a augmenté parallèlement au risque de leur violation. Il est primordial que l'assureur collecte les données utiles actuariellement parlant pour ne pas perdre du temps dans le tri des données massives. Un autre souci que l'assureur fait face est de choisir la technologie télématique la plus efficace pour la collecte des données. Actuellement, le choix est restreint entre les dispositifs temporaires « Dongles » facilement installés et désinstallés du véhicule et les dispositifs permanents « Black Box » un peu plus complexes dans l'installation et donc plus coûteux mais plus avancés en termes de collecte de données [89].

Cela sous-entend que l'introduction de l'assurance connectée requiert un investissement en infrastructure IT et en compétences en Data Science et Big Data pour un meilleur traitement, analyse et suivi en temps réel des données des assurés.

b. Difficultés de confidentialité

Un obstacle considérable au développement de l'assurance comportementale est le souci de la confidentialité de la part des clients et d'un point de vue juridique. Les données liées à la localisation GPS et à la détection des infractions routières et collectées par l'assureur sont considérées par les clients comme atteinte à la vie privée. L'un des plus grands défis de l'assureur est de convaincre les clients et les régulateurs du contraire. L'environnement juridique entourant la question de la confidentialité et de la propriété des données est encore relativement nouveau et très fluide. Par exemple, concernant le respect de la propriété des données collectées par les intermédiaires en bourse, la question est plus facile à répondre que la question concernant la propriété des données de la télématique puisque le cadre juridique est défini pour la catégorie d'information dans le marché financier. Le droit d'utilisation de la télématique par les compagnies d'assurance n'est pas encore traité définitivement par les régulateurs. La loi de partage des données télématiques est en voie de développement et ne suit pas le même rythme rapide de celui du marché de l'assurance automobile. Finalement, les données télématiques sont considérées comme la propriété des conducteurs mais en matière juridique, aucun texte de la loi ne le définit. Tout ce flou juridique crée au sein d'un marché assez régulé comme l'assurance, une certaine méfiance de la part de l'assureur et de l'assuré [82].

Conclusion

A travers ce premier chapitre, il est montré que l'intégration des éléments du comportement au volant dans le modèle conventionnel de tarification permettrait à l'assureur automobile de mieux segmenter sa clientèle, de maîtriser sa sinistralité et de proposer des offres plus innovantes et plus adaptées à leurs besoins. La question qui se pose est la suivante : Comment l'essor des données inexploitées par l'assureur automobile, telles que l'accélérométrie, la vitesse, ou encore la position GPS, collectée par le moyen d'un boîtier installé dans le véhicule des assurés, pourrait être déployé dans sa stratégie, lui permettant de mieux comprendre les comportements au volant et les causes d'accidents et donc de profiter des opportunités précitées? La recherche d'une réponse émane des résultats de l'étude d'impact des nouvelles technologies, et en particulier, de la télématique automobile, sur la logique de l'assurance automobile. Pour être concrétisée, la collecte d'informations sur un ensemble de véhicules fût nécessaire. Le deuxième chapitre de ce mémoire de

fin d'études présente la méthode adoptée d'exploration des données télématiques collectées auprès d'une start-up tunisienne de services télématiques afin d'étudier l'utilité de l'introduction de la télématique dans l'assurance automobile.

Deuxième chapitre

Analyse du comportement au volant par la télématique automobile du produit « Sayartech » d'Ooredoo



Deuxième chapitre : Analyse du comportement au volant par la télématique automobile du produit « Sayartech » d'Ooredoo

Introduction

D'habitude, l'environnement contrôle le comportement humain qui est à son tour une réponse à un stimulus particulier. En psychologie, il existe une technique appelée « conditionnement opérant » dérivée du « Béhaviorisme » qui permet de contrôler le comportement non pas par les stimuli de l'environnement mais par les conséquences du comportement déjà anticipées par l'individu suite au stimulus. L'individu peut changer son comportement et l'adapter pour obtenir ou empêcher les conséquences de sa réponse.

L'approche de l'assurance comportementale se repose sur cette technique ; l'assureur analyse le comportement du conducteur abonné au programme UBI et évalue sa qualité de conduite. L'assuré évalué est conscient des conséquences de son comportement (réduction de prime en cas de conduite prudente ou surprime en cas de conduite risquée). De ce fait, l'assureur agit indirectement sur le niveau de risque individuel de conduite afin de contourner les dérives comportementales et de mieux maîtriser le risque automobile. En Tunisie, le changement de comportement chez les assurés UBI n'est pas encore ressenti vu que le produit est naissant. Par conséquent, pour savoir s'il s'agit d'une réelle opportunité pour l'assureur tunisien, il est plus approprié de commencer par l'étude d'impact de l'introduction des variables comportementales sur la logique de l'assurance en termes de sélection et de tarification du risque automobile.

Le but de ce chapitre est de proposer une méthode statistique d'évaluation du risque individuel de conduite après l'analyse du comportement des conducteurs par le moyen des données télématiques collectées auprès des boîtiers installés dans les véhicules des clients du produit « Sayartech ». Après une brève présentation des partenaires de l'étude et du fonctionnement du produit, une analyse du nouveau processus de sélection de risque d'un point de vue professionnel est exposée. Ensuite, une analyse statistique est effectuée à l'aide d'une régression linéaire multiple pour montrer la corrélation entre le comportement et le risque automobile d'une part et pour estimer les nouveaux paramètres comportementaux intégrant le modèle de tarification du risque d'autre part. Finalement, le chapitre conclut sur l'utilité de l'approche pour l'assureur automobile.

Section 1 Cadre général de l'étude

Les services télématiques à savoir la géolocalisation et le tracking GPS/GPRS sont en train de se développer avec le renforcement des TICs en Tunisie. Par contre, l'assurance comportementale est très récente avec l'initiative de la start-up « Telematik » en collaboration avec l'opérateur téléphonique « Ooredoo » et la première compagnie d'assurance « Maghrebia » qui commence à offrir le produit d'assurance automobile connectée depuis Mai 2017. L'étude de ce nouveau type de produit est une première dans le secteur d'assurance et permettra à tous les assureurs du marché d'évaluer l'opportunité d'offrir un tel produit innovateur. En effet, La COMAR a toujours misé sur la branche automobile avec plusieurs projets novateurs tels que le service de garagiste conventionné, les avantages de l'avance sur recours et l'avance sur réparation. Un intérêt particulier à l'assurance Pay How You Drive est exprimé au sein de l'entreprise et le projet est en cours d'étude. Avant de présenter le produit « Sayartech » offert par Ooredoo, il est primordial de présenter tous les acteurs concernés par cette étude à savoir mon entreprise de parrainage Assurances COMAR, Telematik et Ooredoo.

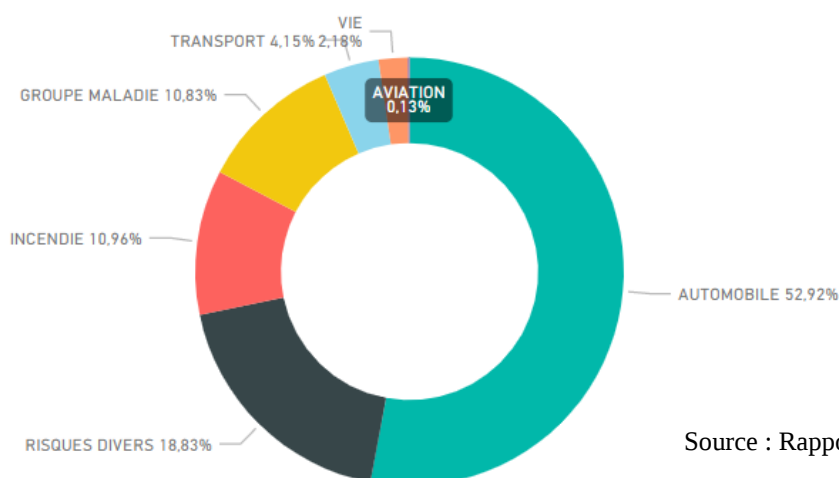
1.1. Entreprise de parrainage

La Compagnie Méditerranéenne d'Assurances et de Réassurances CO.M.A.R opère dans toutes les branches d'assurance. Sa filiale HAYETT est spécialisée en assurance vie et capitalisation. Toutefois, la seule activité vie (assurance temporaire crédit) de la COMAR est gérée par HAYETT.

La COMAR est née suite à la transformation de la représentation de l'ex- compagnie Maghrébine d'Assurances Réunies en société de droit tunisien en 1969. En 1973, Amen Group a pris le contrôle de la compagnie avec les partenaires français (AGP) et allemand (Munich Ré). En 1987, la filiale HAYETT est créée pour se spécialiser en assurance vie et capitalisation. En 2010, la COMAR a procédé à une augmentation du capital social de 40 à 50 millions de Dinars. En 2017, elle a décroché l'agrément IARDT (Incendie-Accidents-Risques Divers-Transport) en Côte d'Ivoire pour démarrer ses activités dans la filiale ivoirienne créée en 2016 sous l'appellation de COMAR CI.

La figure ci-dessous montre que 53% du CA de la COMAR provient de la branche automobile en 2017. Les risques divers représentent 19% du CA suivi par la branche Incendie et groupe maladie avec 11% chacune. Le transport contribue au CA avec 4%. Seul le produit temporaire Décès de l'assurance vie est rentré dans l'activité de la COMAR soit 2% du CA. Le chiffre d'affaires en 2017 a augmenté de 9,54% par rapport à 2016 pour s'établir à environ 179 millions de dinars hors acceptations. Le résultat technique non vie après réassurance s'élève à 32,4 millions de dinars et le

résultat technique vie s'élève à 3,2 millions de dinars. Le bénéfice net au titre de l'année 2017 est de 32,5 millions de dinars contre 30 millions de dinars et 25,911 millions de dinars enregistrés en 2016 et 2015 respectivement.



Source : Rapport annuel 2017, COMAR

Figure 8 Le CA par branche d'activité en 2017

1.2. Les partenaires de l'étude

Cette étude est centrée sur l'analyse des données collectées auprès de la start-up « Telematik » sous l'accord de son partenaire commercial, l'opérateur téléphonique « Ooredoo », dans le cadre de la commercialisation du service de la télématique automobile « Sayartech » depuis Avril 2017.

Telematik est une start-up fondée en 2012 à Zaghouan et qui développe des solutions télématiques à savoir la gestion intelligente de flotte automobile et des solutions UBI ou encore de l'assurance télématique. L'objectif est d'aider les clients (particuliers et entreprises) à optimiser l'utilisation de leurs données de « conduite » afin d'améliorer la gestion de flotte et la sécurité routière. Le détail du fonctionnement du produit UBI « Sayartech » sera exposé plus loin dans ce chapitre.

Ooredoo (ex-TUNISIANA) est l'un des trois opérateurs de téléphonie mobile en Tunisie. Il est devenu depuis 24 Avril 2014, une filiale tunisienne du groupe Ooredoo. TUNISIANA, créée en 2002, proposent une gamme d'offres et de services innovateurs alignés aux normes internationales. En 2012, Tunisiana a lancé les services 3G destinés à plus de 6.6 millions d'abonnés et couvrant 48% de la population. En 2014, le réseau s'est étendu à près de 90 % de la population.

Jusqu'à présent, le produit UBI de Telematik est adopté par une seule compagnie d'assurance sur le marché (Assurances Maghrebria) sous forme de convention entre la dernière et Ooredoo. A travers cette convention, les clients Maghrebria abonnés aux services de Sayartech ou les clients Ooredoo désireux de s'abonner au programme UBI de Maghrebria peuvent bénéficier d'une bonification exceptionnelle allant jusqu'à -30% sur leurs contrats d'assurance automobile.

1.3. Description du produit Sayartech

Sayartech est un produit novateur qui rend les véhicules intelligents et communicants via un boîtier OBD-II et une carte-Sim Ooredoo installés dans le port diagnostic situé dans l'habitacle du véhicule.

Le service Sayartech est composé :

- D'un boîtier installé dans le port OBD-II du véhicule du client. Le dongle OBD-II permet de lire les données liées au véhicule et à la conduite et de les transmettre à une plateforme qui permettra de les analyser et de les afficher à partir de l'application web/mobile. Il est également équipé d'une carte Sim nécessaire à la transmission des données stockées dans le Data center de Ooredoo.
- D'un service de connectivité et d'accès à la plateforme. Il est plafonné à 10 dt/ mois payable à l'avance selon une des modalités de paiement proposées et versé automatiquement le premier jour de chaque mois. Ce service offre une capacité de 100 Mo d'internet dédiée à la localisation GPS et au transfert des données de conduite et du véhicule.

La plateforme mise en place permet :

- La géolocalisation du véhicule 24/7 et l'historique complet de tous les déplacements
- L'émission des alertes en cas de détection d'une défaillance au niveau du véhicule ou d'une situation inhabituelle comme par exemple :
 - la détection du véhicule en mouvement sans être en marche (en cas de mise à la fourrière)
 - des portes non verrouillées (informations disponibles chez certains constructeurs)
 - la détection d'un choc
 - Un vol avéré du véhicule. En effet, les accords de Roaming avec les pays frontaliers permettent la géolocalisation même en dehors du territoire tunisien
- La consultation de l'historique des trajets et de la dernière position enregistrée par le véhicule
- La notification par alerte de certains événements configurables par le client (rappel vidange, entretien, contrôle technique...)
- L'analyse de la conduite et le calcul d'un score (cumulé et par trajet). Un score est calculé à chaque trajet effectué permettant de suivre les statistiques mensuelles et la qualité de conduite dans le temps.

- Des conseils du jour

Les graphiques suivants affichent les captures d'écran de l'application mobile du produit Sayartech :

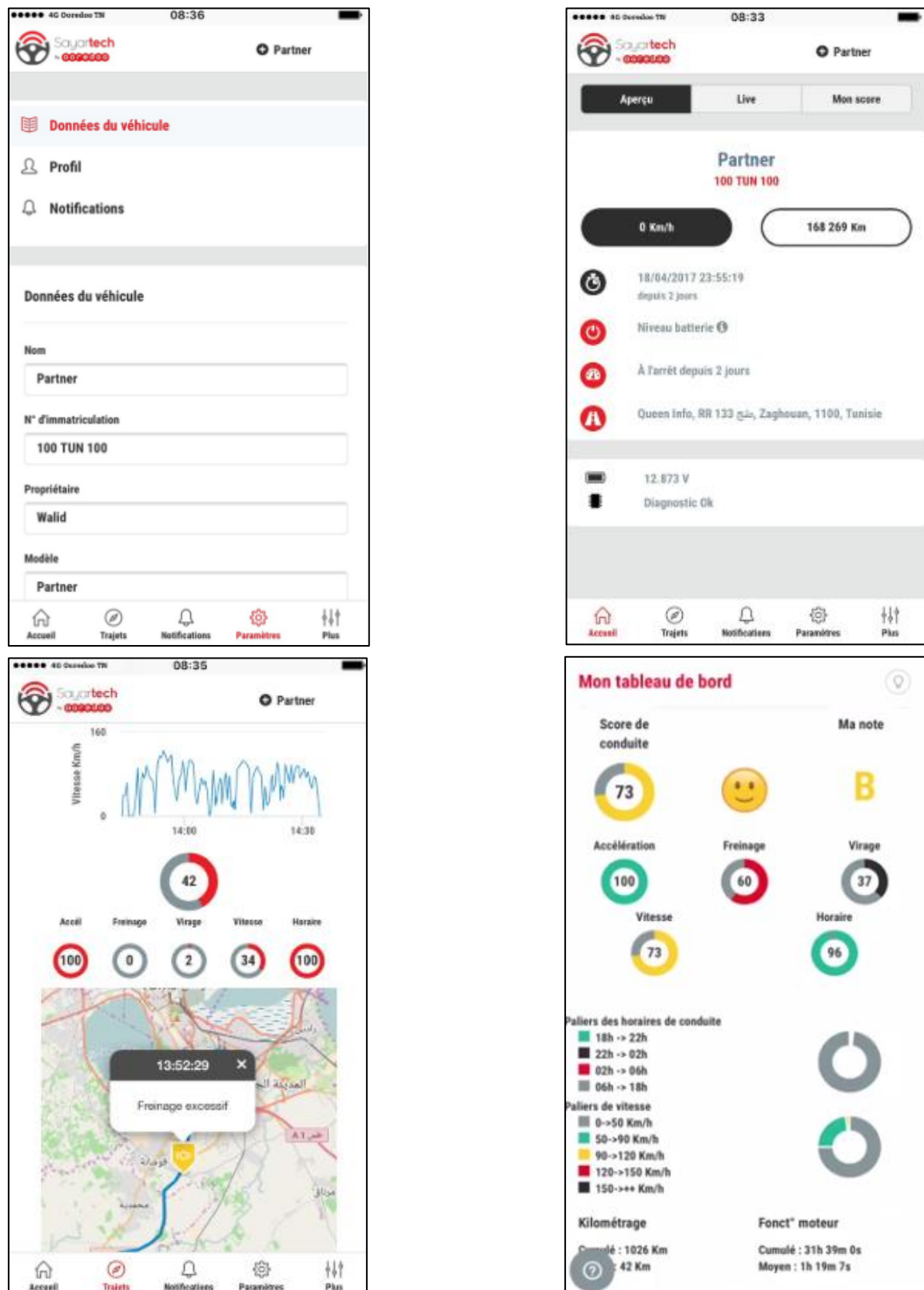
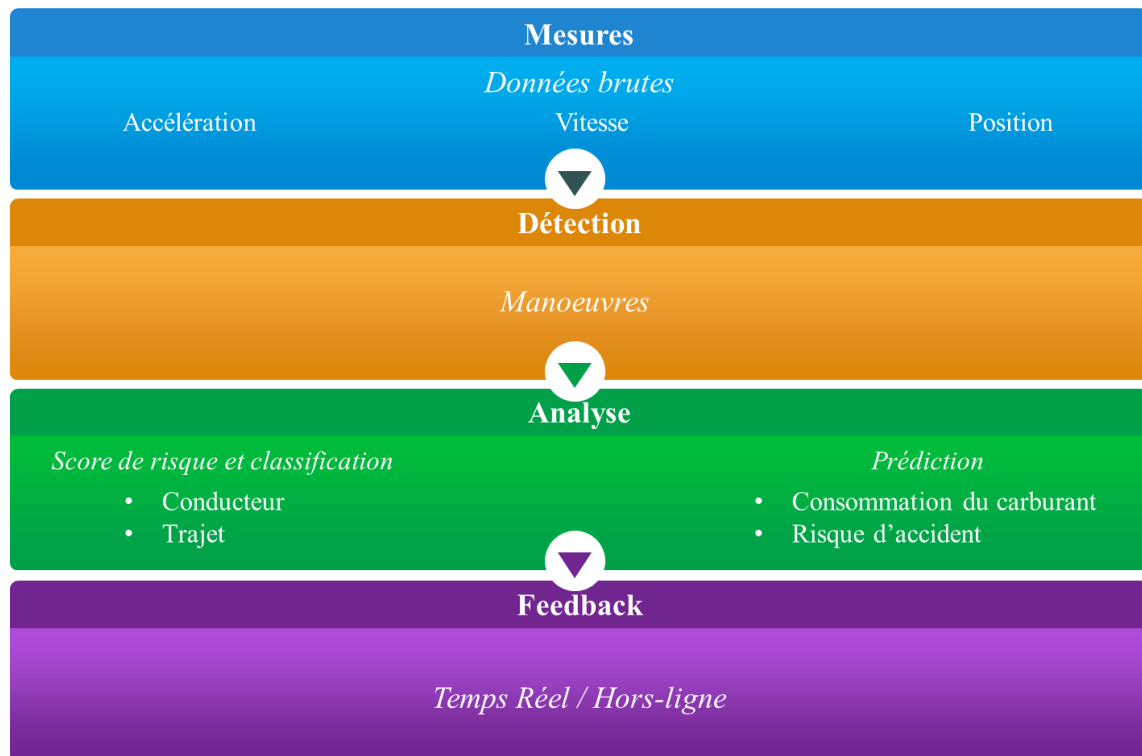


Figure 9 Captures d'écran de l'application mobile de Sayartech

1.4. Fonctionnement du produit Sayartech

Avant de commencer l'étude statistique, il est utile de donner un aperçu sur le fonctionnement du système télématique de Sayartech à partir duquel les données brutes ont été collectées, traitées, agrégées et analysées pour atteindre l'objectif de l'étude. Le fonctionnement du système télématique de Sayartech suit la logique suivante: (Toledo et al 2008)



a. Mesures

Le dispositif télématique collecte les données suivantes :

1. L'identification du véhicule et du conducteur. Le type de véhicule a un impact sur le paramètre « accélération » et la notation des événements de conduite. Pour recevoir des données crédibles et véridiques, il faut sélectionner le type de véhicule correspondant au véhicule dans lequel le boîtier est installé. Chaque type de véhicule disponible dans l'application web et mobile est classé dans une catégorie d'accélération spécifique, ce qui joue sur la sévérité des événements de conduite.
2. Temps du début et de fin des trajets
3. L'accélération du véhicule dans les deux directions longitudinale et latérale mesurée par l'accéléromètre
4. La vitesse du véhicule dérivée des données du récepteur GPS
5. La localisation du véhicule mesurée par le récepteur GPS

6. D'autres informations liées au moteur sont collectées si ce système est connecté au système OBD du véhicule (moteur arrêté/allumé, véhicule en marche/au ralenti/ à l'arrêt...)

b. Détection

Des algorithmes de reconnaissance de formes sont appliqués aux mesures brutes afin de détecter les manœuvres effectuées par un véhicule. Cette étape est primordiale pour réduire le nombre important des données brutes à des observations significatives, au-delà des distributions d'accélération et de vitesse, utilisées pour déduire le comportement du conducteur. Le système OBD identifie, à travers l'accéléromètre, 12 types d'évènements comme le changement de voie, le virage avec ou sans accélération, le freinage brutal, les accélérations excessives... ces évènements sont classés davantage selon leur direction relative (droite ou gauche) et/ou selon trois niveaux de sévérité basés sur le modèle du véhicule et les paramètres de la trajectoire détaillée comme la durée de la manœuvre, l'ampleur des changements brusques de vitesse et d'accélération et à quelle vitesse ces changements ont été effectués. (Voir tableau en dessous)

D'autres notifications sont détectées par le dispositif télématique liées à la batterie, à la température, au carburant et au dysfonctionnement. Le moindre choc du véhicule est également détecté et enregistré dans un fichier log. En total, 46 évènements, alertes et notifications sont détectés et/ou mesurés par le dispositif télématique exposés dans le tableau suivant :

Tableau 9 les évènements détectés par le dispositif télématique

Moteur & mouvement	Température du liquide de refroidissement moteur acceptable
Moteur arrêté	Notification de dysfonctionnement
Moteur allumé	Niveau batterie critique
Véhicule à l'arrêt	Niveau batterie acceptable
Véhicule à l'arrêt sans couverture GPS	Niveau de carburant faible
Véhicule en marche	Niveau de carburant acceptable
Véhicule en marche - virage à gauche	Accéléromètre
Véhicule en marche - virage à droite	Accélération sévère : Seuil Limite 1 dépassé
Véhicule en marche	Accélération sévère : Seuil Limite 2 dépassé
Véhicule au ralenti	Accélération sévère : Seuil Limite 3 dépassé
Fin de l'alerte véhicule au ralenti	Freinage brusque : Seuil Limite 1 dépassé
Mouvement détecté	Freinage brusque : Seuil Limite 2 dépassé
Aucun mouvement détecté	Freinage brusque : Seuil Limite 3 dépassé
Remorquage	Accélération gauche sévère : Seuil Limite 1 dépassé
Début remorquage	Accélération gauche sévère : Seuil Limite 2 dépassé
Véhicule en mouvement (probablement en remorquage)	Accélération gauche sévère : Seuil Limite 3 dépassé
Anomalies GPS	Accélération droite sévère : Seuil Limite 1 dépassé
Pas de couverture GPS	Accélération droite sévère : Seuil Limite 2 dépassé
Calibrage et OBD	Accélération droite sévère : Seuil Limite 3 dépassé
Mise à jour accomplie	LOGS
Autocalibration accomplie	Message de configuration
Protocole OBD détecté	SMS reçu
Temps de détection du protocole OBD dépassé	Impact détecté
Notifications	Fichier log du crash enregistré
Alimentation principale activé	Fichier log du crash non enregistré
Alimentation principale désactivé	Signal GSM brouillé
Température limite du liquide de refroidissement dépassée	

c. Analyse

Les manœuvres détectées sont utilisées pour calculer les différents indices de risque liés au conducteur et au véhicule et les statistiques :

- *Indice de risque conducteur ou encore score de conduite* : il s'agit de la mesure numérique du risque d'implication d'un conducteur dans un accident de route pendant une période donnée. Il est calculé en fonction de fréquence et de sévérité des différents types de manœuvres dites risquées effectués par le conducteur.
- *Classification du risque* : basé sur le score, le conducteur est classé en trois catégories. L'objectif principal de la classification est de fournir un système plus simple pour rapporter les indices de risque compréhensibles par les conducteurs novices. Vert pour le comportement normal, jaune pour le comportement intermédiaire, rouge pour le comportement risqué.
- *Indice de risque et classification par trajet* : le même principe de calcul d'indice et de classification du risque est appliqué pour chaque trajet effectué par le conducteur.
- *Indice de vitesse* : la valeur numérique désigne le comportement de célérité du conducteur. Elle dépend de la durée et l'ampleur de l'excès de vitesse par rapport aux limites de vitesse légales. Bien que la vitesse soit une des variables de mesure de l'indice de risque, la littérature montre que cet indicateur est le plus prédictif du risque d'implication dans un accident et un indice séparé doit être calculé à part.
- *Statistiques de mesure d'exposition au risque automobile* : la distance et la durée du trajet parcouru par le conducteur et leur distribution temporelle (moment de la journée). La distribution spatiale (Type de route) n'est pas encore adoptée par le prestataire du service télématique dans son modèle.

d. Feedback

Le feedback peut être fourni en temps réel et hors-ligne. Dans l'application hors-ligne ou encore la plateforme web, des rapports sont produits et qui feront l'objet d'un résumé sur le niveau du risque du conducteur et de son véhicule.

Le feedback en temps réel, qui inclut les alertes sur les comportements agressifs ou sur les déviations significatives du modèle de la conduite normale pour un conducteur bien déterminé est fourni sous la forme des messages (via Messenger) ou utilisant le dispositif d'affichage du véhicule. Pour réduire la communication, les mesures et les détections sont effectuées en temps réel dans le

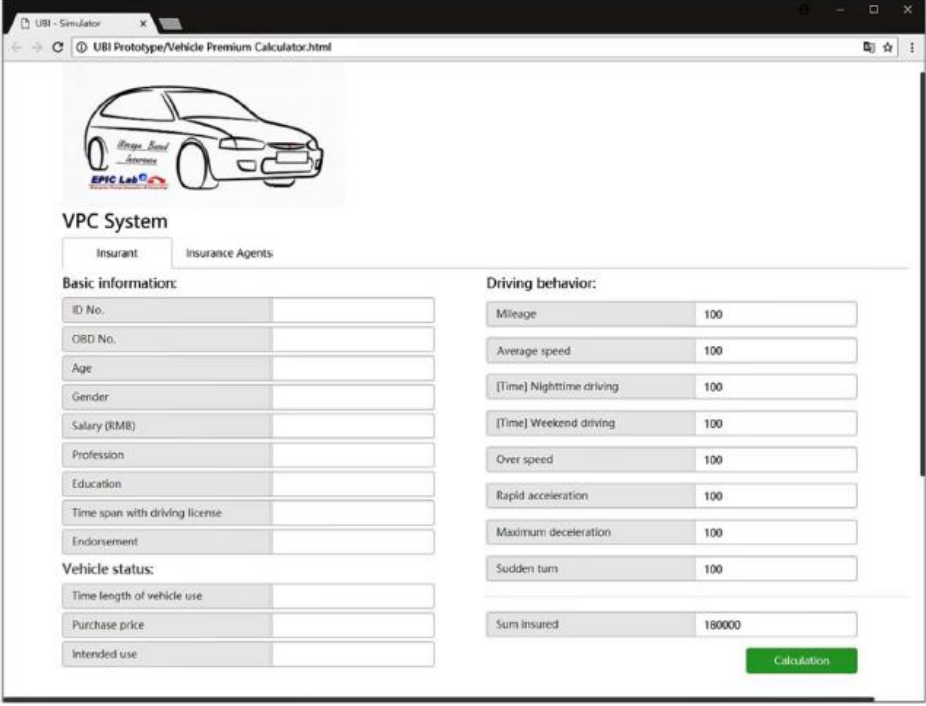
dispositif embarqué dans le véhicule. Les résultats de la détection sont automatiquement transmis au serveur d'application utilisant le réseau sans fil.

La télématique embarquée dans le véhicule est utilisée pour déterminer le score de conduite qui fera d'un input pour trois fins (Warren et al 2007) :

- Servant à calculer les conditions tarifaires d'une nouvelle police d'assurance.
- Servant à modifier les conditions tarifaires d'une police d'assurance existante
- Servant à déterminer l'assurabilité d'un conducteur

Le score reflète un aperçu qualitatif subjectif sur les caractéristiques individuelles de conduite. Il est utilisé pour déterminer le niveau de risque associé à l'assurance d'un conducteur particulier. Un prototype de *pricing* selon le modèle de l'assurance comportementale est présenté dans ce mémoire à titre indicatif (Bian et al 2017). Ce prototype (développé par programmation) permet aux agents d'assurance de déterminer les primes personnalisées par l'identification du niveau du risque de l'assuré et le score de conduite. L'interface de ce prototype comprend trois volets : l'accès de l'utilisateur (Individu/ agent), les inputs et les outputs.

- L'accès de l'utilisateur permet de choisir lequel est l'utilisateur (agent d'assurance ou client).
- Les utilisateurs saisissent leurs informations basiques, l'état du véhicule et la somme assurée dans les cases correspondantes. (Inputs)
- Le prototype permet à l'agent d'assurance de calculer la prime personnalisée, et de modéliser la tarification en analysant la tendance du portefeuille et des catégories de niveau de risque de ses assurés. Pour les utilisateurs individuels (clients), le prototype permet d'analyser et de classer leurs comportements en se basant sur les données comportementales collectées de leurs dispositifs OBD , ensuite de calculer leurs scores de conduite et de simuler la prime personnalisée. (Outputs)



The screenshot shows a web browser window titled 'UBI - Simulator' with the address 'UBI Prototype/Vehicle Premium Calculator.html'. The page features a car icon and the 'VPC System' title. Below the title are two tabs: 'Insurant' (selected) and 'Insurance Agents'. The form is divided into three main sections: 'Basic information:', 'Driving behavior:', and 'Vehicle status:'. Each section contains several input fields. At the bottom right, there is a 'Calculation' button.

Basic information:		Driving behavior:	
ID No.		Mileage	100
OBD No.		Average speed	100
Age		[Time] Nighttime driving	100
Gender		[Time] Weekend driving	100
Salary (RMB)		Over speed	100
Profession		Rapid acceleration	100
Education		Maximum deceleration	100
Time span with driving license		Sudden turn	100
Endorsement			
Vehicle status:		Sum insured	180000
Time length of vehicle use		Calculation	
Purchase price			
Intended use			

Figure 10 Interface client /agent prototypée

1.5. Processus de sélection de risque

Le graphique ci-dessous expose le processus de sélection de risque utilisé pour les nouvelles affaires et le renouvellement des contrats existants. Il est composé de deux étapes : Evaluation du niveau de risque souvent mesuré par un score et la tarification si le client est accepté en se basant sur le score calculé et sur les procédures internes de la compagnie d'assurance.

Le processus d'évaluation du risque inclut les étapes suivantes :

- 1) Entrer les informations basiques : Numéro d'identification de l'assuré, numéro de série de le dispositif OBD, les informations sur le conducteur (âge, expérience, zone de résidence...) et sur le véhicule (usage, marque, modèle, âge du véhicule).
- 2) Analyser le comportement du conducteur. Si le client a un historique d'enregistrements sur le dispositif OBD, le système calcule le score de conduite par l'analyse de la data stockée dans le dispositif. Les variables comportementales comprennent la distance parcourue, la vitesse moyenne, la durée de conduite en jour/nuite, la durée de conduite en weekend/jour de semaine, taux d'excès de vitesse, nombre d'accélération rapides, nombre de freins brusques et de virages rapides. Si le client n'a pas encore d'enregistrements, cette étape va être sautée.
- 3) Le risque automobile est évalué. Le système évalue les clients sur la base de la qualité de conduite, des informations basiques et des données géographiques.

- 4) Basé sur le modèle de l'assurance comportementale, le système génère le score de conduite et la tarification personnalisée du client.

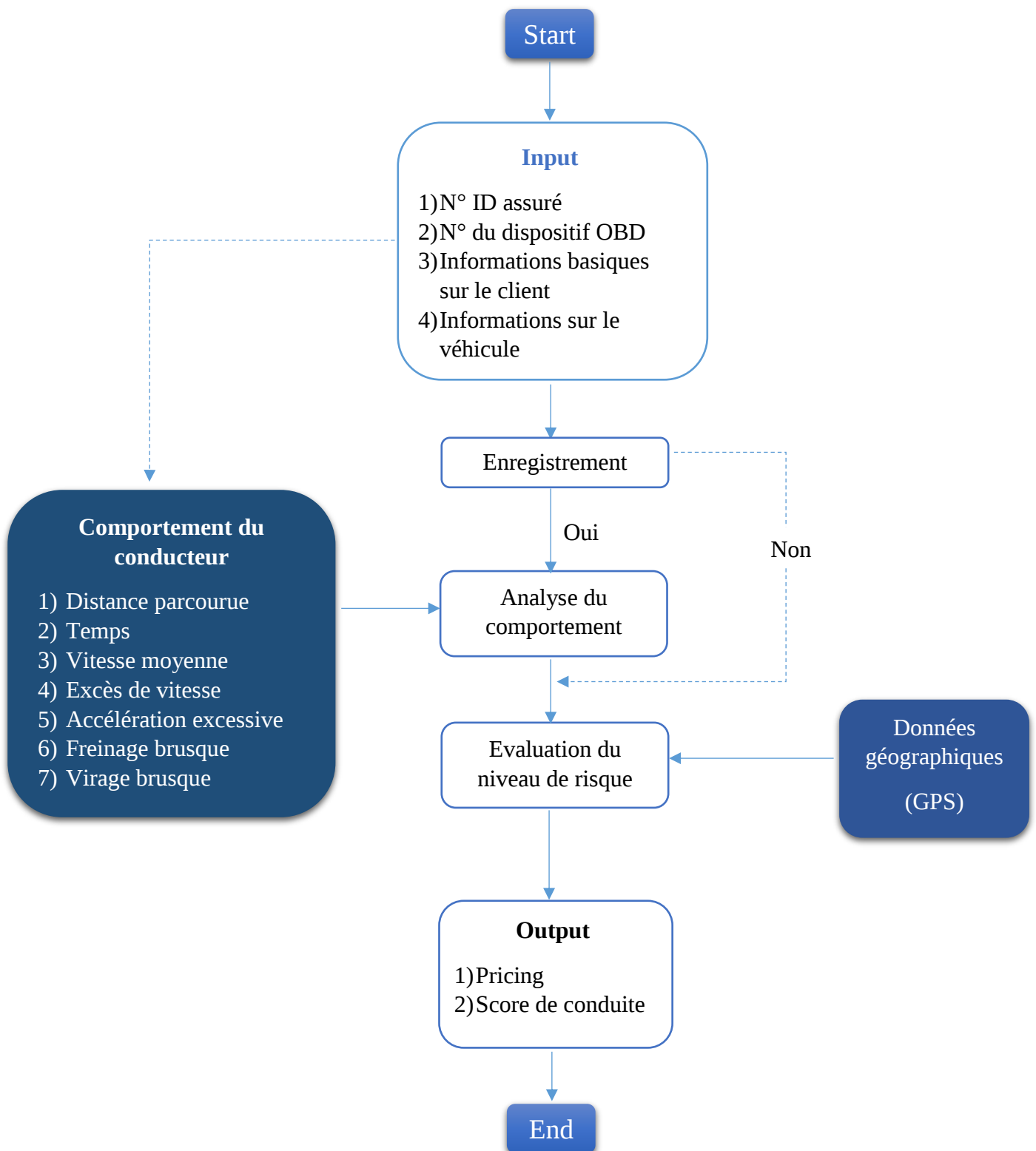


Figure 11 Nouveau processus d'évaluation du risque automobile

Pour mieux comprendre le mécanisme du prototype cité auparavant, des captures d'écran d'exemples de résultat de calcul de primes sont exposées dans les graphiques suivants : Le graphique du radar montre le comportement d'un assuré noté 69.37. Le graphique en 3D montre la distribution globale des observations. Il représente les fonctions du prototype de stockage et de visualisation des données. L'axe X représente la somme assurée totale. L'axe Y représente les primes et l'axe Z expose les scores de conduites.

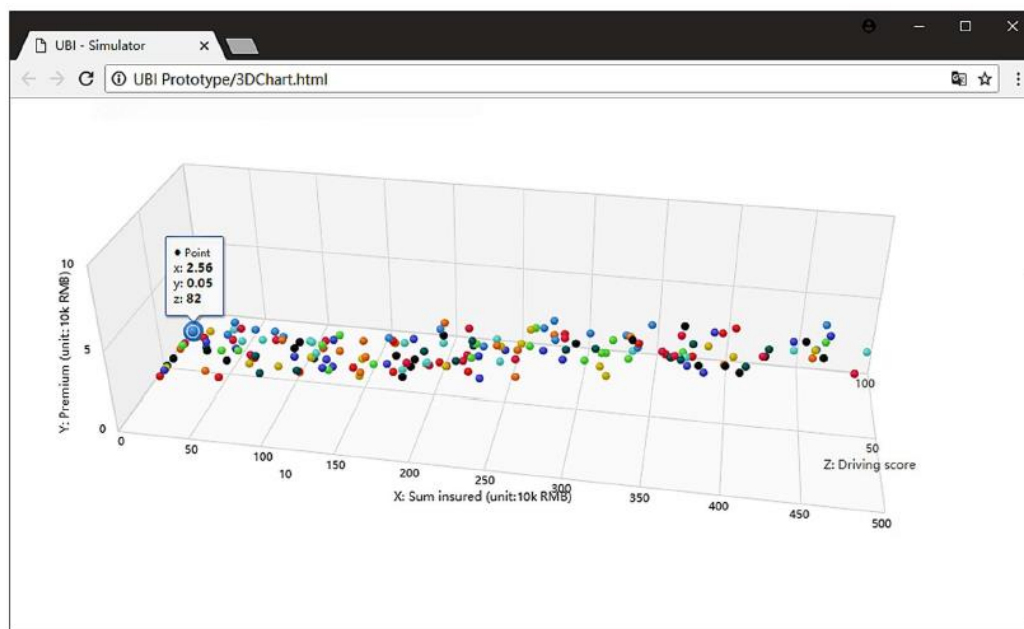
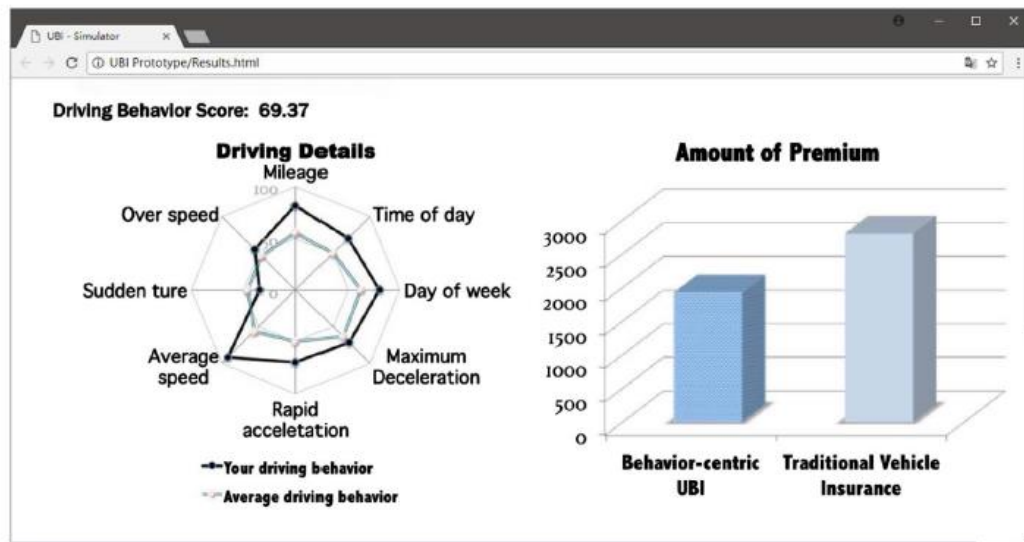


Figure 12 Outputs du prototype d'assurance automobile comportementale

De ce fait, L'assureur peut efficacement suivre son portefeuille ; La situation globale et les tendances du portefeuille aident les assureurs à mieux comprendre la distribution des assurés et de prendre de meilleures décisions stratégiques.

Notre étude se limite à l'évaluation du risque automobile à travers une approche purement scientifique sans recours à la méthode de notation des niveaux de risque de chaque conducteur de l'échantillon objet d'étude décrit dans la deuxième section.

Section 2 Données télématiques

2.1. Description de l'échantillon

La particularité de cette étude comparée à la littérature est que les données ne sont pas collectées auprès d'une compagnie d'assurance offrant le produit d'assurance connectée. En Tunisie, cette innovation est tellement récente (qui date depuis Avril 2017 fondée par la start-up Telematik) qu'un seul et unique prestataire de ce genre de service aura la possibilité de fournir ce type de données. Du mois d'avril 2017 au mois de juin 2018, 51 véhicules ont été surveillés depuis le jour d'installation du boîtier OBD. Les données du comportement du conducteur sont collectées et réduites pour mesurer la distance parcourue mensuelle moyenne, la durée de conduite moyenne mensuelle en nuit et en jour, la durée de conduite mensuelle moyenne sous quatre paliers de vitesse (Allant de 50km/h jusqu'à +150 km/h) et les manœuvres de freinage brusque, accélération brutal et virage rapide. Après traitement et analyse des différentes données collectées, deux véhicules présentent un comportement atypique par rapport à l'échantillon³⁶. Dans l'annexe 4, quatre variables *a priori* explicatives affichent les mêmes véhicules (n°51 et n°46) qui enregistrent des valeurs aberrantes. En statistiques, les observations contenant des valeurs aberrantes ou extrêmes doivent être retirées de l'échantillon pour une meilleure précision de la modélisation. De plus, la régression linéaire multiple (RLM) à développer par la suite a ressorti un 3ème véhicule (N°18) qui présente des valeurs extrêmes par rapport à l'échantillon³⁷ ; Une valeur extrême dans une régression linéaire suppose que l'observation prédite ait une valeur résiduelle de +3,29 ou de -3,29 [90]. Aucune autre valeur extrême n'est éliminée car l'effet sur le modèle de régression linéaire est négligeable. (Distance de Cook<0,5). La base de données finale comporte 48 véhicules avec un peu plus que 32 000 heures de conduite et approximativement 1 450 000 km parcourus. Par contre, aucune information n'est disponible concernant l'historique de la sinistralité des clients UBI par souci de confidentialité. D'autres informations liées au conducteur comme l'âge, le sexe, la catégorie socio-

³⁶ Voir annexe 4 et 5

³⁷ Voir annexe 6 et 7

professionnelle, l'usage... ne sont pas disponibles et ne seront donc pas intégrées dans la modélisation. Le boîtier détecte, certes les éventuelles collisions mais parmi 51 véhicules, un seul accident a été détecté pendant la période d'observation. De ce fait, la variable à prédire du modèle ne sera plus le nombre d'accidents mais plutôt le nombre de quasi-accidents qui font augmenter la probabilité d'être impliqué dans un accident très utile à mesurer pour le contrôle du risque automobile avant la réalisation du sinistre. La deuxième sous-section sera consacrée à l'explication des différents événements routiers et leur importance aux yeux des assureurs automobiles.

2.2. Définition du phénomène d'accident, quasi-accident et incidents critiques

Les données collectées auprès de 51 boîtiers permettent d'identifier trois types d'événements liés à la sécurité routière : accident, quasi-accidents et événements d'incidents critiques ainsi définis par Dingus *et al* (2006). **Un accident** est défini comme un événement de contact entre le véhicule en question et un autre véhicule, un animal, un objet fixé, un piéton ou un cycliste[91]. **Un quasi-accident** est une situation de conflit qui demande une manœuvre rapide, sévère et évasive pour empêcher l'accident. Cette manœuvre comprend le freinage, l'accélération, le virage ou toute combinaison d'actions de contrôle qui s'approche des limites de capacités du véhicule. Par exemple, une collision imminente avec un véhicule lors de la fusion de deux voies, un virage à gauche précipité, un demi-tour inapproprié....[91] En effet, en moyenne, les véhicules dépassant une accélération/décélération de +/- 0.8g (équivalent à +/-28 km/h en une seconde) augmentent le risque d'être impliqué dans un accident. Dans le cas du produit *Sayartech*, les manœuvres dépassant le seuil limite 3 (Freinage, accélération et virage) seront considérées comme conséquentes des situations de quasi-accidents. **Les événements d'incidents critiques** sont des situations de conflit de moindre sévérité que les quasi-accidents. Ils sont détectés par le boîtier OBD quand les capteurs dépassent une valeur spécifique. De ce fait, ces manœuvres pourraient être utilisées pour mesurer l'agressivité du conducteur. Cela sous-entend que les conducteurs relativement prudents, en se basant sur leurs compétences de conduite et leur conscience de sécurité évitent les manœuvres évasives qui pourront contribuer à des scénarii dangereux à savoir les événements d'incidents critiques. La haute fréquence de ces événements reflète un manque de compétences et de conscience de la sécurité routière [91].

L'objectif du mémoire est de permettre à l'assureur de prédire les conducteurs risqués à partir des données télématiques révélatrices du comportement au volant notamment les « événements d'incidents critiques ». Selon Guo et Fang (2013), les incidents critiques se réalisent à une fréquence beaucoup plus élevée que les accidents (100 fois la fréquence des accidents et 10 fois la fréquence des quasi-accidents) [61]. Le tableau ci-après expose le nombre d'événements moyen qu'un

conducteur peut faire face en route pour une vitesse moyenne de 60km/h et un kilométrage annuel de 20 000 km. Ces chiffres sont collectés auprès des données américaines et finlandaises [92]. Même si les accidents sont considérés comme des évènements rares par comparaison aux autres évènements, en moyenne un conducteur est impliqué dans un accident chaque 7 ans et demi ou chaque 150 000 km parcourus. Il s'agit alors d'une opportunité pour l'assureur d'identifier les conducteurs risqués avant même la réalisation des sinistres et de prendre des contre-mesures de sécurité proactives afin d'améliorer la sécurité des conducteurs risqués.

Tableau 10 Nombre d'évènements moyen d'un conducteur par unité de temps ou par km

Evènements	Fréquence (temps)	Fréquence (Distance parcourue)
Informations routières	5 en 1 seconde	300 par 1 km
Observations du conducteur	2 en 1 seconde	120 par 1 km
Décisions du conducteur	40 en 1 minute	40 par 1 km
Actions du conducteur	30 en 1 minute	30 par 1 km
Erreurs de conduite	1 en 2 minutes	1 par 2 km
Situations risquées	1 en 2 heures	1 par 120 km
Quasi-accidents	1 en un mois	1 par 2 000 km
Accidents	1 en 7,5 années	1 par 150 000 km
Domages corporels	1 en 100 ans	1 par 2 000 000 km
Accidents mortels	1 en 2 000 ans	1 par 40 000 000 km

Pour rapprocher les données du tableau 10 au contexte tunisien, le tableau ci-dessous compare le taux d'accident dans les trois pays (Tunisie, Finlande et les Etats- Unis) en 1991.

Tableau 11 Taux d'accident par 1000 habitants par pays en 1991

Pays	Taux d'accident / 1000 habitants
Finlande	9.374 ^[93]
Etats-Unis	20.45 ^[93]
Tunisie	9.411 ³⁸

Il s'avère que le nombre d'accidents en Tunisie est légèrement supérieur à celui de la Finlande sachant que la dernière est beaucoup moins peuplée que la Tunisie (environ 5 millions contre 8

³⁸ Source : Ministère de l'Intérieur, Observatoire national de la circulation, sub-unit of studies and analyses and Annuaire Statistique de la Tunisie.

millions d'habitants respectivement en 1991). Mais cela n'empêche que le tunisien développe un comportement beaucoup plus agressif qu'un finlandais ou encore un américain. En effet en 2015, l'organisation de la santé mondiale WHO, a souligné dans son rapport le degré de gravité du phénomène de la mortalité routière dans les pays africains y compris la Tunisie. Le tableau suivant expose le nombre de victimes d'accidents de circulation mortels par 100 000 habitants par année dans les trois susmentionnés.

Tableau 12 Les victimes de la route par 100 000 habitants par année par pays

Pays	Victimes d'accidents de route
Finlande	4.8
Etats-Unis	10.6
Tunisie	24.4

Dans le même rapport, les principales raisons de la mortalité sont l'excès de vitesse, l'inattention du conducteur et du piéton [22].

A titre de rappel, dans l'échantillon de véhicules à étudier, un seul accident a été détecté par le dispositif télématique. De ce fait, les quasi-accidents seront utilisés comme un substitut aux accidents. En effet, plusieurs études de recherche en évaluation de risque par les données télématiques notamment Klauer *et al* (2006, 2009, 2010) et Guo *et al* (2010,2013) utilisent les quasi-accidents conjointement avec les accidents pour l'évaluation du risque [61] [94] [95] [96] [94].

2.3. Traitement de données

Le boîtier OBD-II ainsi présenté dans le premier chapitre est désigné pour surveiller la performance du véhicule et d'accéder aux informations GPS, aux informations d'accéléromètre etc. Pendant la circulation du véhicule, les données comportementales (vitesse instantanée, état de mouvement du véhicule, accélération, régime moteur...) et les données géographiques (longitude et latitude) sont mises à jour chaque seconde.

Pour l'analyse des comportements des conducteurs, le regroupement des données brutes et l'extraction des variables prédictives est une étape primordiale.

Le volume énorme des données collectées du dispositif télématique ne peut pas être traité par des solutions classiques notamment Excel. Il s'agit de plusieurs centaines de milliers d'observations en seconde pour chaque véhicule sur une période variant entre 3 et 14 mois. La disparité des périodes

d'observation est due à l'abonnement asynchrone des conducteurs aux services télématiques. Chaque évènement détecté parmi les 46 évènements précités compte pour une observation. Un logiciel de Business Intelligence (Power BI de Microsoft) était le premier outil utilisé pour traiter les données de chaque véhicule à part par filtrage³⁹. Cette solution n'était pas efficace vu le nombre élevé d'observations et de véhicules d'où le recours au traitement informatique des données par programmation d'un script SQL et l'extraction de la base de données exploitable pour l'étude. Le traitement comprend l'exploration et l'agrégation des données pour créer de nouvelles variables utiles pour le modèle UBI.

Les données brutes traitées sont réduites d'environ 100 fois leur volume initial par regroupement des observations par trajet et par élimination des données inutiles pour l'assureur. Pour chaque trajet, la distance parcourue est calculée, le nombre de manœuvres avec les trois niveaux de sévérité est compté, la durée de conduite par trajet et par moment de journée est calculée. Enfin, les vitesses auxquelles le conducteur roule sont agrégées en fraction de durée de conduite, c'est-à-dire la durée de conduite fractionnée en paliers de vitesse prédéfinis.

De ce fait, l'exploitation des données est devenue plus simple par les outils classiques de statistique telle qu'Excel pour préparer la base de données finale.

2.4. Variables prédictives

Les variables prédictives du risque d'accident sont généralement les indicateurs d'accident révélés par les autorités de sécurité routière. Comme est évoqué dans le premier chapitre, les accidents de circulation ont de multiples causes, notamment humaines.

Un ensemble des causes des accidents de la circulation réparties en 21 catégories et constatées par la garde nationale et la police de circulation en Tunisie pendant l'année 2018 est relevé dans le graphique 6.

D'après ce graphique, on peut affirmer que les principales causes humaines d'accident sont l'inattention, l'excès de vitesse, l'imprudence des piétons et l'inobservation de la priorité. Ces quatre types d'infraction au code de la route représentent 66% des causes.

³⁹ Voir annexe n°3

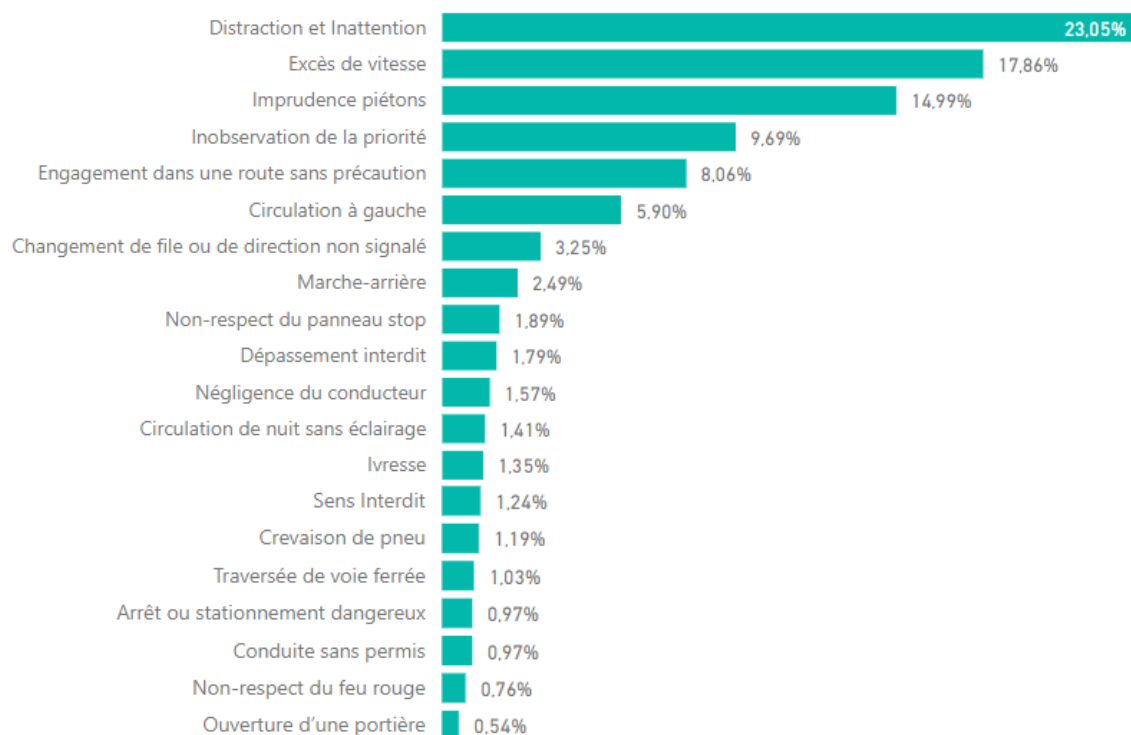


Figure 13 Causes des accidents de la circulation en Tunisie en 2018 (les cinq premiers mois) [97]

Cependant, tout accident de circulation résulte des circonstances de causes humaines, environnementales (infrastructure, conditions météorologiques...) et techniques (liées à l'état du véhicule).

En effet, les statistiques de la sécurité routière par l'ONSR illustrées dans le graphique 14 montrent que les accidents de circulation surviennent le plus en présence d'une « ambiance accidentogène » du trafic intense dans les rues, les avenues, les routes nationales, régionales et locales. La fin de semaine et les heures de pointe représentent un facteur de risque de survenance d'accidents selon la même source. Contrairement à la littérature, la conduite tard la nuit et à l'aube ne constitue pas un facteur de risque dans le contexte tunisien selon les statistiques de l'ONSR [97].

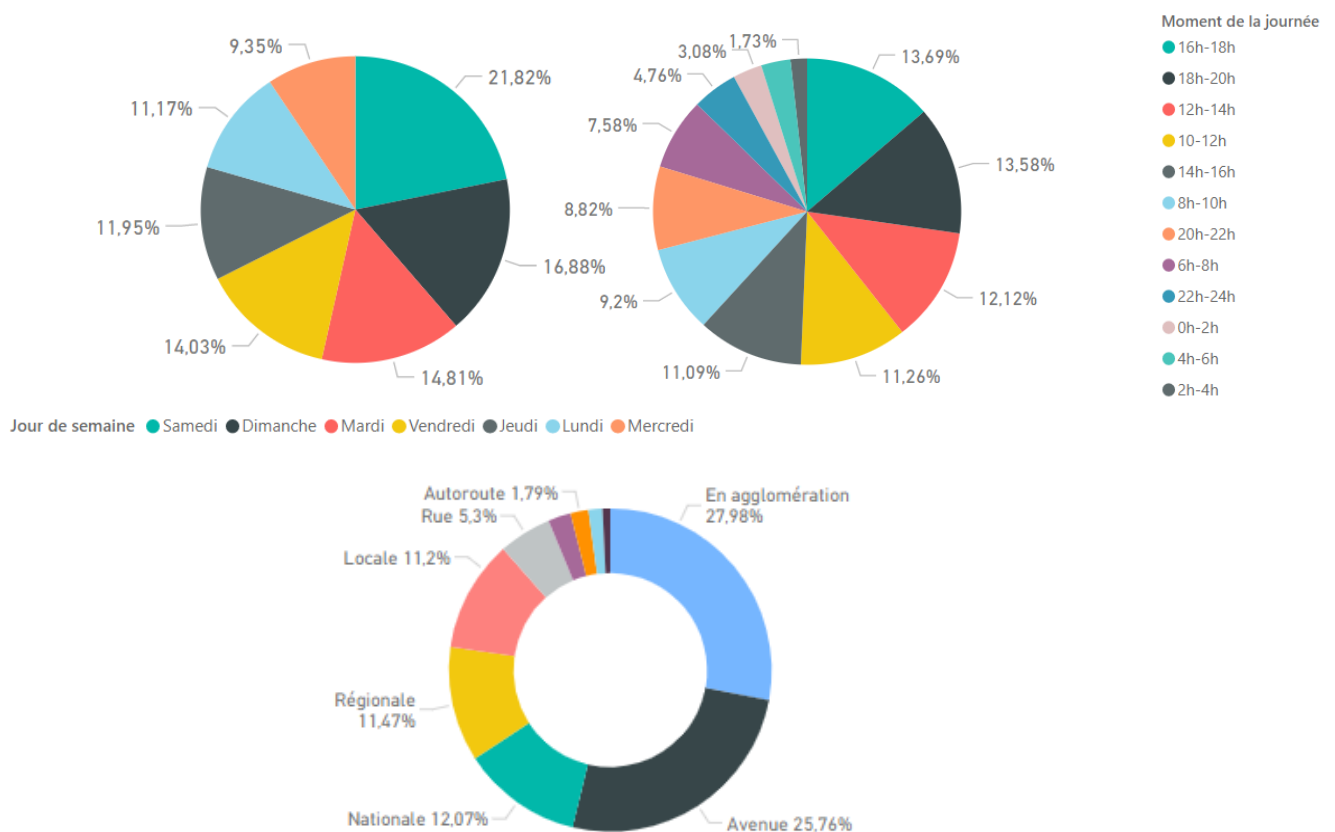


Figure 14 Nombre d'accidents observés en 2018 (5 premiers mois) en fonction du jour de semaine (à gauche), du moment de la journée (à droite) et selon le type de la route (en bas)

C'est la raison pour laquelle l'assureur devrait tenir compte de tous ces aspects dans la tarification de la couverture automobile.

Pour analyser le comportement des différents conducteurs, les variables dérivées des données brutes susceptibles de prédire le risque individuel de conduite sont présentées dans le tableau ci-dessous. Sont absentes du modèle adopté par Telematik, les variables « zone de circulation » et « type de route » à extraire des données brutes de localisation GPS ; La start-up a choisi d'écarter tout facteur autre qu'humain pour l'évaluation du risque automobile notamment la météo et la zone de circulation du véhicule pour se focaliser sur le phénomène pur et dur du comportement du conducteur.

Tableau 13 Liste des données télématiques agrégées

Catégorie de variable	Description	Type	Etiquette
Variable dépendante			
Quasi-Accident	Nombre de manœuvres très sévères d'évitement d'accident	Discrète	QA
Variables indépendantes			
Distance parcourue		Continue	Mileage
Durée totale des trajets		Continue	TT
Moment de la journée	Fraction d'exposition entre 6h et 18h	Continue	DD
	Fraction d'exposition entre 18h et 22h		ND
	Fraction d'exposition entre 22h et 2h		
	Fraction d'exposition entre 2h et 6h		
Vitesse	Fraction d'exposition entre 0 et 50km/h	Continue	Speed0
	Fraction d'exposition entre 50 et 90km/h		Speed1
	Fraction d'exposition entre 90 et 120km/h		Speed2
	Fraction d'exposition entre 120 et 150km/h		Speed3
	Fraction d'exposition plus que 150km/h		Speed4
Nombre de freinage fort et excessif		Discrète	Decc
Nombre d'accélération forte et excessive		Discrète	Acc
Nombre de virage fort et excessif		Discrète	Turn

Les variables télématiques intégrant le nouveau modèle d'assurance sont nouvelles et méritent d'être explicitées :

a. Quasi-accident

Les quasi-accidents ou *near-crashes* en anglais définis auparavant, seront la variable à expliquer comme substitut aux accidents observés durant la période de conduite. Il s'agit des manœuvres de freinage, de virage et d'accélération du 3^{ème} niveau de sévérité (brutal) enregistrées pendant la

période totale de conduite de chaque conducteur. A titre de rappel ; la sévérité est personnalisée selon la marque et le modèle du véhicule du conducteur. C'est-à-dire, pour chaque marque et modèle des véhicules, les 3 niveaux de sévérité des différents événements d'incidents critiques sont prédéfinis pour fiabiliser l'évaluation du risque et l'individualiser davantage.

b. Distance parcourue

La distance totale de tous les trajets parcourus par un conducteur durant la durée de conduite enregistrée par le boîtier. Un trajet est mesuré du moment où le moteur est allumé jusqu'au moment où il est arrêté. De ce fait, un conducteur peut effectuer une dizaine de trajets pendant la journée. L'unité de mesure est le kilomètre.

c. Durée totale des trajets

La durée totale de conduite par un véhicule surveillé pendant une période donnée. Elle est mesurée par la sommation des durées de chaque trajet effectué. L'unité de mesure est l'heure.

d. Moment de la journée

Le moment de la journée au cours duquel le conducteur a utilisé son véhicule. Cinq paliers d'horaire sont prédéfinis par Telematik à savoir :

- T0 : 6h-18h
- T1 : 18h-22h
- T2 : 22h-2h
- T3 : 2h-6h

La durée de conduite totale est répartie sur les paliers d'horaire dans lequel le véhicule a été utilisé. Chaque palier compte pour une variable à part. Afin d'éviter l'effet de multi colinéarité qui biaise les résultats du modèle à développer par la suite, un regroupement de variables fût nécessaire : une variable de durée de conduite en jour (Day Driving –DD-) et une variable de durée de conduite en nuit et à l'aube (Night Driving –ND-). La variable durée totale des trajets ne figurera pas parmi les variables du modèle par souci de multi colinéarité avec les variables durée de conduite fractionnée par palier d'horaire et durée de conduite fractionnée par palier de vitesse.

e. Vitesse

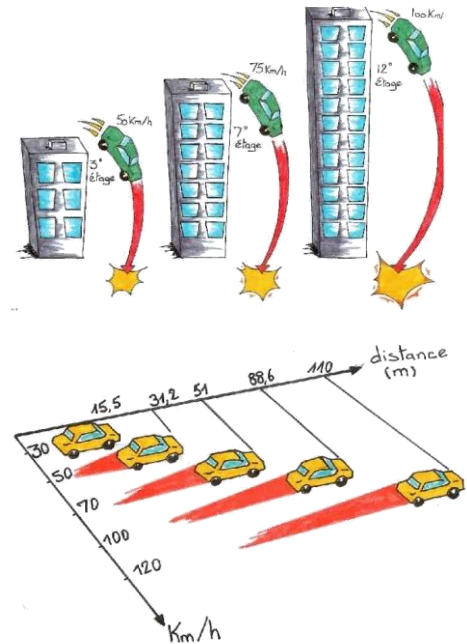
Vu que la vitesse moyenne par trajet ne peut pas être comparée à celle du trafic en temps réel faute d'un système de mesure du trafic routier, l'excès de vitesse ne peut pas être automatiquement détecté

et mesuré par le boîtier. D'où la fixation de cinq paliers de vitesse afin de calculer la fraction d'exposition en termes de durée de conduite pour chaque palier :

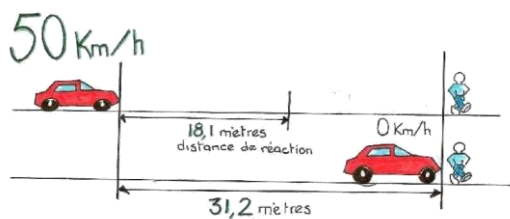
- Speed0 : 0-50 km/h
- Speed1 : 50-90 km/h
- Speed2 : 90-120 km/h
- Speed3 : 120-150 km/h
- Speed4 : +150 km/h

Autrement dit, il s'agit de la fraction de durée de conduite sous une vitesse moyenne qui se situe dans chacun de ces paliers de vitesse.

Les données traitées destinées à être utilisées dans le modèle d'assurance ne comprennent pas le premier palier de vitesse (entre 0 et 50 km/h) vu qu'il s'agit de la vitesse la plus récurrente durant l'exercice de conduite. En effet, l'excès de vitesse devient plus dangereux à partir d'une vitesse de 50km/h c'est-à-dire dans des zones hors agglomération. Le graphique à la droite expose le degré de sévérité de la collision selon la vitesse [98].



La vitesse excessive est parmi les causes les plus fréquentes d'accident. Le graphique à la droite illustre les distances d'arrêt en mètres qui dépend du temps de réaction et du temps de freinage. Plus la vitesse est grande, plus la distance d'arrêt augmente [98]. A une très grande vitesse, la distance de frein augmente et les manœuvres d'évitement d'accident s'accroissent et deviennent inefficaces (Voir graphique ci-dessous) [98].



f. Nombre de freinage brusque

Le freinage brusque est généralement effectué suite à une vitesse excessive par rapport au trafic routier. La haute fréquence de ce genre de manœuvres reflète le manque d'anticipation du conducteur. Plus la vitesse est excessive, plus le freinage est brutal ce qui augmente le risque de

dérapiage du véhicule qu'il soit en route en ligne droite ou en virage. La perte de contrôle du véhicule conduit à un accident inévitable.

Cette variable mesure le nombre de freinage brusque avec les deux niveaux de sévérité (forte et excessive) détecté durant toute la période de conduite.

g. Nombre d'accélération rapide

L'accélération d'un véhicule est principalement liée à son poids et à la puissance maximale de son moteur. Les capacités d'accélération sont considérées comme facteur de sécurité pour les conducteurs notamment lors des dépassements. En situation normale de circulation, il faut en moyenne environ entre 3 à 5 secondes pour passer de 50 km/h à 90km/h. Un véhicule avec une forte capacité d'accélération stimule la prise de risque souvent traduite par un dépassement dangereux au-delà de la vitesse maximale autorisée [99]. Une accélération rapide inadaptée au trafic suscite un dépassement des capacités d'accélération et donc la baisse de la réactivité du conducteur face à une situation de conflit et augmente le risque d'accident.

Cette variable reflète le nombre de fois où le conducteur a effectué des coups d'accélération rapides avec les deux niveaux de sévérité (forte et excessive) pendant la durée totale de conduite.

h. Nombre de virage rapide

Lorsque le virage gauche ou droit est pris à une vitesse élevée, le conducteur aura du mal à anticiper sa sortie voire même perdre le contrôle du véhicule avec un risque de dérapage. Dans les intersections et les ronds-points, ne pas respecter la priorité avec la prise d'un virage rapide augmente le risque de collision.

Cette variable mesure le nombre de fois de prise d'un virage rapide avec les deux niveaux de sévérité (fort et excessif) pendant la durée totale de conduite.

Pour rendre les différents comportements des conducteurs comparables et améliorer la modélisation, toutes les variables précitées seront normalisées en les divisant par le nombre de mois de conduite spécifique à chaque conducteur. Cela permet de calculer une moyenne mensuelle de chaque variable comme est présenté dans le tableau suivant :

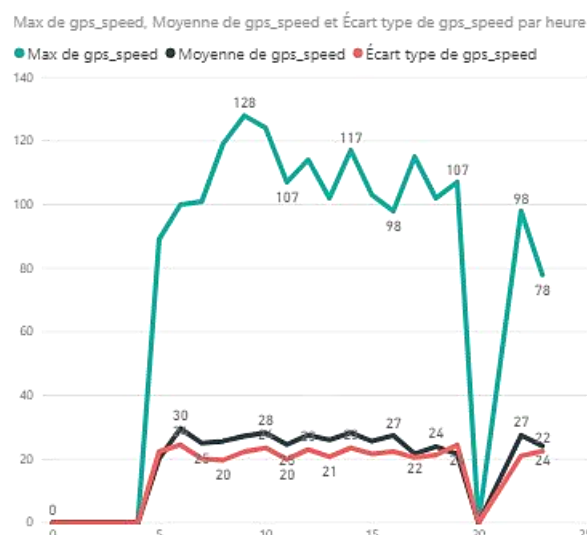
Tableau 14 Liste finale des variables du modèle UBI à étudier

variable dépendante	
QA	Nombre mensuel moyen de manœuvres très sévères niveau 3 (freinage + accélération + virage)
Variables indépendantes	
Acc	Nombre mensuel moyen d'accélération de sévérité niveau 1 et 2
Decc	Nombre mensuel moyen de freinage de sévérité niveau 1 et 2
Turn	Nombre mensuel moyen de virage de sévérité niveau 1 et 2
Mileage	Distance moyenne mensuelle parcourue
Speed1	Durée de conduite moyenne mensuelle à une vitesse entre 50 et 90 km
Speed2	Durée de conduite moyenne mensuelle à une vitesse entre 90 et 120 km
Speed3	Durée de conduite moyenne mensuelle à une vitesse entre 120 et 150 km
Speed4	Durée de conduite moyenne mensuelle à une vitesse supérieure à 150 km
DD	Durée de conduite moyenne mensuelle en jour
ND	Durée de conduite moyenne mensuelle en nuit et à l'aube

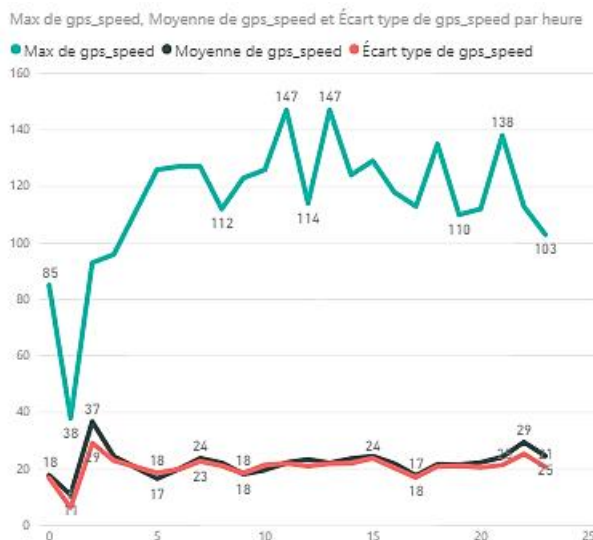
2. Exemple d'analyse du comportement de deux conducteurs

Avant de passer à l'analyse des données réduites, deux véhicules témoins avec leurs données brutes ont été choisis pour démontrer la probabilité de liaison entre le comportement au volant et le risque individuel de conduite.

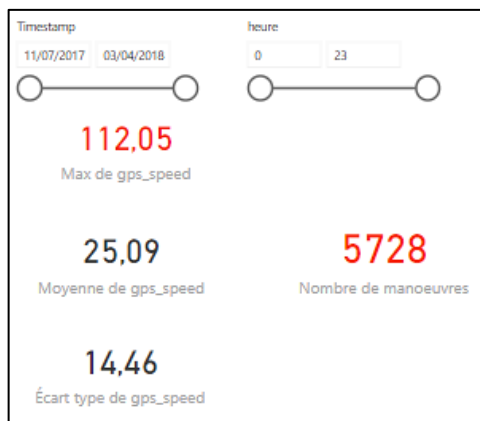
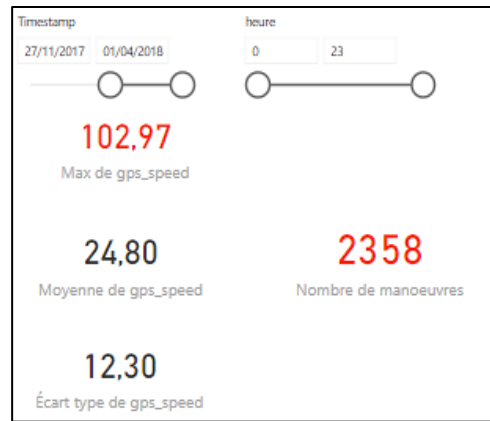
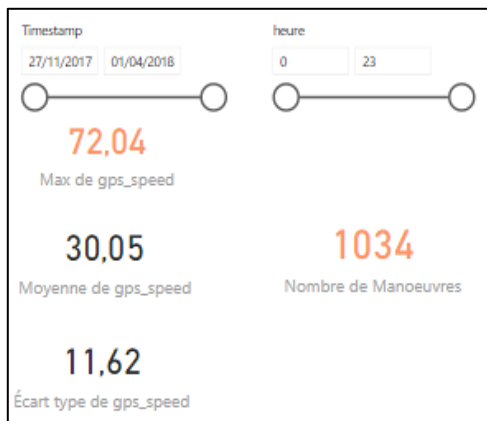
Les observations de vitesse mesurée par GPS du premier véhicule montrent que le conducteur a enregistré une vitesse maximale de 128km/h à 9h du matin. Les plus hautes vitesses sont enregistrées entre 7h et 19h. Pour les contextualiser, la carte ci-dessous retrace l'itinéraire du véhicule en roulant à une haute vitesse (+100 km/h) en parcourant les routes nationales et régionales.



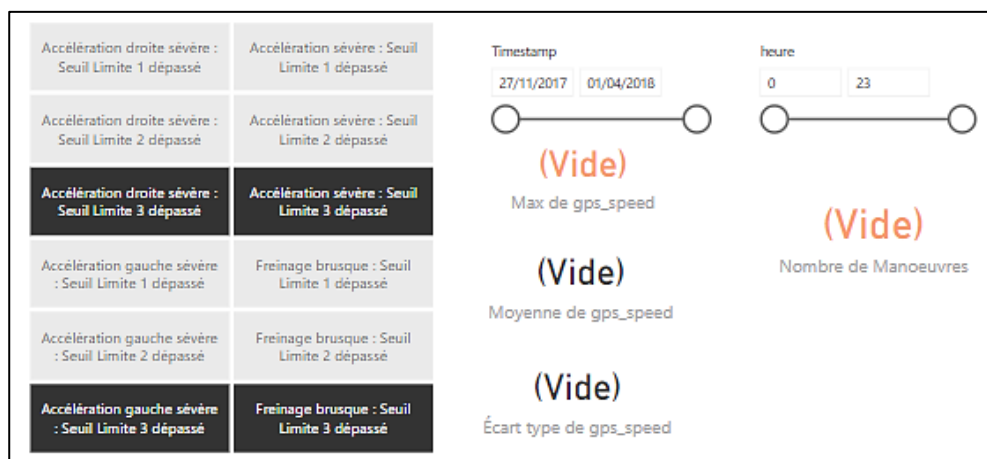
Un deuxième conducteur qui se déplace, à première vue, dans un itinéraire similaire à celui du précédent conducteur à une vitesse supérieure à 100 km/h, enregistre un maximum de vitesse bien plus élevé soit 147 km/h. Il semble a priori, que ce conducteur adopte un comportement au volant plus dangereux.

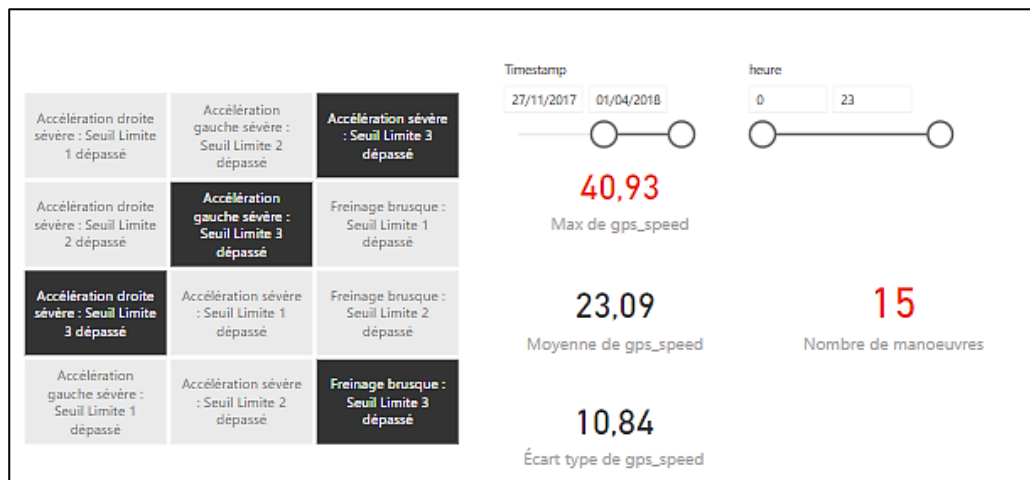


Pour confirmer cette déduction, une analyse des manœuvres dangereuses (Evènements d'incidents critiques) détectées par le boîtier est utile. En fixant la même période de conduite (4 mois et 4 jours), le deuxième conducteur (à la droite) effectue 2358 manœuvres d'accélération, de freinage et de virage avec les deux niveaux de sévérité soit plus que deux fois le rythme du premier conducteur. Il est probable que l'excès de vitesse contribue au comportement agressif au volant. La conduite en nuit peut être également un facteur déterminant du degré d'agressivité au volant. Comme est montré dans les graphiques suivants, le deuxième conducteur effectuant plus de manœuvres dangereuses a conduit à plusieurs reprises dans la nuit. Prenant compte de toute la période au cours de laquelle les observations de conduite du deuxième conducteur sont disponibles (8 mois et 22 jours), le nombre de manœuvres dangereuses s'élèvent à 5728. A priori, plus le conducteur est exposé à l'exercice de conduite (en termes de distance parcourue ou de durée de conduite), plus le risque d'accident augmente.

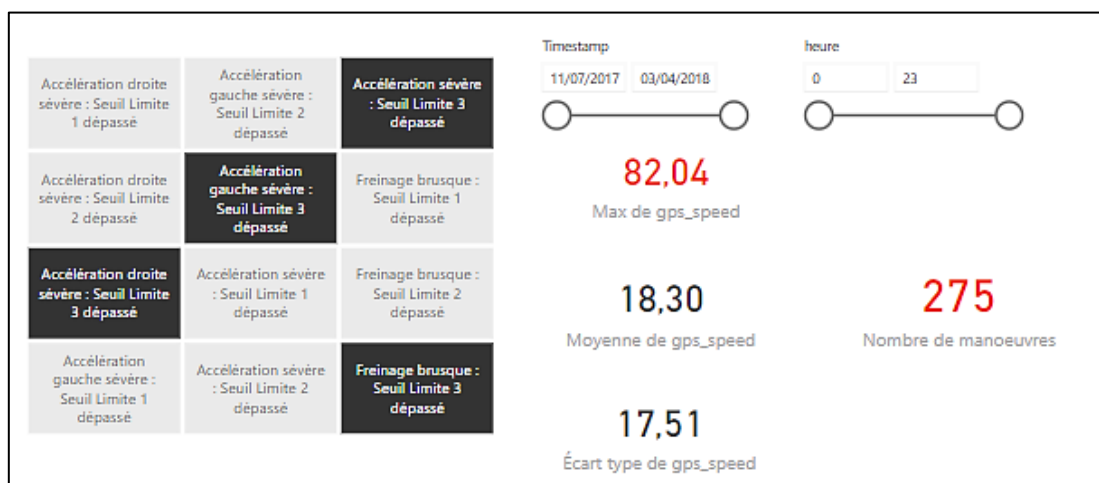


En étudiant les évènements de quasi-accident (manœuvres de 3^{ème} niveau de sévérité), le premier conducteur (le plus prudent) n'enregistre aucun évènement dans ce sens. Par contre, le conducteur le plus dangereux enregistre 15 évènements pendant la même période de conduite.





Prenant compte de toute la période de conduite enregistrée du deuxième conducteur, il est constaté que le nombre d'événements de quasi-accidents augmente pour atteindre 275 événements. Cela confirme qu'un conducteur ayant un comportement dangereux est d'autant plus exposé au risque d'être impliqué dans un accident qu'il est exposé à l'activité de conduite.



A première vue, il est probable qu'une liaison entre le comportement du conducteur et le risque d'accident puisse exister. C'est alors que l'objectif de cette étude est de montrer la corrélation entre le risque d'être impliqué dans un accident et les événements d'incidents critiques reflétant le comportement d'un conducteur.

Le « comportement routier » -introduit dans l'introduction du chapitre- est une étude expérimentale des comportements observables ayant l'objectif de prédire et contrôler le comportement au volant. La prochaine section est consacrée à la présentation de la technique de « comportement » traduite en méthodes statistiques.

Section 3 Méthodologie adoptée

A travers l'application des méthodes purement statistiques, cette étude a pour but d'établir les éléments de comportements permettant à l'assureur de prédire directement le niveau du risque de son client à travers son comportement au volant et de le contrôler s'il s'agit d'un élément dangereux sur la route. Les données ont été saisies et analysées grâce au logiciel SPSS 20.0.

3.1. Etude descriptive

A travers cette étude, les moyennes, les écarts type, les médianes, l'étendue des valeurs extrêmes (min et max) et les pourcentages des variables ont été calculé.

3.2. Etude analytique

Pour l'étude d'une part de corrélations entre la variable à expliquer (QA) et les dix variables explicatives et d'autre part d'interdépendance entre les variables explicatives, une analyse bivariable a été effectuée. Le test de corrélation a pour but de déterminer l'absence ou la présence d'une relation linéaire significative entre les variables. L'étude de la liaison entre deux variables quantitatives (par exemple nombre mensuel moyen de freinage et durée mensuelle moyenne de conduite en nuit) a été effectuée grâce à l'utilisation du coefficient de corrélation de Pearson (r) calculé ainsi :

$$r = \frac{COV(x,y)}{S_x S_y}$$

Cov(x,y) : Covariance entre la variable x et la variable y

S_x : Ecart type pour la première variable

S_y : Ecart type pour la deuxième variable

La corrélation permet de quantifier le degré de linéarité entre des variables continues. Les variables peuvent être :

- associées positivement ($r > 0$) : plus le nombre de freinage augmente, plus le nombre de quasi-accidents augmente;
- associées négativement ($r < 0$) : plus le nombre de freinage augmente, plus le nombre de quasi-accidents diminue;
- non associées ($r = 0$) : le nombre de freinage n'a aucune influence linéaire sur le nombre de quasi-accidents

La force de l'association est déterminée comme suit :

- Coefficient de +1 : corrélation positive parfaite entre les deux variables : l'augmentation se fait dans les mêmes proportions
- Coefficient de -1 : corrélation négative parfaite entre les deux variables : la diminution se fait dans les mêmes proportions
- Coefficient proche de 0 : corrélation faible entre les deux variables
- Coefficient proche de +1 ou -1 : corrélation forte entre les deux variables

Plus la valeur de r est proche de 0 plus grande sera la variation autour de la droite d'ajustement au total. Une valeur de +1 ou -1 signifie que tous les points de données sont exactement sur la droite d'ajustement au total et aucune déviation n'est affichée. Si le nuage de points de données ne forme pas une ligne droite plutôt une tendance sinusoïdale, la relation est non linéaire.

A travers les balises de Cohen (1988), la valeur du coefficient peut être facilement interprétée en dégageant la taille de l'effet d'une variable sur l'autre :

Autour de 0,10	effet de petite taille	corrélacion faible
Autour de 0,30	effet de taille moyenne	corrélacion moyenne
Plus de 0,50	effet de grande taille	corrélacion forte

Pour déterminer le degré de signification d'une relation, une valeur de p (p -value) associée au coefficient de corrélation de Pearson est calculée. Son calcul repose sur le calcul d'une valeur y associée (valeur t) dont la formule est la suivante :

$$t = r \sqrt{\frac{n-2}{1-r^2}}$$

Seule la valeur de p est affichée dans les résultats du logiciel SPSS. La valeur de p est déterminée à un intervalle de confiance de 95%. Cet intervalle permet d'établir la marge d'erreur à 5% entre les données de l'échantillon et celles de la population totale et d'évaluer la précision d'estimation d'un paramètre statistique dans l'échantillon.

Dans tous les cas, le coefficient de corrélation doit être significatif (la valeur de p doit être inférieure à 0,05). Si le coefficient est non significatif, on considère qu'il est similaire à $r = 0$, autrement dit, aucune relation linéaire significative ne peut être considérée.

Pour l'étude multivariée, une régression linéaire multiple a été développée vu que la variable dépendante est quantitative. Cette régression permet d'expliquer (ou de prédire) la variance d'une variable dépendante à l'aide d'une combinaison linéaire de facteurs explicatifs (variables indépendantes) en généralisant l'équation algébrique utilisée dans le module sur la régression simple. En d'autres termes, la régression multiple permet d'identifier les déterminants de la variable Y.

L'équation de régression linéaire multiple est représentée comme suit :

$$Y_i : (b_0 + b_1X_1 + b_2X_2 + \dots + b_kX_k) + \varepsilon_i$$

Il s'agit de la généralisation de l'équation de régression simple suivante :

$$Y_i : (b_0 + b_1X_1) + \varepsilon_i$$

Chaque valeur de la variable dépendante qu'est l'observation i peut être expliquée en partie par le modèle. La partie non expliquée par le modèle est l'erreur spécifique associée à cette valeur.

Chaque coefficient bêta (b) est associé à la variable indépendante qui sous sa forme standardisée correspond à sa contribution relative dans le modèle. La constante (b_0) correspond à la valeur de la variable dépendante lorsque toutes les variables indépendantes égalent 0.

Pour évaluer le modèle, un indice de corrélation multiple R^2 est calculé et qui représente le pourcentage de variance expliquée par le modèle ou encore par la combinaison des variables indépendantes.

L'hypothèse nulle est qu'il n'y a aucune relation entre la variable dépendante (Y) et la combinaison des variables indépendantes ($X_1, X_2, \dots, X_k$). Ce qu'on cherche est l'existence d'une relation significative. Pour fiabiliser le test de cette hypothèse, certaines hypothèses doivent être respectées à savoir [90] :

- Les **types de variables à utiliser** : Indépendantes : continue ou catégorique; Dépendante : continue
- **Pas de variance égale à zéro** : la distribution ne doit pas être constante. Il faut qu'une certaine variance existe.
- **Aucune multi-colinéarité parfaite** : il ne doit pas y avoir de relation linéaire parfaite entre deux ou plusieurs variables indépendantes. De ce fait, si les corrélations entre celles-ci sont trop fortes ($r > 0,5$), le problème de colinéarité peut se poser. Cette hypothèse peut être

vérifiée avec le VIF (Variance Inflation Factor) indiquant si une variable indépendante a une relation linéaire forte avec les autres. Généralement, une valeur de cet indice plus grande que 10 indique la présence d'un tel problème. Pour plus de prudence, une valeur de VIF supérieure à 5 devrait attirer l'attention à la présence d'une forte colinéarité dans le modèle.

- **Homosédasticité (homogénéité des variances des résiduels) :** la variance des valeurs résiduelles doit être identique à tous les niveaux de la variable indépendante.
- **Pas de corrélation entre les variables indépendantes et les variables externes :** toutes les variables susceptibles d'influencer la variable dépendante doivent toutes être incluses dans le modèle.
- **Distribution normale des résiduels :** Il est important que les résiduels suivent une distribution normale contrairement aux variables indépendantes. Les résiduels doivent alors, avoir une moyenne de 0 et la majorité des valeurs doivent être rapprochées.
- **Relation linéaire entre les variables indépendantes et la variable dépendante**

Pour la conception d'un modèle fiable, il faut qu'il soit composé de la combinaison parfaite des variables indépendantes tout en respectant les hypothèses précédentes. Cohen (1992) a démontré que le nombre d'observations détermine la quantité maximale de variables qu'un modèle peut supporter [100]. L'inclusion d'un nombre élevé de variables pourrait avoir un impact négatif sur la qualité de régression. Une sélection fine des variables explicatives fût alors nécessaire.

Six modèles de régression ont été construits par des méthodes différentes définis par la suite. L'inclusion des variables indépendantes dans les modèles de régression était faite sur la base des critères et des seuils de significativité prédéfinis par chaque méthode.

Pour tous les tests statistiques, le seuil de signification « p » a été fixé à 0,05.

L'inclusion des variables indépendantes dans les modèles de régression était faite lorsque leur degré de significativité était inférieur à 0,1.

Quatre méthodes de régression ont été effectuées par la procédure REGRESSION dans SPSS:

- 1- Entrée forcée
- 2- Régression hiérarchique
- 3- Backward Stepwise (méthode pas à pas descendante)
- 4- Stepwise (pas à pas)

Tests et critères de choix du modèle final

Parmi les six modèles construits, on choisit un modèle à l'aide des tests et des critères suivants :

- **Test F de Fischer-Snédecor** : d'analyse de la variance paramétrique (ANOVA à un facteur) : Un test permettant de décider si on rejette ou ne rejette pas $H_0: M_{bi} = 0$ Tel que M_{bi} est la matrice des estimateurs du modèle.
 - Rejeter H_0 , c'est décider que $M_{bi} \neq 0$, c'est-à-dire qu'au moins un estimateur dans M_{bi} n'est pas nul.
 - Ne pas rejeter H_0 , c'est que toutes les composantes de M_{bi} sont nulles
 Si la valeur de F est grande, H_0 est rejetée. C'est à dire, le modèle a un intérêt explicatif.
- **Test de student ou F-partiel** : Un cas particulier du test de Fischer qui teste la nullité individuelle d'un estimateur associé à une variable explicative afin de tester l'influence de chaque variable explicative sur la variance de la variable dépendante.
- **R^2 ajusté** : pour la comparaison de la qualité d'ajustement des modèles construits indépendamment du nombre de variables y incluses. Cet indice peut diminuer en ajoutant au modèle d'autres variables contrairement à l'indice de R^2 . En effet, la valeur de R^2 est souvent trompeuse car elle augmente en fonction en nombre de variables ajoutées au modèle. Plus le nombre de prédicteurs est important, plus est grand le risque de modéliser le bruit aléatoire dans les données, source de sur-ajustement. Le modèle devient inutilement complexe, peu pertinent, difficile à expliquer et peu robuste.
- **Critère d'information d'Akaike (AIC)** : pour la comparaison de la qualité des modèles tout en pénalisant le sur apprentissage de données par l'inclusion d'un nombre important de variables. Le modèle ayant la plus faible valeur d'AIC est retenu. Sa formule de calcul est la suivante :

$$AIC = n \times \log (SSR/n) + 2 \times (k+1)$$

Tel que n : nombre d'observations ; SSR : somme des carrées des erreurs ; k : nombre de prédicteurs

Il est à noter qu'un bon modèle sera constitué de variables ayant une pertinence théorique et qui expliquera une bonne proportion de la variance de la variable dépendante.

Les méthodes de sélection de variables

- **Régression par la méthode *entrée forcée*** : Toutes les variables sont entrées au même moment et le test global de Fischer (test F) évalue le modèle. Le choix des variables à inclure se base sur la théorie de l'assurance comportementale. Par contre, le modèle évalue leur effet combiné indépendamment de leur ordre d'entrée.

- **Régression par la méthode de *régression hiérarchique***: L'ordre d'entrée des variables est déterminé à travers la création des blocs de variables entrés de manière hiérarchisée dans le modèle. cette méthode permet de suivre en détail le comportement du modèle face à l'ajout/suppression des variables. Les résultats indiquent l'apport de chaque bloc en termes de pourcentage de variance expliquée (R^2).
- **Régression par la méthode *Backward Stepwise*** : Les étapes de l'algorithme sont les suivantes:
 - Calculer l'estimateur des moindres carrés pour le modèle incluant toutes les variables explicatives.
 - Effectuer un test de Student pour chacune des variables explicatives. Deux cas de figure se présentent :
 - Les variables sont toutes significatives. Ce modèle est alors retenu. L'algorithme s'arrête.
 - La variable la moins significative et qui ne contribue pas significativement à l'amélioration du R^2 est retirée du modèle. Le processus se recommence avec une variable en moins jusqu'à l'obtention d'un modèle sans aucune variable non significative.

Le seuil de significativité pour l'entrée de variables est 0,05 et celui de la suppression est de 0,1.



Inconvénient de la méthode : Une variable une fois retirée ne peut pas être réintroduite. Ce problème est résolu par la procédure Stepwise précédemment définie.

- **Régression par la méthode *Stepwise*** : cette méthode permet de retenir le modèle composé des variables explicatives les plus corrélées avec la variable dépendante et les moins corrélées entre elles. Le choix de la première variable est basé sur la corrélation la plus élevée et celui des variables suivantes sur la corrélation partielle. L'algorithme consiste à introduire les estimateurs un par un dans la régression et à ne conserver que ceux qui sont les plus significativement associés avec la variable dépendante. Les autres variables explicatives peu significatives sont alors retirées de la régression. De cette manière, il est possible d'éliminer les variables redondantes, source de colinéarité. Le seuil de significativité pour l'entrée de variables est 0,05 et celui de la suppression est de 0,1.

Section 4 Résultats

4.1. Etude descriptive

Le tableau ci-dessous expose les statistiques descriptives de l'ensemble de variables indépendantes où Q1 est le premier quartile (25%), Q2 est le deuxième quartile (médiane), et Q3 est le troisième quartile (75%).

Tableau 15 Statistiques descriptives de l'échantillon

Variable	Max	Min	Moyenne	Q1	Q2 (Médiane)	Q3	Ecart-type
QA	31	0	8	1	4	11	9
Acc	297	0	46	6	30	62	58
Decc	1491	23	434	160	320	610	342
Turn	1323	10	313	115	245	443	273
Mileage	19653,80	438,66	3728,34	1729,08	2940,84	4212,02	3372,70
Speed1	50,14	0,70	15,16	8,82	15,15	18,46	9,43
Speed2	20,56	0,01	4,59	0,99	3,50	6,22	4,42
Speed3	6,80	0,00 ⁴⁰	1,04	0,05	0,30	1,91	1,51
Speed4	0,51	0,00	0,05	0,00	0,00	0,01	0,11
DD	125,28	19,55	70,79	53,91	68,59	93,07	26,84
ND	51,98	0,75	19,29	9,37	17,59	27,39	13,19

L'échantillon étudié se compose de 48 conducteurs dont l'âge varie de 25 à 67 ans, le sexe masculin prédomine et l'usage du véhicule varie entre privé et professionnel. Les situations de quasi-accidents (QA) varient de 0 à 31 situations en moyenne chaque mois. Les 25% premiers conducteurs enregistrent de 0 à 1 QA et les 25% derniers enregistrent 11 QA et plus en moyenne par mois. Le nombre mensuel médian de QA est 4. Le rythme de freinage brusque et de virage rapide est beaucoup plus élevé que celui d'accélération (environ 19 fois le nombre médian d'accélérations et 141 fois le nombre médian de QA).

Les conducteurs de l'échantillon adoptent un comportement assez homogène en termes de distance parcourue (Mileage)⁴¹; 87.5% de la population ont enregistré moins de 5000 km en moyenne chaque mois. Concernant le moment de conduite, tous les conducteurs roulent la nuit et à l'aube (au minimum 0,75h en moyenne par mois). Le maximum enregistré est 52 heures en moyenne par mois en contrepartie 60% de la population roule dans la nuit pendant moins de 20h en moyenne chaque mois.

⁴⁰ 0,00 est très proche de zéro pour toutes les valeurs enregistrées. Autrement dit, il s'agit de valeur en minute ou en seconde plutôt qu'en heure.

⁴¹ Voir annexe 8

La conduite en jour varie de 20 heures à 125 heures du fait de l'usage du véhicule (privé/professionnel).

Quant à la vitesse, les conducteurs adoptent un comportement relativement proche surtout pour les vitesses supérieures à 120km/h (Speed 3 et 4). La durée de conduite sous les paliers de vitesse entre 50 et 120 km/h (Speed 1 et 2) est relative à la distance parcourue qui elle-même, est liée à l'usage du véhicule. Par conséquent, elle enregistre des écarts-type nettement supérieurs à ceux de la durée de conduite sous les autres paliers de vitesse (Speed 3 et 4).

Le tableau suivant expose la répartition des conducteurs en fonction de la durée de conduite sous une vitesse supérieure à 150km/h. Il s'agit d'environ 40% des conducteurs qui roulent à une très haute vitesse. De plus, 23% des conducteurs y roulent avec une durée moyenne mensuelle supérieure à une minute allant jusqu'à trente minutes pour certains. En moyenne, 11 parmi 48 conducteurs roulent à une vitesse supérieure à 150km/h pendant 12 minutes chaque mois.

Tableau 16 Répartition de la populaion selon la durée de conduite sous une vitesse supérieure à 150 km/h

Durée de conduite sous une vitesse+150km/h (en seconde)	Effectifs	Pourcentage (%)
0	29	60,4
< 60s	8	16,7
> 60s	11	22,9
Total	48	100,0

D'après le tableau 9, la vitesse entre 120km et 150km est plus récurrente dans la population soit 91,7% avec une durée qui varie de quelques secondes à un maximum d'environ 7 heures de conduite en moyenne chaque mois. En contrepartie, 4 conducteurs seulement ne dépassent pas la vitesse de 120 km/h lors de l'exercice de conduite.

Tableau 17 Répartition de la population selon la durée de conduite sous une vitesse entre 120 et 150 km/h

Durée de conduite sous une vitesse entre 120 et 150 km/h (en heure)	Effectifs	Pourcentage (%)
0	4	8,3
] 0 ; 1[	30	60,5
[1 ; 2[	5	10,4
[2 ; 3[	5	10,4
Plus de 3	5	10,4
Total	48	100,0

4.2. Etude analytique

a. Analyse bivariée

Une certaine relation linéaire entre la plupart des variables indépendantes susmentionnées et la variable dépendante (QA) est montrée dans le tableau récapitulatif de corrélations dans l'annexe 9. Parmi les dix variables indépendantes, quatre variables ne sont corrélées avec QA. Il s'agit de la distance parcourue, la conduite en jour et la durée de conduite sous les deux premiers paliers de vitesse. Le graphique 9 regroupe les 4 variables les plus significativement corrélées aux QA dont la relation est linéaire. La relation des variables « freinage » et « virage » avec QA est plus forte que celle des variables « accélération » et « conduite en nuit » malgré leur haute significativité ($p=0,000 < 0,05$). Par contre, la variable « conduite en jour » ne représente *a priori* aucune relation statistiquement significative avec les QA (r de $0,079 \approx 0$ et $p\text{-value} = 0,595$). Il est pareil pour la variable « distance parcourue » ($r = 0,009 \approx 0$ et $p\text{-value} = 0,952$). En vérifiant le graphique 10, **la relation entre la variable « distance parcourue » (mileage) et les QA n'est pas proportionnelle dans le sens qu'elle ne semble pas être linéaire**. Ce résultat est conforme aux résultats trouvés par les différentes études mentionnées dans le premier chapitre (Litman, 2007; Boucher *et al.* 2013). L'annexe 10 montre que la relation des variables « speed 3 et 4 » avec QA est statistiquement significative et clairement plus forte que celle des variables « speed 1 et 2 » ($r= 0,469$ et $0,353$ contre $0,132$ et $0,184$ respectivement).

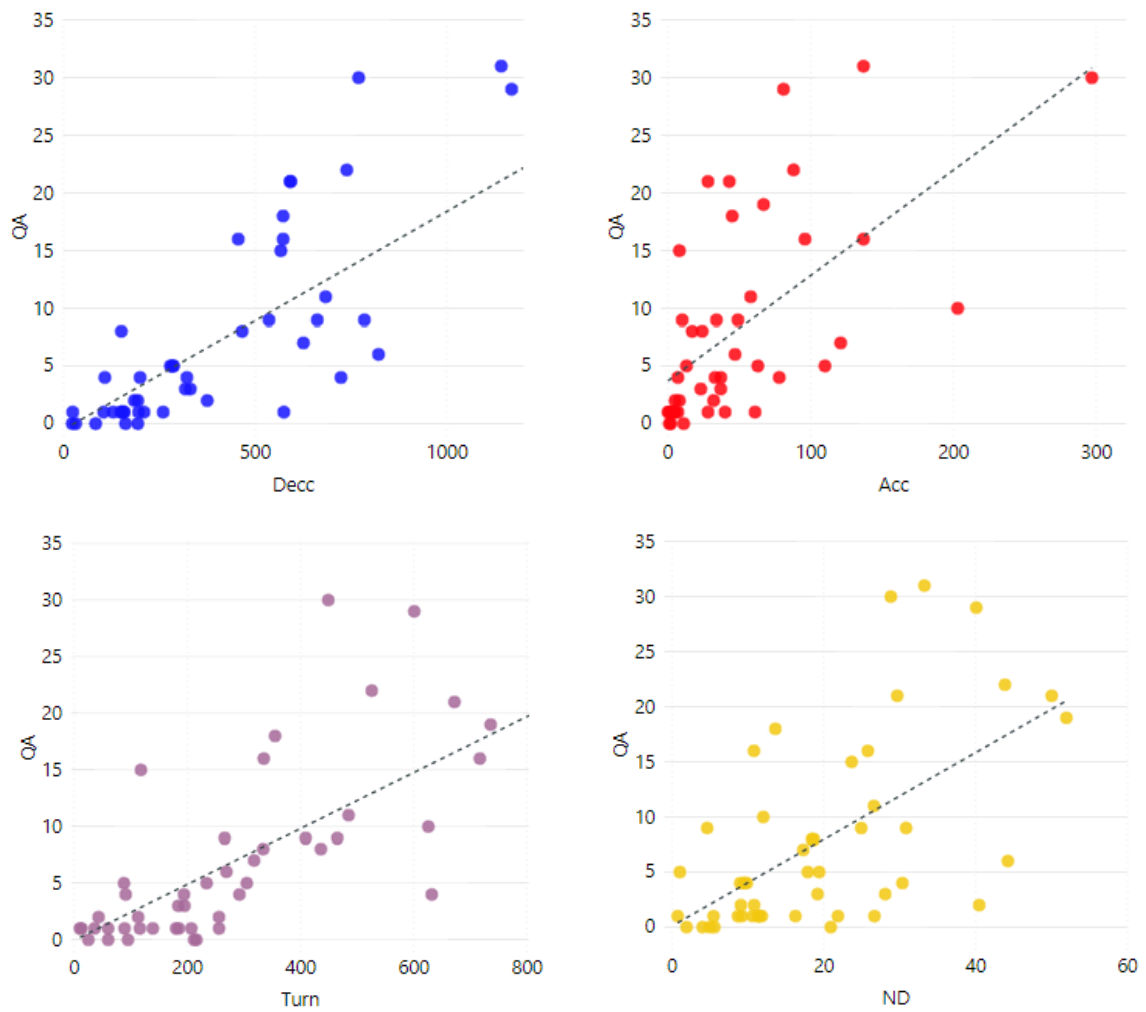


Figure 15 les quatre les plus fortes relations linéaires avec QA ⁴²

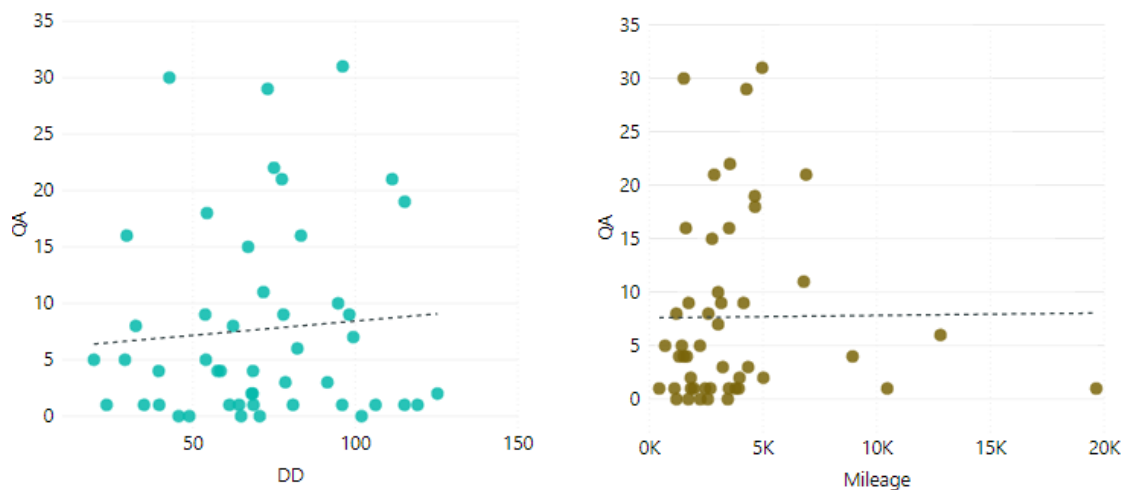


Figure 16 les deux les plus faibles relations avec QA ⁴³

⁴² Bleu : freinage ; Rouge : accélération ; Violet : virage ; Jaune : conduite en nuit

⁴³ Bleu-ciel : Conduite en jour ; Marron: Kilométrage

L'absence d'une corrélation entre une variable indépendante et QA n'affirme pas l'inexistence d'une relation entre elles. En fait, cette variable pourrait mieux expliquer QA en présence d'autres variables indépendantes. Le tableau 10 ci-dessous illustre cette affirmation par un exemple à l'appui.

Tableau 18 Variation du degré d'association de la variable «Speed1 » avec QA en présence d'autres variables

Variables de contrôle			QA
Aucune	Speed1	Corrélation	0,132
		Sig (2-tailed)	0,373
	Decc	Corrélation	0,749
		Sig (2-tailed)	0,000
	DD	Corrélation	0,079
		Sig (2-tailed)	0,595
	Mileage	Corrélation	0,009
		Sig (2-tailed)	0,952
Decc & DD	Speed1	Corrélation	0,360
		Sig (2-tailed)	0,014
Decc & DD & Mileage	Speed1	Corrélation	0,447
		Sig (2-tailed)	0,002

La conduite sous la vitesse entre 50 et 90km/h devient une variable statistiquement significative dès lors que les variables freinage et conduite en jour sont présentes. C'est-à-dire, cette variable, à elle seule, est incapable d'expliquer le phénomène de quasi-accident. D'après le tableau 18, en fixant l'effet de la conduite en jour et du freinage fort et excessif sur la réalisation des quasi-accidents, **plus la durée de conduite sous une vitesse entre 50 et 90km/h augmente, plus le conducteur est susceptible de réaliser des quasi-accidents dans 99,986% des cas mais avec un effet de taille moyenne.** Par contre, la relation entre la conduite sous une vitesse entre 50 et 90km/h et la conduite en jour est fortement significative. A cet effet, avec une forme d'association (Freinage, conduite en jour, distance parcourue, conduite sous une vitesse entre 50 et 90km/h), la relation des quasi-accidents avec la durée de conduite sous la vitesse entre 50 et 90km/h devient possible; un conducteur roulant en jour, ayant un rythme élevé de freinage sous une vitesse moyenne récurrente entre 50 et 90km/h est plus susceptible d'être impliqué dans un accident ($r = 0,360$) malgré l'absence d'une relation directe entre le freinage et la durée de conduite sous une vitesse entre 50 et 90km/h. Cette relation devient plus étroite en ajoutant à l'association la variable distance parcourue. ($r = 0,447$ avec $p\text{-value} = 0,002 \ll 0,01$). Cela sous-entend qu'une association d'un groupe pertinent de variables indépendantes affine la modélisation du « comportement routier ». Ces résultats pourraient

impliquer à part les conditions de congestion routière, l'adoption d'un comportement de talonnage⁴⁴ ou autre facteur à savoir l'usage du téléphone. Brookhuis et al. (1991) ont examiné la relation entre l'usage du téléphone et la performance de conduite de 12 conducteurs mesurée tous les jours de travail pendant trois semaines et ont déduit qu'il y a une augmentation significative du temps de réaction d'arrêt pour suivre un véhicule en ralenti, autrement dit l'augmentation du recours au freinage excessif. Ils ont trouvé également que l'usage du téléphone empêchent les conducteurs à baisser leur vitesse ce qui augmente le risque d'être impliqué dans un accident [101].

Le tableau 19 fait le zoom sur l'interdépendance significative des variables explicatives mesurée par le coefficient de corrélation.

Tableau 19 *L'interdépendance significative des variables explicatives mesurée par le coefficient de corrélation⁴⁵*

Variable	Acc	Decc	Turn	Mileage	Speed1	Speed2	Speed3	Speed4	DD	ND
Acc	1	0,612	0,520							
Decc	0,612	1	0,753				0,413		0,285	0,567
Turn	0,520	0,753	1		0,294		0,367	0,372		0,580
Mileage				1	0,683	0,432			0,522	0,297
Speed1			0,294	0,683	1	0,499			0,735	0,340
Speed2				0,432	0,499	1	0,473		0,363	0,368
Speed3		0,413	0,367			0,473	1	0,392		0,342
Speed4			0,372				0,392	1		
DD		0,285		0,522	0,735	0,363			1	0,361
ND		0,567	0,580	0,297	0,340	0,368	0,342		0,361	1

Les résultats tirés de ce tableau :

- Une forte corrélation entre les trois types de manœuvres (accélération, freinage, virage) avec une relation plus intense entre le freinage et le virage. Cette corrélation pourrait causer un problème de colinéarité.
- La distance parcourue augmente d'autant plus que le conducteur roule longtemps sous les paliers de vitesse [50-90km/h] et [90-120km/h].
- Pas de relation significative entre la distance parcourue et la durée de conduite sous les vitesses > 120km/h, autrement dit, l'excès de vitesse.

⁴⁴ Poursuivre un véhicule de très près

⁴⁵ Cellules vides : corrélation non significative ($p > 0,05$)

- Le rythme de freinage, d'accélération et de virage n'a pas de corrélation significative avec la distance parcourue ni avec la durée de conduite sous une vitesse entre 90 et 120km/h⁴⁶.
- L'accélération n'est pas en relation significative avec l'environnement spatio-temporel de la conduite ni avec le comportement d'excès de vitesse.
- La conduite en nuit et à l'aube est liée plus ou moins fortement à toutes les variables sauf à l'accélération.
- L'évènement de prise d'un virage rapide est susceptible de se produire en nuit plutôt qu'en jour et chez les conducteurs ayant un comportement net d'excès de vitesse (vitesse > 120km/h). La conduite sous une vitesse entre 50 et 90km/h pourrait avoir un effet sur le rythme de ce type d'évènement mais d'une taille moins importante.
- Le freinage fort et excessif ne s'effectue pas nécessairement en conduisant à une vitesse entre 50 et 120km/h ni à une vitesse > 150km/h. Par contre, la conduite sous une vitesse entre 120 et 150 km/h a un effet important sur le rythme de freinage.
- L'excès de vitesse (> 120km/h) est significativement corrélé à la conduite en nuit contrairement à la conduite en jour.
- La vitesse la plus récurrente pendant l'exercice de conduite en jour est celle entre 50 et 90km/h avec une association très forte ($r = 0,735$).

b. Analyse multivariée

Un modèle de régression linéaire multiple a été développé pour ressortir les éléments de comportements de la route. Afin d'améliorer sa précision et d'affiner les résultats de la régression, il est utile de prendre en compte de deux aspects majeurs :

- La relation non linéaire entre la distance parcourue et les QA
- La colinéarité entre les différentes variables indépendantes



Etude du modèle initial et sélection des variables explicatives

• Distance parcourue

Un modèle quadratique⁴⁷ est utile pour étudier l'effet non linéaire de la distance parcourue sur les QA. Comme est montré dans le graphique 17 ci-dessous, le modèle linéaire ne prend pas en compte l'effet non linéaire de la variable distance parcourue et il ne propose aucune relation linéaire avec les QA. Par contre, le modèle quadratique montre que seuls les 9 premiers kilomètres parcourus de

⁴⁶ Voir annexe 11

⁴⁷ Détail statistique dans les annexes 12, 13 et 14

chaque 1000 km enregistrés chez un conducteur constituent un facteur de risque des QA ainsi illustré par le graphique. A partir de ce seuil, le risque de réaliser des quasi-accidents diminue. Ce résultat est conforme aux conclusions de Janke (1991) ; Les conducteurs à bas kilométrage ont tendance à effectuer plus de trajets courts souvent dans les agglomérations alors que les conducteurs à un kilométrage élevé ont tendance à effectuer la majorité des trajets hors agglomération (autoroute, routes nationales...) moins risqués⁴⁸ [102].

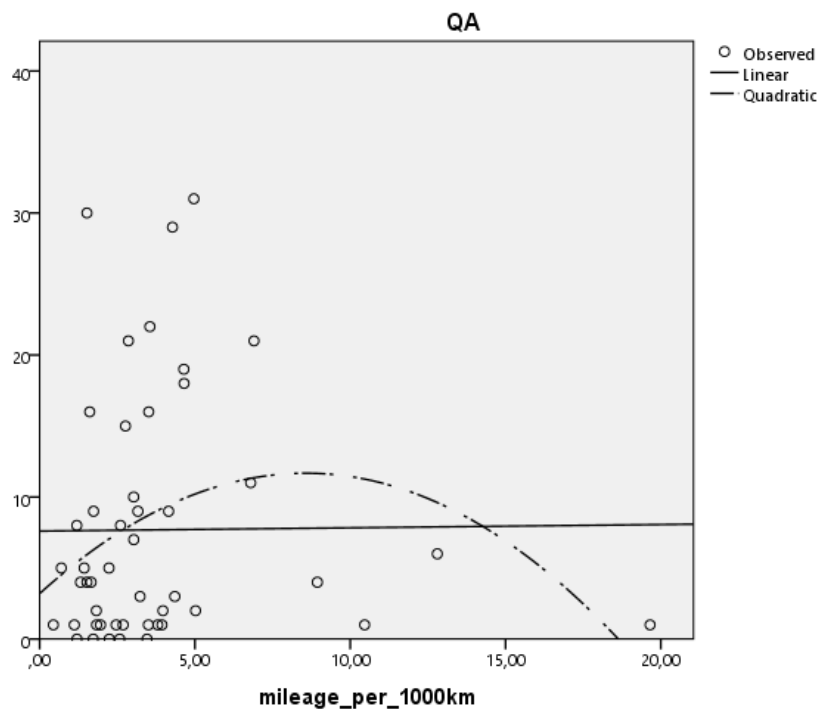


Figure 17 Modèles linéaire et quadratique associant le nombre des QA avec la distance parcourue / 1000km

- **Durée de conduite sous les quatre paliers de vitesse**

Le modèle initial construit est composé de toutes les variables explicatives potentielles afin d'examiner leur effet combiné sur le degré d'évaluation du risque automobile. A première vue, le modèle semble significatif avec un fort pouvoir explicatif ($R^2 = 79,5\%$)⁴⁹ mais les résultats affichés dans le tableau 21 expose quelques problèmes techniques :

- **Un biais de sélection** crée par la variable speed 4 (vitesse >150km/h)⁵⁰ d'où le regroupement des variables speed 3 et speed 4 pour créer une nouvelle variable (overspeed) mesurant la durée de conduite en excès de vitesse (vitesse > 120km/h).

⁴⁸ Conforme à la réalité de l'accidentologie en Tunisie illustrée dans la deuxième section

⁴⁹ Voir annexes 15 et 16

⁵⁰ Variable Speed 4 observée chez 40% de la population ; Variable Speed 3 observée chez 92% de la population

- **Un biais d'estimation des QA** crée par la présence de la forte colinéarité entre les variables indépendantes à savoir la distance parcourue et son carré, le freinage, le virage, la durée de conduite sous une vitesse entre 50 et 90km/h et la durée de conduite en jour (VIF> 3)
- **Un biais de confusion** influençant l'estimation crée par les variable speed 1, speed 2 et distance parcourue vu l'absence des conditions spatio-temporelle permettant de conditionner la relation entre les quasi-accidents d'une part et le kilométrage et la conduite sous les deux premiers paliers de vitesse [50 -90km/h] et [90-120km/h] d'autre part⁵¹.

- *Freinage, accélération et virage*

La taille peu importante de l'échantillon ne permet pas d'intégrer les variables [freinage, virage et accélération] ensemble dans le modèle conformément à la pratique des assureurs par risque de multi colinéarité. La combinaison de variables fortement corrélées peut aboutir à la perte de la significativité d'une variable ayant une forte relation avec les QA. Selon le tableau 20, à 5%, seules les variables conduite en excès de vitesse, conduite en nuit et accélération sont significatives.

Tableau 20 Evaluation des paramètres du modèle initial ⁵²

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
	B	Std. Error	Beta			Tolerance	VIF
(Constant)	,559	2,008		,278	,782		
Decc	,006	,004	,256	1,571	,125	,209	4,774
Acc	,041	,015	,274	2,667	,011	,527	1,897
Turn	,007	,005	,234	1,535	,133	,239	4,189
Mileage	-,001	,001	-,332	-,922	,362	,043	23,361
1 mileage_square	,026	,047	,183	,557	,581	,051	19,560
Speed1	,222	,158	,242	1,400	,170	,185	5,391
Speed2	-,054	,234	-,027	-,230	,819	,388	2,577
Overspeed	1,180	,535	,213	2,205	,034	,593	1,687
DD	-,071	,045	-,220	-1,589	,121	,289	3,465
ND	,169	,069	,259	2,446	,019	,497	2,014

a. Dependent Variable: QA

⁵¹ Similaire aux conclusions de Paefgen et al (2013) et Jensen et Wagner (2011) évoquées dans le premier chapitre

⁵² La couleur rouge : estimateur non significatif

Pour remédier à ce problème, le recours aux algorithmes mathématiques de sélection progressive des variables est utile afin de trouver la combinaison optimale améliorant la précision du modèle et minimisant le risque de son sur-ajustement. Bien que l'efficacité de ces méthodes ne puisse être contestée par la pratique, il ne serait pas acceptable de se contenter des résultats statistiques fournis par un algorithme. En effet, pour décider de supprimer ou d'intégrer une variable dans un modèle, il faut se référer à la logique, à la littérature et au bon sens.

Les résultats de l'algorithme développé *Backward* détaillé dans les annexes 17 conduisent à la sélection des variables intégrant le modèle final sur la base d'un a priori théorique. Une étude réalisée par Klauer *et al* (2010) a mis en évidence que les quasi-accidents sont causés principalement par le freinage brutal, la présence d'une activité secondaire⁵³, l'évitement inapproprié d'un obstacle et le talonnage⁵⁴. Le virage rapide a une faible présence parmi les causes des quasi-accidents. Il est alors, plus approprié de choisir la variable freinage plutôt que la variable virage choisie automatiquement par l'algorithme.

- *Durée de conduite en jour*

Les annexes 22, 23, 24, 25, 26 détaillent la démarche statistique pour l'étude de la variable durée de conduite en jour. Les résultats montrent que :

- La conduite en jour est un facteur statistiquement significatif de prévention contre le risque automobile lorsqu'elle est associée à la conduite sous une vitesse entre 50 et 90km/h et au recours fréquent au freinage brusque ($\beta = -0,133$ avec $p\text{-value} = 0,005 << 0,05$).
➔ Cette relation inverse entre la conduite en jour et les quasi-accidents est probablement expliquée par les conditions de congestion routière, le recours fréquent au freinage du fait du talonnage ou d'un facteur psychologique comme le stress et la nervosité à cause de la conduite en forte intensité de trafic ce qui fait baisser la probabilité d'avoir de situations de quasi-accidents.
- La conduite en jour perd son effet sur le risque des quasi-accidents ($p\text{-value de } \beta = 0,597 >> 0,05$) dès qu'elle est associée avec la conduite en excès de vitesse avec le recours fréquent à l'accélération rapide. En effet, conduire sous une vitesse supérieure à 120km/h est plus probable durant la nuit que durant le jour en prenant en compte du débit du trafic routier dans les autoroutes.

⁵³ Activité secondaire : Usage de téléphone, ajustement de la station radio, manger/boire, parler avec le passager...

⁵⁴ Voir annexe 17

- La variable conduite en jour, ajoutée au modèle final, n'a pas d'effet statistiquement significatif sur les quasi-accidents.

Les annexes 27 et 28 mettent en évidence la non significativité de l'effet des deux variables conduite en jour et distance parcourue, intégrées dans le modèle final, sur la prédiction des quasi-accidents.

Evaluation du modèle final retenu

Les variables finalement retenues dans le modèle sont :

- Nombre de freinage
- Nombre d'accélération
- Durée de conduite en nuit
- Durée de conduite en excès de vitesse

Cette combinaison permet d'expliquer 71% de la variance des QA. En effet, une bonne proportion de la variance des QA est expliquée par ce modèle. (Voir graphique 18).

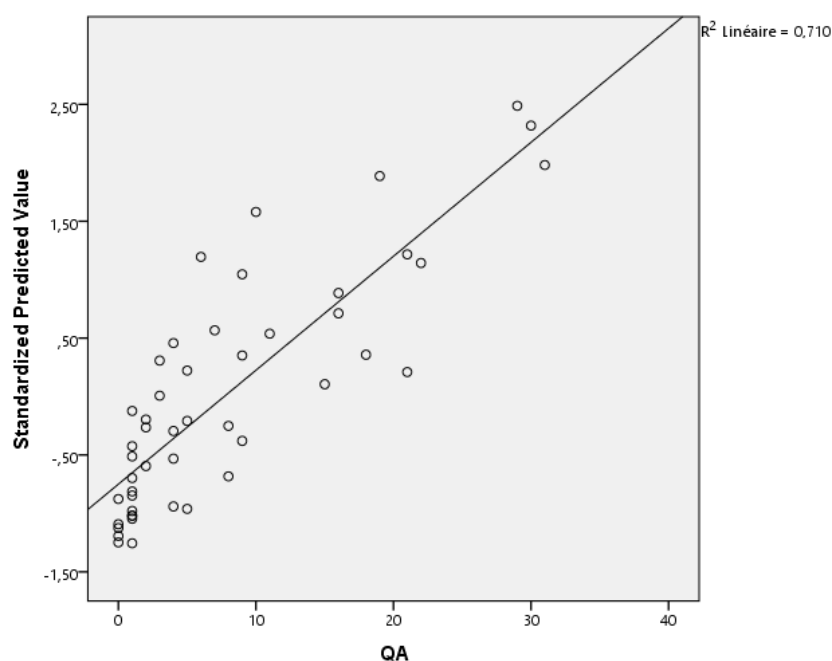


Figure 18 *Qualité d'ajustement du modèle*

Les résultats de la régression sont affichés en détail dans les annexes 29, 30 et 31.

La valeur de $R = 0,843$ reflète une forte relation entre la variable à expliquer et la combinaison des variables explicative précitée. La valeur VIF ne présente pas des valeurs > 3 d'où l'on peut confirmer que l'hypothèse de multi colinéarité est respectée. Le tableau dans l'annexe 32 montre qu'aucune valeur extrême influençant la qualité d'ajustement du modèle aux données. La prédiction des QA n'affiche aucune valeur écartée d'une manière aberrante de la moyenne. De plus, la valeur

des coefficients de toutes les variables est statistiquement significative avec un degré différent de contribution à l'explication des QA comme est montré dans le tableau 21.

Tableau 21 Evaluation des paramètres du modèle final

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	-2,639	1,314		-2,008	,051
1 Decc	,007	,003	,282	2,175	,035
Acc	,054	,016	,363	3,417	,001
Overspeed	1,262	,511	,228	2,468	,018
ND	,175	,066	,267	2,630	,012

a. Dependent Variable: QA

L'équation de régression est comme suit :

$$QA_{\text{prédite}} = (-2,639 + 0,007Decc + 0,054Acc + 0,175ND + 1,262overspeed)$$

Pour un conducteur conduisant en moyenne par mois 19h en nuit et 2h sous une vitesse de 120 km/h et plus, et effectuant 450 manœuvres de freinage fort et excessif et 50 manœuvres d'accélération forte et excessive, le nombre de QA qu'il peut avoir est de 9 par mois en moyenne. Ce résultat est assez proche du résultat observé chez un conducteur de l'échantillon ayant des données très proches de cet exemple (voir la colonne « moyenne » dans le tableau 15). Tous les coefficients d'estimation ont un signe positif. Ceci indique que toutes les variables sont des facteurs d'aggravation du risque. Cependant, chaque prédicteur influence le nombre des QA avec un degré différent :

- *Durée de conduite en excès de vitesse*

D'après l'équation, l'excès de vitesse est la variable la plus contributive à l'explication du risque de QA. Toutes choses égales par ailleurs, deux heures supplémentaires de conduite en excès de vitesse est associée à 2,524 QA de plus en moyenne chaque mois.

- *Durée de conduite en nuit*

La deuxième variable la plus contributive à la prédiction des QA est la conduite en nuit avec un $\beta = 0,175$. Autrement dit, 14 heures supplémentaires de conduite en nuit contribue à l'augmentation du nombre de QA de 2,45 en moyenne chaque mois.

- **Nombre de coups d'accélération excessifs**

Cette variable avec un $\beta = 0,054$ a plus d'impact sur la prédiction des QA que le freinage. En effet, l'augmentation de 60 coups d'accélération excessifs en moyenne par mois conduit à l'augmentation du nombre de quasi-accidents de 3,24 fois.

- **Nombre de coups de frein excessifs**

Malgré l'importance du freinage excessif prouvée par la littérature, le présent modèle montre qu'il s'agit du facteur le moins contributif à l'explication du risque des QA. Toutes choses égales par ailleurs, un conducteur qui effectue 400 coups de freins excessifs supplémentaires en moyenne par mois risque d'augmenter le nombre de QA de 2,8 fois.

N'étant pas aberrants, les résultats liés aux valeurs des coefficients des variables estimés par le modèle linéaire proposé sont confirmés par les travaux de Liu *et al* (2017). Ils ont développé un modèle de scoring de la qualité de conduite intégrant toutes les variables du modèle avec d'autres variables tel qu'évoqué dans le premier chapitre. Ils ont attribué suivant l'avis des experts, plus d'importance à la conduite en excès de vitesse ($>120\text{km/h}$), à la conduite en nuit et à l'accélération. Un conducteur ayant un comportement non risqué aura un maximum de score de 100 points (soit 0 le pire score). Celui qui ne dépasse pas une vitesse de 120km/h , conduit moins d'une heure en nuit et ne fait aucune manœuvre d'accélération brusque pendant un mois gagnerait 26, 10 et 7 points respectivement dans son score. (Un gain total de 43 points). L'absence du freinage brusque permet au conducteur d'avoir un bonus de 4 points seulement.

De plus, Klauer et al (2009) ont montré qu'un conducteur prudent adoptant un comportement de conduite normale peut également effectuer les manœuvres risquées les plus récurrentes tel que le freinage brusque, le dépassement inapproprié et l'inattention ce qui fait augmenter le risque automobile⁵⁵. Ceci affirme la faible valeur des estimateurs des variables freinage et accélération par rapport aux deux autres variables du modèle.

Récapitulatif des résultats de tous les modèles construits

Dans l'annexe 33, un tableau récapitule les résultats obtenus de tous les modèles construits. Le meilleur ajustement est obtenu par le modèle initial mais il s'agit d'un sur-ajustement tenant compte du problème de colinéarité. Le modèle final est choisi sur la base de la combinaison optimale en termes de R^2 et composée des variables significatives seulement. (Plus forte valeur de F-statistics).

⁵⁵ Voir annexe 19

La présence de la variable distance parcourue dans le modèle final améliore l'ajustement (R^2 et R^2 ajusté plus important) plus que la présence de la variable conduite en jour bien qu'elles ne soient pas significatives. Par contre, leur présence ensemble dans le modèle final n'affine pas les résultats. Le paramètre du carré de la distance parcourue n'est pas significatif dans tous les modèles (p -value > 0.1) ce qui signifie que l'hypothèse de la non-linéarité de la relation entre la distance parcourue et les quasi-accidents n'est pas retenue. Le modèle final retenu présente le moins de biais d'estimation en éliminant toutes variables confondantes et non significatives. Le freinage perd son pouvoir contributif à l'explication des QA en présence des variables conduite en excès de vitesse et conduite en nuit ($\beta = 0,007$ contre $\beta = 0,021$ dans le modèle de régression par la combinaison freinage, speed1, DD). Cela pourrait être expliqué par le fait que le recours au freinage fort et excessif est plus fréquent en jour dans un trafic routier plus ou moins intense qu'en nuit dans des voies plus dégagées où l'excès de vitesse relève le caractère principal d'une conduite dangereuse.

Analyse de sensibilité du modèle

Une analyse de sensibilité a été conduite sur le modèle de base pour vérifier sa sensibilité au changement de la combinaison de manœuvres dangereuses effectuées pour éviter un accident (révélatrice d'une situation de quasi-accident).

Les variables dépendantes à étudier sont les suivantes :

- QA1 : Nombre mensuel moyen de manœuvres très sévères niveau 3 de freinage et d'accélération
- QA2 : Nombre mensuel moyen de manœuvres très sévères niveau 3 de virage et d'accélération
- QA3 : Nombre mensuel moyen de manœuvres très sévères niveau 3 de virage et de freinage
- QA_D : Nombre mensuel moyen de freinage très sévère niveau 3
- QA_A : Nombre mensuel moyen d'accélération très sévère niveau 3
- QA_T : Nombre mensuel moyen de virage très sévère niveau 3

Cette analyse exposée dans le tableau 22 ci-dessous met en évidence que la meilleure combinaison pour décrire une situation de quasi-accident est celle utilisée dans la présente étude et qui regroupe les trois types de manœuvres : freinage, accélération et virage. L'ensemble de variables intégrant le modèle de base explique mieux cette combinaison plus que les autres combinaisons (meilleur R^2). En effet, les manœuvres d'accélération très sévères sont assez rares (observées chez 2 conducteurs seulement). Ces manœuvres, seules ou combinées avec les manœuvres de freinage ou de virage, (QA_A, QA1, QA2) ne peuvent pas être expliquées significativement par l'ensemble de variables. Contrairement à l'attendu, un conducteur effectuant des manœuvres fréquentes d'accélération

(niveaux 1 et 2 de sévérité) sont plus susceptibles d'effectuer des manœuvres de virage très sévères plutôt que celles d'accélération. D'ailleurs, la régression a montré qu'il n'existe aucune relation directe significative entre les manœuvres d'accélération de 1^{er} et 2^{ème} niveau de sévérité et les manœuvres d'accélération de 3^{ème} niveau. Par contre, les manœuvres de freinage et de virage très sévères (QA3, QA_D, QA_T) affichent une forte relation significative avec le modèle.

Tableau 22 Analyse de sensibilité du modèle par le changement de la variable dépendante

Variable dépendante	Pouvoir explicatif (R ²) du modèle de base (Toutes variables incluses)	F-statistics	p-value
QA	79,5%	14,332	0.000
QA_T	54.8%	4.480	0.000
QA3	43.5%	2,847	0.01
QA_D	41.9%	2,664	0.015
QA1	37,4%	2,210	0,039
QA_A	34.8%	1.978	0.065
QA2	34,4%	1,941	0.07

Cela est confirmé par la régression de la variable QA_T et QA3 par le modèle finale où la variable « Turn » est absente. Cette variable mesurant le virage fort et excessif ne semble pas utile pour la prédiction des situations de quasi-accidents causés par des événements de virage évasif. En Annexe 33, le modèle expliquant ces événements affiche un R² significatif de 35,1% et celui expliquant la combinaison (virage, freinage) affiche un R² significatif de 27,2% contre un R² de 71% affiché par le modèle final expliquant la combinaison (freinage, virage, accélération) choisie dans l'étude. La meilleure définition d'une situation de quasi-accident la plus expliquée par le modèle retenu est celle choisie par l'étude : *la combinaison de tous les types de manœuvres évasives effectuées afin d'éviter un accident.*

Section 5 Discussion et conclusions

L'approche de la collecte permanente des données télématiques en temps réel offre une opportunité pour évaluer non seulement le risque d'accident mais aussi le comportement au volant dangereux non encore sinistré. Cette étude évalue le risque individuel de conduite et les principaux facteurs associés aux comportements risqués au volant à l'aide d'une base de données de 48 abonnés au produit UBI offert par Telematik et Ooredoo, soit la première tentative en Tunisie pour étudier l'approche de l'assurance comportementale.

L'objectif essentiel de l'étude est de vérifier la possibilité d'utiliser les variables télématiques-révélatrices du comportement au volant- dans l'évaluation du risque automobile et la prédiction des profils de conduite risqués. Les résultats confirment que l'exposition au risque, qu'elle soit en termes d'évènements ou de durée de conduite, a une forte relation avec le risque individuel de conduite. **Un modèle de régression linéaire** utilisant deux types d'évènements d'incidents critiques (accélération et freinage) et la conduite en nuit et en excès de vitesse permet d'optimiser la prédiction du nombre de quasi-accidents. Les évènements d'incidents critiques et les variables d'exposition au risque ont un effet statistiquement significatif sur la prédiction. La variable durée de conduite en excès de vitesse contribue le plus à l'explication du phénomène des quasi-accidents avec $\beta = 1,262$ c'est-à-dire une unité supplémentaire de conduite en excès de vitesse augmente le nombre de QA de plus d'une fois. Ceci est similaire aux résultats de la caractérisation de Jensen et Wagner (2011) détaillée dans le premier chapitre qui montre que les conducteurs ayant un comportement dangereux sont caractérisés en premier degré par l'excès de vitesse. Pour l'augmentation de 400 unités de freins excessives et de 60 unités d'accélération excessives en moyenne par mois, le nombre de quasi-accidents augmente de 2,8 et de 3,24 respectivement. En effet, le freinage fort et excessif contribue mieux à l'explication du risque de quasi-accidents en jour en roulant sous une vitesse entre 50 et 90km/h ($\beta = 0,021$; p-value= 0,000) qu'en nuit en roulant sous une vitesse $> 120\text{km/h}$ ($\beta = 0,007$; p-value= 0,035). Il s'avère que l'exposition réelle au risque conditionne le recours aux manœuvres sévères -conséquentes des évènements d'incidents critiques- autrement dit le style de conduite ainsi exposé dans le premier chapitre par Weidner *et al* (2016). D'autres facteurs humains et environnementaux (fatigue, absence ou faiblesse de luminosité...) pourront aboutir davantage à la survenance de ces évènements dans une situation de conflit particulière. (i.e. l'absence de luminosité dans une route conduit à l'augmentation de la fréquence du freinage fort et excessif d'autant plus que la vitesse est élevée). Ce propos est similaire à l'idée de comportement de base vs comportement occasionnel défendue par Biecheler et Moget (1985) ainsi expliquée dans le premier chapitre.

L'analyse bi-variée a montré que le risque de quasi-accidents est fortement lié aux évènements d'incidents critiques (freinage, accélération, virage) mais également à la conduite en excès de vitesse et à la conduite en nuit comme est démontré par plusieurs chercheurs à savoir Paefgen *et al* (2013) et Guo *et al* (2010). Elle a également affiché une meilleure association des évènements d'incidents critiques avec la conduite en nuit comparée à la conduite en jour. Ces résultats sont confirmés par Musicant *et al* (2010) [103]. Par contre, la distance parcourue n'a pas d'association avec les quasi-accidents ni en uni-varié ni en multivarié contrairement à la littérature. Cela est dû à l'homogénéité

de l'échantillon en termes de kilomètres parcourus. De plus, l'inclusion de la durée de conduite en nuit et en excès de vitesse dans le modèle signifie que l'exposition réelle au risque peut être prise en compte et par conséquent, le calcul de primes individuelles sera plus précis et perspicace. D'ailleurs, appliquer des pénalités de primes aux assurés non respectueux des limites de vitesse contribue au développement des politiques de la sécurité routière et à la collaboration entre les institutions publiques et les assureurs.

L'association forte entre le comportement d'un conducteur et les situations de quasi-accidents peut avoir de fortes implications sur l'évaluation du risque individuel. Puisque le nombre d'accidents individuels des conducteurs est souvent limité, l'utilisation de l'historique de sinistralité pour la prédiction des conducteurs risqués ne sera pas efficace. Les événements d'incidents critiques se réalisent à une plus haute fréquence que les accidents (100 fois plus fréquents) et les quasi-accidents (10 fois plus fréquents) selon les travaux de Klauer *et al* (2010) et Guo et Fang (2013). De ce fait, ce type d'évènement permettra d'identifier les conducteurs risqués d'une façon proactive. Dès lors que les accidents sont des événements rares, il est plus utile d'utiliser des substituts tels que les quasi-accidents, les plus fréquemment exploités par les chercheurs. Afin de bien identifier les quasi-accidents et les expliquer par le modèle d'assurance comportementale, une combinaison des trois types d'évènements d'incidents critiques évasifs (freinage, accélération et virage de 3^{ème} niveau de sévérité) semble la meilleure; L'analyse de sensibilité conduite sur la définition d'un phénomène de quasi-accident implique que la combinaison utilisée dans l'étude est la plus significativement corrélée à l'ensemble des variables comportementales. Elle confirme également que le freinage brusque est le facteur le plus important parmi les incidents critiques pour identifier les situations de quasi-accidents comme est montré par les travaux de Klauer (2006, 2010).

La forte association trouvée entre les événements d'incidents critiques contextualisés⁵⁶ et les quasi-accidents à un niveau individuel de risque implique l'utilité de ces nouvelles variables en tant que substituts valides pour l'analyse du risque des conducteurs.

Toutefois, une des limites des résultats du modèle est l'élimination des variables confondantes suite à l'indisponibilité d'autres variables utiles pour leur contextualisation. Par exemple, les variables durées de conduite sous les paliers de vitesse (50-90 km/h) et (90-120 km/h) devraient être combinées avec la variable type de route pour savoir s'il s'agit bien d'une situation d'excès de vitesse. Une deuxième limite est l'élimination de la variable conduite en jour par manque de précision de l'heure de conduite (i.e. heure de pointe) ou du jour de semaine (weekend/ jour de

⁵⁶ En présence de durée de conduite en nuit et en excès de vitesse

travail). Cette variable, contextualisée, pourrait avoir un meilleur pouvoir explicatif du risque individuel de conduite.

En effet, pour appliquer une prime personnalisée fondée sur un jugement fiable de la qualité de conduite d'un assuré, il est plus pertinent de collecter le maximum d'informations routières à savoir l'effet des conditions météorologiques, les facteurs liés à la personnalité, les facteurs de risque variant dans le temps comme la distraction, la fatigue et les facteurs psychologiques...

Pour conclure, bien que l'étude soit limitée par la petite taille de l'échantillon, une relation significative est observée entre les variables comportementales et le risque individuel de conduite. Le plus important, elle démontre que les données télématiques fournissent des opportunités pour identifier les différences de niveau de risque individuel des conducteurs à partir des événements de moindre sévérité, quand on les observe au fil du temps, permettent d'identifier un profil complet d'un comportement risqué au volant avant même la réalisation d'un sinistre.

Le comportement au volant souvent analysé par les agences de sécurité routière est une opportunité pour les assureurs pour l'évaluation du risque automobile et la prévention contre les dérives comportementales source de sinistralité. Avec l'émergence de la télématique automobile, les voitures intelligentes permettent aux assureurs d'inclure dans le modèle de tarification des facteurs typiquement identifiés par les autorités de la circulation comme étant des indicateurs d'accidents. Les compagnies d'assurance ont toujours fixé les taux de prime par l'analyse des variables liées au conducteur (âge, ancienneté de permis), et liées au véhicule (âge du véhicule, puissance fiscale, usage) qui influencent les probabilités d'être impliqué dans un accident. Ces variables sont déterministes, c'est-à-dire, leurs valeurs sont connues d'avance et statiques (ou elles changent d'une manière contrôlable).

La seule variable ayant un impact sur la prime et susceptible de changer est le nombre d'accidents utilisée pour réviser la prime dans le cadre du système bonus-malus. Cependant, les variables télématiques fournissent de l'information dynamique sur les assurés (comportement de conduite, usage du véhicule en termes de distance parcourue au cours d'une période fixée, style de conduite...) qui affine davantage la tarification et améliore même le système bonus-malus. L'inclusion de ces informations dans le modèle conventionnel pourrait affiner la segmentation des assurés et identifier différents profils de risque au sein même d'une classe d'âge ou de nombre d'années d'expérience (i.e. les jeunes et les novices souvent et automatiquement classés risqués).

Conclusion générale

Le développement de la télématique automobile et des BIG DATA ont permis d'introduire des variables comportementales (distance parcourue, accélérométrie, excès de vitesse, conduite en nuit) dans la modélisation du risque individuel de conduite. Ceci mène à un changement inéluctable dans la logique d'assurance automobile. L'assureur surveille en temps réel les comportements des assurés lors de l'usage de leurs véhicules, ensuite, il les analyse pour évaluer leurs niveaux réels de risque. S'il s'agit d'un nouveau contrat, cette approche lui permet de bien sélectionner les risques sur la base de la qualité de conduite observée chez le client potentiel pendant une période d'essai préfixée. S'il s'agit d'un renouvellement de contrat, l'assureur se permet de réviser la prime en se basant sur la réévaluation du niveau de risque de son assuré. En cas d'amélioration de la qualité de conduite (souvent mesurée par un score de conduite), autrement dit baisse du risque, l'assuré profite d'une réduction de prime. Par contre, en cas de dégradation de la qualité de conduite et pour renforcer le système Bonus-Malus, l'assuré sera soit pénalisé par une surprime soit privé de la bonification déjà accordée. En aucun cas, le contrat PHYD remplace le contrat d'assurance automobile conventionnel. C'est alors que le refus de la surprime implique la résiliation du contrat PHYD et le retour au contrat d'assurance conventionnel.

L'approche d'assurance comportementale semble être utile en Tunisie pour les opportunités suivantes :

- 1- Affiner la sélection et la classification des risques
- 2- Assurer l'équilibre tarifaire par une meilleure précision de la modélisation actuarielle des risques
- 3- Suivre et contrôler l'évolution en temps réel du portefeuille automobile
- 4- Maîtriser le risque automobile (collision / vol) par le système de récompense / punition et via les services de localisation GPS
- 5- Fidéliser les clients par les incitations financières et économiques, les conseils, l'accompagnement en temps réel et l'accélération du processus de gestion de sinistres
- 6- Jouer le rôle social de renforcement de la sécurité routière

A notre connaissance, la présente étude sera la première tentative en Tunisie pour montrer l'utilité des variables comportementales dans l'évaluation du risque automobile et la prédiction des conducteurs risqués. Elle est conduite à l'aide d'une base de données de 48 abonnés au premier produit PHYD en Tunisie conçu par Telematik et commercialisé par Ooredoo depuis Mai 2017. Les

résultats de l'étude indiquent que la modélisation du risque automobile est possible même en absence de l'historique de sinistralité, en utilisant les événements de quasi-accidents⁵⁷, beaucoup plus fréquents que les accidents. Nous avons montré également que ces événements affichent une relation statistiquement étroite avec l'ensemble des variables comportementales. En effet, le modèle de régression linéaire construit permet à l'assureur d'identifier les assurés ayant des comportements à haut risque même s'ils ne sont pas historiquement sinistrés. Ce modèle indique que la durée de conduite en excès de vitesse (p-value = 0,018), la durée de conduite en nuit (p-value = 0,012) et les événements d'incidents critiques, qui sont le freinage fort excessif (p-value = 0,035) et l'accélération forte et excessive (p-value = 0,001) ont un impact significatif sur le risque individuel de conduite. De plus, le modèle présente un fort pouvoir explicatif de ce phénomène malgré la petite taille de l'échantillon. La régression linéaire multiple a démontré que la conduite en excès de vitesse est le facteur le plus important d'aggravation du risque individuel d'accident ($\beta = 1,262$) suivi par la conduite en nuit ($\beta = 0,175$), l'accélération ($\beta = 0,054$) et le freinage ($\beta = 0,007$); Un conducteur roulant pour une heure supplémentaire de conduite en excès de vitesse et une heure supplémentaire en nuit augmente le risque individuel de conduite de 1,262 fois et de 0,175 fois toutes choses égales par ailleurs. Cependant, d'autres variables susceptibles d'améliorer la précision du modèle sont absentes⁵⁸ à savoir le type de la route fréquentée, la conduite pendant les heures de pointe ou pendant les weekends. Contrairement à l'attendu, la variable « distance parcourue » n'a pas montré une relation significative avec le risque individuel de conduite (p-value >> 0,05). Ce résultat est expliqué par la distribution homogène de l'échantillon en fonction du nombre de kilomètres parcourus. Une autre limite de l'étude est l'absence des variables du modèle conventionnel (ancienneté de permis, puissance fiscale, âge du véhicule, valeur assurée). Ces variables devraient être utiles pour mesurer la valeur ajoutée de l'introduction des variables comportementales dans le modèle de tarification du risque automobile.

La présente étude est la première étape et d'autres études avec une plus large base de données seront requises. Il sera intéressant de mettre en évidence le résultat souhaitable de l'approche d'assurance comportementale qu'est le changement du comportement surveillé pendant une plus longue période d'observation (de 24 mois à 60 mois). A travers le changement de comportement, l'assureur sera en mesure d'améliorer la sinistralité et de contourner les dérives comportementales en développant des programmes proactifs de contremesures de sécurité.

⁵⁷ Les événements de quasi-accidents sont identifiés par la réalisation des manœuvres évasives d'accélération, de freinage et de virage

⁵⁸ Non agrégées par le concepteur du produit Sayartech



Annexes

Annexe 1 Critères de tarification a priori selon le pays [104]

Critère	Sous-critère	Pays
Véhicule	Vitesse, puissance, âge, volume moteur et modèle	France, Royaume Uni
	Puissance, âge, poids, volume moteur et modèle	Portugal, Canada, Norvège
	Puissance, âge, volume moteur et modèle	Chypre, Italie
	Puissance, modèle	Liban
	Usage du véhicule et type de combustible	Maroc
Usage	Distance parcourue	RU, Norvège, Italie, Suède, Canada, France
	Activité professionnelle	Italie, Canada, Chypre, RU
	Conduite de jour ou de nuit	RU
Conducteur	Age	Belgique, Suède, Pays-Bas (respectivement en dessous de 23 ans, 24 et par groupe d'âge en dessous de 27 ans)
	Age, genre, statut marital	Etats-Unis et Canada (interdit dans certaines juridictions)
	Nationalité	Suisse
	Nombre de véhicules assurés et nombre de condamnations	Canada
	Nombre d'années de permis de conduire	France, Luxembourg

Annexe 2 modèles du système Bonus-Malus selon le pays [104]

Pays	Modèle
Maroc	Deux ans sans sinistre responsable totalement ou partiellement → réduction de prime de 10% maximum
Royaume Uni	Se baser sur le nombre d'années sans sinistre uniquement
Luxembourg	Un an sans sinistre → 1 point de bonus Un sinistre → un malus de trois points Max de bonus 22 points

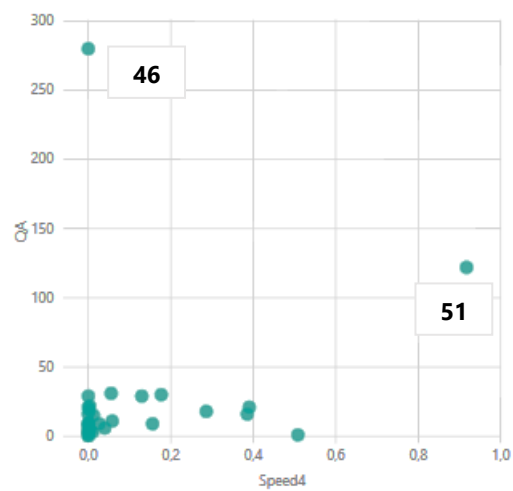
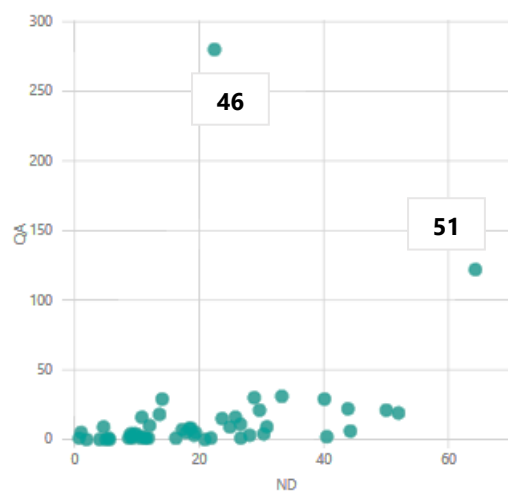
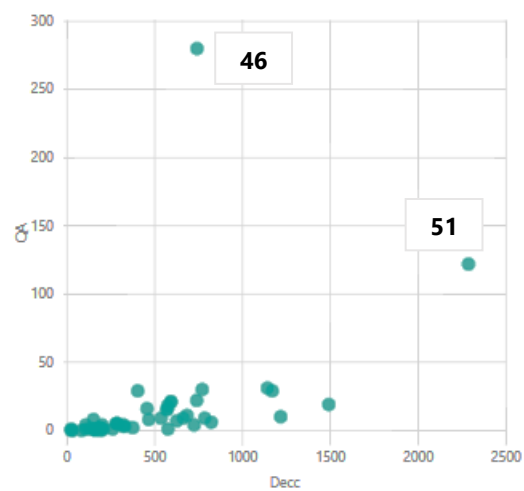
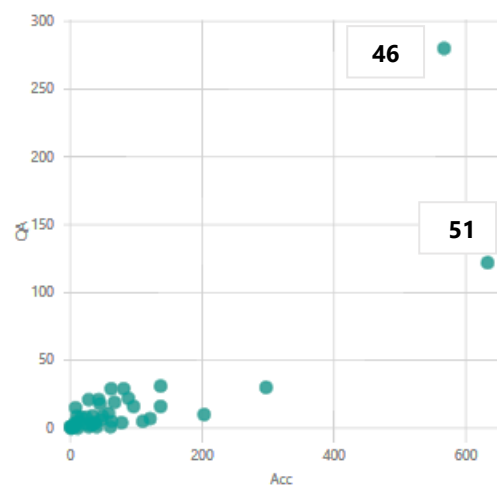
Zone CIMA ⁵⁹	Se baser sur le nombre d'années sans sinistre responsable
Australie	Pas d'application pour la garantie RC (dommages corporels) Le système est non réglementé pour les autres garanties
Norvège et Italie, Chypre, France Portugal et Luxembourg, Zone CIMA	Bonus maximum respectif 75% ; 60% ; 50% ; 30%
Espagne, Suède, Chypre, Vietnam	Système pur bonus
Etats Unis	Système pur malus

Annexe 3 Capture d'écran d'une partie d'observations de circulation d'un véhicule

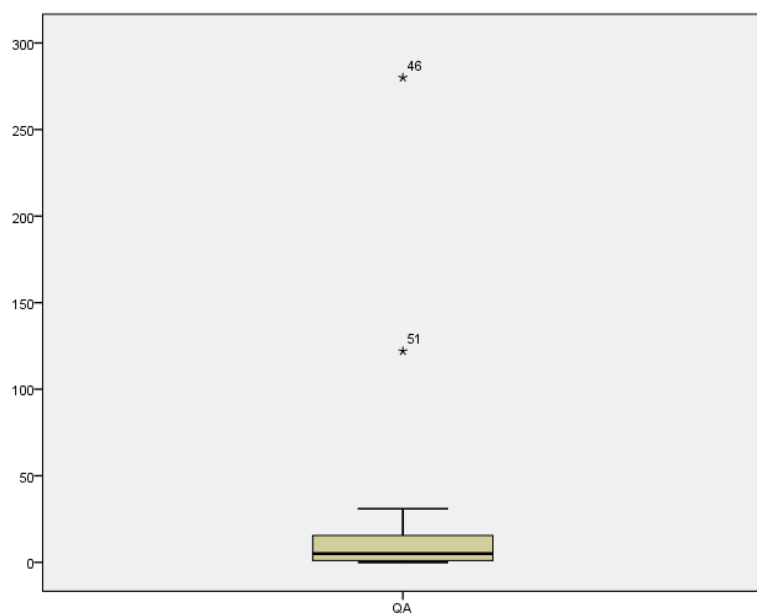
Année	Mois	Jour	time	nom_param	gps_speed	gps_latitude	gps_longitude
2017	décembre	13	17:51:22	Véhicule en marche	52,97	33,14	0,1
2017	décembre	13	17:51:37	Véhicule en marche	52,04	33,14	0,1
2017	décembre	13	17:46:17	Véhicule en marche - virage à droite	57,97	33,00	0,1
2017	décembre	13	17:46:20	Véhicule en marche - virage à droite	57,04	33,00	0,1
2017	décembre	13	16:35:56	Véhicule en marche - virage à gauche	62,04	33,00	0,1
2017	décembre	14	07:22:11	Fin de l'alerte véhicule au ralenti	65,01	33,01	0,1
2017	décembre	14	07:26:27	Freinage brusque : Seuil Limite 1 dépassé	62,04	33,01	0,1
2017	décembre	14	16:32:32	Véhicule en marche	65,01	33,01	0,1
2017	décembre	14	16:32:49	Véhicule en marche	60,00	33,01	0,1
2017	décembre	14	16:33:02	Véhicule en marche	65,93	33,01	0,1
2017	décembre	14	16:33:49	Véhicule en marche	65,01	33,01	0,1
2017	décembre	14	16:33:52	Véhicule en marche	52,97	33,01	0,1
2017	décembre	14	16:34:23	Véhicule en marche	60,00	33,01	0,1
2017	décembre	14	16:35:18	Véhicule en marche	52,04	33,01	0,1

⁵⁹ Zone CIMA : Conférence Interafricaine des Marchés d'Assurances

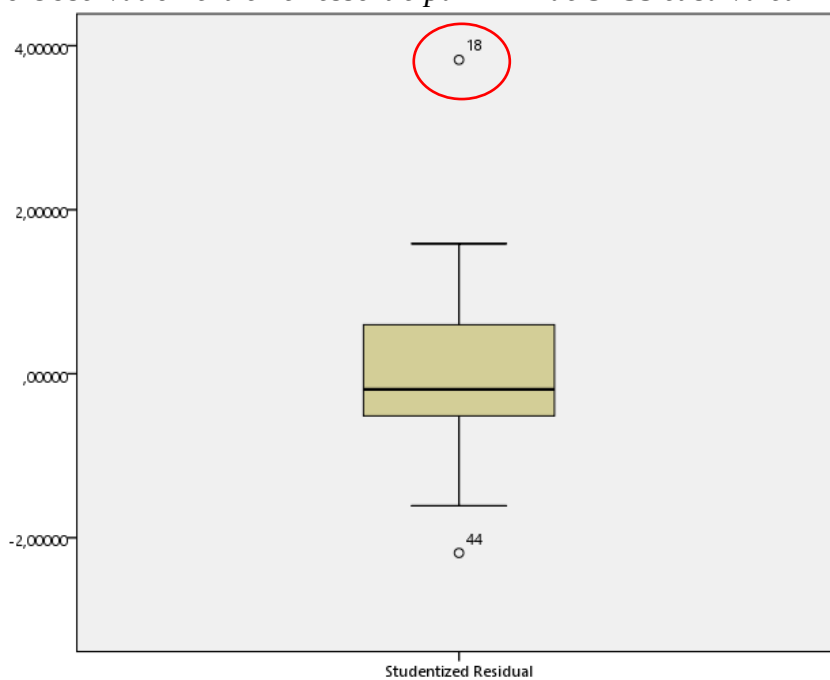
Annexe 4 Graphiques de l'échantillon y compris les deux véhicules atypiques (n° 46 et n°51) (QA



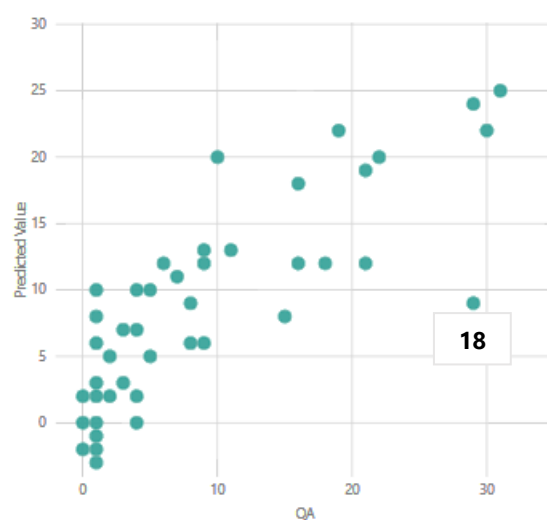
Annexe 5 les observations extrêmes ressorties par SPSS



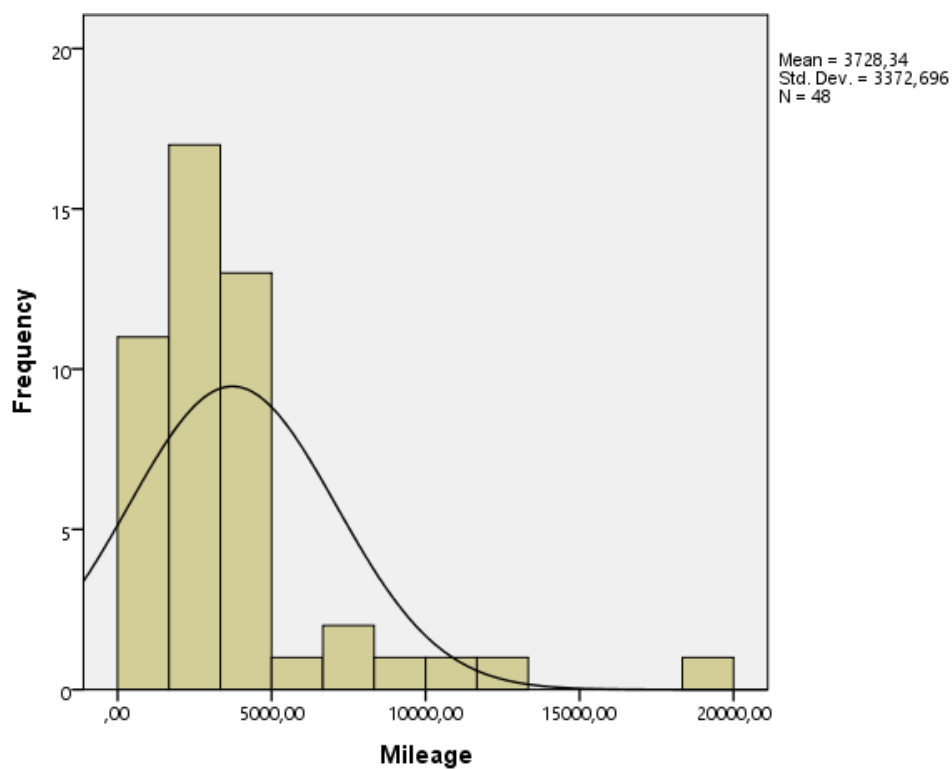
Annexe 6 Observation extrême ressortie par RLM de SPSS et sa valeur résiduelle



Annexe 7 Valeur prédictive de l'observation extrême ressortie par RLM de SPSS



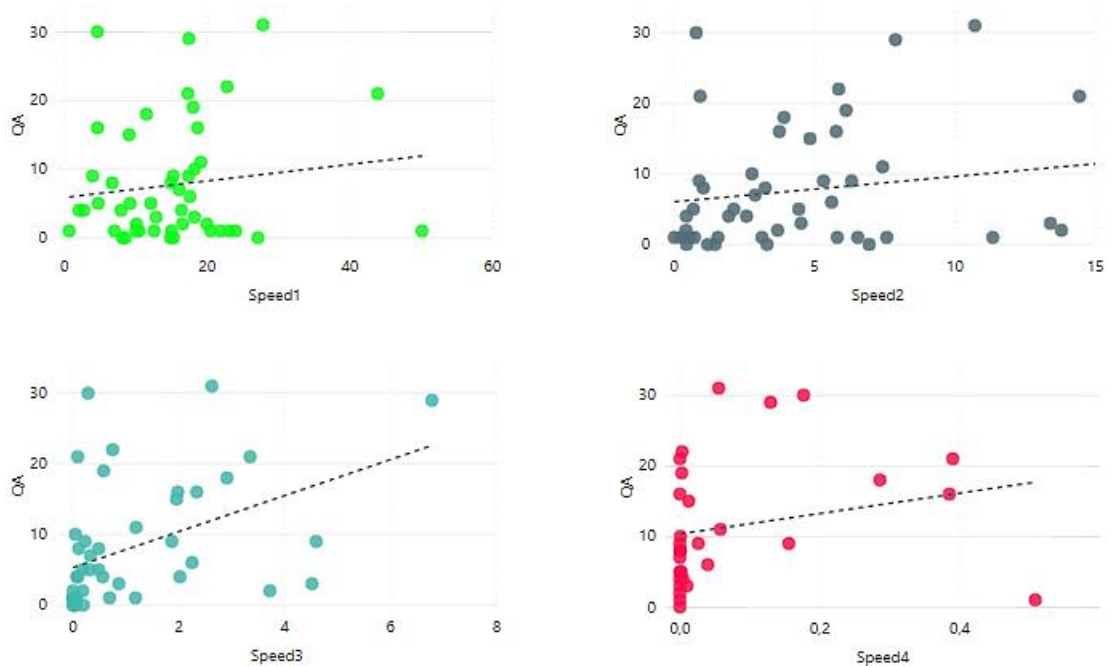
Annexe 8 Distribution de la population en termes de distance parcourue



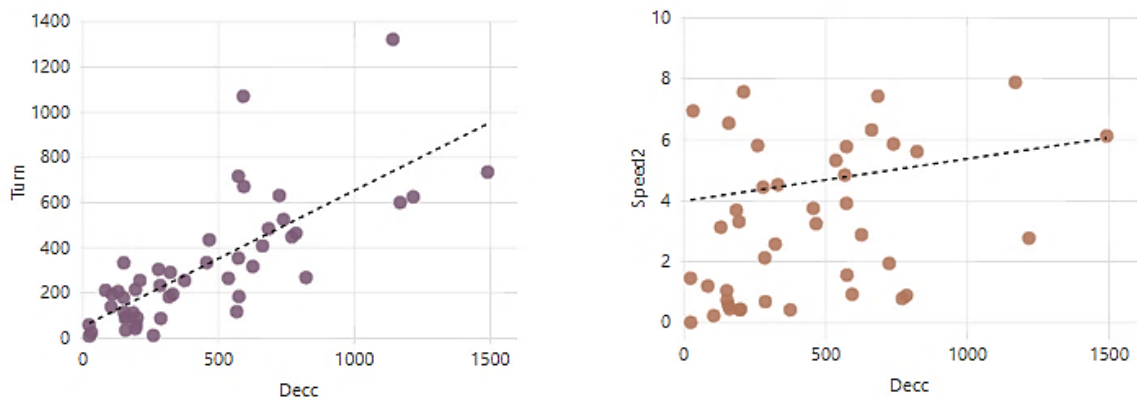
Annexe 9 Tableau récapitulatif de corrélations entre les différentes variables indépendantes et la variable dépendante (QA)

		QA	Acc	Decc	Turn	Mileage	Speed1	Speed2	Speed3	Speed4	DD	ND
Pearson Correlation	QA	1,000	0,632	0,749	0,780	0,009	0,132	0,184	0,469	0,353	0,079	0,601
	Acc		1,000	0,612	0,520	-0,067	-0,108	-0,039	0,108	0,195	-0,039	0,260
	Decc			1,000	0,753	0,136	0,105	0,107	0,413	0,145	0,285	0,567
	Turn				1,000	0,045	0,294	0,252	0,367	0,372	0,241	0,580
	Mileage					1,000	0,683	0,432	0,123	0,042	0,522	0,297
	Speed1						1,000	0,499	0,123	0,170	0,735	0,340
	Speed2							1,000	0,473	0,166	0,363	0,368
	Speed3								1,000	0,392	0,123	0,342
	Speed4									1,000	-0,003	0,242
	DD										1,000	0,361
	ND											1,000
Sig. (2- tailed)	QA	.	0,000	0,000	0,000	0,952	0,373	0,211	0,01	0,014	0,595	0,000
	Acc		.	0,000	0,000	0,652	0,464	0,791	0,466	0,185	0,795	0,074
	Decc			.	0,000	0,356	0,479	0,469	0,004	0,327	0,05	0,000
	Turn				.	0,759	0,043	0,084	0,010	0,009	0,099	0,000
	Mileage					.	0,000	0,002	0,403	0,777	0,000	0,041
	Speed1						.	0,000	0,403	0,248	0,000	0,018
	Speed2							.	0,000	0,006	0,011	0,017
	Speed3								.	0,003	0,202	0,009
	Speed4									.	0,982	0,098
	DD										.	0,012
	ND											.

Annexe 10 Les relations entre QA et les variables de durée de conduite sous les 4 paliers de vitesse



Annexe 11 Les relations entre le freinage et les variables virage et durée de conduite sous une vitesse entre 90 et 120km/h



Annexe 12 Les modèles construits par régression hiérarchique (VI= Distance parcourue)

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	,009 ^a	,000	-,022	8,735	,000	,004	1	46	,952
2	,280 ^b	,078	,037	8,479	,078	3,820	1	45	,057

a. Predictors: (Constant), mileage_per_1000km

b. Predictors: (Constant), mileage_per_1000km, mileage_square

Annexe 13 Analyse de la variance ANOVA du modèle quadratique (VI = distance parcourue)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	,284	1	,284	,004	,952 ^b
	Residual	3510,028	46	76,305		
	Total	3510,313	47			
2	Regression	274,928	2	137,464	1,912	,160 ^c
	Residual	3235,385	45	71,897		
	Total	3510,313	47			

a. Dependent Variable: QA

b. Predictors: (Constant), mileage_per_1000km

c. Predictors: (Constant), mileage_per_1000km, mileage_square

Annexe 14 Evaluation des paramètres du modèle quadratique (VI = distance parcourue)

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	7,602	1,890		4,021	,000
	mileage_per_1000km	,023	,378	,009	,061	,952
	(Constant)	3,204	2,904		1,103	,276
2	mileage_per_1000km	1,978	1,065	,772	1,857	,070
	mileage_square	-,115	,059	-,812	-1,954	,057

a. Dependent Variable: QA

Annexe 15 Modèle initial de régression linéaire multiple

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	,892 ^a	,795	,739	4,412	,795	14,332	10	37	,000

a. Predictors: (Constant), ND, mileage_square, Acc, Overspeed, DD, Speed2, Turn, Decc, Speed1, Mileage

b. Dependent Variable: QA

Annexe 16 Analyse de variance ANOVA du modèle initial

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2790,018	10	279,002	14,332	,000 ^b
	Residual	720,295	37	19,467		
	Total	3510,313	47			

a. Dependent Variable: QA

b. Predictors: (Constant), ND, mileage_square, Acc, Overspeed, DD, Speed2, Turn, Decc, Speed1, Mileage

Annexe 17 Modèle initial construit par la méthode de régression Backward

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	,892 ^a	,795	,739	4,412	,795	14,332	10	37	,000
2	,891 ^b	,795	,746	4,357	,000	,053	1	37	,819
3	,889 ^c	,791	,748	4,335	-,003	,617	1	38	,437
4	,884 ^d	,781	,743	4,381	-,010	1,842	1	39	,182
5	,881 ^e	,776	,743	4,380	-,005	,989	1	40	,326
6	,877 ^f	,769	,741	4,395	-,007	1,284	1	41	,264
7	,870 ^g	,756	,733	4,463	-,013	2,332	1	42	,134

a. Predictors: (Constant), ND, mileage_square, Acc, Overspeed, DD, Speed2, Turn, Decc, Speed1, Mileage

b. Predictors: (Constant), ND, mileage_square, Acc, Overspeed, DD, Turn, Decc, Speed1, Mileage

c. Predictors: (Constant), ND, Acc, Overspeed, DD, Turn, Decc, Speed1, Mileage

d. Predictors: (Constant), ND, Acc, Overspeed, DD, Turn, Decc, Speed1

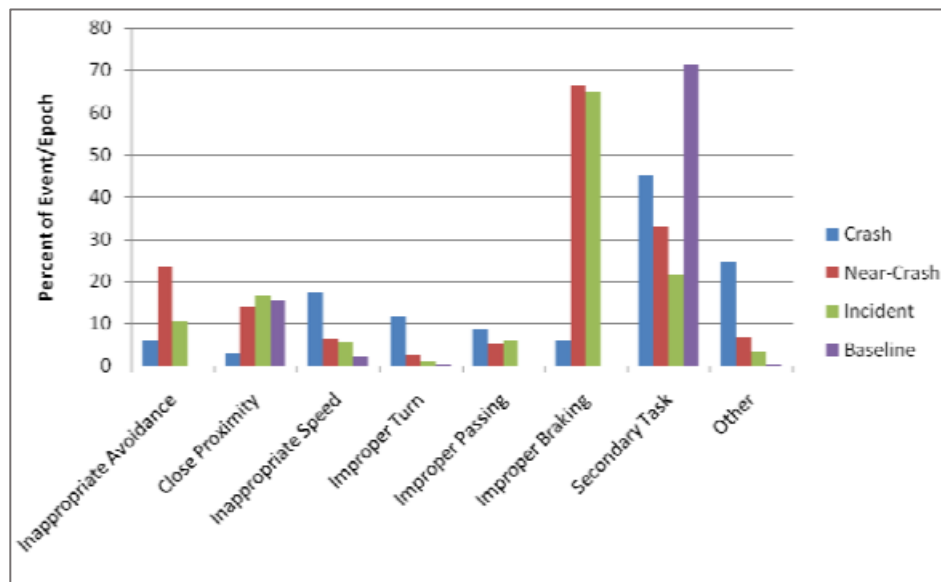
e. Predictors: (Constant), ND, Acc, Overspeed, DD, Turn, Decc

f. Predictors: (Constant), ND, Acc, Overspeed, DD, Turn

g. Predictors: (Constant), ND, Acc, Overspeed, Turn

h. Dependent Variable: QA

Annexe 18 Pourcentage des évènements observés sur un conducteur avec un comportement risqué au volant



Annexe 19 Liste du nombre d'évènements moyen où les comportements risqués étaient le premier facteur contributif

Risky Driving Behavior	Safe Driver	Moderately Safe Driver	Unsafe Driver
Apparent unfamiliarity	0.0	0.2	0.1
Avoiding object	0.0	5.9	5.3
Proximity to other vehicles	0.0	9.9	9.7
Did not see	0.0	1.0	1.5
Speed differential	0.0	2.8	4.6
Failure to signal	0.0	0.9	0.6
Improper backing	0.0	0.0	0.2
Improper turn	0.0	0.5	0.7
Other	0.0	2.6	1.6
Improper parking	0.0	0.0	0.1
Improper passing	4.4	3.3	4.5
Reckless/menacing driving	0.6	0.2	0.9
Traffic signal violation	0.8	0.3	0.7
Improper braking	31.6	47.0	40.6
Inattentive	13.8	14.3	20.1
None	13.9	11.	8.9

Annexe 20 Modèle de régression par la combinaison [Conduite en jour, Freinage, Speed1]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	,762 ^a	,580	,562	5,722	,580	31,116	2	45	,000
2	,797 ^b	,635	,610	5,398	,054	6,554	1	44	,014

a. Predictors: (Constant), DD, Decc

b. Predictors: (Constant), DD, Decc, Speed1

c. Dependent Variable: QA

Annexe 21 Evaluation des paramètres du modèle de régression par la combinaison [Conduite en jour, Freinage, Speed1]

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	2,353	2,394		,983	,331
1 Decc	,020	,003	,790	7,847	,000
DD	-,047	,032	-,146	-1,452	,153
(Constant)	3,171	2,281		1,390	,172
2 Decc	,021	,002	,830	8,620	,000
DD	-,133	,045	-,413	-2,929	,005
Speed1	,319	,125	,348	2,560	,014

a. Dependent Variable: QA

Annexe 22 Evaluation des paramètres du modèle de régression par la combinaison [conduite en jour, Accélération, Overspeed]

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	,021	2,541		,008	,994
1 DD	,017	,032	,053	,532	,597
Acc	,088	,015	,586	5,867	,000
Overspeed	2,236	,556	,404	4,020	,000

a. Dependent Variable: QA

Annexe 23 Evaluation de la variance du modèle de régression par la combinaison [Conduite en jour, Freinage, Speed1]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2037,212	2	1018,606	31,116	,000 ^b
	Residual	1473,100	45	32,736		
	Total	3510,313	47			
2	Regression	2228,198	3	742,733	25,489	,000 ^c
	Residual	1282,115	44	29,139		
	Total	3510,313	47			

a. Dependent Variable: QA

b. Predictors: (Constant), DD, Decc

c. Predictors: (Constant), DD, Decc, Speed1

Annexe 24 Evaluation de la variance du modèle de régression par la combinaison [Conduite en jour, Accélération, Overspeed]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1995,517	3	665,172	19,321	,000 ^b
	Residual	1514,795	44	34,427		
	Total	3510,313	47			

a. Dependent Variable: QA

b. Predictors: (Constant), Overspeed, Acc, DD

Annexe 25 Modèle final construit par la méthode de régression hiérarchique⁶⁰

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	,843 ^a	,710	,683	4,865	,710	26,332	4	43	,000
2	,852 ^b	,726	,693	4,790	,015	2,358	1	42	,132
3	,858 ^c	,737	,691	4,806	,011	,859	2	40	,431

a. Predictors: (Constant), Overspeed, Acc, ND, Decc

b. Predictors: (Constant), Overspeed, Acc, ND, Decc, DD

c. Predictors: (Constant), Overspeed, Acc, ND, Decc, DD, mileage_square, Mileage

d. Dependent Variable: QA

Annexe 26 Analyse de la variance ANOVA du modèle final

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2492,682	4	623,171	26,332	,000 ^b
	Residual	1017,630	43	23,666		
	Total	3510,313	47			
2	Regression	2546,772	5	509,354	22,202	,000 ^c
	Residual	963,540	42	22,941		
	Total	3510,313	47			
3	Regression	2586,434	7	369,491	15,997	,000 ^d
	Residual	923,878	40	23,097		
	Total	3510,313	47			

a. Dependent Variable: QA

b. Predictors: (Constant), Overspeed, Acc, ND, Decc

c. Predictors: (Constant), Overspeed, Acc, ND, Decc, DD

d. Predictors: (Constant), Overspeed, Acc, ND, Decc, DD, mileage_square, Mileage

⁶⁰ Le modèle en vert est celui retenu

Annexe 27 Les étapes d'entrée et de sortie de variables du modèle selon la méthode Backward

Model	Variables Entered	Variables Removed	Method
1	mileage_per_1000km, Acc, Overspeed, ND, DD, Decc, mileage_square ^b	.	Enter
2	.	DD	Backward (criterion: Probability of F-to-remove >= ,100).
3	.	mileage_square	Backward (criterion: Probability of F-to-remove >= ,100).
4	.	mileage_per_1000km	Backward (criterion: Probability of F-to-remove >= ,100).

a. Dependent Variable: QA

b. All requested variables entered.

Annexe 28 Evaluation des paramètres du modèle pour chaque étape de modélisation selon la méthode Backward

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations			Collinearity Statistics	
	B	Std. Error	Beta			Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	,136	2,164	,063	,950					
	Decc	,008	,003	,328	2,473	,749	,364	,201	,375	2,666
	Acc	,046	,016	,310	2,845	,632	,410	,231	,553	1,809
	Overspeed	1,357	,527	,245	2,576	,480	,377	,209	,725	1,379
	ND	,231	,072	,353	3,200	,601	,452	,260	,542	1,846
	DD	-,019	,035	-,060	-,553	,079	-,087	-,045	,552	1,812
	mileage_square	,042	,043	,296	,979	-,088	,153	,079	,072	13,858
	mileage_per_1000km	-1,014	,852	-,396	-1,190	,009	-,185	-,097	,060	16,807
2	(Constant)	-,585	1,713	-,342	,734					
	Decc	,008	,003	,311	2,431	,749	,355	,196	,396	2,528
	Acc	,048	,016	,322	3,032	,632	,428	,244	,574	1,743
	Overspeed	1,411	,514	,255	2,748	,480	,394	,221	,751	1,332
	ND	,231	,072	,352	3,225	,601	,450	,259	,542	1,846
	mileage_square	,050	,040	,355	1,271	-,088	,195	,102	,083	12,084
	mileage_per_1000km	-1,234	,748	-,481	-1,650	,009	-,250	-,133	,076	13,161
3	(Constant)	-1,856	1,401	-1,325	,192					
	Decc	,007	,003	,293	2,286	,749	,333	,185	,401	2,495
	Acc	,050	,016	,338	3,183	,632	,441	,258	,582	1,718
	Overspeed	1,268	,505	,229	2,512	,480	,361	,204	,789	1,267
	ND	,199	,068	,305	2,947	,601	,414	,239	,615	1,627
	mileage_per_1000km	-,325	,221	-,127	-1,473	,009	-,222	-,119	,886	1,129

4	(Constant)	-2,639	1,314		-2,008	,051					
	Decc	,007	,003	,282	2,175	,035	,749	,315	,179	,402	2,487
	Acc	,054	,016	,363	3,417	,001	,632	,462	,281	,597	1,674
	Overspeed	1,262	,511	,228	2,468	,018	,480	,352	,203	,789	1,267
	ND	,175	,066	,267	2,630	,012	,601	,372	,216	,655	1,527

a. Dependent Variable: QA

Annexe 29 Modèle final construit par la méthode de régression Stepwise

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	,749 ^a	,561	,551	5,790	,561	58,710	1	46	,000
2	,780 ^b	,609	,592	5,523	,048	5,550	1	45	,023
3	,818 ^c	,669	,646	5,139	,060	7,990	1	44	,007
4	,843 ^d	,710	,683	4,865	,041	6,093	1	43	,018

a. Predictors: (Constant), Decc

b. Predictors: (Constant), Decc, Acc

c. Predictors: (Constant), Decc, Acc, ND

d. Predictors: (Constant), Decc, Acc, ND, Overspeed

e. Dependent Variable: QA

Annexe 30 Analyse de la variance du modèle final construit par la méthode de régression Stepwise

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1968,206	1	1968,206	58,710	,000 ^b
	Residual	1542,106	46	33,524		
	Total	3510,313	47			
2	Regression	2137,506	2	1068,753	35,033	,000 ^c
	Residual	1372,807	45	30,507		
	Total	3510,313	47			
3	Regression	2348,481	3	782,827	29,647	,000 ^d
	Residual	1161,831	44	26,405		
	Total	3510,313	47			
4	Regression	2492,682	4	623,171	26,332	,000 ^e
	Residual	1017,630	43	23,666		
	Total	3510,313	47			

a. Dependent Variable: QA

b. Predictors: (Constant), Decc

c. Predictors: (Constant), Decc, Acc

d. Predictors: (Constant), Decc, Acc, ND

e. Predictors: (Constant), Decc, Acc, ND, Overspeed

Annexe 31 Evaluation des paramètres du modèle final construit par la méthode de régression Stepwise

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations			Collinearity Statistics	
	B	Std. Error	Beta			Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	-,525	1,359	-,386	,701					
	Decc	,019	,002	,749	7,662	,000	,749	,749	,749	1,000
2	(Constant)	-,565	1,297	-,436	,665					
	Decc	,015	,003	,579	4,910	,000	,749	,591	,458	,625
	Acc	,041	,018	,278	2,356	,023	,632	,331	,220	,625
3	(Constant)	-2,498	1,387	-1,801	,078					
	Decc	,010	,003	,383	2,954	,005	,749	,407	,256	,447
	Acc	,048	,017	,319	2,885	,006	,632	,399	,250	,614
	ND	,197	,070	,300	2,827	,007	,601	,392	,245	,667
4	(Constant)	-2,639	1,314	-2,008	,051					
	Decc	,007	,003	,282	2,175	,035	,749	,315	,179	,402
	Acc	,054	,016	,363	3,417	,001	,632	,462	,281	,597
	ND	,175	,066	,267	2,630	,012	,601	,372	,216	,655
	Overspeed	1,262	,511	,228	2,468	,018	,480	,352	,203	,789

a. Dependent Variable: QA

Annexe 32 Diagnostic des observations du modèle finale

Case Number	Std. Residual	QA	Predicted Value	Residual
1	-,159	1	1,77	-,774
2	,838	9	4,92	4,078
3	-,608	1	3,96	-2,956
4	,440	8	5,86	2,140
5	-,124	11	11,60	-,603
6	-1,889	10	19,19	-9,190
7	1,232	22	16,01	5,992
8	-,317	4	5,54	-1,540
9	,152	1	,26	,741
10	-,268	0	1,30	-1,305
11	-1,191	1	6,79	-5,792
12	,103	0	-,50	,503
13	-,974	3	7,74	-4,739
14	-,774	2	5,76	-3,764
15	,643	16	12,87	3,127
16	-1,424	3	9,93	-6,928
17	1,827	31	22,11	8,888
18	-1,296	9	15,31	-6,306
19	-1,443	4	11,02	-7,019
20	,091	1	,56	,444
21	-,257	9	10,25	-1,249
22	-,886	5	9,31	-4,310
23	,383	16	14,14	1,862
24	1,345	15	8,45	6,545
25	-,499	19	21,43	-2,427
26	,190	1	,08	,924
27	,037	4	3,82	,182
28	,650	4	,84	3,161
29	,505	1	-1,46	2,457
30	-,875	2	6,26	-4,258
31	,205	0	-1,00	,997
32	,055	0	-,27	,269
33	-,738	1	4,59	-3,592
34	,288	0	-1,40	1,402
35	,150	1	,27	,729
36	2,423	21	9,21	11,786
37	,915	21	16,55	4,451
38	-,279	2	3,36	-1,358
39	-,106	1	1,52	-,517
40	-,990	7	11,81	-4,815
41	,886	5	,69	4,310

42	1,086	8	2,72	5,285
43	-,240	5	6,17	-1,166
44	1,584	18	10,29	7,706
45	1,117	30	24,57	5,432
46	-2,135	6	16,39	-10,388
47	,657	29	25,81	3,195
48	-,331	1	2,61	-1,609

a. Dependent Variable: QA

Annexe 33 Récapitulatif des modèles de régression linéaire construits⁶¹

	Modèle initial (toutes les variables)		Modèle final		Modèle final avec DD		Modèle final avec Distance parcourue		Modèle final avec DD et Distance parcourue		Modèle de régression par la combinaison freinage, speed1, DD	
	β	p	β	p	β	p	β	p	β	p	β	p
Intercept	0,559	0,782	-2,639	0,051	-0,177	0,955	-0,585	0,734	0,136	0,950	3,171	0,172
Decc	0,006	0,125	0,007	0,035	0,008	0,015	0,008	0,019	0,008	0,018	0,021	0,000
Acc	0,041	0,011	0,054	0,001	0,048	0,005	0,048	0,004	0,046	0,007		
Turn	0,007	0,133										
Mileage	-0,001	0,362					-1,234	0,107	-0,001	0,241		
Mileage²	0,026	0,581					0,050	0,211	0,042	0,333		
Speed1	0,222	0,170									0,319	0,014
Speed2	-0,054	0,819										
Overspeed	1,180	0,034	1,262	0,018	1,197	0,023	1,411	0,009	1,357	0,014		
DD	-0,071	0,121			-0,045	0,132			-0,019	0,583	-0,133	0,005
ND	0,169	0,019	0,175	0,012	0,199	0,005	0,231	0,002	0,231	0,003		
R² / R² ajusté	79,5% / 73,9%		71% / 68,3%		72,6% / 69,3%		73,5% / 69,6%		73,7% / 69,1%		63,5% / 61%	
F-Statistic	14,332 < 0.001		26,332 < 0.001		22,202 < 0.001		18,993 < 0.001		15,997 < 0.001		25,489 < 0.001	

⁶¹ Les cellules en gris : variable non incluse dans le modèle concerné

AIC	152,006	156 ,595	155,972	156,320	155,954	165,683
Nombre de prédicteurs	10	4	5	6	7	3

Annexe 34 Sensibilité des modèles finaux au changement de variables dépendantes

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	,592 ^a	,351	,291	2,45066	,351	5,814	4	43	,001

a. Predictors: (Constant), ND, Acc, Overspeed, Decc

b. Dependent Variable: QA_T

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	,521 ^a	,272	,204	17,42963	,272	4,013	4	43	,007

a. Predictors: (Constant), ND, Acc, Overspeed, Decc

b. Dependent Variable: QA3



Bibliographie

Bibliographie

Articles de recherche

1. Lorenzi J-H. L'importance économique de l'assurance. Les Tribunes de la santé. 2011(2):31-7.
7. Arrow KJ. Uncertainty and the welfare economics of medical care (American economic review, 1963). Journal of Health Politics, Policy and Law. 2001;26(5):851-83.
10. Hoy M. Categorizing risks in the insurance industry. The Quarterly Journal of Economics. 1982;97(2):321-36.
13. Henriot D, Rochet J-C. La logique des systèmes bonus-malus en assurance automobile: une approche théorique. Annales d'Économie et de Statistique. 1986:133-52.
14. Lambert RA. Long-term contracts and moral hazard. The Bell Journal of Economics. 1983:441-52.
15. Cooper R, Hayes B. Multi-period insurance contracts. International Journal of Industrial Organization. 1987;5(2):211-31.
16. Vasechko OA, Grun-Rehommé M. L'impact de la sinistralité passée sur la sinistralité future (2): une modélisation des classes de risques. : Insurance and Risk Management; 2011.
18. Van Elslande P, Alberton L, Nachtergaële C, Blanchet G. SCENARIOS-TYPES DE PRODUCTION DE " L'ERREUR HUMAINE " DANS L'ACCIDENT DE LA ROUTE: PROBLEMATIQUE ET ANALYSE QUALITATIVE. rapport INRETS. 1997(218).
19. Hydén C. The development of a method for traffic safety evaluation: The Swedish Traffic Conflicts Technique. Bulletin Lund Institute of Technology, Department. 1987(70).
20. Chin H-C, Quek S-T. Measurement of traffic conflicts. Safety Science. 1997;26(3):169-85.
21. Evans L. The dominant role of driver behavior in traffic safety. Am J Public Health. 1996;86(6):784-6.
31. Elander J, West R, French D. Behavioral correlates of individual differences in road-traffic crash risk: an examination method and findings. Psychol Bull. 1993;113(2):279-94.
33. Näätänen R, Summala H. Road-user behaviour and traffic accidents. Publication of: North-Holland Publishing Company. 1976.
34. Duncan J, Williams P, Brown I. Components of driving skill: experience does not mean expertise. Ergonomics. 1991;34(7):919-37.
36. Jensen M, Wagner J, Alexander K. Analysis of in-vehicle driver behaviour data for improved safety. International journal of vehicle safety. 2011;5(3):197-212.
37. Lotan T, Toledo T. Evaluating the safety implications and benefits of an in-vehicle data recorder to young drivers. 2005.

38. Svenson O. Are we all less risky and more skillful than our fellow drivers? *Acta psychologica*. 1981;47(2):143-8.
39. Paefgen V, Doleschel D, Kiessling F. Evolution of contrast agents for ultrasound imaging and ultrasound-mediated drug delivery. *Frontiers in pharmacology*. 2015;6:197.
41. Lehmann G. The features of the accident data recorder and its contribution to road safety. *Proceedings: International Technical Conference on the Enhanced Safety of Vehicles*. 1996;1996:1565-8.
42. Weidner W, Transchel FW, Weidner R. Telematic driving profile classification in car insurance pricing. *Annals of Actuarial Science*. 2017;11(2):213-36.
43. Moget-Monseur M, Biecheler-Fretel M. LE COMPORTEMENT DE BASE DU CONDUCTEUR: UN ESSAI DE CONCEPTUALISATION DU SYSTEME DE NORMES LEGALES ET SOCIALES DE L'USAGER DE LA ROUTE. *CAH ETUD ONSER*. 1985(64).
44. Goyal M. Insurance telematics. *International Journal of Innovative Research and Development*. 2014.
45. Husnjak S, Peraković D, Forenbacher I, Mumdziev M. Telematics system in usage based motor insurance. *Procedia Engineering*. 2015;100:816-25.
51. Chipman ML, MacGregor CG, Smiley AM, Lee-Gosselin M. The role of exposure in comparisons of crash risk among different drivers and driving environments. *Accident; analysis and prevention*. 1993;25(2):207-11.
52. Jun J, Guensler R, Ogle J. Differences in observed speed patterns between crash-involved and crash-not-involved drivers: Application of in-vehicle monitoring technology. *Transportation research part C: emerging technologies*. 2011;19(4):569-78.
53. Paefgen J, Staake T, Thiesse F. Evaluation and aggregation of pay-as-you-drive insurance rate factors: A classification analysis approach. *Decision Support Systems*. 2013;56:192-201.
54. Bian Y, Yang C, Zhao JL, Liang L. Good drivers pay less: A study of usage-based vehicle insurance models. *Transportation Research Part A: Policy and Practice*. 2018;107:20-34.
55. Rhodes N, Pivik K. Age and gender differences in risky driving: the roles of positive affect and risk perception. *Accident; analysis and prevention*. 2011;43(3):923-31.
56. Yannis G, Kanellopoulou A, Aggeloussi K, Tsamboulas D. Modelling driver choices towards accident risk reduction. *Safety science*. 2005;43(3):173-86.
57. Beirão G, Cabral JS. Understanding attitudes towards public transport and private car: A qualitative study. *Transport policy*. 2007;14(6):478-89.
58. Jun J, Ogle J, Guensler R. Relationships between crash involvement and temporal-spatial driving behavior activity patterns: Use of data for vehicles with global positioning systems. *Transportation Research Record: Journal of the Transportation Research Board*. 2007(2019):246-55.

60. Quddus MA, Noland RB, Chin HC. An analysis of motorcycle injury and vehicle damage severity using ordered probit models. *J Safety Res.* 2002;33(4):445-62.
61. Guo F, Fang Y. Individual driver risk assessment using naturalistic driving data. *Accident Analysis & Prevention.* 2013;61:3-9.
62. Gupta PB, Burns DJ, Boyd H. Texting while driving: An empirical investigation of students' attitudes and behaviors. *Information systems management.* 2016;33(1):88-101.
64. Paefgen J, Staake T, Fleisch E. Multivariate exposure modeling of accident risk: Insights from Pay-as-you-drive insurance data. *Transportation Research Part A: Policy and Practice.* 2014;61:27-40.
65. Tselentis DI, Yannis G, Vlahogianni EI. Innovative insurance schemes: pay as/how you drive. *Transportation Research Procedia.* 2016;14:362-71.
66. Ayuso M, Guillén M, Nielsen JP. Improving automobile insurance ratemaking using telematics: incorporating mileage and driver behaviour data. 2016.
68. Langford J, *et al.* Findings from the Candrive/Ozcandrive study: Low mileage older drivers, crash risk and reduced fitness to drive. *Accident; analysis and prevention.* 2013;61:304-10.
69. Litman T. Distance-based vehicle insurance feasibility, costs and benefits. Victoria. 2007;11.
70. Lourens PF, Vissers JA, Jessurun M. Annual mileage, driving violations, and accident involvement in relation to drivers' sex, age, and level of education. *Accident; analysis and prevention.* 1999;31(5):593-7.
72. Liu Z, Shen Q, Ma J. A driving behavior model evaluation for UBI. *International Journal of Crowd Science.* 2017;1(3):223-36.
77. Soleymanian M, Weinberg C, Zhu T. Sensor Data, Privacy, and Behavioral Tracking: Does Usage-Based Auto Insurance Benefit Drivers? 2017.
81. Vaia G, Carmel E, DeLone W, Trautsch H, Menichetti F. Vehicle telematics at an Italian insurer: new auto insurance products and a new industry ecosystem. *MIS Quarterly Executive.* 2012;11(3):113-25.
92. Hakkinen S, Luoma J. Liikennepsykologia (traffic psychology). *Traffic Psychology.* 1991;38.
94. Guo F, Klauer SG, Hankey JM, Dingus TA. Near crashes as crash surrogate for naturalistic driving studies. *Transportation Research Record.* 2010;2147(1):66-74.
96. Klauer SG, Dingus TA, Neale VL, Sudweeks JD, Ramsey DJ. The impact of driver inattention on near-crash/crash risk: An analysis using the 100-car naturalistic driving study data. 2006.
100. Cohen J. A power primer. *Psychological bulletin.* 1992;112(1):155.

101. Brookhuis KA, de Vries G, De Waard D. The effects of mobile telephoning on driving performance. *Accident Analysis & Prevention*. 1991;23(4):309-16.
102. Janke MK. Accidents, mileage, and the exaggeration of risk. *Accident Analysis & Prevention*. 1991;23(2-3):183-8.
103. Musicant O, Bar-Gera H, Schechtman E. Electronic records of undesirable driving events. *Transportation research part F: traffic psychology and behaviour*. 2010;13(2):71-9.

Sites Web

5. Universalis. Le principe de la mutualisation. ; . Available from: <https://www.universalis.fr/encyclopedie/assurance-economie-de-l-assurance/1-le-principe-de-mutualisation/>.
8. lafinancepourtous. Aléa moral [updated 29/11/2017]. ; . Available from: <https://www.lafinancepourtous.com/outils/dictionnaire/alea-moral/>.
9. France U. Antisélection, informations cachées sur les risques ;. Available from: <https://www.universalis.fr/encyclopedie/assurance-economie-de-l-assurance/3-antiselection/>.
17. Axa. Trois exemples de formules assurance auto en ligne ;. Available from: <https://www.axa.fr/assurance-auto/assurance-en-ligne.html>.
25. ABI. ABI average motor insurance premium tracker Q4 2014 data 2015. ; . Available from: <https://www.abi.org.uk/news/industry-data-updates/2015/02/abi-average-motor-insurance-premium-tracker-q4-2014-data/>.
27. Insurance FD. 5 Statistics that illustrate the current auto insurance industry 2016. ; . Available from: <https://www.firstdependable.com/blog/2016/09/5-statistics-that-illustrate-the-current-auto-insurance-industry>.
29. Fraud CAI. Automobile Bodily injury claims 2015. ; . Available from: <http://www.insurancefraud.org/downloads/InsuranceResearchCouncil02-15.pdf>.
30. Insurers AoB. Insurance fraud statistics : insight Investigations; 2017 [January 5, 2017]. ; . Available from: <https://www.investigate.uk/corporate-investigations-news/insurance-fraud-statistics/>
40. Cooper RB. Event data recorders: balancing the benefits and drawbacks 2008. ; . Available from: <https://www.irmi.com/articles/expert-commentary/event-data-recorders-balancing-the-benefits-and-drawbacks>.
46. EY. Introducing ‘Pay How You Drive’ (PHYD) Insurance: Insurance that rewards safe driving : EY; 2016. ; . Available from: [http://www.ey.com/Publication/vwLUAssets/ey-introducing-pay-how-you-drive-insurance/\\$FILE/ey-introducing-pay-how-you-drive-insurance.pdf](http://www.ey.com/Publication/vwLUAssets/ey-introducing-pay-how-you-drive-insurance/$FILE/ey-introducing-pay-how-you-drive-insurance.pdf).
48. InsureTheBox.com. WHAT IS TELEMATICS? ; . Available from: <https://www.insurethebox.com/telematics>.

49. l'assurance Lad. Le développement de la voiture connectée : L'argus de l'assurance; 2015. ; . Available from: <http://www.argusdelassurance.com/jurisprudences/jurisprudence-ja/le-developpement-de-la-voiture-connectee.99528>.
50. IMS. Customer Case Study : Nationwide. ; . Available from: <https://cdn2.hubspot.net/hubfs/604918/PDFs/Case%20Studies/Nationwide-Case-Study.pdf?t=1522874086020>.
74. Telematics O. Telematics and the global auto insurance market : Octo Telematics; 2015. ; . Available from: <https://www.octotelematics.com/focus-on/telematics-and-the-global-auto-insurance-market>.
80. IBM. 5 ways to leverage telematics for insurance 2015. ; . Available from: <http://www.ibmbigdatahub.com/blog/5-ways-leverage-telematics-insurance>.
85. LexisNexis. Usage-Based Insurance (UBI) Research Results for Consumer and Small Fleet Markets 2014. ; . Available from: <https://risk.lexisnexis.com/-/media/files/insurance/white-paper/2014-ubi-research-pdf.pdf>.
86. Confused.com. Car insurance price index The most comprehensive analysis of car insurance pricing in the UK 2017. ; . Available from: <https://azcdubvermedia.azureedge.net/media/themes/fab-four/Price-index/layout/2017-Q4/confused-com-car-insurance-price-index-q4-2017.pdf>.
90. 23 SS. Régression multiple: Rappel théorique. ; . Available from: <http://spss.espaceweb.usherbrooke.ca/pages/stat-inferentielles/regression-multiple.php>.
93. OCDE. Road Accidents. ; . Available from: <https://data.oecd.org/transport/road-accidents.html>
98. France Ddt. La sécurité routière 2009. ; . Available from: <http://www.transports-terrestres.pf/spip.php?article409>.
99. La sécurité routière de A à Z. ; . Available from: <http://www.transports-terrestres.pf/spip.php?article409>.

Ouvrages

2. Albert M. Capitalisme contre capitalisme: Seuil Paris; 1991.
3. Charbonnier J. L'assurance du risque automobile, contrôle et assurance 2012. 5-129 p.
4. Couilbault F, Tommaso SC-D, Huberty V. Les grands principes de l'assurance 2013.
6. Rothschild M, Stiglitz J. Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. Uncertainty in economics: Elsevier; 1978. p. 257-80.
11. S.Chan, J.Landel. Manuel d'assurance automobile. ed 2016.
32. Evans L. Traffic safety and the driver: Science Serving Society; 1991.
35. Porter BE. Handbook of traffic psychology: Academic Press; 2011.

Rapports, livres blancs, études et revues professionnelles

22. santé Omdl. Rapport de situation sur la sécurité routière Organisation mondiale de la santé, 2015.
23. Dorothy K, Michael G, Steve L, Doug T. The Challenge of Auto Insurance Premium Leakage Verisk Insurance Solutions, 2017.
24. Ingenie. Ingenie Young Driver Report. Ingenie, 2014.
26. Hartwig RP, Lynch J, Weisbart S. More Accidents, Larger Claims Drive Costs Higher. Insurance Information Institution, 2016.
28. Europe I. The impact of insurance fraud. Insurance Europe, 2013.
63. Insurance P. Texas Mileage Study: Relationship between Annual Mileage and Insurance Losses. 2005.
71. Klauer SG, Dingus TA, Neale VL, Sudweeks JD, Ramsey DJ. Comparing real-world behaviors of drivers with high versus low rates of crashes and near crashes. 2009.
73. Re S. The future of motor insurance : How car connectivity and ADAS are impacting the market. Swiss Re, 2016.
75. Group PC. L'assurance auto connectée en Europe. PTOLEMUS Consulting Group, 2016.
76. Deloitte. European Motor Insurance Study : The rise of digitally-enabled motor insurance. Deloitte, 2016.
78. Accenture. Insurance telematics: A game-changing opportunity for the industry Accenture, 2014.
79. Watson T. Usage-Based Insurance Consumer Survey Towers Watson, 2013.
82. Reddy AS. The New Auto Insurance Ecosystem: Telematics, Mobility and the Connected Car. 2012.
84. Mai A, Schlesinger D. Connected Vehicle Insurance : The Business of Preventing Crashes. Cisco, 2011.
87. Ingenie. Young Driver Report Ingenie, 2017.
88. Bordoff J, Noel P. Pay-As-You-Drive Auto Insurance: A Simple Way to Reduce Driving-Related Harms and Increase Equity. Brooklings Institute, 2008.
89. Karapiperis D, *et al.* Usage-Based Insurance and Vehicle Telematics: Insurance Market and Regulatory Implications. NAIC, 2015.
91. Dingus TA, *et al.* The 100-car naturalistic driving study, Phase II-results of the 100-car field experiment. 2006.

95. Klauer SG, Guo F, Sudweeks J, Dingus TA. An analysis of driver inattention using a case-crossover approach on 100-car data. 2010.
97. ONSR. Statistiques des accidents de circulation. ONSR, 2018.
104. Collet I, Chouchet Md. Etude sur la tarification de l'assurance automobile à travers le monde. Paris: Fixage, 2017.

Thèses et mémoires

12. Prévot P. Management de l'Innovation dans l'Assurance Automobile [mémoire de recherche]: École Polytechnique ; 2005.

Colloques

47. Tindall J. Telematics - Movements in Behavioural Underwriting & Pricing Hilton Sidney 2013.
59. Zheng Y, *et al.*, editors. Driving risk assessment using cluster analysis based on naturalistic driving data. Intelligent Transportation Systems (ITSC), 2014 IEEE 17th International Conference on; 2014 : IEEE.
67. Boucher J-P, Pérez-Marín AM, Santolino M, editors. Pay-as-you-drive insurance: the effect of the kilometers on the risk of accident. Anales del Instituto de Actuarios Españoles; 2013 : Instituto de Actuarios Españoles.
83. Dou D. UBI - Telematics in Motor Insurance SOA Actuarial Seminar; Seoul, South Korea 2015.