



**HAL**  
open science

# Segmentation automatique multidimensionnelle et classification de motifs temporels : application à la reconnaissance de situation de conduite en virage des deux roues motorisés

Mohamed Diop

► **To cite this version:**

Mohamed Diop. Segmentation automatique multidimensionnelle et classification de motifs temporels : application à la reconnaissance de situation de conduite en virage des deux roues motorisés. Traitement du signal et de l'image [eess.SP]. Université Gustave Eiffel, 2022. Français. NNT : 2022UEFL2046 . tel-04072516

**HAL Id: tel-04072516**

**<https://theses.hal.science/tel-04072516v1>**

Submitted on 18 Apr 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE DE DOCTORAT

Spécialité: Génie informatique

présentée par

MOHAMED DIOP

---

Segmentation automatique multidimensionnelle et classification  
de motifs temporels : application à la reconnaissance de  
situations de conduite en virage des 2 Roues motorisés

---

Soutenue publiquement le 02 Décembre 2022 devant le jury composé de :

M. Mustapha LEBBAH	Professeur des Universités, (UVSQ),	Rapporteur
M. Niccolò BALDANZINI	Associate Professor, (Université de Florence),	Rapporteur
Mme. Zoi CHRISTOFOROU	Assistant Professor , (Université de Patras),	Examinatrice
M. Samer Mohammed	Professeur des Universités, (UPEC),	Examinateur
M. Paul HONEINE	Professeur des Universités, (Université de Rouen-Normandie),	Examinateur
Mme. Latifa OUKHELLOU	Directrice de Recherche, (Univ-Eiffel),	Directrice
M. Stéphane ESPIÉ	Directeur de Recherche, (Univ-Eiffel),	Co-Directeur
M. Abderrahmane BOUBEZOUL	Chargé de Recherche, (Univ-Eiffel),	Encadrant



---

*À mes parents,  
À ma femme,  
À mon fils,*



## Remerciements

Il me sera très difficile de remercier tout le monde car c'est grâce à l'aide de nombreuses personnes que j'ai pu mener cette thèse à son terme.

Je tiens tout d'abord à exprimer mes plus profonds remerciements à ma directrice de thèse Mme Latifa OUKHELLOU et mon Co-Directeur de thèse M. Stéphane ESPIÉ pour avoir accepté de co-diriger cette thèse. Je leur exprime toute mon estime et leur remercie pour leurs précieux conseils et encouragements. J'aimerais également remercier mon encadrant M. Abderrahmane BOUBEZOUL de m'avoir offert un encadrement certes exigeant, mais doublé d'une grande disponibilité.

Je tiens également à remercier les membres du jury : Mme Zoi CHRISTOFOROU Professeure Assistante à l'Université de Patras, M. Samer Mohammed Professeur des Universités à l'UPEC, M. Paul HONEINE Professeur des Universités à l'Université de Rouen-Normandie ainsi qu'aux deux rapporteurs M. Niccolò BALDANZINI Professeur Associé à l'Université de Florence et M. Mustapha LEBBAH Professeur des Universités à l'UVSQ. Merci pour votre lecture attentive de la thèse.

J'adresse ma sympathie à l'ensemble des personnes que j'ai pu côtoyées durant cette thèse plus particulièrement celles avec qui j'ai partagé le même bureau : Aquib, Youness, Maria, Hadj, Najeh, Mathias...

Je tiens à exprimer mes plus profonds remerciements à ma famille particulièrement à mon père et ma mère pour le soutien et la confiance qu'ils m'ont accordé tout au long de mes études.

Enfin et pas la moindre, je tiens à exprimer mes profonds sentiments et mes remerciements à ma femme pour tout son soutien.

Mohamed DIOP  
Paris, le 02 Décembre 2022



---

## Résumé

La thématique « deux-roues motorisés » (2RM) est un sujet à forte attente sociétale en termes de sécurité routière. Les statistiques d'accidents montrent que les conducteurs de 2RMs font partie des usagers de la route les plus vulnérables. L'accident en virage représente, pour le 2RM, plus de 50% des cas d'accidents impliquant un « véhicule seul » d'où la nécessité d'apporter des solutions. Cette thèse s'attache à l'analyse des signaux issus d'expérimentations réelles conduites dans le projet ANR VOROLO++. Ces dernières ont été menées sur une moto instrumentée dans un contexte de conduite réelle incluant des conduites sous différentes instructions avec des sujets ayant différents niveaux de l'expérience de conduite : novices et expérimentés. Dans la première partie de la thèse, des méthodes d'apprentissage non supervisé ont été développées pour l'analyse globale de la trajectoire d'un 2RM lors de la prise de virage. Cette méthodologie a été mise en œuvre pour la classification des différents comportements de conduite observés. Les résultats obtenus à l'issue de la mise en œuvre de cette méthodologie sont analysés puis interprétés. En outre, une analyse approfondie du comportement de conduite d'un conducteur en particulier a été menée, et ce dans le but d'étudier les actions du conducteur nécessaires au contrôle de l'inclinaison du véhicule dans les différentes phases du virage. La seconde partie de cette thèse porte sur le développement d'outils de segmentation et de classification de séquences de conduite d'un 2RM lors de la prise de virage. Pour ce faire, l'approche proposée repose sur un modèle de régression logistique à processus cachés. Ce modèle sert à la segmentation non supervisée des séquences de conduite. Des caractéristiques statistiques extraites de chaque segment sont utilisées comme entrées d'un classifieur pour classifier les comportements de conduite à l'entrée du virage. Les résultats obtenus sur une base de données réelle ont permis de montrer l'efficacité de la méthodologie proposée à la fois pour la détection de l'entrée du virage et pour la détection de styles de conduite à l'approche du virage.

---

## Abstract

The topic of Powered Two-Wheelers (PTWs) is a subject with high societal expectations regarding road safety. Accident statistics show that PTWs drivers are among the most vulnerable road users. Accidents on bends account for more than 50% of accidents involving a single vehicle, hence the need to find solutions. This thesis considers the analysis of signals from real experiments conducted in the ANR VOROLO++ project. These experiments were conducted on an instrumented motorcycle in a real driving context, including driving under different instructions with subjects having different levels of driving experience: novice and experienced. In the first part of the thesis, unsupervised learning methods were used for the global analysis of the trajectory of a PTW during bend taking. This methodology has been proposed to classify the different driving behaviors observed. The results obtained from the implementation of this methodology are analyzed and interpreted. In addition, an in-depth analysis of a particular driver's driving behavior was conducted, aiming to study the driver's actions necessary to control the vehicle's inclination in the different turn phases. The second part of this thesis deals with developing tools for segmenting and classifying driving sequences of PTWs during cornering. The proposed approach is based on a hidden process logistic regression model to achieve this aim. This model is used for the unsupervised segmentation of driving sequences. Statistical features extracted from each segment are used as inputs to a classifier to classify the driving behaviors at the turn entry. The results obtained on a real database have shown the effectiveness of the proposed methodology both for detecting the turn entry and for detecting driving styles at the approach of the turn.

---

## Table des matières

---

<b>Table des figures</b>	<b>ii</b>
<b>Liste des tableaux</b>	<b>v</b>
<b>Glossaire des abréviations</b>	<b>vii</b>
<b>1 Introduction générale</b>	<b>1</b>
1.1 Contexte et motivations	2
1.1.1 Particularités et difficultés de la conduite d'un 2RM en virage	4
1.1.2 La trajectoire de sécurité	5
1.2 Problématiques de recherche et contributions	7
1.3 Organisation du manuscrit	8
<b>2 Protocole expérimental et acquisition de données</b>	<b>11</b>
2.1 Introduction	12
2.2 Protocole expérimental	13
2.2.1 Description des consignes de conduite	13
2.2.2 Description de l'instrumentation et des capteurs utilisés	15
2.3 Acquisition de données	19
2.3.1 Pré-traitement des données collectées	19
2.3.2 Représentation sous formes de base de données des mesures fournies par les différents capteurs	21
2.3.3 Aperçu de quelques signaux collectés lors de l'expérimentation	21
2.3.4 Annotation manuelle des trajectoires	23
2.4 Conclusion	23
<b>3 Étude globale de la trajectoire lors de la prise de virage</b>	<b>25</b>
3.1 Introduction	26

3.2	État de l'art . . . . .	26
3.2.1	Modélisation du pilotage de 2RMs . . . . .	27
3.2.2	Étude du comportement de conduite dans le cas d'un 2RM . . . . .	28
3.3	Méthodologie de classification des différents profils de conduite . . . . .	29
3.3.1	Description et représentation des données utilisées . . . . .	31
3.3.2	Méthodologie de classification des participants . . . . .	32
3.4	Application de la méthodologie proposée : Résultats et Discussions . . . . .	37
3.4.1	Analyses des classes fournies par les dendrogrammes . . . . .	39
3.4.2	Interprétation des classes fournies par les dendrogrammes . . . . .	39
3.4.3	Analyse approfondie des comportements de conduite du participant P4 . . . . .	42
3.5	Conclusion . . . . .	49
<b>4</b>	<b>Segmentation et analyse de la prise de virage</b>	<b>51</b>
4.1	Introduction . . . . .	52
4.2	Analyse approfondie de la prise de virage dans le cas de 2RMs . . . . .	52
4.2.1	Formalisation du problème de la segmentation et analyse de la prise de virage . . . . .	52
4.2.2	Segmentation de signaux multidimensionnels avec les modèles MHMMR et MRHLP . . . . .	53
4.2.3	Stratégie de détection et de classification des segments de virage . . . . .	57
4.3	Application de la méthodologie proposée et résultats . . . . .	58
4.3.1	Résultats de la détection et la classification des segments de conduite dans un virage . . . . .	59
4.3.2	Caractérisation statistique des segments <b>S1</b> et <b>S2</b> . . . . .	60
4.4	Conclusion . . . . .	75
	<b>Conclusion générale et perspectives</b>	<b>77</b>
<b>5</b>	<b>Conclusion et Perspectives</b>	<b>77</b>
	<b>Bibliographie</b>	<b>81</b>

---

## Table des figures

---

1.1	Part dans le trafic routier des véhicules motorisés en (2019). . . . .	2
1.2	Répartition des personnes tuées et blessées selon le mode de déplacement utilisé. . . . .	3
1.3	Illustration de la trajectoire de sécurité pour un virage à droite. . . . .	6
2.1	Circuit de la piste « La Ferté-Gaucher » le rectangle rouge matérialise le virage étudié. . . . .	14
2.2	Instrumentation réalisée au niveau de la moto (a) et du motard (b). . . . .	16
2.3	Logiciel de visualisation / export multi-sources et multi-canaux. . . . .	20
2.4	Évolution : de la position du véhicule (a) temporelle de la vitesse (b) et de l'angle guidon (c) tout au long du virage pour tous les sujets pendant les trois tours la consigne CL. . . . .	22
2.5	Un exemple d'annotation manuelle d'une trajectoire en combinant les données vidéos, GPS et données capteurs. . . . .	24
2.6	Représentation de l'angle et la vitesse de roulis ainsi que l'angle de guidon, et la vitesse longitudinale du véhicule pour une trajectoire segmentée en 4 segments : Segment1 (S1) : rouge, Segment2 (S2) : vert, Segment3 (S3) : bleu et Segment4 (S4) : violet. . . . .	24
3.1	Résumé de la méthodologie de classification des profils de conduite lors de la prise de virage. . . . .	30
3.2	Processus de mise en place de la base de données extraite. . . . .	31
3.3	Un exemple d'application de l'algorithme PAA sur la séquence de conduite $\mathbf{x}_{k,m}$ de taille $M=124$ correspondant au deuxième tour du participant P6 lors de la consigne de conduite CL. . . . .	33
3.4	Comportement moyen de conduite des trois tours du participant P2 pour la consigne de conduite libre. . . . .	35

3.5	Résultats de la classification obtenues par l'algorithme Classification ascendante hiérarchique (CAH) avec les métriques : déformation temporelle dynamique (DTD) et corrélation de Pearson. . . . .	38
3.6	Distribution de l'écart-type par classe pour les 5 variables de conduite : (a) Angle du guidon, (b) Vitesse, (c) Accélération Normale, (d) Jerk et (e) Courbure. . . . .	41
3.7	Conventions utilisées pour (a) le Roulis ( $\phi$ ) et (b) les angles de guidon ( $\delta$ ) et les forces ( $Fg_x, Fd_x$ ) appliquées sur le guidon. . . . .	43
3.8	Angle du guidon $\delta(\dots)$ et Angle de roulis $\phi(-)$ exprimés en degré ( $^\circ$ ) du participant P4 pour les instructions CL, SC et AC. La zone grisée en vert représente l'entrée du virage, le début de cette zone correspond à l'instant où l'angle de roulis franchit le seuil fixé à $0^\circ$ . . . . .	44
3.9	Forces appliquées en Newton, suivant l'axe horizontal, sur les demi-guidons droit ( $Fdx$ ) et gauche ( $Fgx$ ) et l'angle du guidon ( $\delta$ ) en degré ( $^\circ$ ) lors des instructions CL, SC et AC. . . . .	45
3.10	Forces des repose-pieds gauche ( $Fg_{r-p}$ ) et droit ( $Fd_{r-p}$ ), des fesses gauche ( $FgF$ ) et droite ( $FdF$ ) exprimées en Newton (N) et l'angle de roulis ( $\phi$ ) en degré ( $^\circ$ ) lors de la consigne de conduite CL. . . . .	46
3.11	Forces des repose-pieds gauche ( $Fg_{r-p}$ ) et droit ( $Fd_{r-p}$ ), des fesses gauche ( $FgF$ ) et droite ( $FdF$ ) exprimées en Newton (N) et de l'angle roulis ( $\phi$ ) en degré ( $^\circ$ ) lors de la consigne de conduite AC. . . . .	47
3.12	Forces des Repose-pieds gauche ( $Fg_{r-p}$ ) et droit ( $Fd_{r-p}$ ), des fesses gauche ( $FgF$ ) et droite ( $FdF$ ) exprimées en Newton (N) et de l'angle roulis ( $\phi$ ) en degré ( $^\circ$ ) lors de la consigne de conduite SC. . . . .	48
4.1	Résumé de la méthodologie de segmentation et d'analyse de la prise de virage. . . . .	58
4.2	P-valeurs du test statistique de Mann-Whitney entre les scores moyens des modèles MRHLP, MHMMR et GHMM sur chacune des transitions : (a) <b>S1</b> $\mapsto$ <b>S2</b> , (b) <b>S2</b> $\mapsto$ <b>S3</b> et (c) <b>S3</b> $\mapsto$ <b>S4</b> . . . . .	60
4.3	Un exemple de caractérisation des segments S1 (rouge) et S2 (verte) pour une séquence donnée $\mathbf{Y}^m$ . . . . .	62
4.4	Résultats de la classification des participants pour la consigne de conduite libre (CL). . . . .	64
4.5	L'évolution de l'angle roulis $\phi$ sur le segment S1 (à gauche) et la distribution des dernières valeurs de $\phi$ sur le segment S1 pour tous les participants (à droite). . . . .	65
4.6	Projection en 3D des données de la consigne de conduite CL. . . . .	67
4.7	Valeurs des « loadings » de l'ACP . . . . .	68
4.8	Les résultats de la classification des données collectées durant l'instruction CL. La carte de chaleur est triée selon la valeur $\phi_1$ . La dynamique du roulis : angle de roulis et vitesse de roulis sont représentés de gauche à droite, respectivement. Chaque ligne représente une classe. . . . .	70

## Table des figures

---

4.9	Évolution de l'angle du guidon dans le segment S2 pour les séquences constituant les deux classes : ■ et ■ . . . . .	71
4.10	Comportement dynamique des deux séquences P2.T1 et P5.T3 pendant la manœuvre de prise de virage segmentée en différentes phases : de gauche à droite ( S1 : en rouge, S2 : en vert, S3 : en bleu, S4 : en violet). . . . .	73
4.11	Résultats de la classification des tours de conduite pour chacun des 7 participants.	74
4.12	Trajectoires représentées sur Google Earth des trois séquences de conduite réalisées par le participant P4. . . . .	74



---

## Liste des tableaux

---

1.1	Caractéristiques des différentes catégories de 2RMs en France. . . . .	2
2.1	Données déclaratives des différents participants et ordre des instructions. . . . .	14
2.2	L'ensemble des capteurs utilisés ainsi que le nom du paramètre mesuré et la fréquence d'acquisition. . . . .	17
3.1	$p$ -valeurs du test d'Anderson-Darling pour toutes les instructions et variables. . .	34
3.2	Résultats de la classification obtenues par l'algorithme Classification ascendante hiérarchique (CAH) avec les métriques : déformation temporelle dynamique (DTD) et corrélation de Pearson. . . . .	37
3.3	$p$ -valeurs du test d'AD entre les classes 2 et 3. . . . .	40
4.1	Score moyen de la métrique « indice de Rand » sur les 63 séquences de conduite en virage pour les modèles MRHLP, MHMMR et GHMM. . . . .	59
4.2	Description des différents indicateurs statistiques. . . . .	61

**Glossaire des abréviations :**

2RMs	Deux Roues Motorisés.
VL	Véhicule légers.
ONISR	Observatoire national interministériel de la sécurité routière.
V-I-C	Véhicule-Infrastructure-Conducteur.
GPS	Global Positioning System.
CL	Conduite Libre.
AC	Conduite avec Corps.
SC	Conduite sans Corps.
PAA	Piecewise Aggregate Approximation
AD	Test d'Anderson-Darling
CAH	Classification Ascendante Hiérarchique
DTD	Déformation Temporelle Dynamique
MHMMR	Multiple Hidden Markov Model Regression model”
RHLP	Regression model with a Hidden Logistic Process.
MRHLP	Multiple Regression Hidden Logistic Process
GHMM	Gaussian Hidden Markov Model
EM	Espérance-Maximisation
ACP	Analyse en Composantes Principales
GMMs	Gaussian Mixture Model.

---

## Introduction générale

---

### Sommaire

---

<b>1.1</b>	<b>Contexte et motivations</b>	<b>2</b>
1.1.1	Particularités et difficultés de la conduite d'un 2RM en virage	4
1.1.2	La trajectoire de sécurité	5
<b>1.2</b>	<b>Problématiques de recherche et contributions</b>	<b>7</b>
<b>1.3</b>	<b>Organisation du manuscrit</b>	<b>8</b>

---

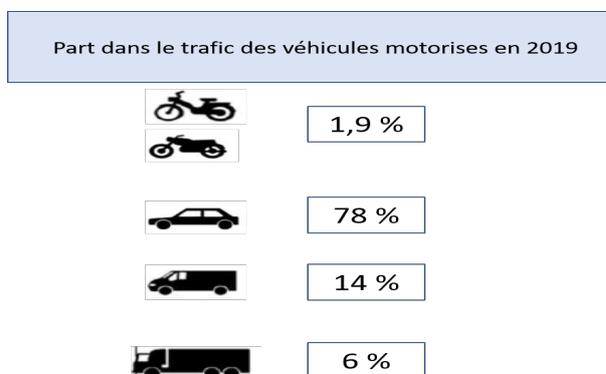
## 1.1 Contexte et motivations

Les modèles de deux-roues motorisés (2RMs) peuvent être catégorisés en deux grandes familles : les cyclomoteurs et les motocyclettes. Les motocyclettes peuvent être aussi subdivisées en deux catégories : motocyclette légère et motocyclette lourde. Le Tableau 1.1 donne les différentes caractéristiques des catégories de 2RMs circulant en France [52].

Catégorie	Vitesse (km/h)	Cylindrée (cm <sup>3</sup> )	Puissance (kW)	Poids moyen (kg) avec carburant
Cyclomoteur	<45	<50	<4	50 à 100
Motocyclette légère	Pas de limite	50 à 125	4 à 11	125 à 170
Motocyclette lourde	Pas de limite	>125	>11	>170

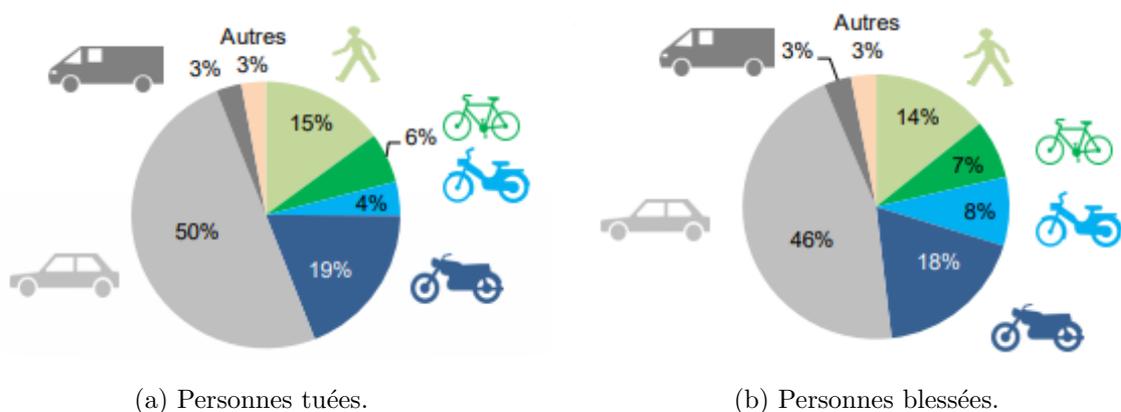
**Tableau 1.1** : Caractéristiques des différentes catégories de 2RMs en France.

Les Motocyclettes légères sont particulièrement adaptées aux trajets quotidiens surtout dans les grandes villes où la circulation est très dense aux heures de pointe. Même si elles sont plus utilisées en ville, la prévisibilité des temps de parcours, la facilité de stationnement et une faible consommation de carburant comparés aux voitures, et pour des cycles urbains, ainsi qu'une facilité d'entretien les rendent aussi économiques que pratiques. Pour toutes ces raisons, ce n'est pas surprenant de voir de plus en plus d'automobilistes se tourner vers ces catégories de 2RMs pour les déplacements quotidiens ou de loisirs. L'estimation du nombre de 2RMs en circulation reste une tâche très difficile. En effet, il n'est pas rare de voir certains propriétaires conserver leurs précédents engins lors de l'acquisition de nouveaux. Néanmoins, le volet consacré aux 2RMs de l'enquête annuelle « Parc Auto » publiée par KANTAR TNS (anciennement TNS Sofres) mentionne qu'en 2019 le parc de 2RMs est estimé à plus de 2,9 millions de véhicules [48]. Cette enquête réalisée de Mars à Avril 2020 sur un panel de 10000 foyers représentatifs de la population française a montré que 8,7% des ménages sont équipés d'au moins un 2RM en 2019. Par ailleurs, selon l'Observatoire National Interministériel de la Sécurité Routière (ONISR), les 2RMs représentent 1,9% dans le trafic routier contre 78% pour les véhicules de tourisme en France en 2019, voir figure 1.1.



**Figure 1.1** : Part dans le trafic routier des véhicules motorisés en (2019).

Avec cette estimation à moins de 2% du trafic routier, le sur-risque d'être tué des usagers de 2RM est particulièrement élevé. Ce sur-risque, ramené aux kilomètres parcourus, est évalué respectivement à 17 et 22 fois supérieur pour un cyclomotoriste et un motocycliste que pour un conducteur de véhicule de tourisme. Selon l'ONISR, en 2019 la catégorie de conducteurs classée dans la rubrique 2RM ; à savoir : motocyclistes et cyclomotoristes, représentaient 23% (4% pour les cyclomotoristes et 19% pour les motocyclistes) des tués et 26% des blessés graves (8% pour les cyclomotoristes et 18% pour les motocyclistes), cf figure 1.2a. Pour un cyclomotoriste, les accidents sont 6 fois plus mortels hors agglomération contre 1,3 fois en agglomération. Du côté des motocyclistes, les accidents mortels surviennent le plus souvent en agglomération particulièrement en intersection. Les accidents de 2RM sont ceux qui laissent le plus de séquelles lourdes aux blessés avec majoritairement des lésions orthopédiques : membres inférieurs et/ou bassin (64%), puis membres supérieurs (50%), voir figure 1.2b.



**Figure 1.2** : Répartition des personnes tuées et blessées selon le mode de déplacement utilisé.

Ces statistiques d'accidentologie classent cette catégorie d'usagers du trafic routier comme étant les plus vulnérables. Par conséquent, il est plus que nécessaire d'approfondir nos connaissances sur les règles comportementales adoptées par les conducteurs de 2RM ou par les autres usagers de la route vis-à-vis de ces véhicules. Cela semble d'autant plus nécessaire que la littérature abordant la question des deux-roues s'est essentiellement développée sur le thème de l'accidentologie, on peut citer comme exemples de ces études les projets : **RIDER**<sup>1</sup> (2003 – 2005), et **MAIDS**<sup>2</sup> (2004 – 2008) qui ont en particulier permis de cartographier les situations d'accidents, et sont à la base de travaux de recherche tels que **DAMOTO** (ANR/PREDIT) (2008 – 2011) pour l'étude et le développement de gilets « airbags » sans fils, les projets européens **SAFE-RIDER**<sup>3</sup> (2008-2011) (7ème PCRD) avait pour objectif le développement des systèmes d'assistance à la conduite ou d'aide à la navigation , **2BESAFE**<sup>4</sup> (7ème PCRD) (2009 – 2011) pour les questions de saillance visuelle mais aussi de comportements de conduite. Le projet **SIM2CO+**

<sup>1</sup>RIDER : Recherche sur les accidents Impliquant un DEux-Roues motorisé

<sup>2</sup>MAIDS : Motorcycle Accidents In Depth Study : <https://www.maids-study.eu/>

<sup>3</sup>SAFERIDER : Advanced Telematics for enhancing the safety and comfort of motorcycle riders

<sup>4</sup>2BESAFE : 2-Wheeler BEhaviour and SAFEty. <https://www.2besafe.eu/about/>

(ANR/PREDIT) (2010-2013) pour l'étude des difficultés des novices et la question de la formation aux habiletés cognitives de conduite.

Si la majeure partie des accidents a lieu en interaction avec d'autres usagers, les accidents sans tiers identifiés représentent un enjeu conséquent pour les 2RMs. En effet, d'après l'étude **MAIDS** en 2004, près de 30% des accidents de 2RM surviennent dans les virages. De même, l'ONISR dans son bilan de l'accidentalité de l'année 2019 rapporte que les motocyclistes décèdent en courbe dans 49% des cas en dehors de l'agglomération. Ces statistiques assez alarmantes montrent un besoin assez urgent pour sensibiliser les conducteurs de 2RMs sur la conduite en virage à travers une prise de virage « **sécuritaire** ». Ce concept de prise de virage sécuritaire repose sur deux piliers, le premier étant, la prise en compte des particularités et difficultés de la conduite d'un 2RMs, et le deuxième pilier, étant le choix de la trajectoire optimale au sens « sécuritaire ». Aborder les virages de façon « optimale », permettrait au pilote de mieux réagir face aux obstacles se trouvant sur sa trajectoire et de garder une plus grande distance de sécurité par rapport aux véhicules venant en sens inverse. Savoir conduire en toute sécurité dans les virages, c'est réduire considérablement le risque d'accidents.

### 1.1.1 Particularités et difficultés de la conduite d'un 2RM en virage

Conduire un véhicule de type 2RM est tâche difficile due à la recherche permanente de stabilité. Étant donné le caractère intrinsèquement instable du 2RM, les différences de caractéristiques des véhicules classés sous la rubrique « 2RM » ainsi que la variabilité intra et inter individuelle des conducteurs, le problème est particulièrement complexe. Pour la prise de virage, nous pouvons noter une grande variabilité dans les pratiques réelles des conducteurs. Si les trajectoires produites permettent aux différents conducteurs de franchir le virage, les moyens et les stratégies mises en oeuvre par les conducteurs diffèrent. La prise d'un virage dans le cas d'un véhicule de type mono-voie « pendulaire », un 2RM par exemple, implique que le conducteur doit incliner son véhicule pour compenser l'inertie engendrée par un changement de direction [15].

Pour ce faire, il faut provoquer un basculement sur l'axe du roulis vers le côté où on veut aller en entreprenant le changement de direction. Ce changement de direction peut être initié, selon les pratiques de chaque conducteur, par un changement de la position du centre de gravité du conducteur, ou bien par un phénomène caractéristique des véhicules deux-roues qui est le contre braquage et qui fait que l'on peut braquer à droite pour incliner le véhicule à gauche et inversement. En plus de ces particularités et difficultés de la conduite propres à ce type de véhicule, nous avons une autre particularité bien spécifique au 2RM qui est le positionnement multiple sur la chaussée. En effet, le faible gabarit des 2RMs a une incidence sur leur positionnement latéral sur la chaussée. Le conducteur d'un 2RM dispose de plus de degrés de liberté qu'un conducteur de quatre roues dans ce positionnement.

### 1.1.2 La trajectoire de sécurité

La trajectoire de sécurité telle qu'elle pratiquée en conduite sur route est différente de la trajectoire « optimale » pratiquée en compétition sur piste, car elle ne vise pas à maximiser la vitesse de passage mais à proposer un compromis entre la vitesse de passage et la sécurité pour le conducteur. En effet, pour la trajectoire de sécurité, on prend en compte les éventuels obstacles et défauts de chaussée présents dans ou en sortie de virage (autres véhicules, nids de poule, mais aussi graviers, sable, ...). Dans cette optique, on va chercher à ouvrir l'angle de vision et permettre la découverte du virage et de sa sortie, et ce afin de permettre au conducteur d'anticiper au plus tôt sur les éventuels obstacles. La prise de virage sur piste n'a pas ces contraintes dans la mesure où la visibilité est assurée ainsi que la garantie d'une chaussée exempte de défauts ou dépôts divers (ces derniers étant signalés par un commissaire de piste le cas échéant).

La [figure 1.3](#) illustre une trajectoire de sécurité lors d'une prise de virage à droite.

On peut y remarquer que la trajectoire est divisée en 4 phases matérialisées par les couleurs : vert (■), orange (■), bleu (■) et jaune (■), labellisées par les lettres E, D, S et R. Chacune de ces quatre phases est décrite comme suit :

- **Phase d'entrée (E)** : C'est la phase d'adaptation de la vitesse où le conducteur adapte son allure lorsqu'il s'apprête à entrer dans le virage ;
- **Phase de découverte (D)** : Ici, le conducteur est appelé à se positionner le plus possible vers l'extérieur du virage, c'est-à-dire vers la ligne médiane pour un virage à droite ou en restant sur sa droite pour prendre un virage vers la gauche de sorte à porter son regard le plus loin possible ;
- **Phase de sollicitation (S)** : Pour cette phase le conducteur incline son 2RM en s'assurant toujours de porter son regard le plus loin possible
- **Phase de reprise (R)** : Enfin, dans cette dernière phase de la trajectoire de sécurité le conducteur se repositionne sur la droite de la chaussée.

En plus de permettre aux usagers de 2RMs de réduire les risques d'accident lors de la prise de virage, une trajectoire de sécurité oblige les usagers à faire montre des 5 points suivants :

- **Observer**, il s'agit d'être attentif en permanence à la lecture de la route ;
- **Anticiper**, recherchant l'information du regard le plus loin possible ;
- **Analyser**, en étant en mesure de s'adapter aux situations rencontrées ;
- **Adapter sa vitesse**, avec l'objectif de maîtriser sa moto en toutes circonstances ;
- **Concentré**, en maîtrisant sa manœuvre par des actions sur les commandes.

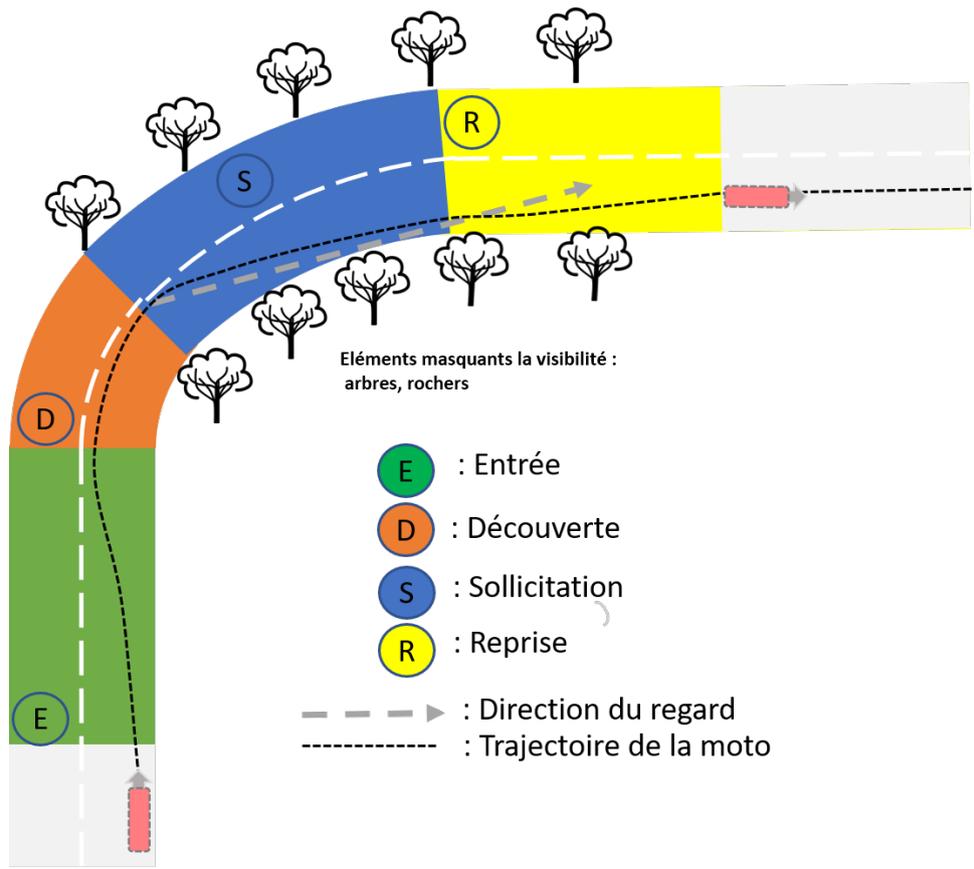


Figure 1.3 : Illustration de la trajectoire de sécurité pour un virage à droite.

Il est à noter que le principal inconvénient de cette trajectoire de sécurité réside dans le fait que le conducteur doit retarder le plus possible l'entrée dans le virage, l'obligeant ainsi à réaliser une manœuvre plus « brutale » afin d'incliner la moto pour créer une rupture, "un point d'inflexion" dans la trajectoire, et ce pendant la phase sollicitation. Et par conséquent, la mise en oeuvre de cette dernière repose essentiellement sur l'expérience de conduite chez le conducteur. Par ailleurs, il n'existe pas de vérités scientifiques fondées sur l'efficacité de cette trajectoire de sécurité. Aussi, sa mise en pratique n'est pas chose facile d'autant plus qu'elle dépend de plusieurs paramètres tels que la maîtrise du véhicule, une expérience de conduite ainsi qu'une certaine connaissance des situations de conduite. C'est ainsi que le projet **VIROLO++** (2016–2020)<sup>5</sup>, s'est fixé comme objectif de combler le manque de connaissances sur les pratiques réelles de conduite en virage par le biais de travaux sur les stratégies de prise de virage et le développement d'outils pour une évaluation de cette prise de virage « sécuritaire ».

Le développement d'outils ou de systèmes dans la perspective d'améliorer la sécurité des 2RMs, est une tâche complexe. Cette tâche ne peut se faire sans la phase d'analyse et de compréhension de la conduite (pratiques réelles des conducteurs de 2RMs), car il subsiste de nombreuses inconnues et particularités dans les interactions du triptyque Véhicule-Infrastructure-Conducteur (V-I-C).

## 1.2 Problématiques de recherche et contributions

Afin de mieux appréhender ces difficultés inhérentes à la conduite d'un 2RM en particulier lors de la prise de virage, il est nécessaire d'avoir des observations réelles de comportements de conduite mais également de développer des outils et des méthodes opérationnelles permettant de mesurer objectivement et finement les paramètres de la conduite et les interactions entre le conducteur et son véhicule tenant compte du contexte dans lequel il évolue. Cette problématique a été appréhendée durant le projet VIROLO++ qui a permis de collecter des données relatives aux actions du conducteur nécessaires au contrôle du véhicule dans les différentes phases du virage. Notre objectif opérationnel d'analyse de cette interaction entretenue entre le conducteur d'un 2RM et son véhicule durant les différentes phases de la prise de virage soulève de nombreuses difficultés techniques ainsi que des verrous scientifiques qui structurent ce travail de thèse. Nous visons deux objectifs dans le cadre de cette thèse :

1. Après avoir recueilli les données, la question scientifique que nous abordons est la suivante : les comportements des conducteurs lors de la prise de virage peuvent-ils être réduits à un « profil » de conducteur, ou catégorisés dans des classes ? Nous faisons le postulat que l'utilisateur peut être « réduit » à un problème sensori-moteur et que les différentes stratégies observées devraient par conséquent être identifiées à l'aide de marqueurs comportementaux discriminants.

---

<sup>5</sup><http://www.virolol2plus.universite-paris-saclay.fr/?cat=5>

2. L'analyse de la prise de virage proposée dans l'item ci-dessus concerne le virage dans sa globalité. Pour cette deuxième partie, nous proposons d'effectuer une segmentation du virage, en 4 segments, comme illustrée sur la sur la [figure 1.3](#), afin d'effectuer une analyse plus fine des interactions du motard avec sa moto sur chacune des différentes phases du virage, et plus particulièrement, à l'entrée du virage.

La prise d'un virage chez un conducteur de 2RM, comme illustrée sur la [figure 1.3](#) correspondante à la trajectoire de sécurité, peut être considérée comme un changement de régime (segment). Chaque étape décrite dans la trajectoire de sécurité correspond à un régime. Lors du passage d'un segment à un autre, les pratiques mises en œuvre (vitesse, actions sur la moto avec le guidon ou le corps, emplacement sur la chaussée), lors de la prise de virage, diffèrent d'un conducteur de 2RM à un autre. En ce qui nous concerne, nous disposons des enregistrements vidéo de conduite en virage nous permettant de délimiter ces différents régimes. Cette tâche peut s'avérer trop lourde, voire impossible (dans le cas d'un nombre de séquences trop important). Elle nécessite, par ailleurs, une attention particulière et une très bonne connaissance de la conduite moto. Une alternative à cela consiste en l'utilisation de modèles de segmentation existants dans la littérature qui tiennent compte de ces changements de régimes. Cette thèse a comme objectif principal le développement d'outils amont pour l'analyse des signaux issus des expérimentations réelles conduites dans le projet VIROLO++. Elle s'inscrit également dans la continuité des travaux de la thèse de F. Attal (2011-2014) [2] réalisée au sein de l'IFSTTAR et qui ne traitaient pas de l'analyse de la prise de virage.

### 1.3 Organisation du manuscrit

Ce manuscrit débute par le chapitre 1 qui décrit le contexte applicatif et la problématique de l'étude de la prise de virage d'un 2RM. Dans un premier temps, nous aborderons les enjeux de l'accidentologie des 2RMs ainsi que les particularités et les difficultés liées à la conduite de ce type de véhicule. Ensuite, nous présenterons un état de l'art des travaux de recherche liés à la conduite des 2RMs. Enfin, sur la base de ces éléments, nous décrirons une démarche fondée sur l'exploitation des données de mesures issues d'expérimentations réelles via les capteurs installés sur la moto et sur le conducteur afin d'étudier d'une manière approfondie les différentes modalités d'interactions entre le conducteur et son véhicule lors de la prise de virage.

Le chapitre 2 présente les moyens expérimentaux mis en œuvre dans le cadre de nos travaux afin d'étudier le comportement des conducteurs de 2RMs lors de la prise de virage. Après avoir décrit les différents capteurs utilisés dans le cadre de l'expérimentation et la base de données recueillie, la suite du chapitre introduit les pré-traitements des données qui consistent au remplacement des valeurs manquantes et aberrantes, au filtrage de bruit des données de mesure. Ces pré-traitements spécifiques sont effectués pour augmenter le pouvoir descriptif et discriminant

du phénomène étudié : « l'interaction entretenue entre le conducteur et son véhicule » durant les différentes phases de la prise de virage, et notamment, l'initiation du virage.

Le chapitre 3 concerne la problématique de l'analyse globale de la trajectoire d'un 2RMs lors de la prise de virage, à l'aide d'approches fondées sur l'apprentissage non supervisé. La première partie du chapitre s'attache à présenter la mise en œuvre de la méthodologie développée pour la reconnaissance des comportements de conduite lors de la prise de virage. Dans la deuxième partie de ce chapitre, nous présenterons la mise en œuvre de cette méthodologie sur les données réelles collectées dans le cadre du projet VIROLO++. Une analyse approfondie du comportement de conduite d'un conducteur en particulier dans les deux consignes : conduite avec corps « AC » et conduite sans corps « SC » clôturera ce chapitre.

Une méthodologie permettant d'analyser d'une manière « fine » les actions entretenues par le conducteur et son véhicule lors de la prise de virage fait l'objet du chapitre 4. La première partie est consacrée au problème de segmentation de la trajectoire d'un 2RM en quatre segments décrits dans le chapitre 2. Quant à la seconde partie, elle traite le problème de l'étude de l'initiation « déclenchement » du virage dans le cas d'un 2RM, en focalisant notre analyse sur l'entrée du virage. Dans la dernière partie du chapitre, nous présentons la mise en œuvre de la méthodologie proposée pour identifier les profils, ainsi que les comportements de conduite dans la phase « entrée du virage ». Les données utilisées dans ce chapitre sont les données collectées durant la consigne libre (CL), car ces données mettent en évidence la variabilité intra et inter individuelle des conducteurs.

Une conclusion de ce mémoire rappellera les différentes étapes du travail effectué ainsi que les résultats obtenus et les perspectives de recherche découlant de ces travaux de thèse.

Les travaux réalisés dans cette thèse ont été valorisés dans dans deux revues de rang A, listées ci-dessous :

1. **Diop, M.** ; Boubezoul, A. ; Oukhellou, L. ; Espié, S. Powered Two-Wheeler Riding Profile Clustering for an In-Depth Study of Bend-Taking Practices. Sensors 2020, 20, 6696. <https://doi.org/10.3390/s20226696>
2. **Diop, M.** ; Boubezoul, A. ; Oukhellou, L. ; Espié, S. ; Bouaziz, S Powered Two-Wheelers Bend-taking Maneuver Study using Segmentation and Data Mining Approaches. IEEE Transactions on Intelligent Transportation Systems 2022 (DOI : 10.1109/TITS.2022.3222421)



---

**Protocole expérimental et acquisition de données**

---

**Sommaire**

---

<b>2.1</b>	<b>Introduction</b>	<b>12</b>
<b>2.2</b>	<b>Protocole expérimental</b>	<b>13</b>
2.2.1	Description des consignes de conduite	13
2.2.2	Description de l'instrumentation et des capteurs utilisés	15
<b>2.3</b>	<b>Acquisition de données</b>	<b>19</b>
2.3.1	Pré-traitement des données collectées	19
2.3.2	Représentation sous formes de base de données des mesures fournies par les différents capteurs	21
2.3.3	Aperçu de quelques signaux collectés lors de l'expérimentation	21
2.3.4	Annotation manuelle des trajectoires	23
<b>2.4</b>	<b>Conclusion</b>	<b>23</b>

---

## 2.1 Introduction

Pour mieux appréhender l'accidentologie des 2RMs, il est nécessaire de comprendre le comportement des acteurs routiers notamment celui des conducteurs de 2RMs. En effet, dans les études [32, 50] et [55], les auteurs ont montré que les comportements de conduite constituent le principal facteur des accidents routiers.

Pour recueillir des données relatives aux comportements de conduite, il existe 4 différentes méthodes dans la littérature [60]. La première, la plus utilisée, concerne l'acquisition de données à travers les enquêtes et questionnaires. Dans ces enquêtes et questionnaires, il est souvent question de recueillir des informations concernant l'âge du conducteur, l'année d'obtention du permis, le nombre de kilomètres parcouru entre autres. Comme exemple de travaux utilisant des données relatives au comportement de conduite en 2RM, nous pouvons citer [18, 44] et [14]. Cependant, la validité de ces études fondées sur des données issues d'enquête ou de questionnaires a souvent été critiquée dans la littérature. Cette méthode repose sur l'idée d'évaluer qualitativement les facteurs ayant un impact sur le risque d'accident ou en identifiant les causes d'accidents d'après les réponses aux questionnaires et enquêtes. L'inconvénient majeur de cette méthode est que cette dernière est sujette à un biais « social », car la personne sondée peut contrôler sa réponse, en orienter la tonalité, afin de la rendre socialement acceptable [1, 5] et [66].

La deuxième méthode consiste à faire une analyse approfondie des processus de dysfonctionnement du système routier (Véhicule-Infrastructure-Conducteur) et à déterminer les causes et les conséquences des accidents, à partir de recueils de données sur la scène des accidents. Néanmoins, cette méthode ne permet pas de faire une analyse approfondie du comportement du conducteur. Cette dernière ne permet pas d'extraire des informations précises et détaillées sur la manière dont les conducteurs se comportent sur la route et surtout avant, pendant et après les situations critiques ou d'accident. En outre, cette méthode repose sur le recueil de données après la survenue de l'accident ce qui peut constituer une vision partielle et biaisée de l'interaction (Véhicule-Infrastructure-Conducteur) [64].

La troisième méthode passe par une utilisation d'un simulateur de conduite. Les simulateurs de conduite permettent d'étudier le comportement humain dans différentes conditions déterminées expérimentalement « un environnement contrôlé ». L'avantage le plus important à l'utilisation des simulateurs de conduite réside dans le fait que ce dernier n'expose pas le sujet à un risque objectif et dans les conditions de test qui sont « facilement » mises en œuvre et contrôlables. Ceci justifie l'utilisation de simulateurs de conduite, aussi bien dans le cas des véhicules à 4 roues que 2 roues, dans la plupart des études concernant le comportement des conducteurs [13, 17, 34, 47, 49]. Néanmoins, l'utilisation des simulateurs de conduite pose la question de la généralisation et de la transposition des comportements observés en simulation à la situation réelle de conduite. il est en effet impératif de s'assurer que le simulateur mesure l'objet d'étude

sans le modifier ou l'influencer. Dans les travaux [56,67] et [74], il a été constaté des différences significatives selon que l'on conduit sur simulateur ou en situation réelle.

Une quatrième méthode pour étudier le comportement de conduite est la conduite en situation naturelle ou réelle dite « *Naturalistic Driving* » en anglais. Les auteurs de [65] ont défini l'étude de comportements réels de conduite comme étant : « Une étude entreprise pour fournir un aperçu du comportement du conducteur lors de déplacements quotidiens en enregistrant des détails sur le conducteur, le véhicule et l'environnement au moyen d'un équipement de collecte de données non-intrusive et sans contrôle expérimental ». On distingue étude en situation naturelle « classique » et approfondie. Dans cette dernière, outre les données de conduite, on collecte par le biais d'entretiens d'auto confrontation, des éléments sur les situations vécues par les conducteurs, les éléments du contexte qu'ils ont pris en compte pour prendre leurs décisions, voire les motifs sous-jacents [3].

Il est à rappeler que l'un des objectifs du projet **VIROLO++** est d'approfondir la question de la prise de virage. Il s'agit de combler un manque de connaissance sur les pratiques de conduite de virage pour une population de conducteurs professionnels et non professionnels, et en particulier de mieux comprendre ces pratiques de conducteurs lors des phases d'entrée, de conduite et de sortie de virage. Afin d'atteindre cet objectif, un protocole expérimental a été défini dans le cadre du projet **VIROLO++** pour collecter des données relatives aux actions du conducteur nécessaires au contrôle du véhicule dans les différentes phases du virage. Une instrumentation du véhicule à des endroits stratégiques notamment : selle, poignées, réservoir, repose-pieds, angle guidon a été aussi proposée.

Nous allons dans les sections suivantes exposer le déroulement de cette expérimentation ainsi que l'acquisition de données.

## 2.2 Protocole expérimental

Le protocole expérimental s'est déroulé sur la piste de La Ferté-Gaucher en Seine-et-Marne dans la banlieue parisienne en France. L'expérimentation s'est déroulée sur une piste d'une longueur de 2km caractérisée par quatre lignes droites et des virages diverses, voir [figure 2.1](#).

### 2.2.1 Description des consignes de conduite

Un des objectifs poursuivis dans l'expérimentation était d'acquérir des connaissances sur la façon dont le conducteur interagit avec la moto pour assurer le contrôle de la trajectoire dans les virages. Les données pour remplir ces objectifs ont été obtenues lors d'expérimentations sur la piste de « La Ferté-Gaucher » présentée ci-dessus. Nous avons cherché à déterminer en particulier s'il est possible d'opérer une distinction entre motocyclistes favorisant l'inclinaison du corps et



**Figure 2.1** : Circuit de la piste « La Ferté-Gaucher » le rectangle rouge matérialise le virage étudié.

ceux favorisant la manœuvre dite du contrebraquage pour contrôler leurs manœuvres de prise de virage.

Pour ce faire, dans cette expérimentation, des professionnels de la gendarmerie nationale française ont eu à effectuer trois tours de pistes sous trois consignes différentes avec une moto instrumentée avec différents capteurs. Un questionnaire, réalisé en amont, a permis de recueillir diverses informations sur ces participants, cf. [tableau 2.1](#). Parmi celles-ci nous pouvons citer l'âge du permis de moto (Permis), le nombre de kilomètres parcourus en moto l'année précédente l'expérimentation (Km). Il a été également demandé aux participants d'indiquer quelle condition parmi les deux consignes de conduite contrainte (avec le guidon / avec le corps) celle qui avait semblé la plus naturelle ou la moins difficile à suivre (Préférence). Et enfin, l'ordre de passage des instructions (ordre des instructions).

Participants	Permis	Km	Préférence	Ordre des instructions
P1	2	8000	guidon	CL, SC, AC
P2	7	3000	guidon	CL, AC, SC
P3	0	0	corps	CL, SC, AC
P4	11	0	corps	CL, SC, AC
P5	2	20,000	guidon	CL, AC, SC
P6	16	0	guidon	CL, SC, AC
P7	18	2000	guidon	CL, AC, SC
P8	12	3000	guidon	CL, SC, AC

**Tableau 2.1** : Données déclaratives des différents participants et ordre des instructions.

Les différentes consignes de conduite sont décrites comme suit :

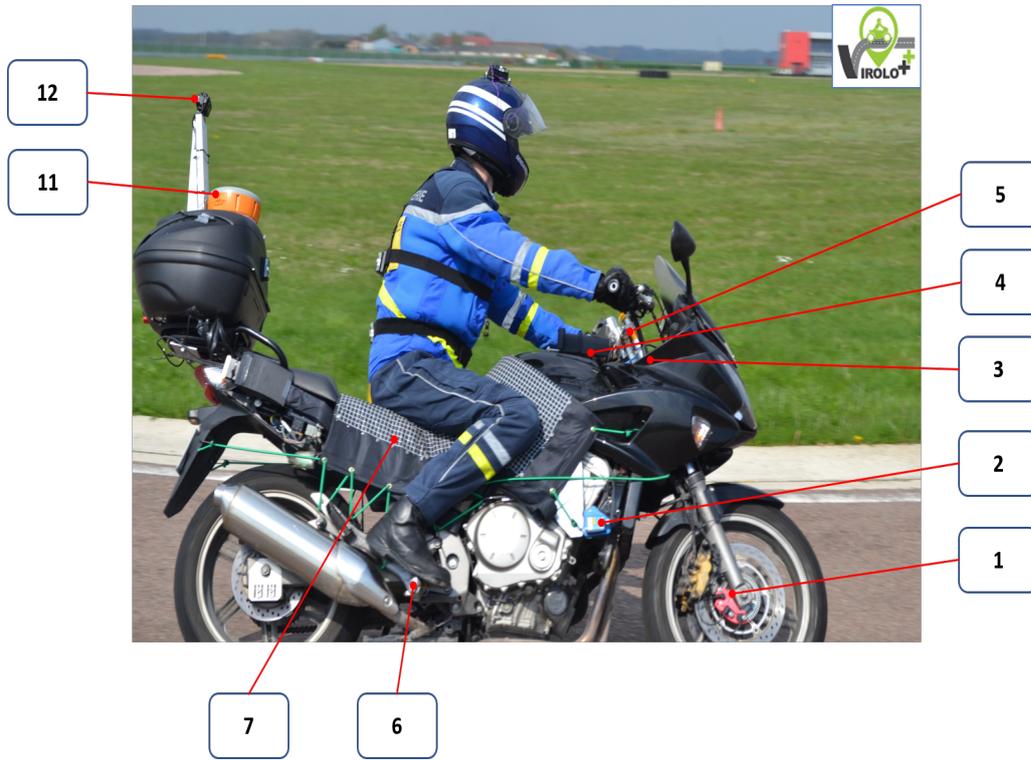
- **Conduite libre (CL)** : Dans cette consigne de conduite, les participants sont appelés à conduire de façon habituelle, et notamment à diriger la moto comme ils ont l'habitude de le faire ;
- **Conduite avec corps (AC)** : Ici, il est demandé aux participants d'utiliser le plus possible les actions de leur corps plutôt que celles sur le guidon pour incliner la moto afin d'aborder le virage.
- **Conduite sans corps (SC)** : Les actions demandées dans cette consigne sont tout simplement le contraire de la consigne précédente. En effet, ici le participant doit mettre l'accent sur le guidon plutôt que d'utiliser son corps pour diriger la moto lors de la prise du virage.

Pour éviter un éventuel biais d'adaptation liée à la moto ou aux équipements portés par les conducteurs, trois tours de conduite ont été réalisés dans chacune des consignes de conduite. Néanmoins, les consignes AC et SC peuvent éventuellement présenter un "biais de consigne" contrairement à celle de la conduite libre. En effet, ce biais est lié la compréhension du participant de la consigne qui lui a été donnée et sa mise en oeuvre.

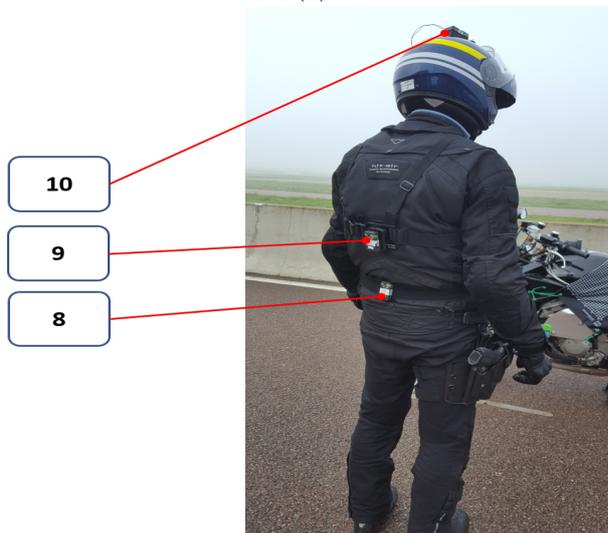
### 2.2.2 Description de l'instrumentation et des capteurs utilisés

Les données collectées dans le projet **VIROLO++** ont été obtenues grâce à une instrumentation sur différents niveaux. Une instrumentation réalisée au niveau de la moto a permis de collecter des mesures sur la dynamique de la moto ainsi que des informations sur le contexte de la conduite. Par ailleurs, une instrumentation a été réalisée sur le conducteur permettant ainsi de recueillir des mesures sur ses actions et mouvements.

La figure 2.2 montre l'emplacement de l'ensemble des capteurs utilisés matérialisés par des numéros allant de 1 à 12.



(a)



(b)

Figure 2.2 : Instrumentation réalisée au niveau de la moto (a) et du motard (b).

## 2.2. Protocole expérimental

Le [tableau 2.2](#) donne des informations sur les capteurs en indiquant la nature physique du capteur, la grandeur qu'il mesure et la fréquence d'acquisition de ces dernières.

<b>Mesures relatives à la dynamique de la moto</b>			
<i>N°</i>	<b>Capteur</b>	<b>Mesure</b>	<b>Fréquence</b>
1	Capteurs à effet Hall et disques ABS	Rotation de la roue $\mapsto$ Vitesse de la moto	1KHz
2	Capteurs Laser SICK DT35 <sup>1</sup>	Distance au sol $\mapsto$ Angle de roulis	100Hz
3	Capteur magnétique AS5047P <sup>2</sup>	Aimant permanent et capteur magnétique $\mapsto$ Angle du guidon	1KHz
4	Centrale Inertielle (MTi-Xsens) <sup>3</sup>	Accélérations et vitesses de rotation selon les axes (X-Y-Z)	100Hz
<b>Mesures relatives aux actions du conducteur</b>			
5	Jauges de contrainte <sup>4</sup>	Forces appliquées sur le guidon	10 Hz
6	Capteur de force Mesurex D2 piézo-électrique	Forces appliquées sur les repose-pieds	100 Hz
7	Tapis de pression : XSENS-SOR <sup>5</sup> LX100 et PX100	Pression gauche et droite de la fesse et des genoux	10 Hz
<b>Mesures relatives aux mouvements corporels du conducteur</b>			
8	CAPTIV Motion de TEA <sup>6</sup>	Angle de roulis du bas du corps.	128 Hz
9	CAPTIV Motion de TEA	Angle de roulis du haut du corps.	128 Hz
10	CAPTIV Motion de TEA	Angle de roulis de la tête.	128 Hz
10b	CAPTIV Motion de TEA	Angle de roulis de la moto.	128 Hz
<b>Informations relatives au contexte</b>			
11	GPS-RTK Altus APS3G <sup>7</sup>	Positions centimétrique en Latitude et Longitude.	25Hz
12	Caméra	Videos du contexte routier	---

**Tableau 2.2** : L'ensemble des capteurs utilisés ainsi que le nom du paramètre mesuré et la fréquence d'acquisition.

<sup>1</sup><https://www.sick.com/fr/fr/capteurs-de-distance/capteurs-de-distance-a-moyenne-portee/dx35/dt35-b15251/p/p295353>

<sup>2</sup><https://ams.com/AS5047P>

<sup>3</sup><https://www.xsens.com/products/mti-100-series>

<sup>4</sup><https://www.mesurex.fr/en/>

<sup>5</sup><https://www.teaergo.com/products/tea-captiv-t-sens-motion-imu/>

<sup>6</sup><https://www.septentrio.com/en/products/gnss-receivers/rover-base-receivers/smart-antennas/aps3g>

Chacun des capteurs utilisés dans cette expérimentation est décrit comme suit :

- **Capteurs N°1** : deux capteurs à effet hall placés en quadrature pour la mesure de l'angle de rotation de la roue arrière du véhicule (même dispositif que pour la roue avant). Ce capteur permet d'avoir des mesures de la vitesse longitudinale de la moto ;
- **Capteurs N°2** : de type laser, ils sont placés des deux côtés (droit et gauche) de la moto pour mesurer l'angle d'inclinaison de la moto "angle de roulis", par rapport, à la route dérivée par triangulation de la mesure par les 2 lasers de la distance au sol ;
- **Capteur N°3** : ce capteur magnétique permet de mesurer l'angle absolu de rotation de la colonne de direction de la moto, dit angle guidon ;
- **Capteur N°4** : une centrale inertielle (MTi-Xsens) pour l'acquisition des mouvements inertiels selon les axes (X-Y-Z) : accélérations (longitudinale, latérale et verticale) ; vitesses et angles de rotation (tangage, lacet et roulis).
- **Capteurs N°5** : Il s'agit de jauges de contraintes placées sur les demi-guidons (droite et gauche) de la moto pour mesurer les forces appliquées par le conducteur sur chaque demi-guidon selon les axes « X/Y », dont les valeurs sont récupérées et enregistrées par un système industriel propriétaire ;
- **Capteurs N°6** : constitués de jauges de contrainte placées sur les repose-pieds (droit et gauche) de la moto pour mesurer les forces appliquées par le conducteur sur chaque repose-pied « X/Y », dont les valeurs sont récupérées et enregistrées par un système industriel propriétaire ;
- **Capteur N°7** : tapis de pression pour l'acquisition des forces de pression des fesses et des genoux du conducteur sur la selle et le réservoir respectivement, dont les valeurs sont récupérées et enregistrées par un système industriel propriétaire ;
- **Capteurs N°8 et N°9, et N°10** : 4 centrales inertielles pour mesurer les orientations du tronc et de la tête.
- **Capteur N°11** : un (GPS-RTK) qui a pour but d'acquérir un positionnement cinématique précis en temps réel de la moto.
- **Capteur N°12** : caméra d'action placée sur le boîtier supérieur, regardant vers l'avant (champ de vision incluant le dos du conducteur).

### 2.3 Acquisition de données

La mesure fine de la dynamique du véhicule et des interactions entretenues entre le conducteur et le 2RM a nécessité une instrumentation dite « lourde ». Cette instrumentation s'appuie sur une architecture utilisant un bus CAN 1Mbps (Controller Area Network) (bus de terrain), des noeuds de mesure autonomes (acquisition et mise en forme de données), ainsi que d'un noeud en charge de l'enregistrement des données mises sur bus CAN, et de la synchronisation des systèmes propriétaires additionnels (i.e. systèmes industriels pour lesquels les données sont enregistrées de manière locale). Les systèmes développés et intégrés sur la moto sont compatibles avec les contraintes d'un usage sur 2RMs à savoir, compacité, poids, robustesse aux vibrations et consommation électrique limitée. Même si la partie instrumentation a été traitée en dehors de ces travaux de thèse, [61], il nous a semblé intéressant d'en présenter les principales orientations avant d'aborder la partie pré-traitements de données à proprement parler.

#### 2.3.1 Pré-traitement des données collectées

Les données utilisées dans l'étude des interactions entre le conducteur et son véhicule proviennent d'expérimentations réelles. Ces données de mesures hétérogènes (provenant de systèmes de mesures propriétaires différents) sont par nature incomplètes (présence de valeurs manquantes), et entachées de bruit issu de l'environnement d'expérimentation (bruit de hautes fréquences qui est causé principalement par les vibrations du moteur, par exemple).

Comme cela peut être le cas avec d'autres méthodes d'analyse de données, les méthodes d'analyse fondées sur l'apprentissage automatique sont particulièrement sensibles à la qualité des données. Le contrôle de cette qualité repose notamment sur des techniques permettant de réduire le bruit dans ces données. Le pré-traitement de ces données s'avère alors une étape capitale en amont de l'étape d'analyse « fine » de ces interactions.

La fusion de ces données hétérogènes, le remplacement des valeurs manquantes ainsi que le filtrage du bruit sont les différentes étapes constituant cette phase de pré-traitement.

1. **Recueil et fusion des données brutes** Les données ont été recueillies par des systèmes d'acquisition propriétaires différents qu'il a fallu par la suite fusionner. Or, la fusion de données est toujours une opération délicate qui exige de nombreuses vérifications pour être menée à bien. Il est donc indispensable, avant de fusionner des données, de déterminer si les données décrivent bien les situations observées, en visualisant pour chaque système les canaux d'acquisition jugés pertinents. Cette première étape de vérification « visuelle » de la cohérence des données acquises a été faite avec un logiciel de traitement de données développé au sein de l'Université Gustave Eiffel (Ex. IFSTTAR), voir [figure 2.3](#). Par la suite, nous avons utilisé ce logiciel pour synchroniser à la même fréquence tous les signaux collectés. En effet, un choix d'avoir des données avec une fréquence d'acquisition égale à 10Hz a été fait. Pour ce faire, un sous-échantillonnage et un sur-échantillonnage ont été respectivement effectués, depuis ce logiciel, pour les capteurs avec une fréquence supérieure

et inférieure à 10Hz. Par ailleurs, grâce à ce logiciel, il est possible de mettre en forme les différentes données acquises (transformation des données du format binaire au format texte). Par la suite, les données ont été exportées depuis ce logiciel avec la fréquence choisie sous la forme d'un fichier avec l'extension CSV. Ainsi, ces données pourront être importées dans un logiciel de traitement d'analyse de données notamment Python.

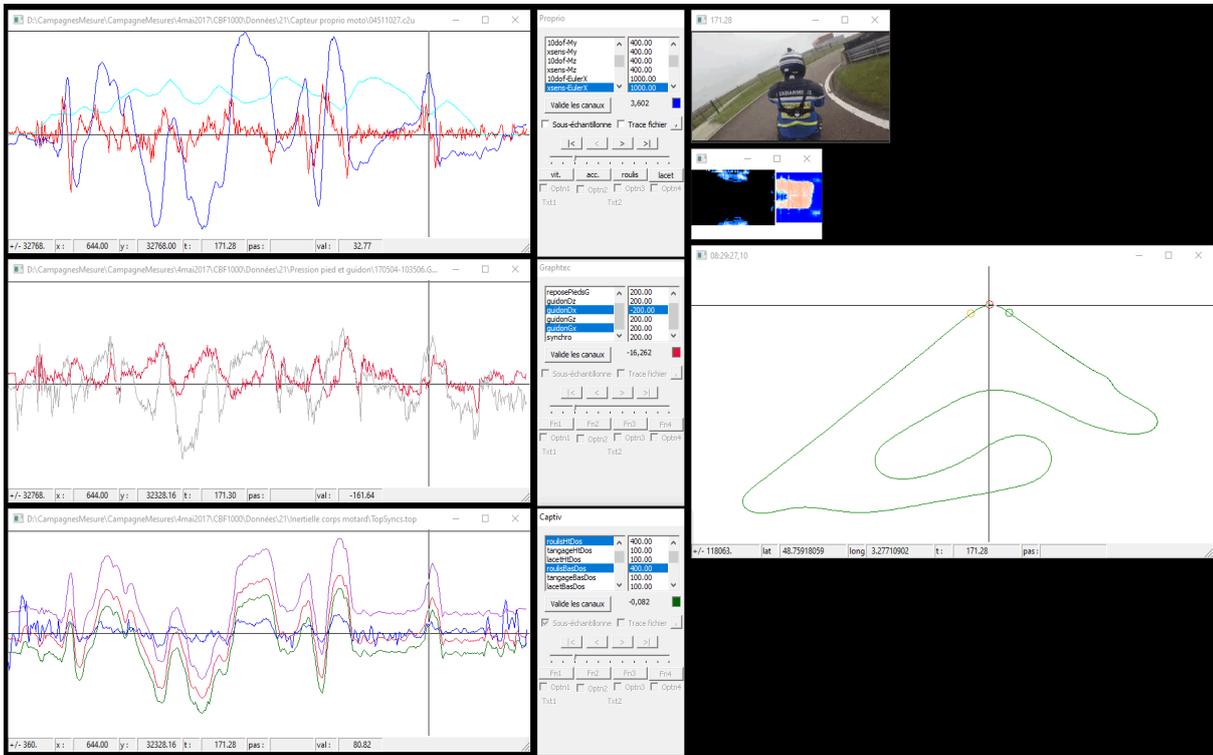


Figure 2.3 : Logiciel de visualisation / export multi-sources et multi-canaux.

## 2. Remplacement des valeurs manquantes et aberrantes

Nous avons choisi de remplacer les valeurs manquantes par la moyenne empirique calculée sur un intervalle de 2s.

## 3. Filtrage des données

En raison de la complexité et de la nature bruitée des données collectées, une étape de filtrage nous est apparue nécessaire. Les variations des signaux sont très lentes, dues à la nature mécanique de la moto. Pour cela, nous nous n'intéressons pas aux hautes fréquences, qui représentent le bruit qui pourrait causer des problèmes à nos expériences ultérieures. Nous proposons un filtrage de données utilisant un filtre médian de taille 5 pour éliminer ce bruit.

### 2.3.2 Représentation sous formes de base de données des mesures fournies par les différents capteurs

Pour faciliter la manipulation des données issues de cette expérimentation, une représentation sous forme de base de données a été mise en place. Ainsi, cette base de données est notée de la manière suivante :

$$\mathcal{D} = \langle \mathcal{D}^l \rangle_{l \in \{CL, SC, AC\}},$$

avec  $\mathcal{D}^l = (\mathbf{x}_k)_{\{k=1, \dots, N\}}$ , où  $N$  est le nombre total de séquences pour l'instruction  $l \in \{CL, SC, AC\}$ . Chaque séquence  $\mathbf{x}_k = (\mathbf{x}_{k,m})_{m=1, \dots, M}$  est un ensemble de valeurs observées sur l'intervalle de temps  $(t_m)_{m=1, \dots, M}$ ,  $M$  (un entier entre 82 et 124) est le nombre total d'observations et  $\mathbf{x}_{k,m}$  est donné par :

$$\mathbf{x}_{k,m} = \{C_{k,m}^1, C_{k,m}^2, \dots, C_{k,m}^{11}\},$$

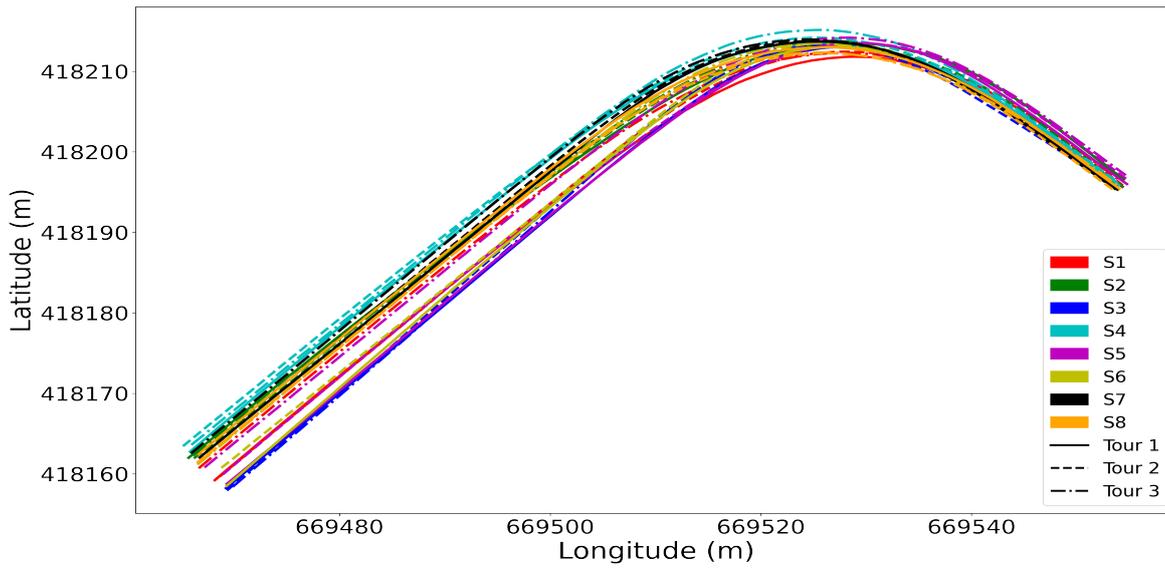
où  $C_{k,m}^r$  est la valeur mesurée à  $t_m$  lors de la séquence  $k$  pour le capteur numéro  $r$  avec  $r = 1, 2, \dots, 11$ .

### 2.3.3 Aperçu de quelques signaux collectés lors de l'expérimentation

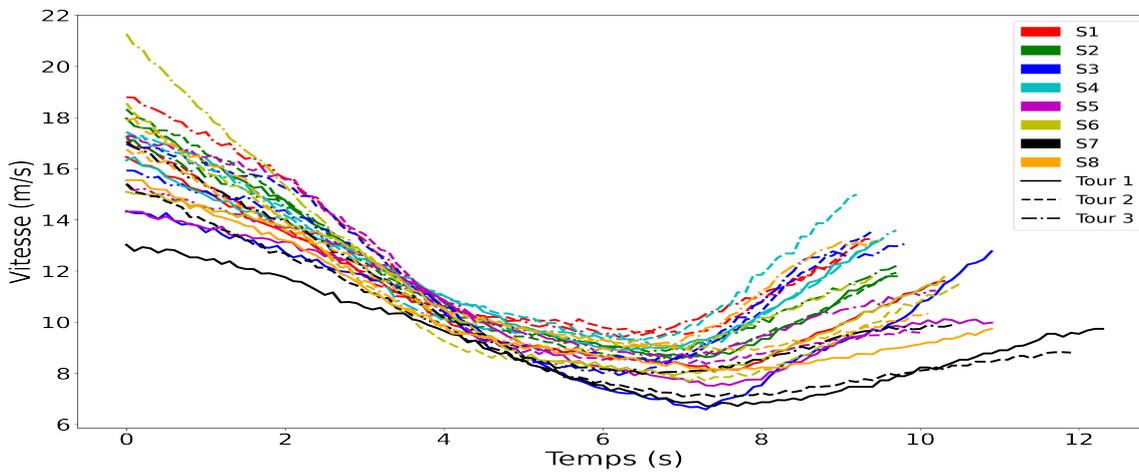
La [figure 2.4](#), représente les variations de trois capteurs constituant la base de donnée décrite ci-dessus, à savoir : angle guidon, vitesse longitudinale et la trace (latitude et longitude) de chaque sujet pour l'instruction Conduite Libre (CL).

Nous avons opté pour la visualisation des capteurs : vitesse et guidon afin d'illustrer la variabilité qui existe entre les conducteurs et leurs véhicules en terme de contrôle du véhicule et de dynamique lors de la prise de virage. Le guidon et la vitesse traduisent bien les actions entreprises par le conducteur pour agir sur les organes de contrôle et de commande de la moto. Enfin, un dernier élément pour illustrer cette variabilité dans les comportements de conduite pendant la prise de virage est le positionnement sur la chaussée. Cet élément est donnée par la trace « trajectoire » du véhicule.

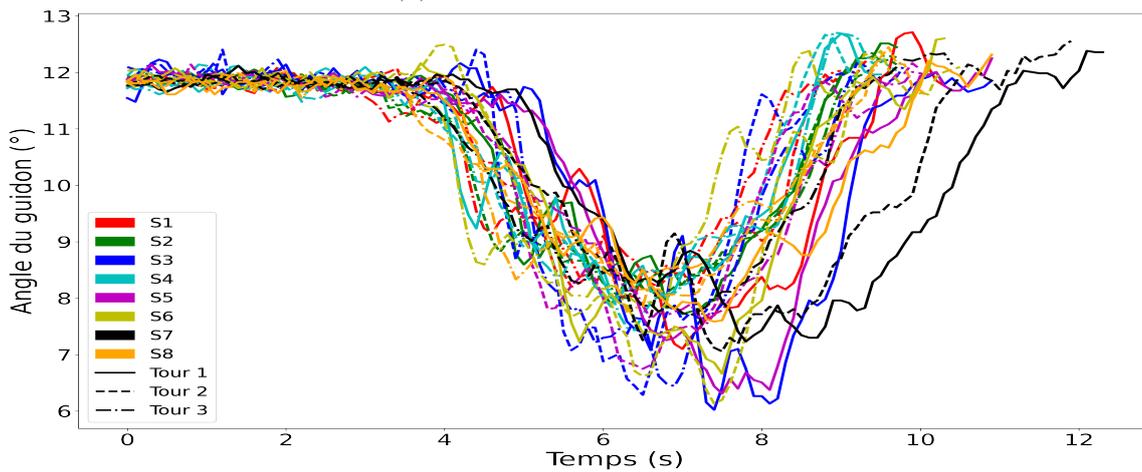
A première vue, on note sur la [figure 2.4](#) une grande variabilité intra-individuelle (d'un tour à un autre) et inter-individuelle des conducteurs, notamment en terme de vitesse adoptée à l'approche du virage, de positionnement sur la chaussée et de contrôle du véhicule. Cette variabilité intra et inter individuelle des conducteurs rend le problème auquel nous nous intéressons dans ce travail de thèse particulièrement complexe.



(a) Évolution de la position du véhicule (Latitude, Longitude).



(b) Évolution temporelle de la vitesse.



(c) Évolution temporelle de l'angle guidon.

**Figure 2.4** : Évolution : de la position du véhicule (a) temporelle de la vitesse (b) et de l'angle guidon (c) tout au long du virage pour tous les sujets pendant les trois tours la consigne CL.

### 2.3.4 Annotation manuelle des trajectoires

Dans le but d'étudier de façon plus fine les pratiques de conduite lors des phases d'entrée, de conduite et de sortie de virage, une segmentation des trajectoires a été effectuée manuellement grâce au logiciel propriétaire décrit précédemment. Ce logiciel permet de visualiser et de synchroniser les données des capteurs. La détermination des différentes plages temporelles de chaque segment (temps de début et fin) a été effectuée en visualisant au ralenti les données vidéos acquises grâce à la caméra de contexte.

Chaque trajectoire a été segmentée en 2 arcs (segments 2 et 3) et deux (2) lignes droites qui correspondent à l'entrée/sortie du virage (segments 1 et 4), voir [figure 2.5](#). La description des différents segments est donnée comme suit :

- **Segments (S1) et (S4)** qui correspondent à une conduite en ligne droite à l'entrée et la sortie du virage, respectivement. Ils sont caractérisés par un angle de roulis proche de zéro, ainsi que par de faibles variations à hautes fréquences de la vitesse de roulis autour de zéro, cette dynamique correspond au maintien de la moto en position verticale [42]. En outre, nous pouvons constater des fluctuations de l'angle guidon autour d'une valeur constante pendant ces phases. La principale différence entre ces 2 segments se situe au niveau de l'accélération longitudinale moyenne. Dans le segment 1, l'accélération longitudinale moyenne est négative en raison de la décélération du véhicule à l'entrée du virage, alors qu'elle est positive dans le segment 4 en raison de l'accélération du véhicule en sortie de virage ;
- **Segments (S2) et (S3)** sont principalement caractérisés par un écart important par rapport à zéro des valeurs d'angle de roulis ainsi que par des valeurs de vitesse de roulis avec des variations faibles mais de haute fréquence autour de zéro [42].

Dans le cadre de cette thèse, nous travaillons à développer une méthodologie permettant de proposer une segmentation automatique de séries temporelles multivariées, cf. chapitre 4.

## 2.4 Conclusion

Dans ce chapitre, nous avons décrit l'outil de recueil de données installé sur une moto. Nous avons décrit également le protocole expérimental ainsi que les séquences de conduite réalisées qui constituent notre base de données. Pour la mise en forme de la base de données, une étape de pré-traitement a été effectuée. Cette étape a permis le traitement de valeurs manquantes, le filtrage des données à l'aide d'un filtre médian ainsi que la segmentation et l'annotation manuelles des trajectoires.

Ce travail préliminaire a permis de poser les briques de base qui nécessitent les étapes plus avancées de traitement que nous décrirons dans les prochains chapitres.

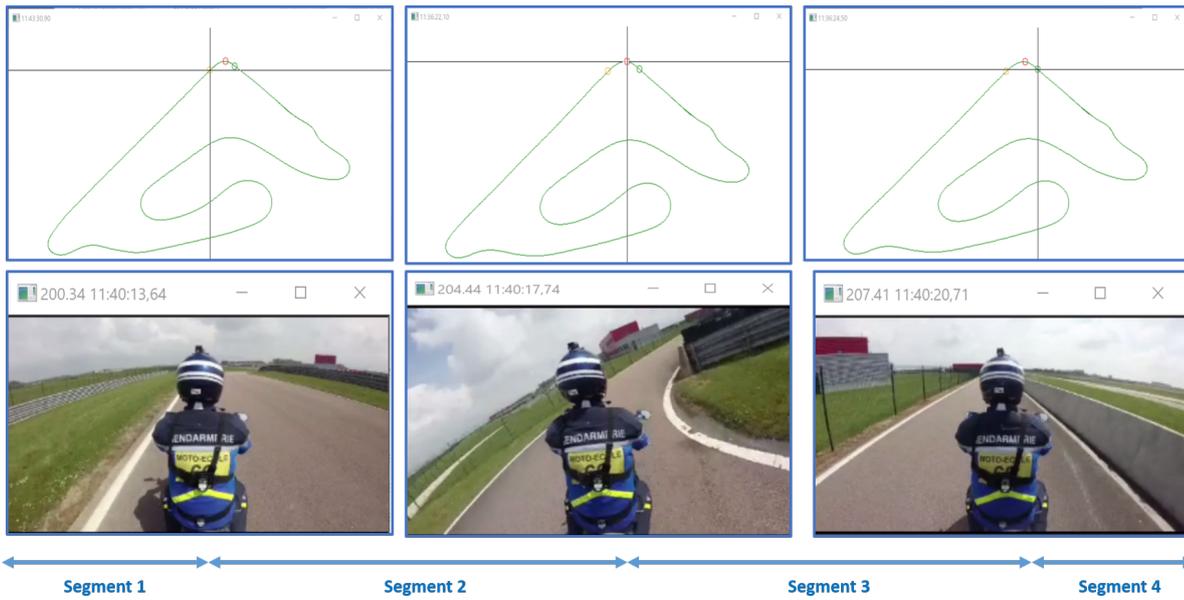


Figure 2.5 : Un exemple d'annotation manuelle d'une trajectoire en combinant les données vidéos, GPS et données capteurs.

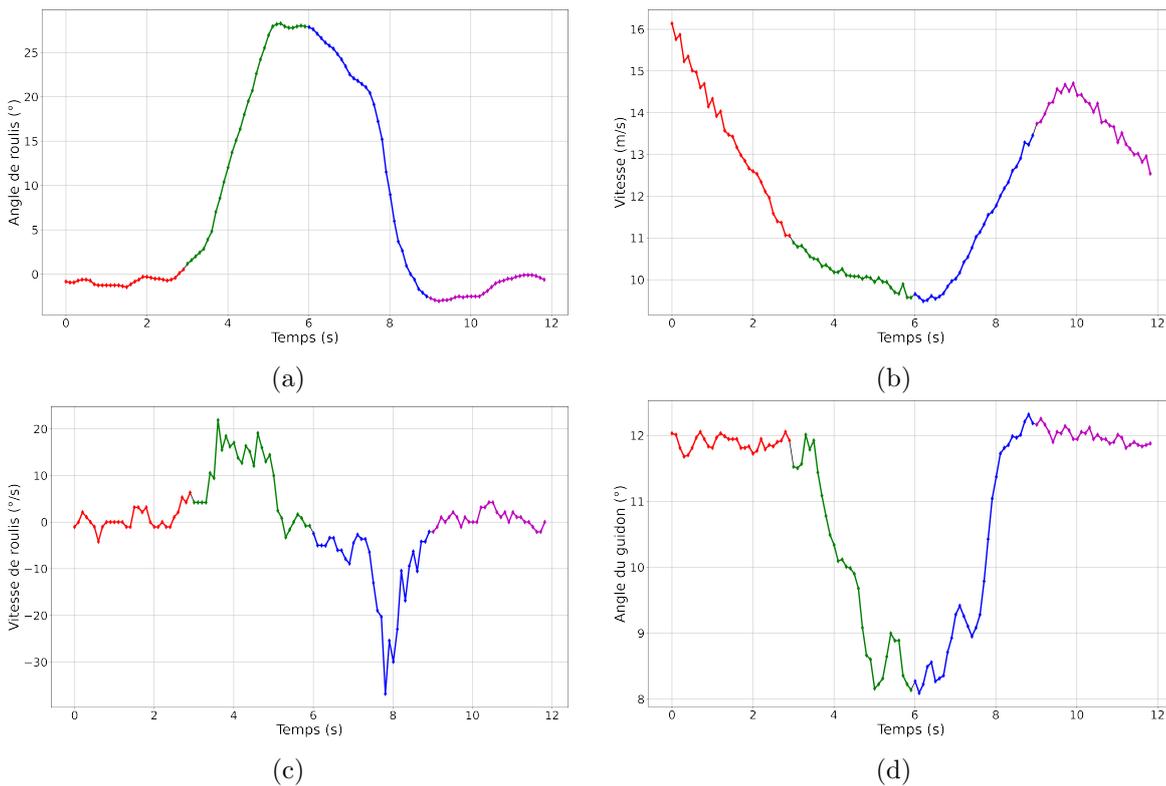


Figure 2.6 : Représentation de l'angle et la vitesse de roulis ainsi que l'angle de guidon, et la vitesse longitudinale du véhicule pour une trajectoire segmentée en 4 segments : Segment1 (S1) : rouge, Segment2 (S2) : vert, Segment3 (S3) : bleu et Segment4 (S4) : violet.

---

**Étude globale de la trajectoire lors de la prise de virage**

---

**Sommaire**

---

<b>3.1 Introduction</b> . . . . .	<b>26</b>
<b>3.2 État de l’art</b> . . . . .	<b>26</b>
3.2.1 Modélisation du pilotage de 2RMs . . . . .	27
3.2.2 Étude du comportement de conduite dans le cas d’un 2RM . . . . .	28
<b>3.3 Méthodologie de classification des différents profils de conduite</b> . . . . .	<b>29</b>
3.3.1 Description et représentation des données utilisées . . . . .	31
3.3.2 Méthodologie de classification des participants . . . . .	32
<b>3.4 Application de la méthodologie proposée : Résultats et Discussions</b> . . . . .	<b>37</b>
3.4.1 Analyses des classes fournies par les dendrogrammes . . . . .	39
3.4.2 Interprétation des classes fournies par les dendrogrammes . . . . .	39
3.4.3 Analyse approfondie des comportements de conduite du participant P4 . . . . .	42
<b>3.5 Conclusion</b> . . . . .	<b>49</b>

---

### 3.1 Introduction

Dans ce chapitre, nous nous intéressons à l'analyse globale de la trajectoire d'un 2RMs lors de la prise de virage, à l'aide d'approches fondées sur l'apprentissage non supervisé. Pour ce faire, nous mobiliserons une partie des données présentées dans le chapitre 2 ainsi que les données déclaratives. La première partie de ce chapitre est consacrée à l'état de l'art relatif au comportement des conducteurs de 2RMs lors de la prise de virage. Par la suite, nous présentons la méthodologie mise en place pour la classification des différents comportements de conduite observés dans une expérimentation réalisée dans le cadre du projet VIROLO++. Les résultats obtenus à l'issue de la mise en œuvre de cette méthodologie sont analysés puis interprétés. Une analyse approfondie du comportement de conduite d'un conducteur en particulier, dans les deux consignes : conduite avec corps « AC » et conduite sans corps « SC » clôturera ce chapitre.

### 3.2 État de l'art

Les 2RM constituent une solution de déplacement de plus en plus convoitée, notamment pour leur capacité à éviter les congestions du trafic. Mais les 2RM restent encore à ce jour, malgré une baisse générale de l'accidentalité, un mode de transport particulièrement risqué qui classe ses utilisateurs dans la catégorie des usagers très vulnérables. Les circonstances d'occurrence des accidents impliquant des 2RM sont multiples. On peut citer le cas des intersections avec des problèmes d'interactions avec les autres usagers (notamment refus de priorité par les automobilistes) mais aussi la perte de contrôle du véhicule liée à une vitesse inadaptée (accident sans tiers). Les pertes de contrôle en virages (ronds-points et changement de direction en intersection inclus) sont courantes (plus de 50% des accidents véhicule seuls), la situation étant particulièrement complexe à gérer du fait de la dynamique complexe du véhicule et de son instabilité intrinsèque. La question de la prise de virage est un enjeu à la fois du point de vue des connaissances scientifiques et pour la sécurité routière. Une meilleure connaissance permettrait de proposer des mesures de formation (objet du projet VIROLO++), et à terme d'identifier des pistes pour la conception et/ou la validation de systèmes d'aides à la conduite dédiés aux 2RMs.

Les travaux de recherche sur les comportements des conducteurs de 2RMs, d'une façon générale, et en particulier sur la question de la prise de virage, sont relativement récents et limités, comparés à ceux sur les conducteurs d'automobiles, et les pratiques réelles de conduite largement méconnues. Cette recherche sur les comportements des conducteurs de 2RM a été abordée initialement par le prisme de la modélisation du pilotage de 2RMs.

Dans le domaine de modélisation du pilote, le principal défi scientifique à relever réside dans la conception d'un modèle de contrôle de conduite approchant le plus possible de celui réalisé par les conducteurs de 2RMs. Pour atteindre cet objectif, il est essentiel de représenter les processus sensorimoteurs utilisés par un conducteur de 2RMs pour contrôler la trajectoire

de son véhicule. Il s'agit donc, de représenter les différentes modalités d'actions qu'entreprend le conducteur pour contrôler la trajectoire de son véhicule et assurer la stabilisation de ce dernier.

Dans ce qui suit, nous avons fait le choix de présenter que les travaux de recherche qui ont trait à cette question de modélisation des pilotes pour contrôler la trajectoire de la moto.

### 3.2.1 Modélisation du pilotage de 2RMs

Des travaux ont été proposés dans les années 70 par exemple par Weir [69] et Rice [54], avec comme une principale préoccupation de déterminer dans quelles proportions le couple de braquage et l'inclinaison du buste du pilote intervenaient dans le pilotage de la moto, en termes de suivi de trajectoire et régulation de l'angle d'inclinaison du véhicule.

Cette thématique de la modélisation et de la compréhension dynamique des pilotes de 2RM connaît un regain d'intérêt depuis le début des années 2000, comme en témoigne la synthèse bibliographique de Popov et al. [51]. Une motivation reste toujours la compréhension fondamentale de la stratégie de pilotage. Quel moyen d'action est-il prédominant, parmi le couple de braquage, l'inclinaison du haut du corps, ou l'action du bas du corps à travers les repose-pieds, par exemple?. Comme le souligne Popov et al. [51], et comme cela se retrouve dans la littérature sur la modélisation de conducteur, deux approches cohabitent pour mener à bien cette modélisation des pilotes moto. Une première cherche à établir des modèles type « boîte noire », en exploitant les outils de la commande optimale. S'appuyant sur le fait que pilote de 2RM ou conducteur automobile construisent leur stratégie de pilotage sur l'anticipation de la courbure de la route en différents points, l'objectif est de concevoir un régulateur supposé avoir le même comportement « entrées – sorties » que l'humain.

Une autre approche, cherche à modéliser les différentes fonctions sensorielles, cognitives et neuro-musculaires de l'humain [54], [6] et [73]. Récemment, dans la même lignée Loiseau et al. dans [39], [38] se sont intéressés au problème d'identification d'un modèle cybernétique du conducteur de 2RM à partir de données obtenues lors d'expérimentations sur piste d'essai. L'objectif était de représenter la dynamique sensorimotrice sous-tendant le contrôle de la moto. L'idée est de déterminer en particulier s'il était possible d'opérer une distinction entre motocyclistes favorisant l'inclinaison du corps et ceux favorisant la manœuvre dite du contre-braquage pour contrôler leurs manœuvres de prise de virage.

Dans ce travail de thèse, nous nous intéressons à cette question de catégorisation des conducteurs de 2RMs dans ces deux catégories à savoir : ceux favorisant l'inclinaison du corps et ceux favorisant le guidon pour contrôler leur moto pendant la manœuvre de prise de virage. La question scientifique qui en découle que nous aborderons dans ce chapitre est : Les comportements des conducteurs lors de la prise de virage peuvent-ils être réduits à un « profil/style » de conduite, ou classés dans des classes ? Nous faisons le postulat que le conducteur peut être « réduit » à un problème sensori-moteur et que les différentes stratégies observées devraient par conséquent être identifiées à l'aide de marqueurs discriminants sur les actions du conducteur.

L'atteinte de cet objectif passe par l'étude du comportement de conduite et la caractérisation du style de conduite de conducteurs de 2RMs en calculant des indicateurs statistiques, d'une façon générale, et en particulier lors de la prise de virage.

### 3.2.2 Étude du comportement de conduite dans le cas d'un 2RM

Divers travaux de recherche portant sur l'étude du comportement de conduite et la caractérisation du style de conduite existent dans la littérature. Nous ne présenterons ici que les études fondées sur l'apprentissage automatique qui nous ont paru les plus pertinentes pour nos travaux. Dans ce qui suit, nous allons citer les travaux les plus significatifs de notre point de vue.

Dans [30], les auteurs ont proposé une méthodologie fondée sur deux étapes à savoir, une étape de caractérisation des trajectoires, et une étape de classification. Cette première étape de caractérisation consiste à calculer des caractéristiques décrivant le comportement de conduite issues des données capteurs pour chacune des trajectoires. Par la suite, une étape de classification en utilisant des algorithmes de classification de type k-moyennes, et CAH a été mise en œuvre pour la classification des styles de conduite. Les auteurs ont montré des résultats probants quant à l'utilisation de cette méthodologie pour la reconnaissance des situations de conduite.

Dans une autre étude [28], un algorithme en deux étapes a été utilisé pour la segmentation et le regroupement des comportements de conduite. Les données analysées dans cette article proviennent d'une expérimentation de conduite naturelle qui s'est intéressée à l'étude de comportements de conduite dans le cas d'une situation de suivi entre véhicules. Tout d'abord, une période de suivi entre deux véhicules est divisée en différents segments de durée qui tiennent compte de leur distribution temporelle. Ensuite, les segments produits par l'étape précédente sont regroupés sur la base de leur similarité. Des variantes de l'algorithme k-means ont été utilisées dans ce processus. Les centroïdes des segments obtenus ont été regroupés par la suite. Les résultats ont révélé que chaque conducteur présente une distribution unique de comportements, mais certains comportements existent chez plus d'un conducteur mais à des fréquences différentes.

Toutefois, il existe peu d'études dans la littérature liée à la prise de virage dans le cas d'un 2RMs. Les auteurs ont appliqué les modèles de Markov cachés (HMMs) sur des données collectées sur route ouverte [42]. L'objectif de l'étude était de proposer un indicateur pour l'évaluation des compétences de conduite des conducteurs de 2RMs lors de la prise de virage, en décomposant cette manoeuvre en deux phases : phase dynamique et phase stationnaire. Dans l'implémentation du modèle HMM, les auteurs ont utilisé seulement la dynamique du roulis (angle, vitesse et accélération de roulis).

Ces différentes études sur l'étude du comportement de conduite nous ont inspiré afin de développer une méthodologie pour l'étude du comportement de conduite pour les 2RMs tout en tenant compte des particularités et spécificités de la conduite de ce mode de transport. L'approche présentée dans ce chapitre est une méthodologie fondée sur une exploration ou fouille de données. Celle-ci, très simple à utiliser, est adaptée pour l'étude du comportement des conducteurs ainsi que pour une bonne compréhension de la façon dont ces derniers interagissent avec leur véhicule lors de la prise d'un virage. Cette dernière, nous permettra d'analyser, de découvrir et de vérifier certaines hypothèses émises sur les pratiques réelles de conduite en 2RMs.

Par ailleurs, cette première approche nous aidera à apporter des solutions aux problèmes suivants :

- Comment mettre en place une visualisation rapide et robuste de la classification des conducteurs de 2RMs en se fondant sur des variables quantitatives ;
- Comment interpréter le résultat de cette classification en utilisant les données déclaratives issues de l'enquête, en amont, effectuée chez les participants à l'expérimentation ;
- Comment mettre en évidence les variables quantitatives les plus discriminantes pour chacun des groupes identifiés.

### 3.3 Méthodologie de classification des différents profils de conduite

Dans cette section, nous allons présenter la méthodologie utilisée pour effectuer la classification des 8 participants à l'expérimentation suivant les paramètres de conduite notamment : l'angle guidon, reflétant les actions du pilote sur la moto, ainsi que celles décrivant la trajectoire de la moto : positionnement sur la chaussée (courbure) [37], la dynamique de la moto à l'approche du virage, à savoir : la vitesse, l'accélération et le jerk de la moto à l'entrée du virage [46]. Notre choix s'est porté sur ces variables car nous pensons que ces dernières refléteraient si les sujets ont été influencés par la condition expérimentale (consigne donnée : « corps » / « guidon »), ont été essentiellement déterminés par la géométrie de la route, ou par les deux facteurs. Cette méthodologie est illustrée sur La [figure 3.1](#).

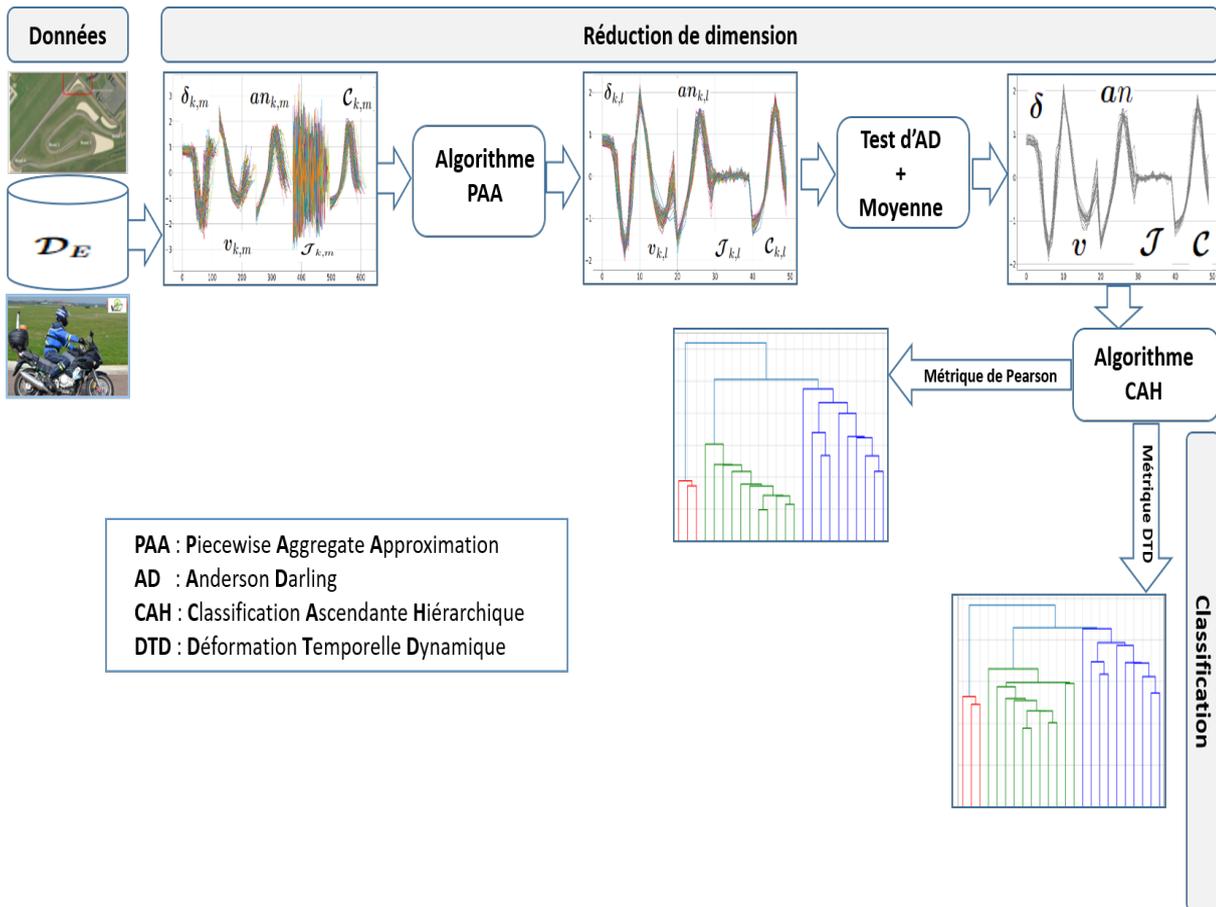


Figure 3.1 : Résumé de la méthodologie de classification des profils de conduite lors de la prise de virage.

### 3.3.1 Description et représentation des données utilisées

Les données utilisées dans ce chapitre ont été extraites depuis la base de données globale  $\mathcal{D}$  présentée au chapitre 2. Ces données sont issues des capteurs C1, C3 et C11 mesurant respectivement l'angle du guidon ( $\delta$ ), la vitesse ( $v$ ) ainsi que les coordonnées GPS (longitude et latitude). À partir des données de la vitesse, du guidon et des coordonnées GPS, nous avons obtenu les mesures d'accélération normale ( $an$ ), de jerk ( $\mathcal{J}$ ) et de la courbure ( $\mathcal{C}$ ). L'accélération normale et le jerk sont calculés comme étant la dérivée première et seconde par rapport au temps de la vitesse [9]. Quant à la courbure, elle a été calculée grâce aux coordonnées GPS (longitude et latitude) [71]. La figure 3.2 illustre le processus d'obtention de la base de données  $\mathcal{D}_E^l$ , obtenue à partir de la base de données globale  $\mathcal{D}$ , composée des mesures  $\delta$ ,  $v$ ,  $an$ ,  $\mathcal{J}$  et  $\mathcal{C}$ .

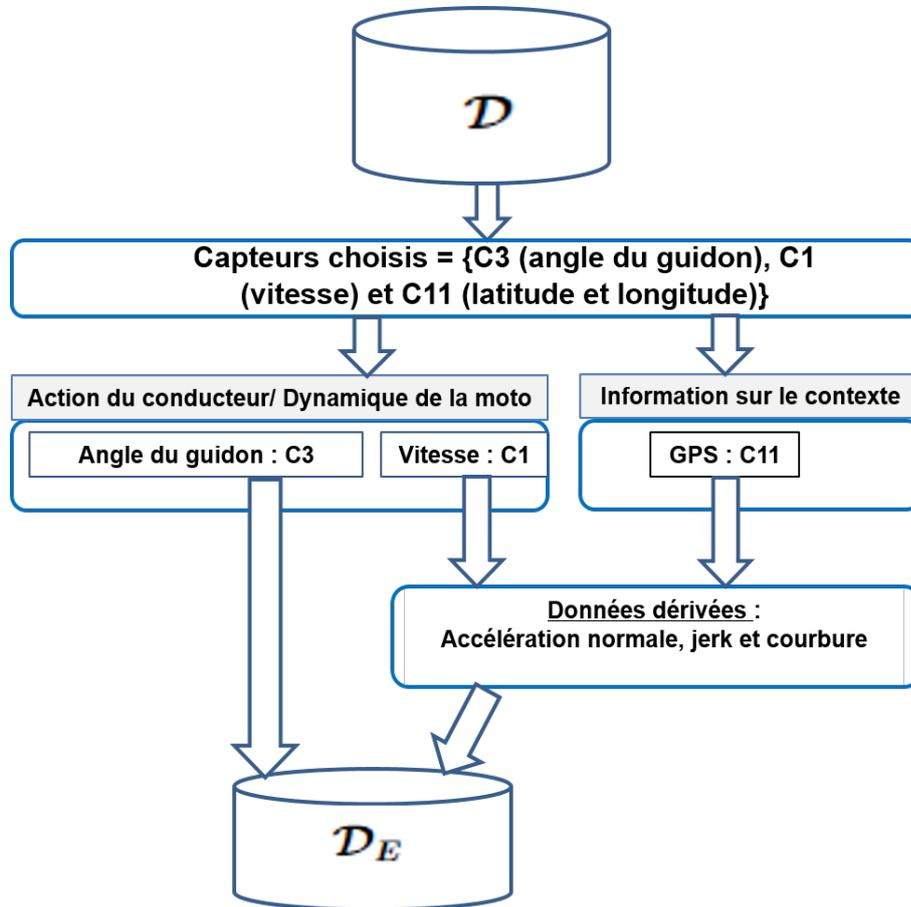


Figure 3.2 : Processus de mise en place de la base de données extraite.

Pour mieux appréhender la base de données  $\mathcal{D}_E^l$  utilisée dans ce chapitre, nous avons mis en place les notations suivantes :

$$\mathcal{D}_E = \langle \mathcal{D}_E^l \rangle_{l \in \{CL, SC, AC\}},$$

avec  $\mathcal{D}_E^l = (\mathbf{x}_k)_{\{k=1,\dots,N\}}$ , où  $N$  est le nombre total de séquences, la consigne de conduite  $l \in \{CL, SC, AC\}$ . Chaque séquence  $\mathbf{x}_k = (\mathbf{x}_{k,m})_{m=1,\dots,M}$  est un ensemble de valeurs observées sur l'intervalle de temps  $(\mathbf{t}_m)_{m=1,\dots,M}$ ,  $M$  (un entier entre 82 et 124) est le nombre total d'observations et  $\mathbf{x}_{k,m}$  est donné par :

$$\mathbf{x}_{k,m} = \{\delta_{k,m}, v_{k,m}, \mathbf{an}_{k,m}, \mathcal{J}_{k,m}, \mathcal{C}_{k,m}\},$$

où  $\delta_{k,m}$ ,  $v_{k,m}$ ,  $\mathbf{an}_{k,m}$ ,  $\mathcal{J}_{k,m}$  et  $\mathcal{C}_{k,m}$  sont l'angle du guidon en degré; la vitesse en  $m/s$ , l'accélération normale en  $m/s^2$ , le jerk (troisième dérivée de la vitesse) en  $m/s^3$  et la courbure en  $m^{-1}$  mesurés au temps  $t_m$  pour la séquence  $k$ , respectivement.

### 3.3.2 Méthodologie de classification des participants

La méthodologie mise en oeuvre à cet effet se résume en trois grandes étapes :

- (1) : Réduction de dimension avec l'utilisation de l'algorithme *Piecewise Aggregate Approximation* (PAA);
- (2) : Utilisation du test d'Anderson-Darling (AD) pour l'étude de l'homogénéité des trois tours de conduite lors des différentes consignes;
- (3) : Classification des participants à l'expérimentation à l'aide de l'algorithme de Classification Ascendante Hiérarchique (CAH).

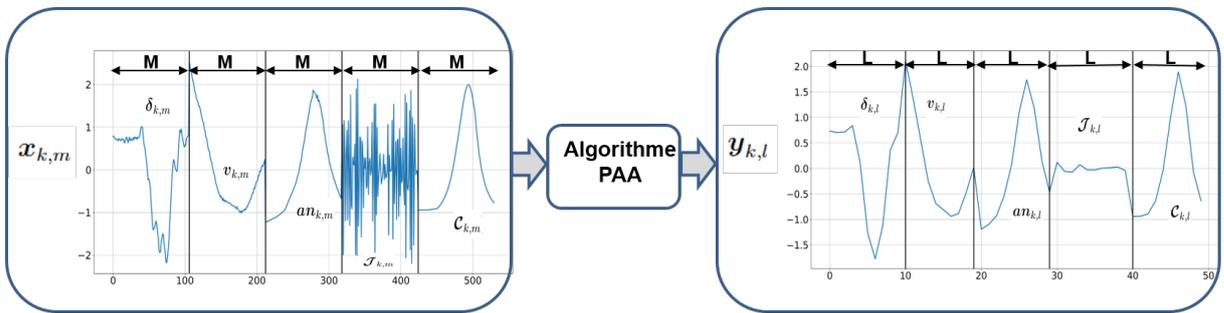
#### 3.3.2.1 Réduction de la dimensionnalité avec l'algorithme d'approximation agrégée par morceaux (« regroupement *Piecewise Aggregate Approximation* » (PAA))

L'algorithme de Classification Ascendante Hiérarchique (CAH) repose sur des métriques mesurant la similarité entre deux échantillons. Dans pareil cas, comme mentionné dans [31], les solutions les plus prometteuses suggèrent d'effectuer, en amont, une réduction de la dimensionnalité des données. Dans [31], les auteurs montrent de façon théorique et empirique que l'algorithme PAA donne de meilleurs résultats que les autres techniques de réduction de dimensionnalité telles que la décomposition en valeur singulière (SVD), la transformée de Fourier discrète (DFT) et la transformée en ondelettes discrètes (DWT).

L'algorithme PAA est très souvent utilisé pour la recherche de comportements similaires en raison de sa simplicité d'utilisation. La réduction de la dimensionnalité avec l'algorithme PAA passe par une représentation en valeur réelle d'une série de données à l'instar de la moyenne, la variance, le kurtosis, etc. Contrairement à ces derniers, l'algorithme PAA permet la conservation des principales informations d'un ensemble de données [31]. En d'autres termes, l'idée de base de la méthode PAA est de réduire la dimensionnalité des séries temporelles en les divisant en segments de taille égale, ensuite en calculant la moyenne des valeurs dans ces segments, et en les remplaçant par cette dernière. La méthode PAA a été utilisée dans plusieurs domaines d'études,

notamment la détection d'anomalies [53], l'analyse financière [8] et l'étude de l'activité cardiaque [35].

La méthode PAA comporte dans la pratique un paramètre essentiel nommé taille-PAA. Ce paramètre est choisi de façon empirique. Dans notre cas, après avoir effectué une normalisation des données de type Z-score<sup>1</sup> comme suggéré dans [31] sur tous les  $\mathbf{x}_{k,m}$ , nous avons appliqué l'algorithme PAA sur ces dernières en choisissant le nombre 10 comme taille-PAA. Nous obtenons avec une telle taille-PAA approximativement, 10 valeurs, dans chacun des segments, coïncidant avec la fréquence de nos données (10Hz). Sur une séquence  $\mathbf{x}_{k,m}$  donnée, nous obtenons ainsi une séquence réduite  $\mathbf{y}_{k,l}$  avec  $l = 1, \dots, L$  où  $L=10$ . La figure 3.3 montre un exemple d'application de l'algorithme PAA sur une séquence de conduite  $\mathbf{x}_{k,m}$ .



**Figure 3.3 :** Un exemple d'application de l'algorithme PAA sur la séquence de conduite  $\mathbf{x}_{k,m}$  de taille  $M=124$  correspondant au deuxième tour du participant P6 lors de la consigne de conduite CL.

#### 3.3.2.2 Étude de l'homogénéité entre les trois tours de conduite pour chacune des consignes

Pendant la phase de pré-analyse des données, cf. chapitre 2, nous avons constaté une variabilité intra-sujet sur les données mesurées, en terme de dynamique d'approche du virage, et de positionnement sur la chaussée. Cette variabilité pourrait être expliquée par l'hétérogénéité des profils des conducteurs (cf. Table 2.1), et aussi, par le fait que les motards n'étaient pas habitués à conduire la moto instrumentée. Par conséquent, les mesures des variables :  $\delta$ ,  $v$ ,  $an$ ,  $\mathcal{J}$  et  $\mathcal{C}$  utilisées en entrée de cette méthodologie peuvent différer d'un tour à l'autre. Dans le but de vérifier l'hypothèse selon laquelle ces mesures diffèrent (hétérogènes) ou non (homogène) d'un tour à un autre, des tests statistiques, notamment ceux de Kolmogorov-Smirnov [29] ou d'Anderson-Darling [58] peuvent être employés. Cependant, pour les petits échantillons, le test d'Anderson-Darling donne de meilleurs résultats comparés à ceux fournis par celui de Kolmogorov-Smirnov si l'on se réfère à l'étude [19]. C'est la principale raison pour laquelle nous utilisons, dans ce chapitre, le test d'Anderson-Darling (AD) pour l'étude de l'homogénéité des

<sup>1</sup>Cette méthode de normalisation consiste à soustraire la moyenne et à la diviser par l'écart type. Dans ce cas, chaque valeur refléterait la distance par rapport à la moyenne en unités d'écart-type. Les données résultantes seraient centrées-réduites, caractérisées par une moyenne de 0 et un écart-type de 1.

tours de conduite dans les différentes consignes.

Le test statistique d'Anderson-Darling est un test non paramétrique. Autrement dit, il n'existe aucun paramètres à estimer dans la distribution de probabilité. Dans ce sens, on dit que le test et l'ensemble de ses valeurs critiques sont indépendants de la distribution de probabilité. Dans la pratique, ce test est souvent utilisé pour tester si 2 ou plusieurs échantillons de données proviennent de la même distribution. C'est ainsi qu'on parle, dans la littérature, de test d'Anderson-Darling à  $k$  échantillons (avec  $k \geq 2$ ). C'est d'ailleurs sous cette forme que nous avons utilisée ce test dans le cadre de ce chapitre avec comme nombre d'échantillons 3 représentant le nombre de tours effectué dans chacune des consignes de conduite.

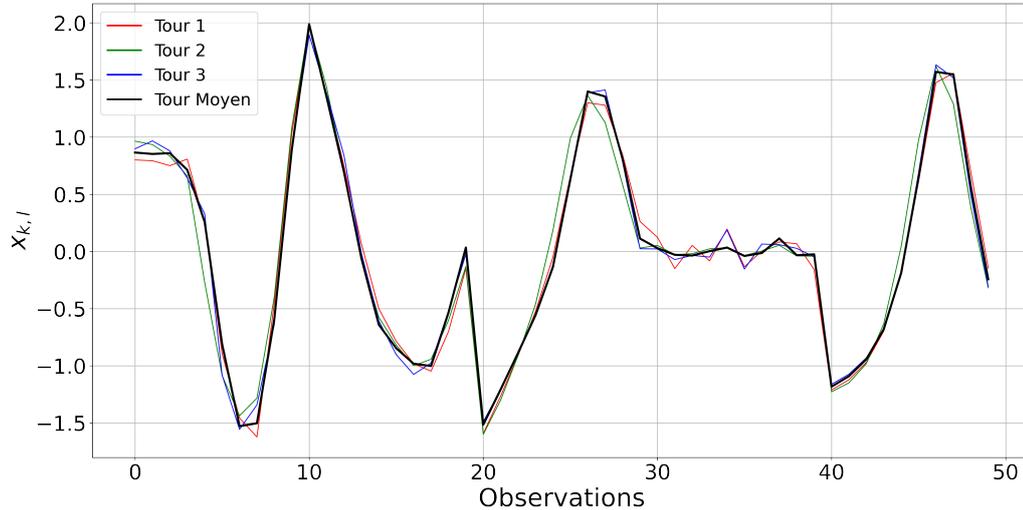
En pratique, les hypothèses du test d'Anderson-Darling sont les suivantes : une hypothèse nulle ( $H_0$ ) et une hypothèse alternative ( $H_1$ ). Sous l'hypothèse  $H_0$ , on admet que tous les échantillons sont issus de la même probabilité de distribution, tandis que sous  $H_1$ , au moins un des échantillons a une probabilité de distribution différente des autres. Dans notre cas, le test d'AD est effectué sur les  $\mathbf{y}_{k,l}$  obtenu au terme de la réduction de la dimensionnalité avec l'algorithme PAA. Par conséquent, pour chaque instruction et une variable donnée, nous pouvons tester s'il existe une différence significative entre les distributions des trois tours de conduite. Dans le test d'Anderson-Darling, nous rejetons l'hypothèse nulle ( $H_0$ ) si la  $p$ -valeur est inférieure au seuil de référence  $\alpha$ , qui est généralement fixé à 5 %. En d'autres termes, la valeur  $p$  est le plus grand seuil pour lequel l'hypothèse  $H_0$  est acceptée. Dans [57], les auteurs ont fait un exposé détaillant le calcul de la statistique d'AD ainsi que ses valeurs critiques.

Participants	$\delta$			$v$			$an$			$\mathcal{J}$			$\mathcal{C}$		
	CL	SC	AC	CL	SC	AC	CL	SC	AC	CL	SC	AC	CL	SC	AC
P1	0,69	0,51	0,96	0,99	0,99	0,91	0,99	0,99	0,99	0,70	0,30	0,74	0,99	0,99	0,99
P2	0,97	0,90	0,95	0,99	0,99	0,99	0,99	1	1	<b>0,24</b>	0,87	0,99	1	1	0,99
P3	0,96	0,66	0,70	0,99	0,99	0,99	1	0,99	0,99	0,84	0,90	0,92	1	0,99	0,99
P4	0,99	0,98	0,99	1	0,99	0,91	0,99	1	0,99	0,85	0,94	0,99	0,99	0,99	0,99
P5	0,74	0,92	0,99	0,90	0,99	0,99	1	0,99	0,99	0,76	0,92	0,77	1	0,99	1
P6	0,55	0,74	0,81	0,88	0,99	0,99	0,99	0,99	1	0,99	0,64	0,98	0,99	0,99	1
P7	0,99	0,97	0,92	0,99	0,99	0,99	1	1	0,99	0,96	0,85	0,88	0,99	1	0,99
P8	0,95	0,98	0,90	0,99	0,94	0,99	1	0,99	0,99	0,48	0,26	0,99	0,99	0,99	0,93

Tableau 3.1 :  $p$ -valeurs du test d'Anderson-Darling pour toutes les instructions et variables.

Le [tableau 3.1](#) fournit les  $p$ -valeurs du test d'Anderson-Darling pour toutes les instructions et variables des 8 participants. Dans ce tableau, la valeur minimale, qui est indiquée en gras, est de 0,24, ce qui est supérieur à  $\alpha$  (0,05). Par conséquent, pour toutes les instructions de conduite et les variables, nous ne pouvons rejeter l'hypothèse nulle. Ainsi, la différence existante entre les trois tours n'est pas significative pour un sujet donné dans toutes les instructions et pour toutes les variables.

En conclusion, les résultats du test statistique d'AD nous permettent de confirmer l'hypothèse selon laquelle cette hétérogénéité intra-sujet n'est pas significative statistiquement, et par conséquent, nous pourrions réduire les trois tours effectués dans chaque instruction de conduite en un seul tour en prenant leur moyenne. Ainsi, pour une consigne de conduite donnée, nous pouvons caractériser chaque sujet par un comportement moyen de conduite lors de la prise de virage, voir [figure 3.4](#).



**Figure 3.4 :** Comportement moyen de conduite des trois tours du participant P2 pour la consigne de conduite libre.

Pour les huit participants et les trois instructions de conduite, nous obtenons un ensemble de données regroupé dans une matrice de taille  $24 \times 50$ . Le nombre de lignes 24 est obtenu en multipliant le comportement moyen des 8 participants par le nombre 3 d'instructions de conduite. Quant au nombre de colonnes 50, il est obtenu en multipliant  $L = 10$  la taille de la fenêtre choisie dans l'algorithme PAA par le nombre de variables de conduite qui est égal à 5.

#### 3.3.2.3 Classification CAH des participants à l'expérimentation

La dernière étape de la méthodologie consiste à utiliser un algorithme de classification automatique. Les algorithmes de classification automatique reposent sur le même principe de partitionnement d'un ensemble de données en classes (groupes) homogènes. Nous pouvons distinguer deux catégories de méthodes : approches non probabilistes et approches probabiliste. Une approche est dite non probabilistes lorsqu'elle ne fait pas d'hypothèse concernant la distribution de probabilité des données à classifier. Comme exemples de ce type d'algorithmes, nous pouvons citer : l'algorithme CAH [72], la méthode des K-moyennes [41], les cartes auto-organisatrices de Kohonen [33] et la classification floue [4].

Les algorithmes CAH et K-moyennes font partie des algorithmes de partitionnement les plus utilisés, et ce, dans divers domaines de recherches et d'applications. Ces derniers sont faciles à mettre en oeuvre et partagent un principe d'optimisation commun où il s'agit de minimiser l'inertie intra-classe et de maximiser celle d'inter-classes. L'algorithme CAH est une méthode de classification itérative [36] où les classes sont regroupées itérativement avec comme point de départ un partitionnement élémentaire (un élément par classe), puis en fusionnant successivement les classes qui se ressemblent jusqu'à l'obtention d'une seule classe. Ces regroupements successifs de classes sont matérialisés dans un dendrogramme représentant une hiérarchie des partitions. Ce dendrogramme constitue un outil adapté pour une visualisation rapide, claire et précise des différentes classes. C'est la raison pour laquelle, dans ce chapitre, nous utiliserons l'algorithme CAH pour la classification des participants à l'expérimentation de conduite lors des différentes consignes de conduite.

Dans notre cas, l'algorithme CAH est utilisé sur la base de données des comportements moyens de conduite lors des différentes consignes que nous noterons  $Y_{ij}$ . Pour utiliser cet algorithme, nous avons eu recours à deux méthodes pour le calcul de la distance existante entre deux séries temporelles mais aussi celle entre deux classes (critère de liaison). Pour mesurer la distance entre deux séries d'observations  $Y_{ij}$  et  $Y_{mj}$  ( $\forall i \neq m$ ), nous utilisons les métriques de Pearson [21] et celle de la déformation temporelle dynamique (DTD) [59]. Quant à la distance entre deux classes, le critère de lien unique désignant la distance minimale entre deux classes a été utilisé. Les étapes principales de la méthode sont décrites dans le pseudo-code 1.

---

**Algorithm 1** Pseudo-code de l'algorithme CAH appliqué sur la matrice de données  $Y_{ij}$ .

---

**Initialiser:** Matrice de données  $Y_{ij}$

- 1: **Calculer la similarité** : Utilisation des métriques de la corrélation de Pearson et de la déformation temporelle dynamique entre chaque paire de lignes de la matrice de données  $Y_{ij}$ .
  - 2: **Calculer** la matrice de Similarité.
  - 3: **Fusionner** les classes les plus proches en utilisant les résultats fournis par la matrice de Similarité.
  - 4: **Mettre-à-jour** la Similarité entre les nouvelles et anciennes classes.
  - 5: **Répéter** les étapes 3 et 4 jusqu'à l'obtention d'une seule classe.
-

### 3.4 Application de la méthodologie proposée : Résultats et Discussions

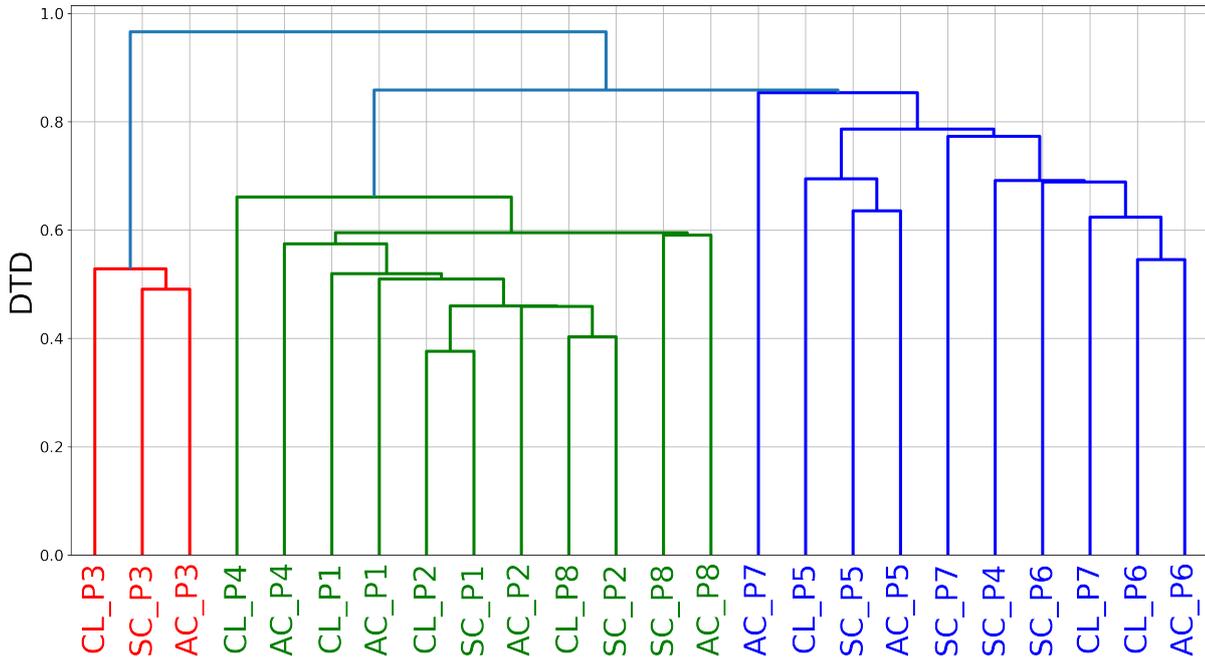
Dans cette section, nous présentons et discutons les résultats obtenus après application de l'algorithme CAH sur la matrice de données  $Y_{ij}$ . Comme tout algorithme de classification non-supervisée, le choix du nombre de classes constituent la tâche la plus difficile. D'ailleurs, cet exercice est à l'origine de plusieurs discussions dans la littérature conduisant à la mise en place de plusieurs critères de choix. Parmi ceux-ci, nous avons le Critère d'Information Bayésien [68], la règle du coude et la statistique de l'écart [70]. Cependant, ces critères sont souvent difficiles à mettre en oeuvre dans la pratique où la plupart des cas le nombre de classes est choisi en se fondant sur l'expérience ou la connaissance métier. Ainsi, dans notre cas, le nombre de classes est choisi de façon intuitive pour rendre plus aisée l'interprétation des résultats. En effet, si chaque participant à l'expérimentation respectait les consignes de conduite tout en étant en adéquation avec ses déclarations de préférence de conduite (SC ou AC), le nombre idéal de classes serait de 2.

La figure 3.5 donne les résultats après application de la méthodologie en utilisant les métriques de la corrélation de Pearson et de la déformation temporelle dynamique (DTD). Une première analyse des dendrogrammes montre qu'une coupure pratiquée sur ces derniers permettant une partition en deux classes n'est guère une solution optimale. En effet, cette coupure opposerait le participant P3 (CL, SC et AC) aux autres participants restants, toutes consignes confondues. Ainsi, pour une solution optimale, nous avons coupé les dendrogrammes de façon à avoir 3 classes, voir tableau 3.2. Ce tableau montre que nous obtenons les mêmes résultats pour les métriques DTD et de corrélation de Pearson.

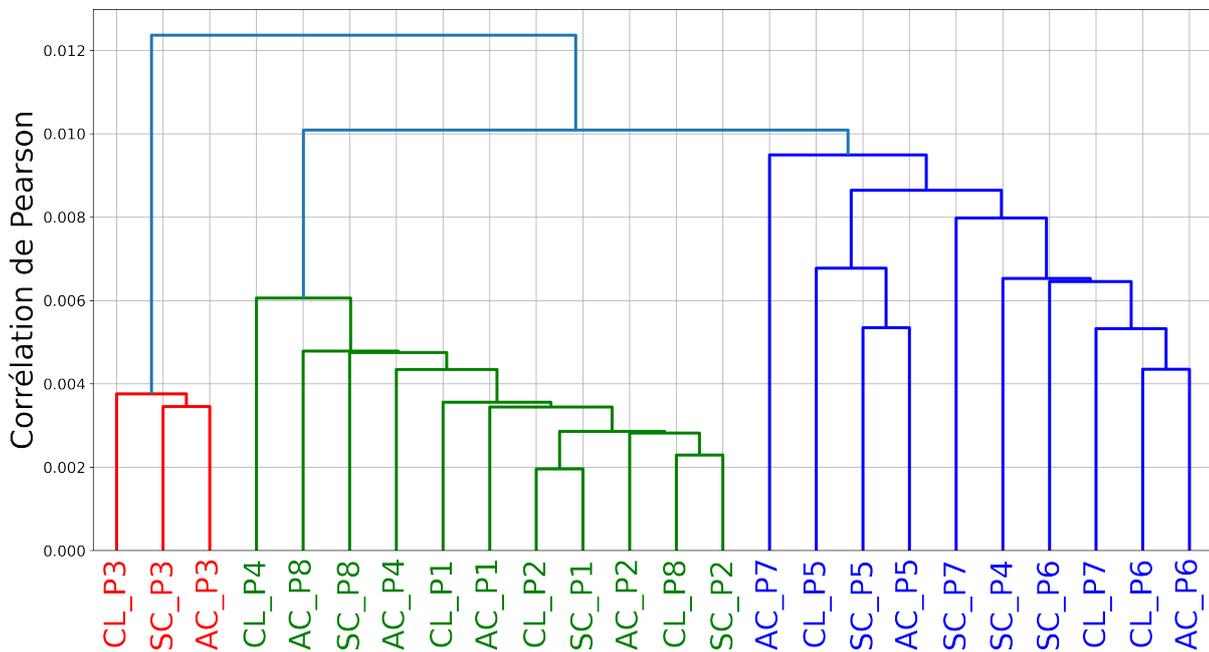
Classes	Métriques	
	DTD	Corrélation de Pearson
Classe 1 (Rouge)	P3 (CL, SC, AC)	P3 (CL, SC, AC)
Classe 2 (Verte)	P1 (CL, SC, AC)	P1 (CL, SC, AC)
	P2 (CL, SC, AC)	P2 (CL, SC, AC)
	<b>P4 (CL, AC)</b>	<b>P4 (CL, AC)</b>
Classe 3 (Bleue)	P8 (CL, SC, AC)	P8 (CL, SC, AC)
	<b>P4 (SC)</b>	<b>P4 (SC)</b>
	P5 (CL, SC, AC)	P5 (CL, SC, AC)
	P6 (CL, SC, AC)	P6 (CL, SC, AC)
	P7 (CL, SC, AC)	P7 (CL, SC, AC)

**Tableau 3.2** : Résultats de la classification obtenues par l'algorithme Classification ascendante hiérarchique (CAH) avec les métriques : déformation temporelle dynamique (DTD) et corrélation de Pearson.

(a) Résultats de l'algorithme de Classification ascendante hiérarchique (CAH) avec la métrique déformation temporelle dynamique (DTD).



(b) Résultats de l'algorithme de Classification ascendante hiérarchique (CAH) avec la métrique de la corrélation de Pearson.



**Figure 3.5** : Résultats de la classification obtenues par l'algorithme Classification ascendante hiérarchique (CAH) avec les métriques : déformation temporelle dynamique (DTD) et corrélation de Pearson.

#### 3.4.1 Analyses des classes fournies par les dendrogrammes

Dans cette sous-section, nous effectuons une confrontation des résultats d'analyse du comportement de conduite extraits de l'exploration de données (cf. [tableau 3.2](#)) et les préférences de conduite déclarées en amont de l'expérimentation (conduite SC ou AC). Cette analyse croisée nous permettra de déterminer dans quelle mesure les participants ont respecté ou non les instructions de conduite données.

- **Analyse de la Classe 1** : Cette classe est essentiellement composée des trois instructions de conduite du participant P3. Une analyse approfondie des dendrogrammes montre que la similarité est davantage observée entre les instructions SC et AC de ce participant. Nous pouvons conclure donc que ce participant semble avoir eu plus de difficultés à moduler son comportement en fonction de la consigne donnée. Ainsi, pour ce participant, des contradictions peuvent être observées au niveau de la préférence de conduite (AC) et du respect des instructions SC et AC.
- **Analyse de la Classe 2** : Cette classe regroupe, d'une part, toutes les instructions de conduite des participants P1, P2 et P8, et d'autre part, les instructions CL et AC de P4. Par conséquent, les participants P1, P2 et P8 ont eu plus de difficultés à moduler leurs comportements en fonction de la consigne donnée. Par ailleurs, le participant P4 qui a déclaré avoir une préférence pour l'instruction AC reste cohérent avec sa déclaration tout en respectant les instructions de conduite.
- **Analyse de la Classe 3** : Pour les mêmes raisons que celles mentionnées dans l'analyse de la classe 2, les participants P5, P6 et P7 ont eu plus de difficultés à moduler leurs comportements en fonction des consignes de conduite données SC et AC. Ce qui n'est pas le cas du participant P4 qui a son instruction SC dans cette classe et qui semble avoir pu moduler son comportement et respecter cette consigne de conduite.

En conclusion, il convient de noter que le participant P3 montre un comportement atypique (attitude très différente des autres). Ce comportement pourrait être expliqué par son niveau d'expérience en conduite moto. En effet, ce participant est en stage en vue de l'obtention de son permis moto (âge permis est 0). Par ailleurs, le participant P4 est le seul à avoir respecté les consignes de conduite sans corps (SC) et avec corps (AC). Sur la base de ce constat, nous pouvons conclure que la classe 2 est composée de participants qui utilisent les actions du corps plus que celles sur le guidon. La classe 3, quant à elle, regroupe des participants utilisant plus les actions sur le guidon que celles du corps.

#### 3.4.2 Interprétation des classes fournies par les dendrogrammes

Dans cette partie, nous nous intéressons à l'interprétation quantitative des classes obtenues par l'algorithme CAH. Pour ce faire, nous posons les hypothèses suivantes : les participants avec une

préférence pour la conduite sans corps (SC) présentent une variabilité importante de l'angle de guidon, et par analogie, les conducteurs avec une faible variabilité de l'angle de guidon sont de potentiels candidats à la préférence de la conduite avec corps (AC). Pour quantifier ces précédentes hypothèses sur la variabilité de l'angle du guidon, l'écart-type est un indicateur simple d'utilisation et efficace. Par conséquent, nous proposons de calculer l'écart-type sur les signaux moyens des trois tours de conduite.

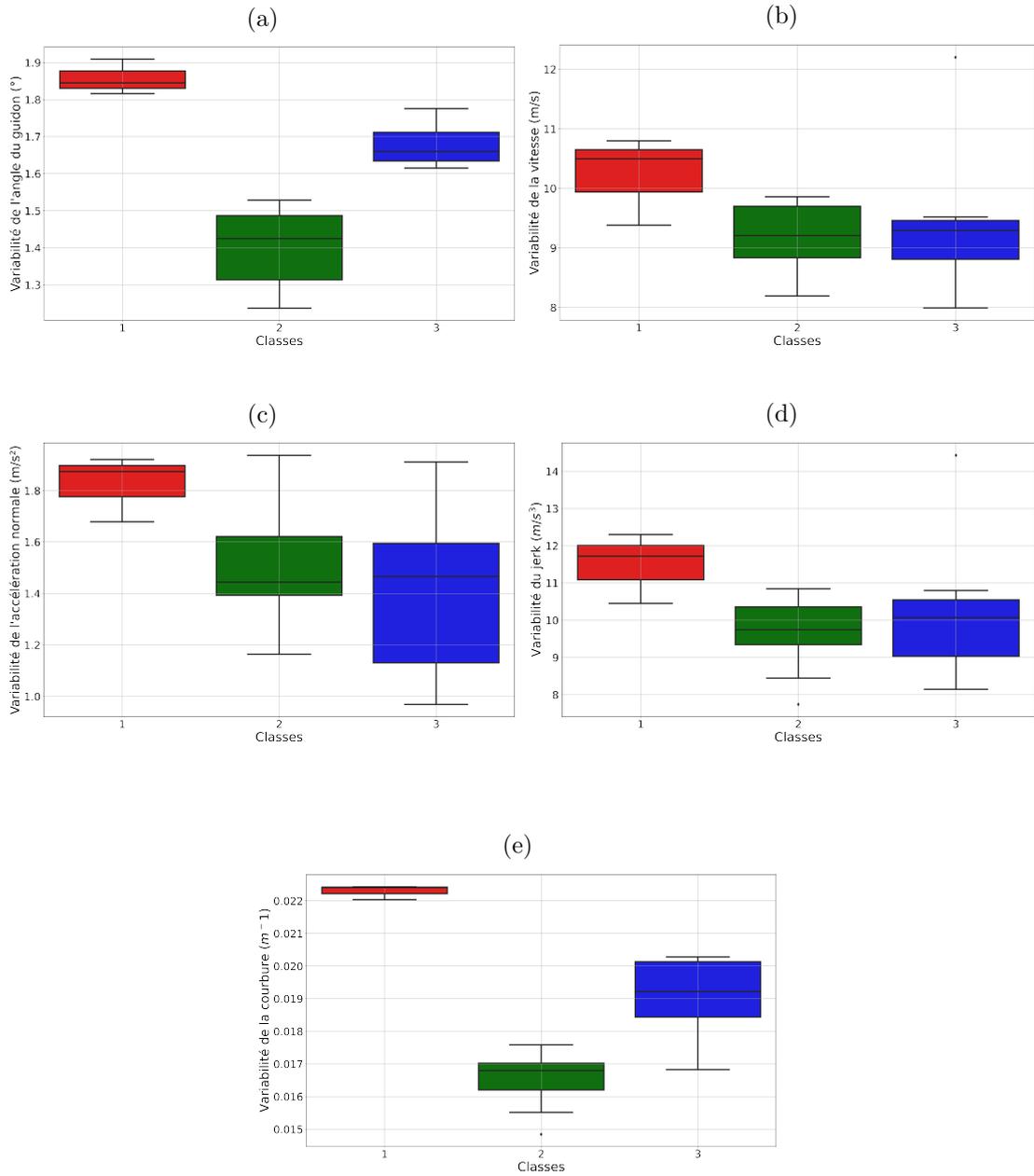
La [figure 3.6a](#) montre que la variabilité de l'angle du guidon contribue le plus dans la distinction des deux principales classes (2 et 3). Pour vérifier ce constat, nous effectuons le test d'AD entre les distributions d'écart-type des classes 2 et 3. Pour les variables de conduite de l'angle du guidon et de la courbure, nous rejetons l'hypothèse nulle selon laquelle la différence n'est pas significative au seuil  $\alpha = 5\%$ . Ainsi, la différence existante entre les classes 2 et 3 est significative pour la variabilité des mesures de conduite de l'angle du guidon et de la courbure, comme nous pouvons le constater dans le [tableau 3.3](#).

Variables de conduite	<i>p</i> -Valeurs
Angle du guidon	<b>0.001</b>
Vitesse	0.250
Accélération Normale	0.250
Jerk	0.250
Courbure	<b>0.001</b>

**Tableau 3.3** : *p*-valeurs du test d'AD entre les classes 2 et 3.

Les éléments de la classe 2 ont une faible variabilité de l'angle du guidon par rapport à ceux de la classe 3. Ainsi, nous pouvons en déduire que la classe 2 regroupe les participants utilisant davantage l'inclinaison du corps que le guidon tandis que ceux de la classe 3 font l'inverse. Par conséquent, les participants P1, P2 et P8 sont en contradiction avec leur préférence de conduite alors que P5, P6 et P7 sont en accord avec leurs déclarations. Ceci n'est pas le cas du participant P4 qui semble être le seul à se conformer aux instructions sans corps (SC) et avec corps (AC).

En se fondant sur les résultats précédents, nous pouvons conclure que notre hypothèse de départ a été vérifiée : les comportements des participants résultant de l'interaction entre le conducteur-moto-infrastructure peuvent être classifiés. Ces classes représentent différentes stratégies de contrôle de l'inclinaison de la moto à savoir au guidon (SC) et avec le corps (AC). Ces différentes stratégies ont été identifiées à l'aide de marqueurs comportementaux discriminants qui ont été calculés sur la base de données obtenues à partir de capteurs physiques comme l'angle du guidon et la vitesse ainsi que d'autres caractéristiques calculées telles que l'accélération normale, le jerk et la courbure.



**Figure 3.6 :** Distribution de l'écart-type par classe pour les 5 variables de conduite : (a) Angle du guidon, (b) Vitesse, (c) Accélération Normale, (d) Jerk et (e) Courbure.

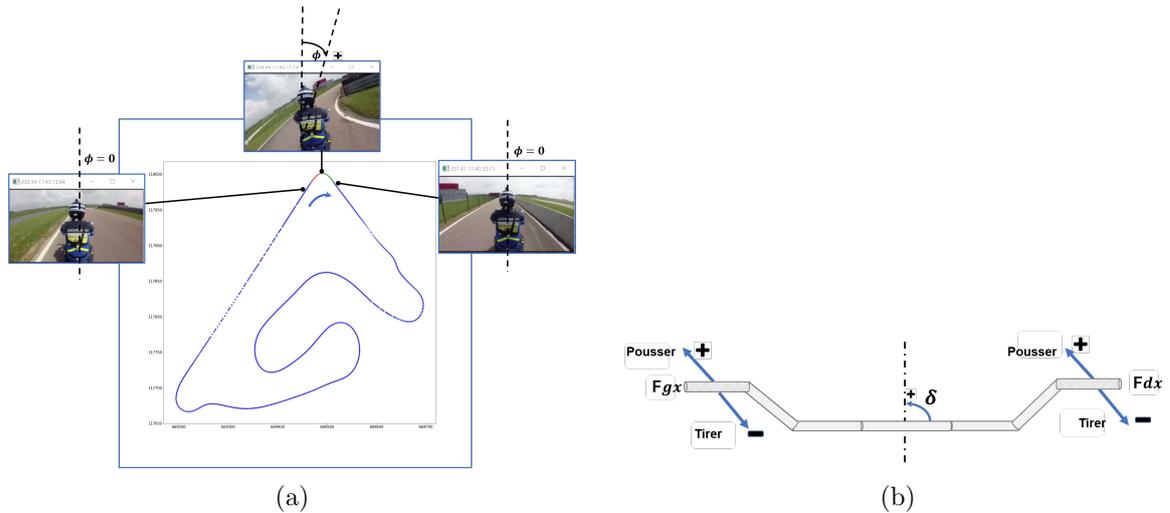
À travers cette caractérisation du comportement du conducteur, nous avons pu capturer les différences existantes dans les processus sensorimoteurs utilisés par le pilote pour contrôler son véhicule par sa gestion du guidon et les adaptations posturales nécessaires au contrôle de l'inclinaison du véhicule. Ces processus sensorimoteurs peuvent différer d'un conducteur à un autre. Cette différence existante entre ces styles de contrôle entre conducteurs est principalement liée à l'expérience de la conduite de chaque conducteur. Néanmoins, à un certain niveau, on assiste à l'existence d'un comportement commun. Cependant, il est important d'indiquer que les limites de ces résultats sont essentiellement relatives au nombre de participants à cette étude. Le nombre réduit des participants dans cette expérimentation ne permet pas d'élargir nos conclusions au vu des résultats enregistrés. Il faut donc dire que la portée de cette étude pourrait s'étendre à un nombre plus grand de participants afin de fournir des résultats susceptibles d'être généralisables. Néanmoins, nous tenons à rappeler que l'objectif de ce travail est de proposer une méthodologie d'analyse de données issues de ce type d'expérimentations et que les résultats obtenus semblent très prometteurs sur la pertinence de cet outil fondée sur l'analyse de données, et notamment, l'apprentissage non-supervisé pour mener ce type de recherche.

Par ailleurs, l'analyse de ces données peut soulever une autre question à savoir : existe-t-il une différence entre les méthodes du contrôle de l'inclinaison de la moto par le guidon ou par l'inclinaison du corps dans l'initiation du virage ?. La partie suivante de ce chapitre est consacrée à la mise en évidence des différences entre les méthodes de contrôle de l'inclinaison de la moto par le guidon ou par l'inclinaison du corps en effectuant une analyse approfondie des données fournies par les capteurs de pression installés sur le guidon, réservoir, les repose-pieds, et la selle, et ce dans le cas du participant P4.

### 3.4.3 Analyse approfondie des comportements de conduite du participant P4

Dans le cas des 2RMs, le conducteur est fortement impliqué dans la tâche de conduite. En effet, ce dernier doit contrôler la trajectoire de son véhicule, et dans le même temps, assurer la stabilité de son véhicule pour contrer l'effet gyroscopique qui joue un rôle essentiel dans la conduite d'une moto, particulièrement dans les virages [16].

Pour une meilleure compréhension dans cette étude approfondie, nous allons présenter les conventions adoptées concernant l'orientation des angles de roulis, des angles de braquage et des forces appliquées sur la direction du guidon, voir [figure 3.7a](#) et [figure 3.7b](#).



**Figure 3.7 :** Conventions utilisées pour (a) le Roulis ( $\phi$ ) et (b) les angles de guidon ( $\delta$ ) et les forces ( $Fg_x, Fd_x$ ) appliquées sur le guidon.

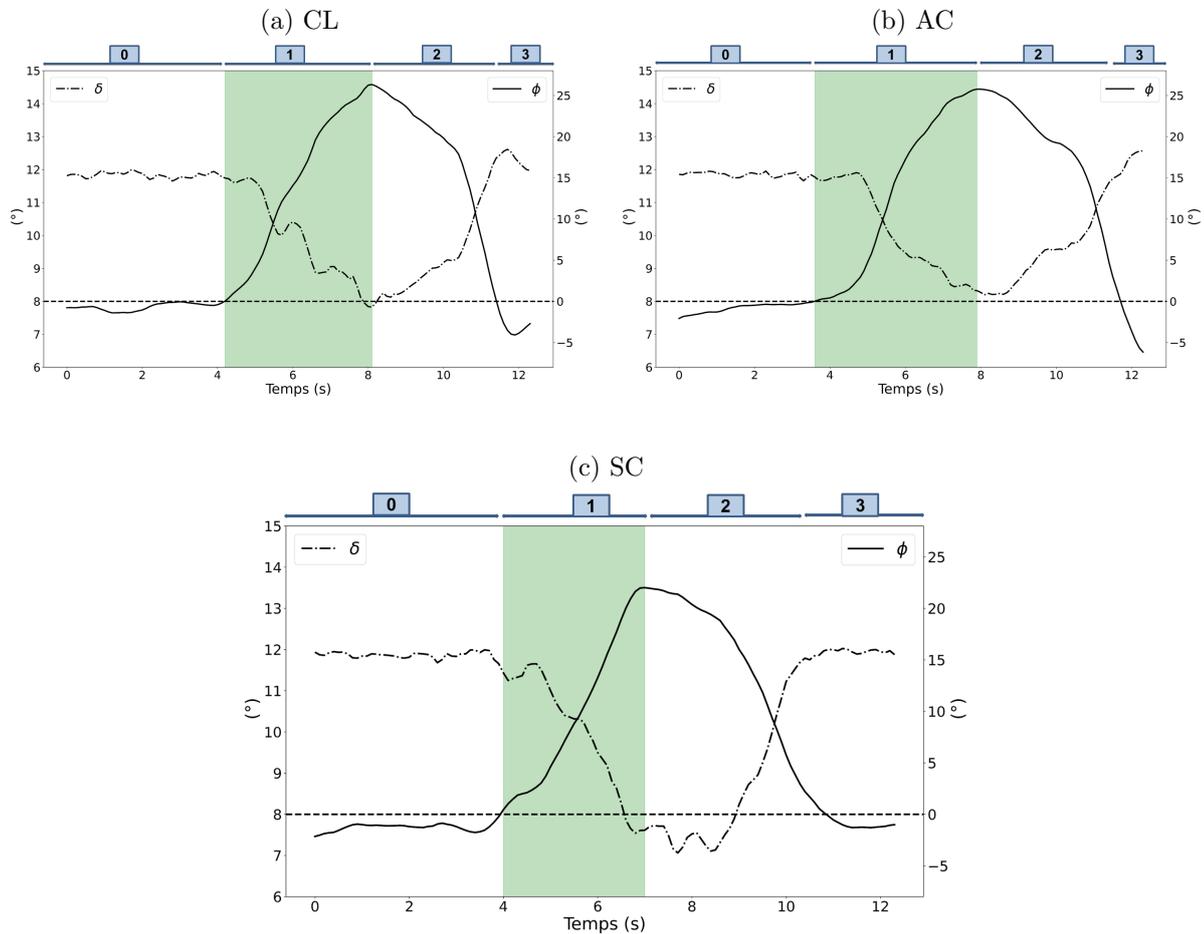
Dans ce qui suit, une analyse approfondie de la manœuvre de virage est illustrée sur le participant P4, et ce, lors des trois instructions données aux conducteurs à savoir : la conduite libre (CL), celle où il a été demandé au conducteur d'utiliser principalement les mouvements de son corps pour incliner la moto (AC), et celle où il a été demandé d'utiliser principalement le guidon (SC) pour incliner la moto, pendant la phase d'initiation du virage. Notre choix s'est porté sur ce participant, qui, selon nous, semble le mieux adapté pour illustrer l'objectif de la méthodologie proposée, car il ressort de l'étape précédente de cette étude que ce participant a respecté les instructions de conduite qui lui ont été fournies, d'une part. Et d'autre part, nous constatons également que sa conduite (consigne CL) coïncide avec celle déclarée, une préférence pour la conduite (AC) cf. Table 2.1. Cette analyse approfondie de la façon dont le participant P4 conduit une moto est faite sur le comportement moyen de ses trois tours de conduite durant les instructions CL, SC et AC. Par conséquent, dans cette analyse approfondie, les résultats attendus sont comme suit :

- ce participant initie le virage en appliquant un couple de forces sur le guidon lors de la réalisation de l'instruction SC. Il est à rappeler que nous avons opté pour des capteurs permettant de mesurer les forces appliquées par le conducteur sur les demi-guidons gauche et droit à la place d'un capteur mesurant le couple appliqué par le conducteur directement.
- dans le cas des consignes de conduite CL et AC, le participant P4 initie le virage par le mouvement de son corps, soit en déplaçant son centre de gravité soit en appuyant sur les repose-pieds.

Notre objectif n'est pas simplement d'identifier les différentes phases du virage, mais plutôt d'analyser l'évolution du comportement du conducteur au cours de ces différentes phases, et

notamment la phase d'initiation du virage.

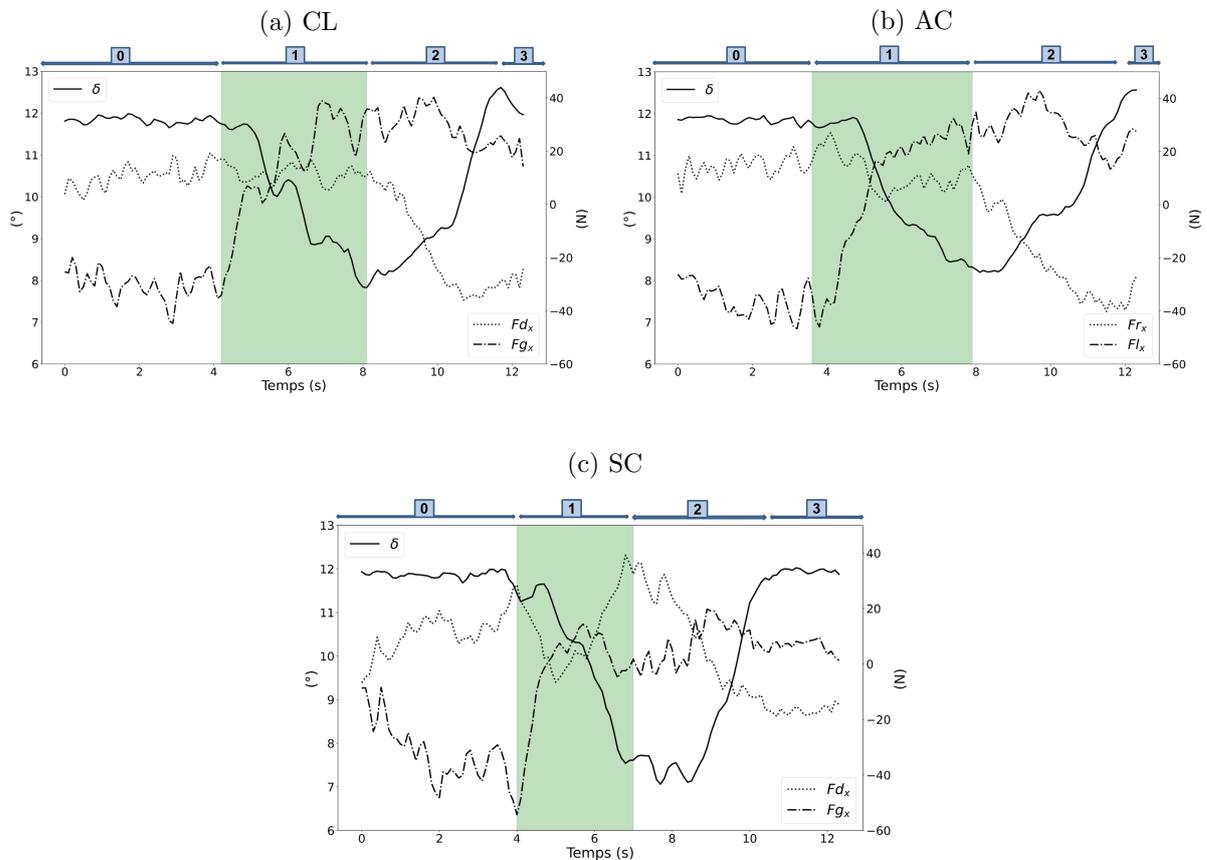
La première étape de cette analyse approfondie, consiste à analyser l'évolution de l'angle du guidon et de l'angle de roulis tout au long du virage. La [figure 3.8](#) montre les angles du guidon et de roulis du comportement moyen du participant P4 lors des différentes consignes de conduite.



**Figure 3.8** : Angle du guidon  $\delta$ (---) et Angle de roulis  $\phi$ (—) exprimés en degré (°) du participant P4 pour les instructions CL, SC et AC. La zone grisée en vert représente l'entrée du virage, le début de cette zone correspond à l'instant où l'angle de roulis franchit le seuil fixé à  $0^\circ$ .

D'après la [figure 3.8](#), nous pouvons constater que le virage peut être décomposé en quatre phases : de 0 à 3. À la **Phase 0**, l'angle de roulis est approximativement égal à 0. Ceci indique que la moto est en position verticale. Au début de la **Phase 1**, la moto commence à rouler dans la direction du virage. Ainsi, on note un angle de roulis qui augmente jusqu'à atteindre un maximum. Ce maximum, correspondant au milieu de la courbe, marque aussi la fin de la **Phase 1** et le début de la **Phase 2**. À partir de ce moment, la force centrifuge devient de plus en plus importante entraînant la moto vers l'extérieur du virage et qui a pour effet de redresser la moto vers sa position de la **Phase 0** (position verticale).

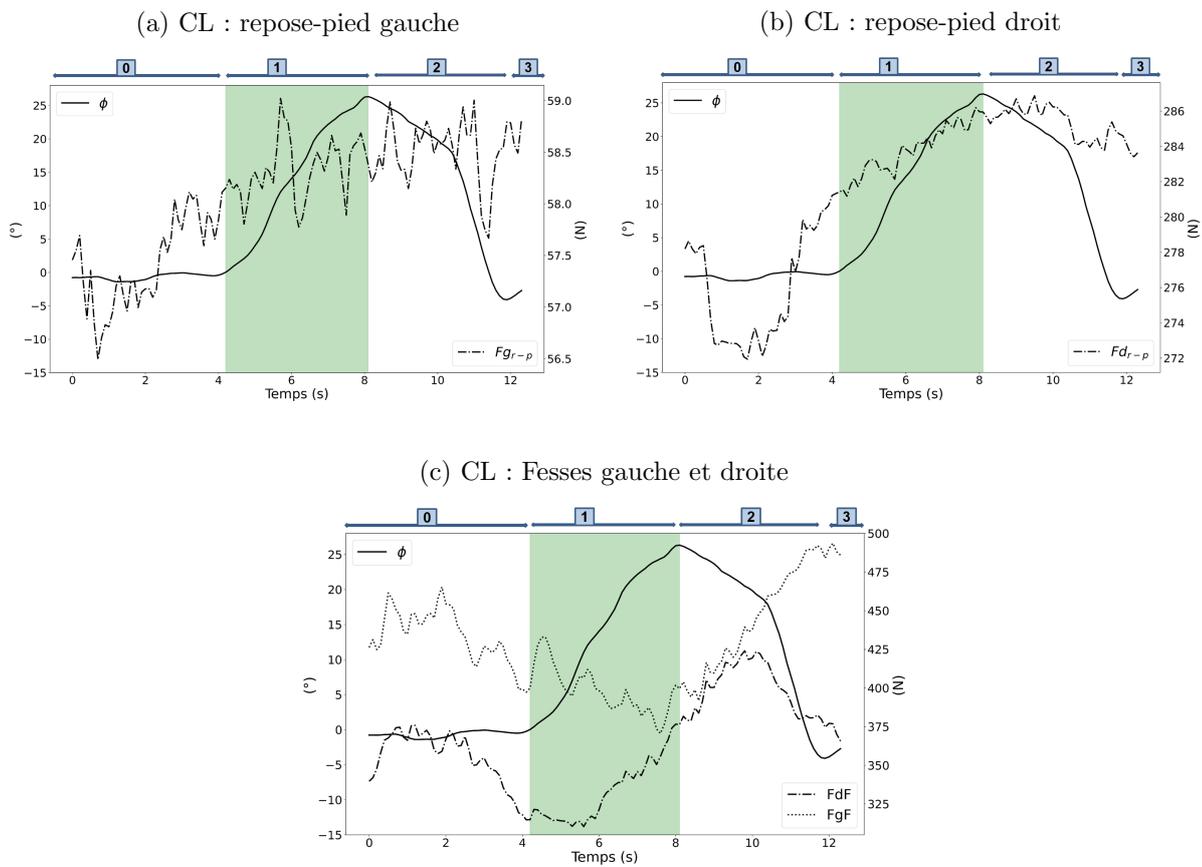
Dans le cas de l'instruction de conduite SC (voir [figure 3.9c](#)), on remarque une diminution de l'angle du guidon avant le début de la **Phase 1** (moment d'amorçage de la courbe). Ce phénomène n'a pas été observé dans les autres instructions où l'angle du guidon reste constant. Dans ce cas, nous pouvons émettre l'hypothèse selon laquelle le conducteur initie le virage en appliquant un couple de forces sur le guidon. Le conducteur crée ce couple en appliquant des forces sur le guidon le long de l'axe longitudinal, à savoir  $F_{dx}$  et  $F_{gx}$  comme le montrent [figure 3.9a](#), [figure 3.9b](#) et [figure 3.9c](#). C'est seulement dans le cas de la consigne SC que la force  $F_{gx}$  diminue de manière significative au début de la phase 1 (initiation de la courbe).



**Figure 3.9** : Forces appliquées en Newton, suivant l'axe horizontal, sur les demi-guidons droit ( $F_{dx}$ ) et gauche ( $F_{gx}$ ) et l'angle du guidon ( $\delta$ ) en degré ( $^\circ$ ) lors des instructions CL, SC et AC.

Il est important de souligner que les données sont issues d'un virage à droite. Par conséquent, si le conducteur essaie d'amorcer le virage avec les actions de son corps et non avec celles sur le guidon, il n'aura que deux possibilités : déplacer son centre de gravité ou appuyer sur le repose-pied droit. Nous nous intéresserons ici uniquement aux actions entreprises par le conducteur avant le début de la **Phase 1**.

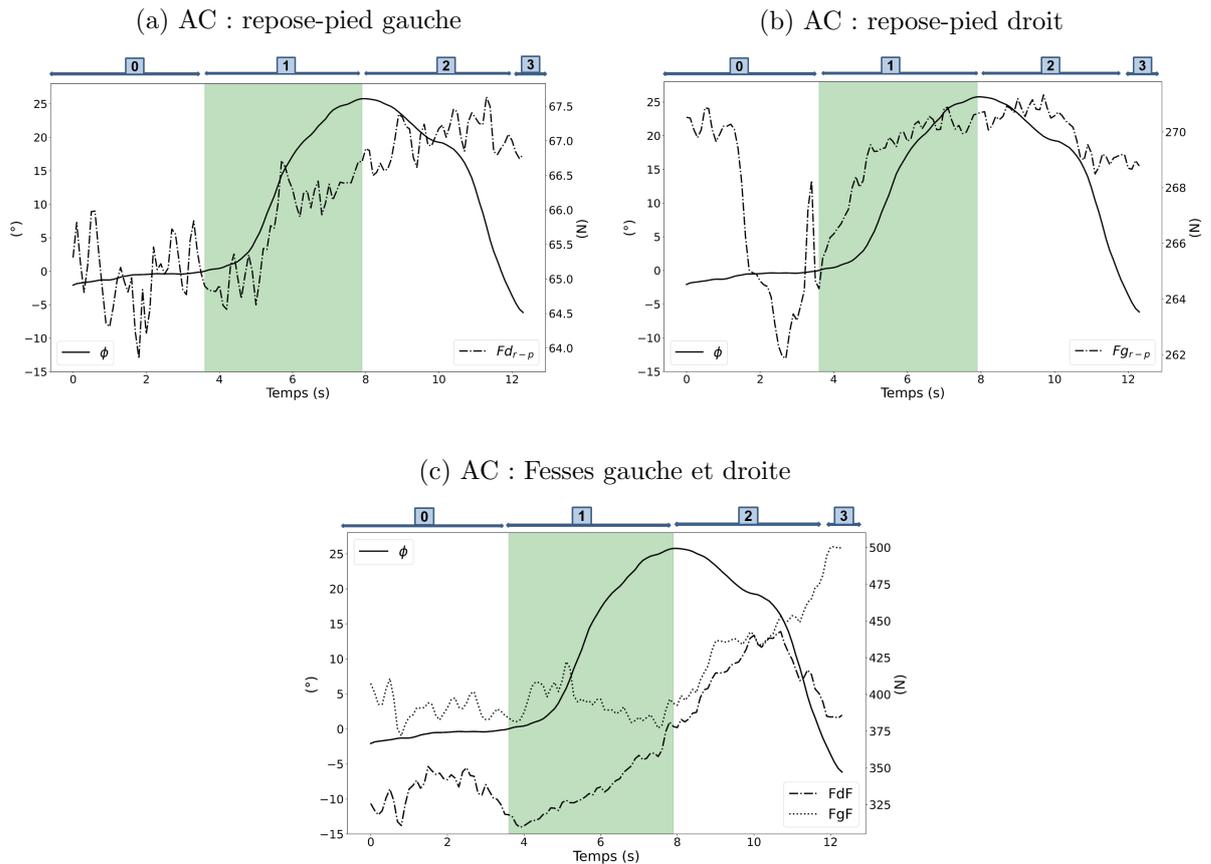
Comme le montrent les [figure 3.10a](#), [figure 3.10b](#) et [figure 3.10c](#), on observe un changement de la valeur de la force appliquée par le conducteur sur le repose-pied droit lors de la consigne CL. La variation est d'environ 274 Newton, ce qui correspond à l'instant  $t = 2$  secondes. À partir de cet instant, cette force augmente continuellement jusqu'à la fin de la **Phase 0** et le début de la **Phase 2** où elle atteint la valeur de 286 Newton. Pendant ce temps, il n'y a pas de variation significative de la force appliquée par le conducteur sur le repose-pied gauche. Nous remarquons le même comportement sur la force appliquée par la fesse droite. Cela signifie que le conducteur essaie de déplacer son centre de gravité tout en appuyant sur le repose-pied droit. Ceci peut être expliqué par le fait que les données utilisées dans cette analyse sont issues d'un virage à droite.



**Figure 3.10** : Forces des repose-pieds gauche ( $F_{gr-p}$ ) et droit ( $F_{dr-p}$ ), des fesses gauche ( $F_{gF}$ ) et droite ( $F_{dF}$ ) exprimées en Newton (N) et l'angle de roulis ( $\phi$ ) en degré (°) lors de la consigne de conduite CL.

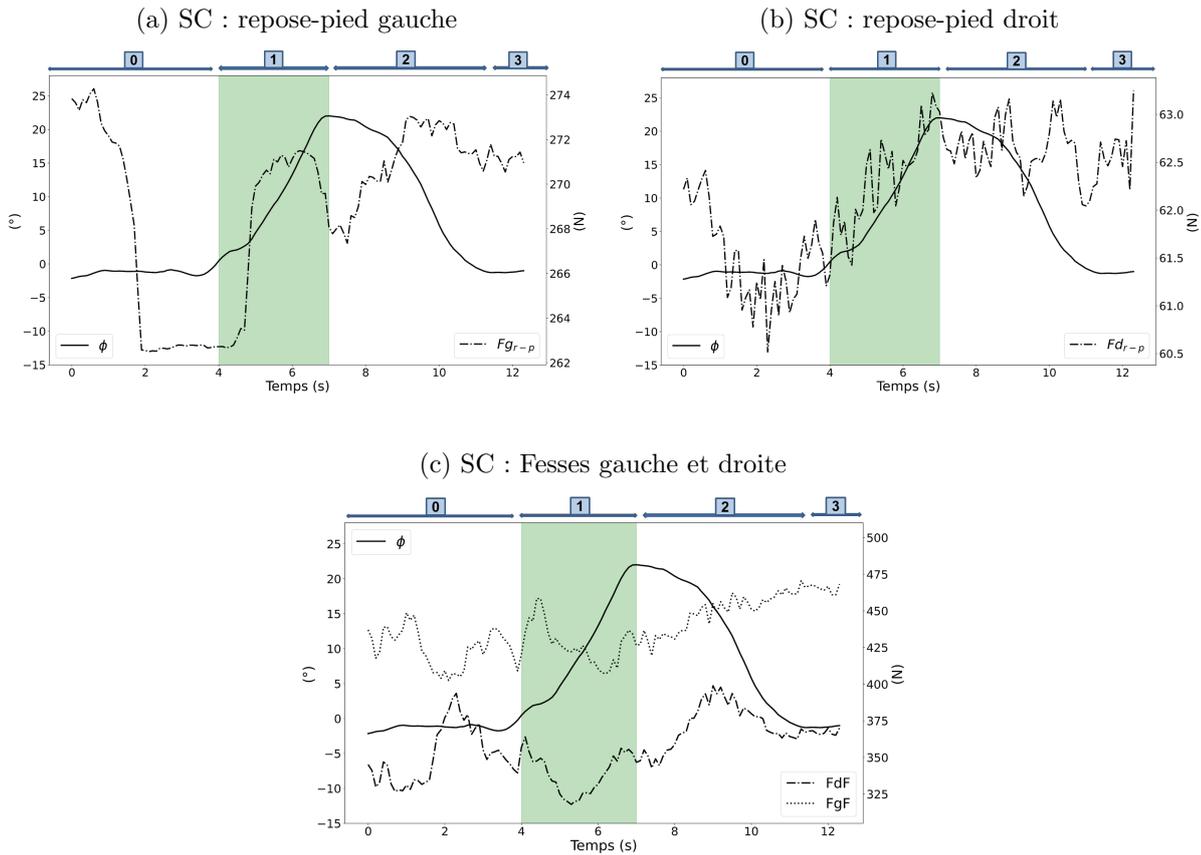
### 3.4. Application de la méthodologie proposée : Résultats et Discussions

Dans le cas de la consigne avec corps (AC), comme le montrent les [figure 3.11a](#), [figure 3.11b](#) et [figure 3.11c](#), on constate un changement de la valeur de la force appliquée par le conducteur sur le repose-pied droit à l'instant  $t = 3s$  avant le début de la **Phase 1**. À partir de ce moment, cette force augmente continuellement jusqu'à la fin de la **Phase 1** coïncidant avec le début de la **Phase 2**. En outre, il n'y a pas eu de variation importante de la force appliquée par le conducteur sur le repose-pied gauche. Par ailleurs, on remarque également le même comportement sur la force appliquée par la fesse droite. Ce qui implique que le conducteur essaie de déplacer son centre de gravité tout en appuyant sur le repose-pied droit.



**Figure 3.11** : Forces des repose-pieds gauche ( $Fg_{r-p}$ ) et droit ( $Fd_{r-p}$ ), des fesses gauche ( $FgF$ ) et droite ( $FdF$ ) exprimées en Newton (N) et de l'angle roulis ( $\phi$ ) en degré (°) lors de la consigne de conduite AC.

Dans le cas de la consigne sans corps (SC), comme le montrent [figure 3.12a](#), [figure 3.12b](#) et [figure 3.12c](#), nous n’observons aucun changement au niveau de la valeur de la force appliquée par le conducteur sur le repose-pied droit. Elle est même maintenue constante et durant la phase 1 nous n’observons aucune tendance (croissante et décroissante) concernant son évolution. Un comportement similaire est noté à l’endroit de l’évolution de la force appliquée par la fesse droite. Ainsi, il n’y a pas d’action volontaire de la part du conducteur dans cette consigne de conduite.



**Figure 3.12** : Forces des Repose-pieds gauche ( $Fg_{r-p}$ ) et droit ( $Fd_{r-p}$ ), des fesses gauche ( $FgF$ ) et droite ( $FdF$ ) exprimées en Newton (N) et de l’angle roulis ( $\phi$ ) en degré ( $^{\circ}$ ) lors de la consigne de conduite SC.

En somme, nous avons mis en évidence la relation existante entre la représentation comportementale des pratiques de conduite mesurée par les capteurs et les actions entreprises par le conducteur.

## 3.5 Conclusion

Dans ce chapitre, nous avons comme objectifs d'utiliser les données collectées lors d'une expérience réelle de conduite en virage sur piste pour :

- Analyser le comportement de différents conducteurs de 2RMs ;
- Identifier les conducteurs ayant respecté les consignes de conduite sans corps (SC) et avec corps (AC) ;
- Vérifier les conducteurs qui sont en concordances avec leurs préférences de conduite déclarées en amont lors d'un questionnaire.

Pour atteindre ces différents objectifs, nous avons développé une approche d'exploration de données fondée sur une extraction de caractéristiques (en utilisant une technique de réduction de la dimension) et de classification automatique. En utilisant l'écart-type, nous avons montré que la variabilité de l'angle du guidon est un élément discriminant dans la classification des participants, et un élément essentiel dans l'interprétation des classes obtenues par une approche non-supervisée. La classification résultante est de 2 classes qui s'opposent dans la manière dont les conducteurs interagissent avec leurs véhicules via le guidon. Par ailleurs, un participant au comportement atypique a été mis en évidence.

Enfin, nous avons entrepris une analyse approfondie des pratiques de prise de virage d'un participant. Les résultats préliminaires présentés dans ce chapitre sont prometteurs pour apporter une réponse à cette question de l'étude de l'initiation du virage.

L'analyse de la prise de virage dans ce travail considère le virage dans son intégralité et dépend d'un paramétrage des signaux se fondant sur la connaissance de la dynamique moto. Une alternative possible est de formaliser le problème comme une segmentation automatique de séries temporelles multivariées. Cette segmentation permettra une analyse différenciée et approfondie pour l'entrée, le milieu et la fin du virage. C'est ce que nous verrons dans le chapitre suivant.



---

## Segmentation et analyse de la prise de virage

---

### Sommaire

---

<b>4.1</b>	<b>Introduction</b>	<b>52</b>
<b>4.2</b>	<b>Analyse approfondie de la prise de virage dans le cas de 2RMs</b>	<b>52</b>
4.2.1	Formalisation du problème de la segmentation et analyse de la prise de virage	52
4.2.2	Segmentation de signaux multidimensionnels avec les modèles MHMMR et MRHLP	53
4.2.3	Stratégie de détection et de classification des segments de virage	57
<b>4.3</b>	<b>Application de la méthodologie proposée et résultats</b>	<b>58</b>
4.3.1	Résultats de la détection et la classification des segments de conduite dans un virage	59
4.3.2	Caractérisation statistique des segments <b>S1</b> et <b>S2</b>	60
<b>4.4</b>	<b>Conclusion</b>	<b>75</b>

---

## 4.1 Introduction

Ce chapitre s’inscrit dans la continuité du précédent. Son objectif est de présenter une méthodologie permettant d’analyser, d’une manière plus « fine », les actions entretenues par le conducteur et son véhicule lors de la prise de virage. Pour atteindre cet objectif, nous utilisons les données issues des mesures d’angle et vitesse de roulis. La première partie est consacrée au problème de segmentation de la trajectoire d’un 2RM en quatre segments : **S1**, **S2**, **S3** et **S4**, segments décrits précédemment dans le chapitre 2. Quant à la seconde partie, elle traite le problème d’étude de l’initiation « déclenchement » du virage dans le cas d’un 2RM. Du fait que ce phénomène d’initiation du virage intervient lors de la transition : **S1**  $\mapsto$  **S2**, notre analyse se focalisera sur ce passage. Dans la dernière partie du chapitre, nous présentons la mise en œuvre des différentes parties de la méthodologie pour identifier les profils de comportements de conduite dans la phase « entrée du virage ». Les données utilisées dans cette dernière partie du chapitre sont celles collectées durant la consigne libre (CL). L’utilisation de ces données se justifie par leur mise en évidence, sans biais de consigne, de la variabilité intra et inter-individuelle des différents participants à l’expérimentation.

## 4.2 Analyse approfondie de la prise de virage dans le cas de 2RMs

La méthodologie d’analyse de la prise de virage proposée dans le chapitre 3 comporte des limites. En effet, le virage est analysé dans sa globalité à travers une caractérisation reposant sur la dynamique, le positionnement sur la chaussée de la moto ainsi que les actions du conducteur. Une solution alternative consiste à effectuer, en amont, une segmentation automatique du virage en 4 segments représentant chacun les différentes phases de la trajectoire de sécurité, voir [figure 1.3](#). Dans ce chapitre, il est important de noter que nous n’utiliserons que les données correspondantes aux 7 participants parmi les 8 utilisés dans le chapitre 3. En effet, au regard des résultats obtenus dans le chapitre 3, les données du participant P3 (le stagiaire) ont été écartées de cette analyse.

### 4.2.1 Formalisation du problème de la segmentation et analyse de la prise de virage

La prise d’un virage chez un conducteur de 2RM, comme illustrée sur la [figure 1.3](#) correspondante à la trajectoire de sécurité, peut être considérée comme un changement de régime (segment). Chaque étape décrite dans la trajectoire de sécurité correspond à un régime. Lors du passage d’un segment à un autre, les pratiques mises en œuvre (vitesse, actions sur la moto avec le guidon ou le corps, emplacement sur la chaussée), lors de la prise de virage, diffèrent d’un conducteur de 2RM à un autre. En ce qui nous concerne, nous disposons des enregistrements vidéo de conduite en virage nous permettant de délimiter ces différents régimes. Cette tâche peut s’avérer trop lourde, voire impossible (dans le cas d’un nombre de séquences trop important). Elle nécessite, par ailleurs, une attention particulière et une très bonne connaissance de la conduite moto. Une alternative à cela consiste en l’utilisation de modèles de segmentation existants dans la littérature qui tiennent compte de ces changements de régimes.

Ainsi, nous formulons la détection des changements de régime comme un problème de segmentation de séries temporelles multidimensionnelles. En effet, pour un participant et une consigne de conduite donnés, chaque tour de conduite réalisé peut être considéré comme une série temporelle multivariée où, à chaque pas de temps, différentes mesures simultanées sont enregistrées.

Dans ce chapitre, les données utilisées pour la segmentation sont l'angle de roulis de la moto ( $\phi$ ) ainsi que sa vitesse ( $\dot{\phi}$  première dérivée de  $\phi$ ). Par conséquent, pour les 7 participants, nous aurons à segmenter 63 séquences de conduite grâce aux algorithmes et aux mesures  $\phi$  et  $\dot{\phi}$ . Le nombre 63 le résultat de  $7$  (*participants*)  $\times$   $3$  (*consignes*)  $\times$   $3$  (*tours*).

Pour la suite, de ce chapitre nous adopterons les notations suivantes : nous noterons par  $\mathcal{Y}$  l'ensemble des séquences de conduite défini comme suit :

$$\mathcal{Y} = \{\mathbf{Y}^m\}_{m=1,\dots,63}$$

où  $\mathbf{Y}^m = (\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_{n_m})$  est une série temporelle multivariée aux instants  $\mathbf{t} = (t_1, t_2, \dots, t_{n_m})$ . Pour chaque  $i = 1, \dots, n_m$ ,  $\mathbf{y}_i = (\mathbf{y}_i^{(\phi)}, \mathbf{y}_i^{(\dot{\phi})})^T \in \mathbb{R}^2$  est le vecteur contenant les mesures à l'instant  $t_i$  de  $\phi$  et  $\dot{\phi}$ .

À l'issue de cette étape de segmentation, nous analysons la prise de virage chez un conducteur de 2RM, qui nécessite la compréhension des différentes pratiques adoptées par ces derniers, particulièrement lors de la transition entre les différents segments. Pour capter ces différentes pratiques, nous caractérisons les différents segments à l'aide d'indicateurs statistiques, notamment la moyenne, l'écart-type, le minimum et le maximum, entre autres. Cette caractérisation statistique des segments nous permettra, grâce à un algorithme de classification comme le CAH, de regrouper les conducteurs en différentes classes selon ces caractéristiques statistiques.

#### 4.2.2 Segmentation de signaux multidimensionnels avec les modèles MHMMR et MRHLP

Dans cette sous-section, nous présentons deux modèles de segmentation de séries temporelles fondés sur une approche de régression. La segmentation de séries temporelles a été, et continue à être appliquée dans plusieurs champs d'applications. Par exemple, elle est utilisée pour la détection d'anomalies ou la détection de changements de comportements ou l'analyse d'une activité physique [25] [10].

Dans notre cas, la segmentation de séries temporelles est utilisée dans le cadre de la détection de changements de régimes (passage d'un segment à un autre) comme une première étape dans le développement d'outils d'analyse de données pour l'étude de la prise de virage dans le cas d'un 2RMs. Le développement d'outils d'analyse de données adaptés à cette problématique est une étape nécessaire pour approfondir nos connaissances sur les comportements de conduite adoptés par les conducteurs de 2RMs dans leurs conduites et leurs interactions avec leurs véhicules. Cette connaissance se révélera primordiale dans tous développements de systèmes dédiés à la sécurité

des 2RMs. Pour cela, nous avons utilisé des algorithmes de segmentation de séries temporelles fondés sur une approche de régression et prenant en compte le changement de régime comme « Multiple Hidden Markov Model Regression model » (MHMMR) [22] et « Multiple Regression Hidden Logistic Process (MRHLP) » [12].

Dans les paragraphes suivants, nous allons décrire ces modèles en posant  $\mathbf{Y} = (\mathbf{y}_1, \dots, \mathbf{y}_n)$  une série temporelle multivariée observée à des instants  $\mathbf{t} = (t_1, \dots, t_n)$  avec  $\mathbf{y}_i = (\mathbf{y}_i^{(1)}, \dots, \mathbf{y}_i^{(d)})^T \in \mathbb{R}^d$ ,  $i = 1, \dots, n$ .

#### 4.2.2.1 Présentation des modèles MHMMR et MRHLP pour la segmentation de série temporelle multivariée

Les modèles MHMMR et MRHLP sont adaptés pour une segmentation de séries temporelles multivariées avec un changement de régimes. En effet, dans ces modèles, chaque régime est associé à un modèle de régression polynomiale et le passage d'un modèle de régression à un autre est régi par un processus logistique caché. Par ailleurs, les modèles MHMMR et MRHLP constituent tout simplement les généralisations des cas univariés (une variable  $d = 1$ ) des modèles « Hidden Markov Model Regression model (HMMR) » et « Regression Hidden Logistic Process (RHLP) » proposés dans [23] et [12], respectivement.

Pour chaque  $i = 1, \dots, n$ , les modèles MHMMR et MRHLP sont décrits comme suit :

$$\begin{aligned} y_i^{(1)} &= \beta_{z_i}^{(1)T} \mathbf{t}_i + \sigma_{z_i}^{(1)} \epsilon_i \\ y_i^{(2)} &= \beta_{z_i}^{(2)T} \mathbf{t}_i + \sigma_{z_i}^{(2)} \epsilon_i \\ &\vdots \\ y_i^{(d)} &= \beta_{z_i}^{(d)T} \mathbf{t}_i + \sigma_{z_i}^{(d)} \epsilon_i \end{aligned}$$

où pour chaque  $j = 1, \dots, d$ , le vecteur  $\beta_{z_i}^{(j)T} = (\beta_{z_i 0}^{(j)}, \dots, \beta_{z_i p}^{(j)})$  regroupe les coefficients de la régression polynomiale d'ordre  $p$ ,  $\mathbf{t}_i = (1, t_i, t_i^2, \dots, t_i^p)^T$  est un vecteur de covariables de dimension  $p + 1$  à l'instant  $t_i$  et  $\epsilon_j \sim \mathcal{N}(0, 1)$  représentant le bruit additif aux modèles.

La différence entre les modèles MHMMR et MRHLP, réside au niveau de la nature de la distribution du processus caché contrôlant le passage d'un modèle de régression à un autre. Ce processus caché associé aux observations  $(\mathbf{y}_1, \dots, \mathbf{y}_n)$  est noté  $\mathbf{z} = (z_1, \dots, z_n)$  où la variable  $z_i \in \{1, \dots, K\}$ , pour  $i = 1, \dots, n$ , est une valeur discrète cachée qui contrôle le passage d'un segment (associé à un modèle de régression polynomiale) à un autre parmi les  $K$  segments à l'instant  $t_i$  [62]. La description de la nature de la distribution du processus caché de chacun des modèles MHMMR et MRHLP est donnée, ci-après :

- **Dans le modèle MHMMR**, pour chacun des modèles HMMR  $y_i^{(j)} = \beta_{z_i}^{(j)T} \mathbf{t}_i + \sigma_{z_i}^{(j)}$ ,  $j =$

$1, \dots, d$ , la distribution du processus caché  $\mathbf{z} = (z_1, \dots, z_n)$  est une chaîne de Markov homogène de premier ordre [7]. Dans un processus de Markov simple, les principaux paramètres à estimer sont la loi de la distribution initiale  $\boldsymbol{\pi}$  et la matrice de transition  $\mathbf{A}$  ;

- **Dans le modèle MRHLP**, pour chacun des modèles RHLP  $y_i^{(j)} = \beta_{z_i}^{(j)T} \mathbf{t}_i + \sigma_{z_i}^{(j)}$ ,  $j = 1, \dots, d$ , on suppose que  $\mathbf{z} = (z_1, \dots, z_n)$  est un processus logistique caché [12]. Dans un processus logistique, pour un vecteur temps donné  $\mathbf{t} = (t_1, \dots, t_n)$ , les variables  $z_i$  ( $i = 1, \dots, n$ ) sont générées indépendamment suivant une distribution multinomiale  $\mathcal{M}(1, \pi_1(t_i; \mathbf{w}), \dots, \pi_K(t_i; \mathbf{w}))$  où :

$$\pi_k(t_i; \mathbf{w}) = p(z_i = k | t_i; \mathbf{w}) = \frac{\exp(\mathbf{w}_k^T \mathbf{v}_i)}{\sum_{l=1}^K \exp(\mathbf{w}_l^T \mathbf{v}_i)}$$

est la transformation logistique d'une fonction linéaire dépendante du vecteur de covariables  $\mathbf{v}_i = (1, t_i, \dots, t_i^u)^T$ ,  $\mathbf{w}_k = (w_{k0}, w_{k1}, \dots, w_{kq})^T$  est un vecteur (de dimension  $u + 1$ ) de coefficients associé aux  $\mathbf{v}_i$  and  $\mathbf{w} = (\mathbf{w}_1, \dots, \mathbf{w}_K)$ . Dans [12], les auteurs ont démontré que le processus logistique est flexible. Ceci constitue un atout majeur pour la détection de changements de régimes brusques ou lisses.

Les modèles MHMMR et MRHLP, présentés ci-dessus, peuvent être écrits dans une forme compacte (plus lisible et compréhensible) à l'aide de matrices comme suit :

$$\underbrace{\begin{bmatrix} y_i^{(1)} \\ y_i^{(2)} \\ \vdots \\ y_i^{(d)} \end{bmatrix}}_{\mathbf{y}_i} = \underbrace{\begin{bmatrix} \beta_{z_i 0}^{(1)} & \dots & \beta_{z_i p}^{(1)} \\ \beta_{z_i 0}^{(2)} & \dots & \beta_{z_i p}^{(2)} \\ \vdots & \dots & \vdots \\ \beta_{z_i 0}^{(d)} & \dots & \beta_{z_i p}^{(d)} \end{bmatrix}}_{\mathbf{B}_{z_i}^T} \underbrace{\begin{bmatrix} 1 \\ t_i \\ \vdots \\ t_i^{(p)} \end{bmatrix}}_{\mathbf{t}_i} + \underbrace{\begin{bmatrix} \sigma_{z_i}^{(1)} \epsilon_i \\ \sigma_{z_i}^{(2)} \epsilon_i \\ \vdots \\ \sigma_{z_i}^{(d)} \epsilon_i \end{bmatrix}}_{\mathbf{e}_i \sim \mathcal{N}(\mathbf{0}; \boldsymbol{\Sigma}_{z_i})}$$

où  $\mathbf{y}_i = (\mathbf{y}_i^{(1)}, \dots, \mathbf{y}_i^{(d)})^T \in \mathbb{R}^d$ , ( $i = 1, \dots, n$ ) est la  $i$ -ème observation de la série temporelle multivariée,  $\mathbf{B}_{z_i}^T$  est une matrice de dimension  $d \times (p + 1)$  regroupant les paramètres du modèle de régression multiple associé à la classe  $z_i = k$  et  $\boldsymbol{\Sigma}_{z_i}$  est une matrice de covariance.

#### 4.2.2.2 Estimation des paramètres dans les modèles MHMMR et MRHLP

Dans les modèles MHMMR et MRHLP, la distribution de probabilité  $p(\mathbf{y}_i | t_i)$  des observations  $\mathbf{y}_i$  à la date  $t_i$  dépend des vecteurs de paramètres inconnus  $\boldsymbol{\theta}_{\text{MHMMR}} = (\boldsymbol{\pi}, \mathbf{A}, \mathbf{B}_1, \dots, \mathbf{B}_K, \boldsymbol{\Sigma}_1, \dots, \boldsymbol{\Sigma}_K)$  et  $\boldsymbol{\theta}_{\text{MRHLP}} = (\mathbf{w}, \mathbf{B}_1, \dots, \mathbf{B}_K, \boldsymbol{\Sigma}_1, \dots, \boldsymbol{\Sigma}_K)$ . Ces paramètres inconnus peuvent être estimés en utilisant la méthode d'estimation du maximum de vraisemblance [45]. L'idée de base de la méthode d'estimation du maximum de vraisemblance est d'estimer les paramètres  $\boldsymbol{\theta}_{\text{MHMMR}}$  et  $\boldsymbol{\theta}_{\text{MRHLP}}$  maximisant les fonctions de vraisemblance permettant d'obtenir les données observées

$\mathbf{y}_i$ . Les fonctions de vraisemblance des modèles MHMMR et MRHLP sont détaillées, respectivement, dans [62] et [11].

En pratique, la fonction de vraisemblance est de nature complexe. Par conséquent, sa maximisation est une tâche extrêmement difficile à réaliser. Comme alternative, la maximisation est effectuée sur le logarithme de la fonction de vraisemblance (log-vraisemblance). Les fonctions de log-vraisemblance des modèles MHMMR et MRHLP de données observées sont définies de la manière suivante :

$$\mathcal{L}(\boldsymbol{\theta}_{\text{MHMMR}}; \mathbf{Y}, \mathbf{z}, \mathbf{t}) = \log \sum_{\mathbf{z}} p(z_1; \pi) \prod_{i=2}^n p(z_i | z_{i-1}; \mathbf{A}) \prod_{i=1}^n \mathcal{N}(\mathbf{y}_i; \mathbf{B}_{z_i} \mathbf{t}_i, \Sigma_{z_i})$$

$$\mathcal{L}(\boldsymbol{\theta}_{\text{MRHLP}}; \mathbf{Y}, \mathbf{z}, \mathbf{t}) = \sum_{i=1}^n \sum_{k=1}^K (\mathbb{1}_{z_i=k}) \log[\pi_k(t_i; \mathbf{w}) \mathcal{N}(\mathbf{y}_i; \mathbf{B}_k^T \mathbf{t}_i, \Sigma_k)]$$

Pour estimer ces fonctions de log-vraisemblance, l'algorithme le plus couramment utilisé est l'algorithme d'Espérance-Maximisation (EM). EM est un algorithme itératif d'estimation paramétrique qui a été principalement utilisé pour les modèles de mélanges et les modèles de Markov cachés (HMM).

Dans les lignes qui suivent, nous allons décrire les deux principales étapes de l'algorithme EM pour les modèles MHMMR et MRHLP. Notons par  $\boldsymbol{\theta} \in \{\boldsymbol{\theta}_{\text{MHMMR}}; \boldsymbol{\theta}_{\text{MRHLP}}\}$  le paramètre commun à estimer.

L'algorithme EM commence avec un paramètre initial  $\boldsymbol{\theta}^0$  et répète les étapes E et M suivantes jusqu'à la convergence :

- **Étape E (Espérance)** : Dans cette étape, on calcule l'espérance de la log-vraisemblance conditionnellement aux données observées  $\mathbf{Y}$ , au temps  $\mathbf{t}$  et la valeur courante du paramètre à estimer  $\boldsymbol{\theta}^{(q)}$  :

$$Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(q)}) = \mathbb{E} \left[ \mathcal{L}(\boldsymbol{\theta}; \mathbf{Y}, \mathbf{z}, \mathbf{t}) | \mathbf{Y}, \mathbf{t}; \boldsymbol{\theta}^{(q)} \right] \quad (4.1)$$

Dans le cas où  $\boldsymbol{\theta} = \boldsymbol{\theta}_{\text{MHMMR}}$ , dans [62], les auteurs ont montré que le calcul de l'espérance conditionnelle définie dans l'équation 4.1 est lié au calcul en amont des probabilités a posteriori et des probabilités a posteriori conjointes. Le détail de ces calculs est clairement exposé dans [62]. Quant au cas où  $\boldsymbol{\theta} = \boldsymbol{\theta}_{\text{MRHLP}}$ , le calcul de l'espérance conditionnelle précédemment définie nécessite uniquement le calcul des probabilités a posteriori dont on peut retrouver le détail dans [11].

- **Étape M (Maximisation)** : Dans cette étape, la valeur du paramètre  $\theta$  est mise à jour en calculant le paramètre  $\theta^{(q+1)}$  maximisant l'espérance conditionnelle  $Q$  par rapport à  $\theta$  :

$$\theta^{(q+1)} = \arg \max_{\theta} Q(\theta, \theta^{(q)}) \quad (4.2)$$

Les détails de cette étape de maximisation pour chacun des paramètres  $\theta_{\text{MHMMR}}$  et  $\theta_{\text{MRHLP}}$  peuvent respectivement être consultés dans [62] et [11].

### 4.2.3 Stratégie de détection et de classification des segments de virage

Pour atteindre notre but, nous avons mis en place une méthodologie en deux étapes : une première qui consiste à effectuer une segmentation de chacune des séquences de conduite réalisée dans l'expérimentation en 4 segments, et une seconde, où l'on réalise une classification des segments obtenus.

#### 4.2.3.1 Détection des segments dans un virage

Comme énoncé précédemment, la segmentation d'une séquence de conduite en virage est formulée comme un problème de segmentation multidimensionnelle. En effet, la prise de virage, lors de la conduite en 2RM est divisée en 4 segments de conduite où le passage entre 2 segments consécutifs est le résultat des interactions du triptyque Véhicule-Infrastructure-Conducteur (V-I-C), car la spécificité principale de la conduite des 2RMs est la recherche continue de l'équilibre et de la stabilité de la trajectoire. Chaque segment représente une zone de stabilité et le conducteur doit rompre cet équilibre pour passer d'un segment à un autre, tout en maintenant la stabilité de sa trajectoire pendant ces moments de transitions. Ce qui place cette problématique de recherche dans le contexte de détection de changements de régimes « lisses ». En ce sens, les modèles MHMMR et MRHLP sont adaptés pour « délimiter » ces 4 segments de conduite. Ces modèles de segmentation ont été appliqués, indépendamment, sur chacune des 63 séquences de conduite en virage de l'ensemble  $\mathcal{Y}$ .

Comme c'est le cas pour tout modèle, il existe des paramètres à définir ou à optimiser le cas échéant. Dans les modèles MHMMR et MRHLP, il existe deux paramètres essentiels à fixer, notamment le nombre de segments  $K$  et  $p$  le degré du polynôme dans la régression polynomiale. Le nombre de segments  $K$  est fixé à 4 en se fondant sur les connaissances « métier » des différentes phases de la prise d'un virage en 2RM. Quant au paramètre  $p$ , un entier fixé entre 2 et 4, il sera défini en utilisant le critère d'information Bayésien [68].

Afin d'évaluer les performances des modèles MHMMR et MRHLP, nous avons comparé les résultats de segmentation de chacun de ces deux modèles à ceux proposés par l'expert (segmentation à travers les vidéos de séquences de conduite), et ce en utilisant le score de « l'indice de Rand » défini dans [63]. Plus proche est la segmentation automatique de la segmentation de

l'expert, meilleure est la qualité de la segmentation. Cet indice varie dans un intervalle de  $[0, 1]$ , plus élevée est la valeur de cette indice meilleure est la qualité de la segmentation. En d'autres termes, cette évaluation nous permet de déterminer lequel de ces modèles s'approche le mieux de la segmentation manuelle proposée par l'expert.

Par ailleurs, afin de mettre en évidence l'importance de l'approche de régression dans les modèles MHMMR et MRHLP, nous les avons comparés avec un autre modèle plus simple qui est le modèle de Markov cachés « Hidden Markov Model (HMM) ». Pour cela, nous avons utilisé le modèle HMM droite-gauche décrit dans [2] avec des émissions de probabilité suivant une distribution Gaussienne. Ainsi, pour différencier ce modèle avec la version standard, nous le noterons GHMM pour « Gaussian Hidden Markov Model ». Comme les modèles MHMMR et MRHLP, l'un des paramètres essentiels de GHMM, est le type de la matrice de covariance, qui a été optimisé suivant le critère de l'information Bayésienne.

### 4.3 Application de la méthodologie proposée et résultats

Dans cette partie du chapitre, nous allons appliquer la méthodologie proposée sur l'ensemble  $\mathcal{Y}$  comportant les 63 séquences de conduite en virage. Le résumé de cette méthodologie est représentée sur la figure 4.1.

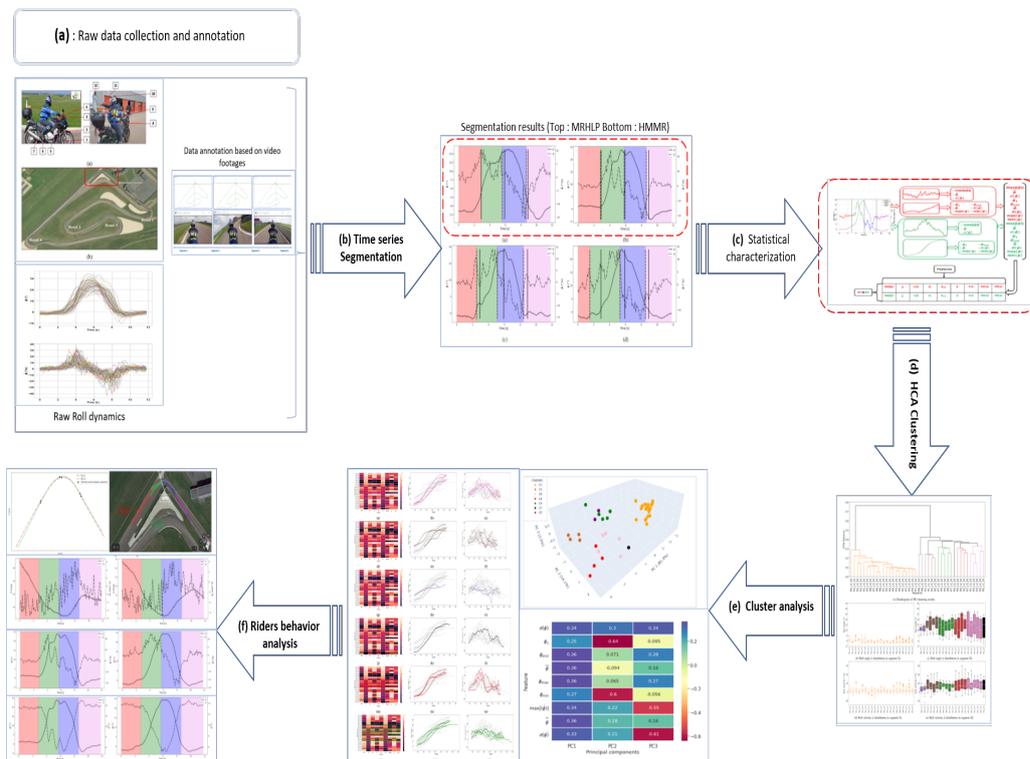


Figure 4.1 : Résumé de la méthodologie de segmentation et d'analyse de la prise de virage.

### 4.3.1 Résultats de la détection et la classification des segments de conduite dans un virage

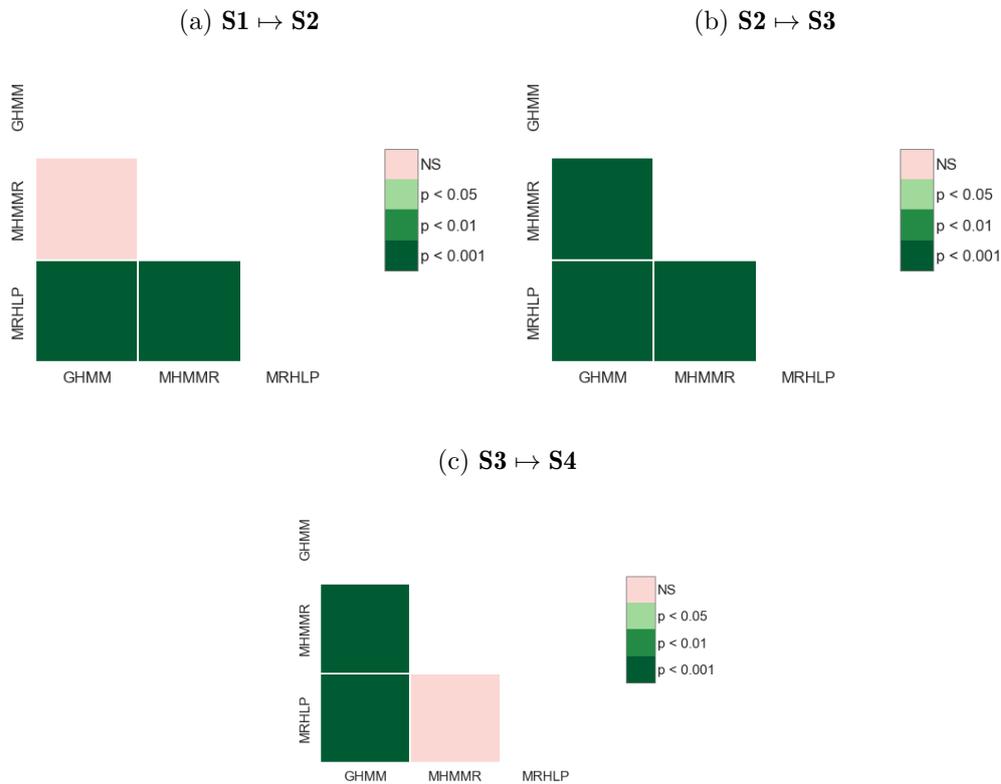
Dans cette sous-section, nous présentons les résultats de l'application des modèles MHMMR, MRHLP et GHMM sur chacune des séquences  $\mathbf{Y}^m$  de l'ensemble  $\mathcal{Y}$ . Pour avoir une performance globale des différents modèles, nous avons calculé la moyenne de l'indice de Rand sur chacune des 63 séquences lors des différentes transitions entre les 4 segments ( $\mathbf{S1} \mapsto \mathbf{S2}$ ,  $\mathbf{S2} \mapsto \mathbf{S3}$  et  $\mathbf{S3} \mapsto \mathbf{S4}$ ). Pour une séquence donnée, l'indice de Rand représente la similarité moyenne entre les résultats de la segmentation d'un modèle (entre MRHLP, MHMMR et GHMM) et ceux proposés par l'expert grâce aux enregistrements vidéo. Intuitivement, il est égal au nombre de points concordants entre deux résultats de segmentation. Le score de l'indice de Rand est normalisé entre 0 (parfaite discordance) et 1 (parfaite concordance).

	$\mathbf{S1} \mapsto \mathbf{S2}$	$\mathbf{S2} \mapsto \mathbf{S3}$	$\mathbf{S3} \mapsto \mathbf{S4}$
MRHLP	<b>0.969</b> ( $\pm 0.029$ )	<b>0.956</b> ( $\pm 0.028$ )	<b>0.947</b> ( $\pm 0.029$ )
MHMMR	0.934( $\pm 0.052$ )	0.899( $\pm 0.069$ )	0.932( $\pm 0.041$ )
GHMM	0.904( $\pm 0.087$ )	0.767( $\pm 0.092$ )	0.857( $\pm 0.107$ )

**Tableau 4.1** : Score moyen de la métrique « indice de Rand » sur les 63 séquences de conduite en virage pour les modèles MRHLP, MHMMR et GHMM.

Le [tableau 4.1](#) donne la moyenne (et l'écart-type), sur les 63 séquences segmentées, la métrique « indice de Rand » des modèles de segmentation MRHLP, MHMMR et GHMM lors des différentes transitions entre les segments  $\mathbf{S1} \mapsto \mathbf{S2}$ ,  $\mathbf{S2} \mapsto \mathbf{S3}$  et  $\mathbf{S3} \mapsto \mathbf{S4}$ . Nous pouvons y constater que les modèles avec une approche de régression (MRHLP et MHMMR) obtiennent les meilleurs résultats caractérisés par des moyennes se rapprochant de 1, avec des écart-types faibles sur les différentes transitions. En particulier, le modèle MRHLP, au vu des résultats, est considéré comme étant le modèle reproduisant une segmentation la plus proche de celle proposée par l'expert. Afin de vérifier si la différence, entre la distribution des scores moyens des différents modèles, est significative nous avons utilisé un test non paramétrique à deux hypothèses, en l'occurrence celui de Mann-Whitney [43]. Dans la première hypothèse ( $H_0$ ) on suppose qu'il n'y a pas de différence significative entre les scores moyens sur chacune des transitions. Et dans la seconde hypothèse ( $H_1$ ) (hypothèse alternative), on admet que la différence entre les scores moyens est significative. En utilisant ce test, nous avons calculé les p-valeurs lors de chacune des transitions entre les modèles de segmentation pris deux à deux, voir [figure 4.2](#). Une p-valeur inférieure aux seuils standards fixés à 5%, 1% et 0,1% est synonyme d'un rejet de l'hypothèse nulle au profit de l'hypothèse alternative. Par conséquent, la [figure 4.2](#) nous montre que les scores moyens obtenus précédemment ([tableau 4.1](#)) sont significativement différents pour la plupart. Cependant, lors de la transition  $\mathbf{S3} \mapsto \mathbf{S4}$ , la différence de scores entre les modèles MRHLP et

MHMMR n'est pas significative (p-valeur=0,08).



**Figure 4.2** : P-valeurs du test statistique de Mann-Whitney entre les scores moyens des modèles MRHLP, MHMMR et GHMM sur chacune des transitions : (a) **S1**  $\mapsto$  **S2**, (b) **S2**  $\mapsto$  **S3** et (c) **S3**  $\mapsto$  **S4**.

### 4.3.2 Caractérisation statistique des segments S1 et S2

La sous-section précédente avait pour but de trouver un modèle de segmentation s'approchant au mieux de celui proposé par l'expert. Dans cette partie de la thèse, nous utilisons ces résultats en choisissant parmi les modèles celui ayant le plus grand score global de l'indice de Rand. Ainsi, nous utilisons la segmentation avec le modèle MRHLP afin de faire une caractérisation statistique des segments S1 et S2. En effet, rappelons que dans ce chapitre, nous nous intéressons aux comportements des conducteurs de 2RMs lors de la prise de virage, et notamment, lors de l'initiation du virage (transition **S1**  $\mapsto$  **S2**) pour la consigne de conduite naturelle (CL).

#### 4.3.2.1 Indicateurs statistiques pour la caractérisation des segments S1 et S2

La manœuvre de la prise d'un virage sécuritaire peut être décomposée en deux phases de conduite. Une première phase a lieu en ligne droite et une seconde lors de l'amorçage de la courbe. Ces deux phases sont décrites comme suit :

- **Conduite en ligne droite**, le segment **S1** est caractérisé par un angle de roulis avoisinant 0 et de légères variations dynamiques au niveau de la vitesse (du roulis). En d'autres termes, cela signifie que, sur le segment **S1**, le conducteur essaie de stabiliser la moto autour une position d'équilibre (position verticale).
- **Amorçage de la courbe**, dans cette phase représentée par le segment **S2**, la moto commence à rouler dans la direction du virage. L'angle de roulis augmente lentement jusqu'à atteindre son maximum. Ce segment est caractérisé par des valeurs d'angle de roulis qui s'écartent significativement de 0. Ce segment est aussi caractérisé par des valeurs de la vitesse du roulis élevées autour de 0 mais à une fréquence faible.

Dans ce chapitre, la décomposition de la manœuvre de la prise de virage est faite en se fondant essentiellement sur la dynamique du roulis ( $\phi, \dot{\phi}$ ). Nous posons l'hypothèse selon laquelle il existe des seuils au niveau des mesures d'angle et de vitesse du roulis pouvant indiquer les intentions du conducteur lors d'un changement d'orientation de la moto. Ainsi, dans ce chapitre, nous allons essayer de délimiter ces seuils en utilisant des indicateurs statistiques dont le choix est fondé sur les travaux effectués dans [42]. La description de ces indicateurs statistiques, sur chaque segment, est représentée dans le [tableau 4.2](#) :

Indicateurs Statistiques	Description
$\phi_1$	Première valeur observée de l'angle roulis $\phi$
$\phi_d$	Dernière valeur observée de l'angle de roulis $\phi$
$\bar{\phi}$	Moyenne de l'angle roulis $\phi$
$\sigma(\phi)$	Écart-type de l'angle roulis $\phi$
$\min(\phi)$	Valeur minimale de l'angle roulis $\phi$
$\max(\phi)$	Valeur maximale de l'angle roulis $\phi$
$\max \dot{\phi} $	Valeur maximale de la valeur absolue de la vitesse de roulis ( $\dot{\phi}$ )
$\bar{\dot{\phi}}$	Moyenne de la vitesse du roulis $\dot{\phi}$
$\sigma(\dot{\phi})$	Écart-type de la vitesse du roulis $\dot{\phi}$

**Tableau 4.2** : Description des différents indicateurs statistiques.

La [figure 4.3](#) illustre le processus décrivant l'extraction des indicateurs statistiques à partir des segments S1 et S2 pour une séquence de conduite  $\mathbf{Y}^m$ . La séquence de conduite  $\mathbf{Y}^m$  correspond à un tour de conduite réalisé par un participant dans une consigne de conduite donnée.

En répétant le processus d'extraction des indicateurs statistiques sur les 63 séquences de conduite, nous arrivons à construire une base de données sous la forme d'une matrice de taille  $126 \times 9$ . Le nombre 126 est la résultante de la multiplication entre 63 (nombre total de séquences) et 2 (nombre de lignes par séquence). Ainsi, nous noterons par  $\mathcal{X}$  cette matrice qui est une union de trois sous-matrices correspondantes à chacune des consignes de conduite CL, SC et

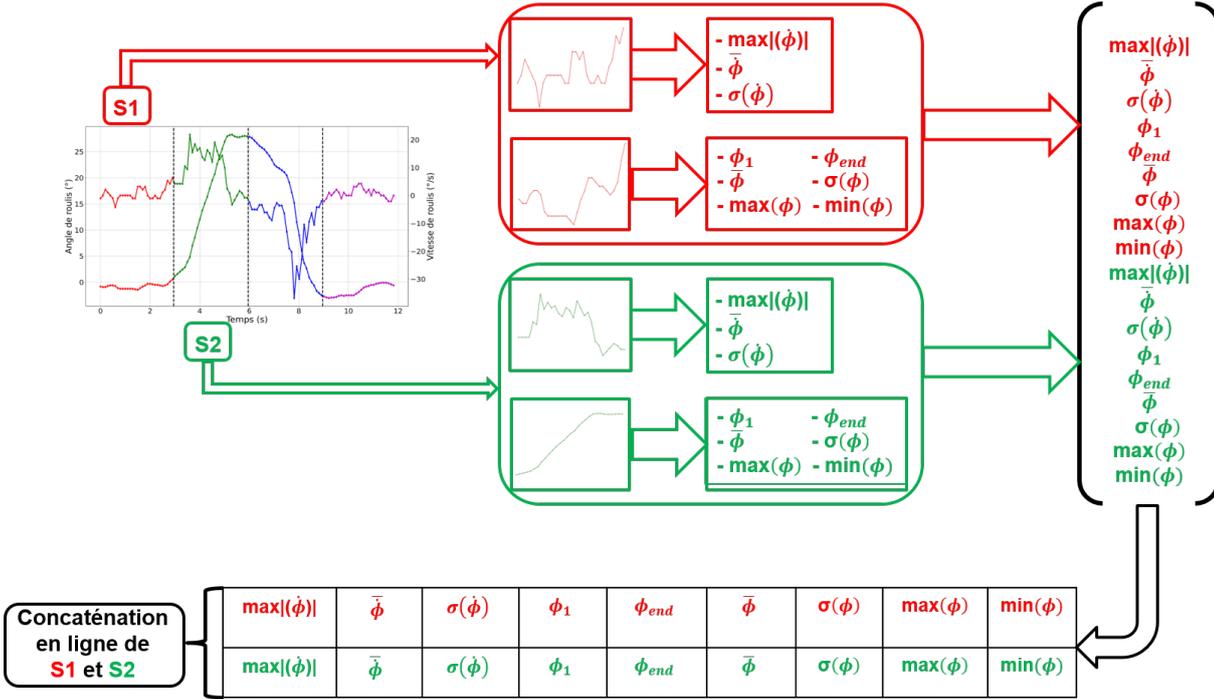


Figure 4.3 : Un exemple de caractérisation des segments S1 (rouge) et S2 (verte) pour une séquence donnée  $\mathbf{Y}^m$ .

AC. Concrètement, la matrice  $\mathcal{X}$  s'écrit sous la forme suivante :

$$\mathcal{X} = \mathcal{X}_{CL} \cup \mathcal{X}_{SC} \cup \mathcal{X}_{AC}$$

Rappelons que dans ce chapitre, nous nous intéressons uniquement à l'analyse de la sous-matrice  $\mathcal{X}_{CL}$ . Les autres sous-matrices  $\mathcal{X}_{SC}$  et  $\mathcal{X}_{AC}$  comportent un biais de consigne. En effet, les libellés de ces consignes sont, dans la pratique, favorables à plusieurs interprétations chez les conducteurs. Dans des soucis de se défaire du biais de consigne, nous n'utiliserons que la sous-matrice correspondante à la consigne CL où il s'agit d'effectuer une conduite naturelle.

Pour analyser la sous-matrice d'indicateurs  $\mathcal{X}_{CL}$ , nous utilisons une technique de classification, cette méthode nous permettra de mettre en évidence les différences entre les pratiques mises en œuvre par les conducteurs de 2RMs lors de la transition  $S1 \mapsto S2$ . Pour ce faire, nous avons, dans la littérature, le choix entre plusieurs méthodes comme la méthode des K-moyennes [40], DBSCAN ("Density-Based Spatial Clustering of Applications with Noise") [20], GMM ("Gaussian Mixture Model") [27] et le CAH [72]. Cependant, dans cette étude notre choix s'est porté sur la classification ascendante hiérarchique (CAH), car c'est une méthode de classification qui présente les avantages suivants :

- On travaille à partir des dissimilarités entre les objets que l'on veut regrouper. On peut donc choisir un type de dissimilarité ou similarité adapté au sujet étudié et à la nature

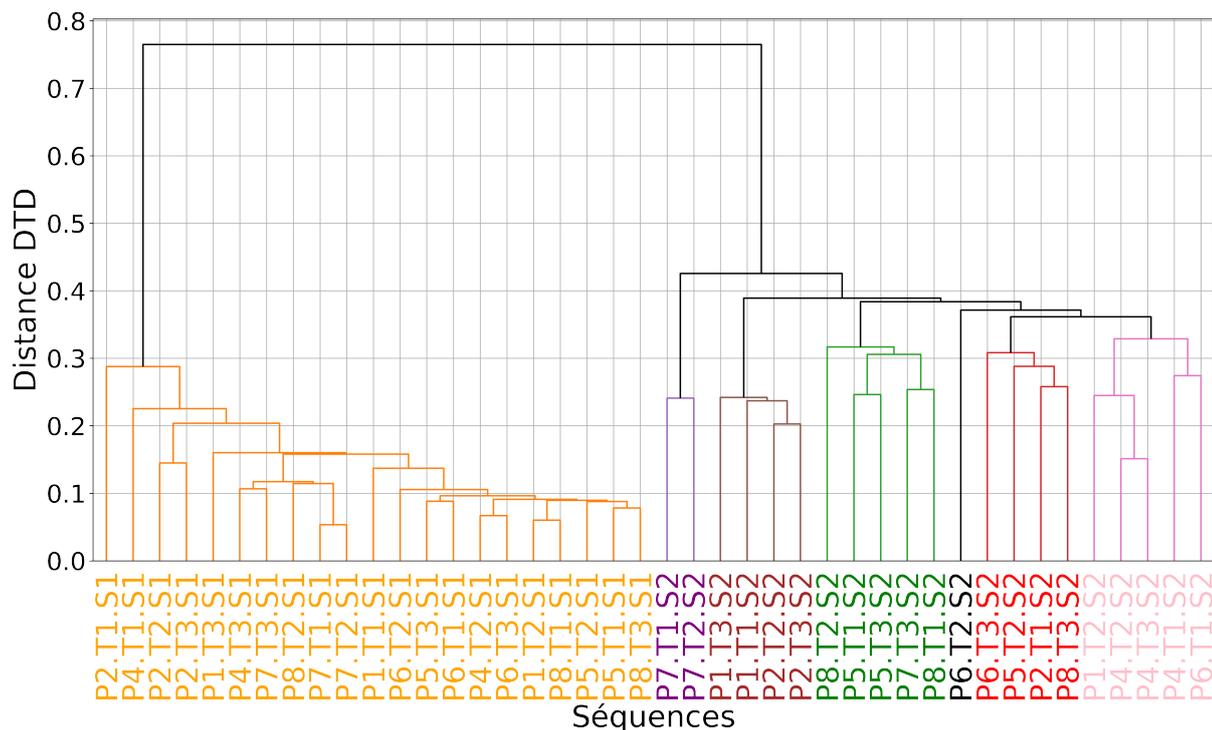
des données. Dans le cas de cette étude, les méthodes de similarité privilégiées sont celles permettant de mesurer le degré de similitude entre des séries temporelles multidimensionnelles, la Déformation Temporelle Dynamique (DTD), en est un parfait exemple ou du moins, la plus utilisée.

- L'un des résultats est le dendrogramme, qui permet de visualiser le regroupement progressif des données. On peut alors se faire une idée d'un nombre adéquat de classes dans lesquelles les données peuvent être regroupées. Il est à rappeler que l'un des objectifs opérationnels de cette étude est de développer un outil d'aide à l'analyse de données à destination de chercheurs dans le domaine des transports, ces derniers ne disposant pas forcément d'information à priori sur la structuration des données à analyser. Dans ce contexte, une méthode disposant d'un outil de visualisation des résultats de la classification pourrait se révéler fort utile.

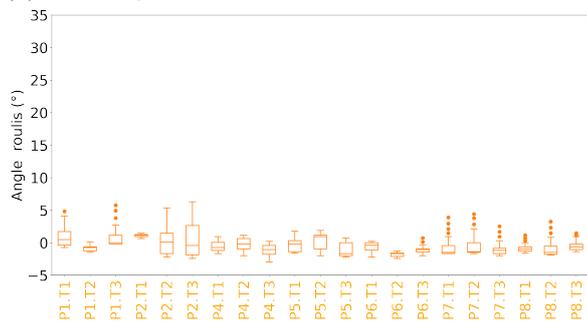
La même CAH utilisée dans le chapitre 3 sera reconduite dans ce chapitre avec comme distance la DTD.

#### 4.3.2.2 Analyse des résultats de classification par rapport aux indicateurs statistiques extraits de S1 et S2

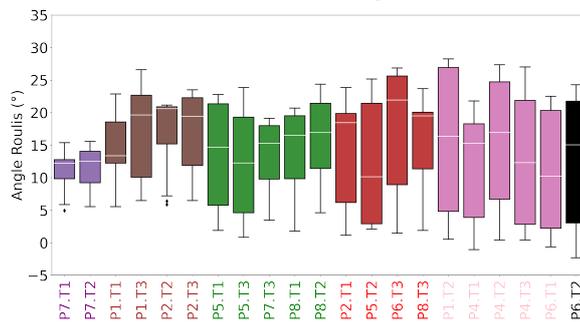
En appliquant la CAH sur la sous-matrice  $\mathcal{X}_{CL}$ , on produit le dendrogramme illustré sur la [figure 4.4a](#). De ce dendrogramme, nous pouvons clairement identifier deux grandes classes. La première et la deuxième classe regroupant des éléments des segments S1 et S2, respectivement. Cette classification en deux classes est une évidence dans la mesure où, par rapport aux caractéristiques statistiques choisies, les comportements de conduite adoptés sur la phase S1 diffèrent naturellement de ceux pratiqués sur S2. Ainsi, il serait plus intéressant d'effectuer, grâce au dendrogramme, une classification en sous-classes de ces deux principales classes. En analysant le dendrogramme, on peut voir que les sous-classes sont difficilement identifiables dans la classe correspondante au segment S1 contrairement à celle du segment S2. Ceci peut simplement s'expliquer par le fait que les participants adoptent sensiblement les mêmes comportements par rapport à la dynamique du mouvement de roulis de la moto sur le segment S1. En effet, sur ce segment les conducteurs de 2RMs maintiennent, le long de l'axe de déplacement, leur véhicule ( $\phi = 0$ ) tout en assurant la stabilité de ce dernier. Cette stabilité maintenue le long de l'axe de déplacement se caractérise par une variation presque nulle de la vitesse du roulis de la moto ( $\dot{\phi}$ ). Ces résultats se confirment à travers les boîtes à moustaches représentant les dynamiques du mouvement de roulis de la moto, voir les [figure 4.4b](#) et [figure 4.4d](#).



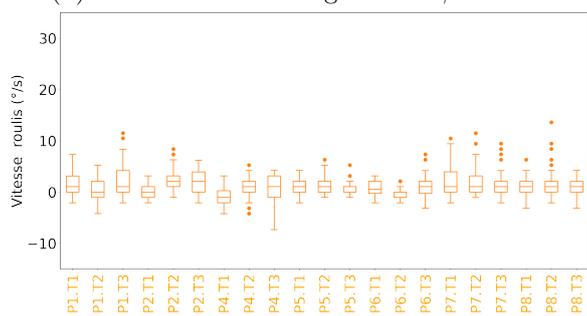
(a) Dendrogramme des résultats de la classification des participants lors de la consigne de conduite CL.



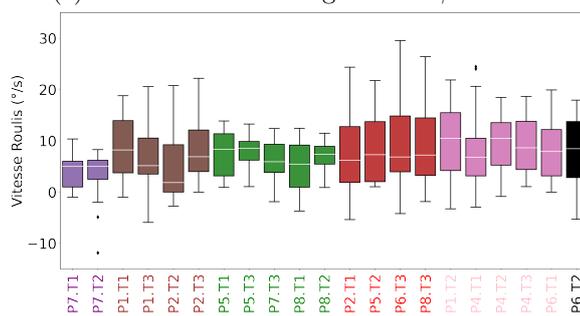
(b) Distribution de l'angle roulis  $\phi$  dans S1.



(c) Distribution de l'angle roulis  $\phi$  dans S2.



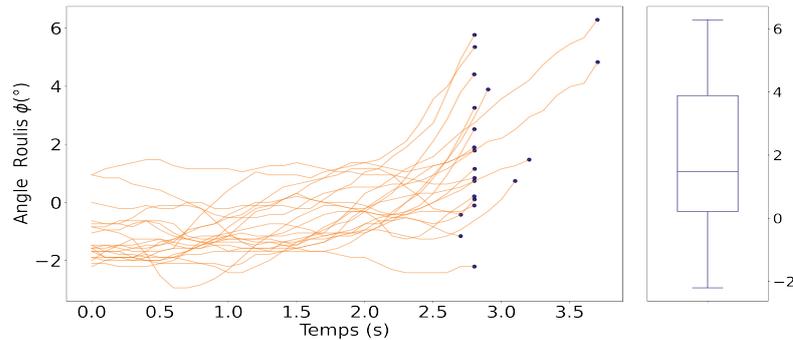
(d) Distribution de la vitesse de roulis  $\dot{\phi}$  dans S1.



(e) Distribution de la vitesse de roulis  $\dot{\phi}$  dans S2.

Figure 4.4 : Résultats de la classification des participants pour la consigne de conduite libre (CL).

Ces premiers résultats montrent que nous avons pu apporter des éléments de réponse à la problématique posée en début de ce chapitre qui consiste à effectuer un regroupement des comportements de conduite similaires en termes d'actions de contrôle sensori-moteur. Ceci, malgré la présence de diversités tant au niveau d'un même participant (d'un tour à l'autre) qu'entre participants. Cette diversité peut se voir à travers la [figure 4.5](#) où l'on peut constater que la dernière valeur de l'angle roulis diffère, d'une part, d'un tour à l'autre pour un participant donné et, d'autre part, d'un participant à un autre lors de la conduite sur le segment S1.



**Figure 4.5 :** L'évolution de l'angle roulis  $\phi$  sur le segment S1 (à gauche) et la distribution des dernières valeurs de  $\phi$  sur le segment S1 pour tous les participants (à droite).

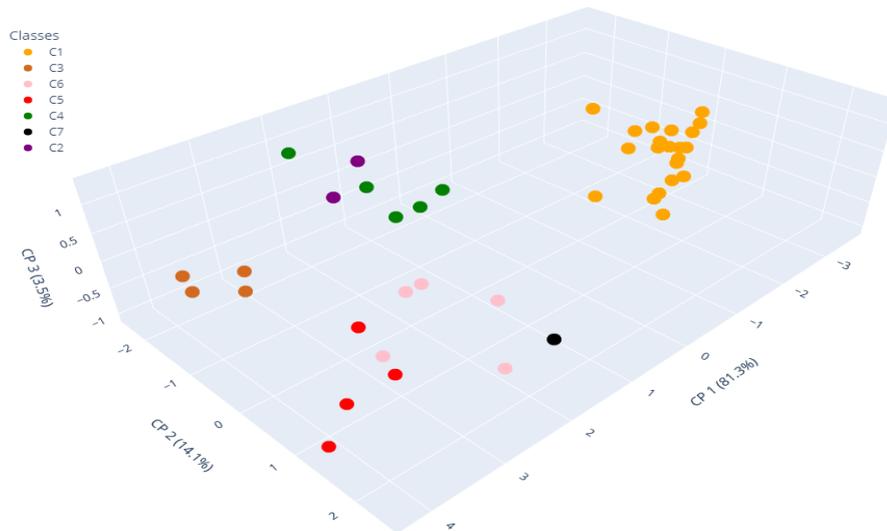
Par ailleurs, on note aussi une différence entre les valeurs de départ de l'angle roulis sur le segment S1 qui pourrait s'expliquer par une différence de positionnement sur la moto chez les différents participants. D'une part, cela signifie que le franchissement de la courbe (segment S2) résulte de l'interaction entre le conducteur sa moto et l'infrastructure dans laquelle il évolue. D'autre part, il sera difficile, voire impossible, de choisir un seuil au niveau de l'angle roulis pour détecter le franchissement de la courbe (début du segment S2). Ainsi, une meilleure approche serait de s'orienter vers une solution multivariée prenant en compte la dynamique du roulis (angle roulis et vitesse roulis).

Comme mentionné plus haut, la manœuvre de la prise d'un virage peut être considérée comme une transition séquentielle entre deux états stables (segments S1 et S2). L'instant de transition entre ces deux segments, est considéré comme l'entrée du virage que nous cherchons à déterminer dans ce chapitre, et qui diffère d'une séquence à une autre. La raison derrière cette différence réside dans la dynamique adoptée par le conducteur lors de la prise de virage. Cette dynamique est caractérisée par la vitesse et l'accélération pratiquées par le conducteur sur la ligne droite (segment S1), son positionnement sur la chaussée ainsi que sa perception visuelle. Ainsi, une analyse approfondie des comportements de conduite dans le segment S2 permettrait de mettre en exergue les différences existantes dans les actions entretenues entre le conducteur et son véhicule lors de la manœuvre prise de virage, et de mettre en évidence, des problèmes de contrôle la trajectoire pour certains conducteurs. Par conséquent, dans la partie qui suit, nous effectuons une interprétation de l'amorçage du virage (segment S2) à travers les différentes

caractéristiques statistiques extraites dans le segment S2. Par ailleurs, cette interprétation nous permettra d'affiner les résultats de la segmentation afin de mettre en évidence la diversité de pratiques des conducteurs de 2RMs.

#### **4.3.2.3 Interprétation des différentes sous-classes du segment S2**

La [figure 4.4a](#) montre que la principale classe regroupant des séquences de conduite issues du segment S2 peut encore être divisée en 6 sous-classes. Ainsi, l'objectif de cette partie est d'analyser et interpréter le comportement des participants dans ces différentes sous-classes. Pour ce faire, nous utilisons l'Analyse en Composantes Principales (ACP) [24]. En général, l'ACP est utilisé pour la recherche de motifs dans des données à grande dimension. L'idée principale de l'ACP consiste en une transformation des nombreuses caractéristiques de départ, généralement corrélées entre elles, en de nouvelles moindres caractéristiques décorrélatées les unes des autres. Ces nouvelles caractéristiques, appelées composantes principales, sont des combinaisons linéaires des caractéristiques de départ. Dans notre cas, la dimension de départ de nos caractéristiques est égale à 9. Même si cette dimension n'est pas trop importante, l'utilisation de l'ACP est nécessaire pour la visualisation des résultats de la classification avec l'algorithme CAH. Cette visualisation nous permettra de mettre en évidence la séparabilité existante entre les différentes classes obtenues avec le dendrogramme [figure 4.6](#). Par ailleurs, l'ACP nous facilitera l'interprétation de ces différentes classes à travers l'utilisation des « loadings » qui décrivent le niveau d'importance de chacune des caractéristiques de départ.



**Figure 4.6** : Projection en 3D des données de la consigne de conduite CL.

À travers la [figure 4.6](#), nous pouvons noter que les trois premières composantes principales (CP1, CP2 et CP3) représentent 98.9% de la variance totale des données de départ. Par conséquent, une représentation en 3D est largement suffisante pour avoir une idée sur la structure des données de départ de dimension 9. De même, de cette figure, nous pouvons constater une bonne séparation entre les différentes classes.

La [figure 4.7](#) donne la valeur des « *loadings* » de l'ACP. Une interprétation des composantes principales n'a pas réellement une signification dans la mesure où ces dernières ont été construites comme des combinaisons linéaires des caractéristiques de départ. Un « *loadings* » est, en d'autres termes, un score établi entre une composante principale et une caractéristique, permettant de mesurer le degré d'importance d'une caractéristique dans la construction d'une composante principale. La valeur d'un « *loading* » varie entre  $-1$  et  $1$ . Un « *loading* » proche de 1 en valeur absolue signifie que la caractéristique possède une forte influence sur la composante principale. De même, un « *loading* » proche de 0 signifie une faible influence de la variable sur la composante principale. Par ailleurs, le signe  $\pm$  d'un « *loading* » montre que la variable est positivement ou négativement corrélée à la composante principale.

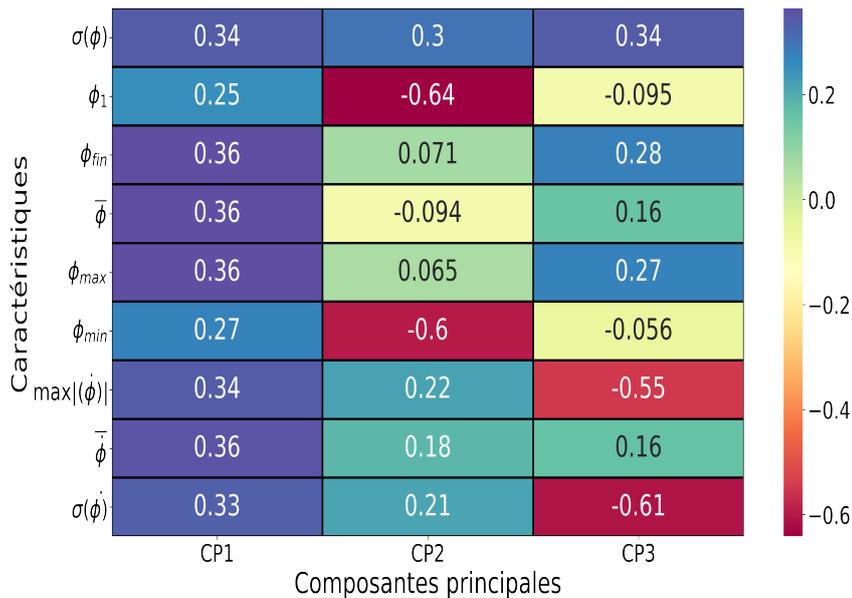


Figure 4.7 : Valeurs des « loadings » de l'ACP

Par exemple, dans la [figure 4.7](#), la première composante principale CP1, avec 81.3% de variance expliquée, est positivement corrélée avec les 9 caractéristiques de départ. La seconde composante principale ayant 14.1% de variance expliquée de la variance totale est en valeur absolue fortement corrélée avec les caractéristiques  $\phi_1$  and  $\phi_{min}$ . La troisième composante principale, quant à elle, se retrouve avec un faible pourcentage (3.5%) de variance expliquée, néanmoins cette dernière est, en valeur absolue, assez bien corrélée avec les caractéristiques  $\sigma(\dot{\phi})$  et  $\max|(\dot{\phi})|$ . Du fait que les trois premières composantes principales expliquent au total 98.9% de la variance totale présente dans les données de départ, nous les utiliserons pour les futurs besoins d'interprétations. En effet, une part de 98.9% est assez représentative de la structuration des données de départ.

Cependant le choix du nombre de caractéristiques utilisées pour l'interprétation des classes obtenues ne se résume pas qu'à des arguments statistiques. Il est également dicté par l'interprétation que l'on peut faire de ces regroupements. Dans le cas de cette étude, l'accent est mis sur la compréhension des classes et leur lecture en lien avec la problématique posée qui est de décomposer la manœuvre de la prise de virage en se fondant essentiellement sur la dynamique du roulis ( $\phi$ ,  $\dot{\phi}$ ), d'une part, et d'autre part, de détecter différents styles de conduite en approche du virage. Il est à noter que la caractéristique ( $\phi_1$ ) permettrait de capter les différents groupes avec différentes actions d'initiation du virage, et les caractéristiques ( $\sigma(\dot{\phi})$  et  $\sigma(\phi)$ ) refléteraient une variabilité sur la dynamique du roulis définie par ses deux composantes ( $\phi$ ,  $\dot{\phi}$ ). À cet égard, les caractéristiques  $\phi_1$ ,  $\sigma(\dot{\phi})$  et  $\sigma(\phi)$  seront utilisées dans le but d'interpréter les cartes de chaleur représentées sur la [figure 4.8](#). Cette analyse nous permettra, d'une part, de donner une interprétation des différentes classes et, d'autre part, de mettre en évidence les comportements

adoptés par les conducteurs en approche du virage, cette analyse sera détaillée dans les sections suivantes.

Il est important de préciser que ces cartes de chaleur des 7 classes sont triées par ordre croissant suivant la première observation de l'angle roulis ( $\phi_1$ ) dans le segment S2.

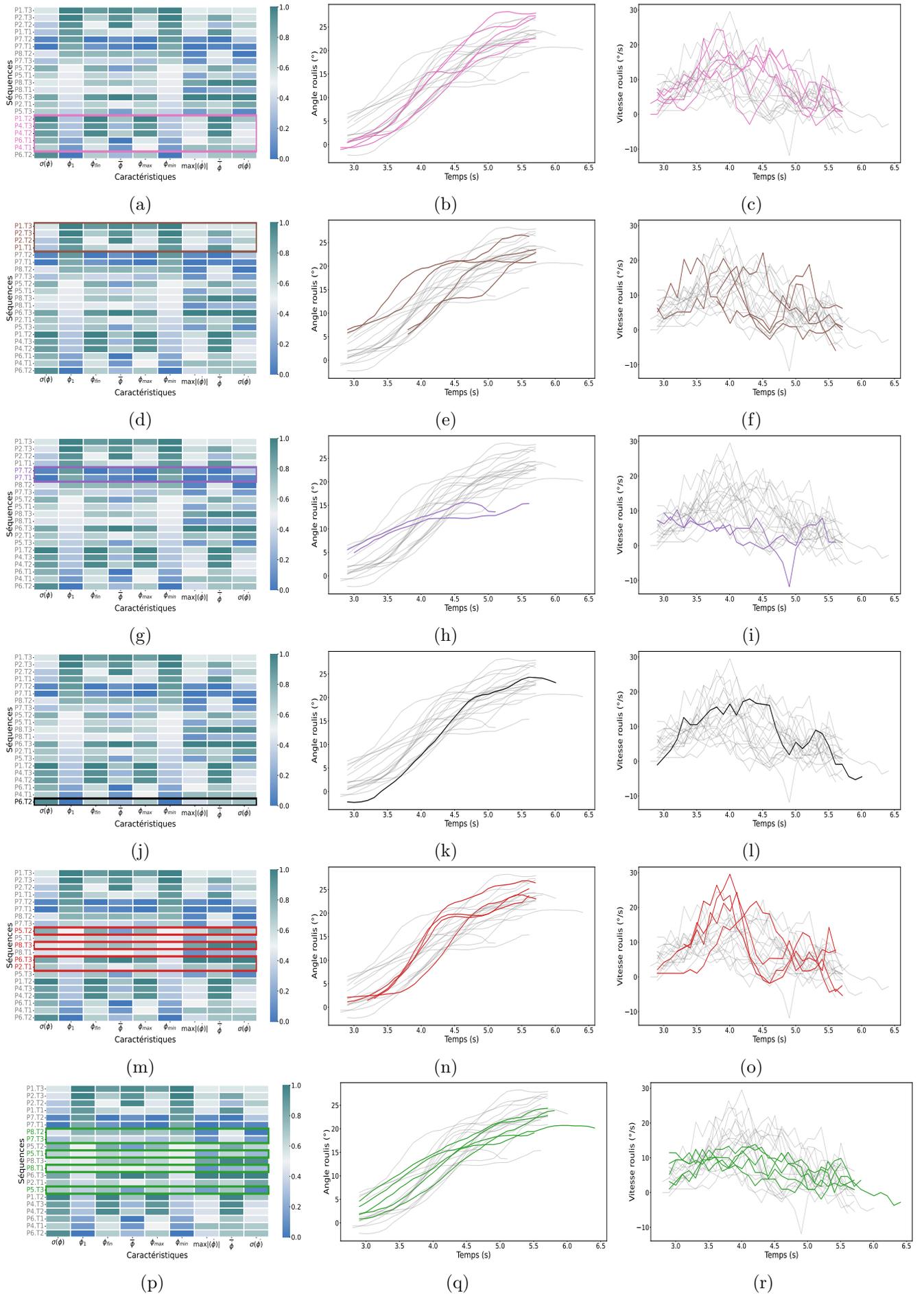
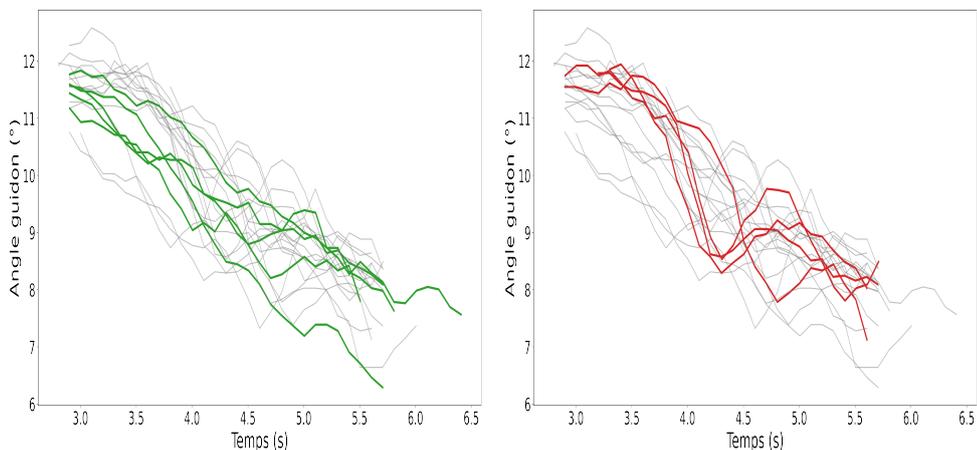


Figure 4.8 : Les résultats de la classification des données collectées durant l’instruction CL. La carte de chaleur est triée selon la valeur  $\phi_1$ . La dynamique du roulis : angle de roulis et vitesse de roulis sont représentés de gauche à droite, respectivement. Chaque ligne représente une classe.

Notons que la séquence de conduite du participant P6 au deuxième tour de la consigne CL (P6.L2) a la plus faible valeur de  $\phi_1$ , voir [figure 4.4c](#). Par ailleurs, cette séquence est la seule à être classifiée dans ■. Une valeur négative de  $\phi_1$  conduit à la seule explication selon laquelle qu'il y a une erreur lors de la segmentation de cette séquence de conduite. Les séquences classifiées dans ■ sont caractérisées par des valeurs élevées de  $\phi_1$  accompagnées d'une dynamique de roulis intermédiaire. Les séquences classifiées dans ■ sont caractérisées par une forte valeur de  $\phi_1$  avec une faible dynamique du roulis. Dans cette classe, les seules séquences obtenues sont le premier et second tour du participant P7. Ceci est considéré comme une phase d'adaptation dans laquelle le participant P7 découvre le circuit et la moto avec un très faible niveau de familiarisation. Dans la classe ■, on retrouve les trois tours de conduite du participant P4. Cette classe est caractérisée par des valeurs élevées de  $\phi_1$  et une dynamique de roulis intermédiaire.

Les séquences de la classe ■ peuvent être comparées avec celles de ■. En d'autres termes, les éléments de la classe ■ présentent une forte variabilité de la dynamique du roulis comparée à ceux de ■. Cette forte variabilité de la dynamique du roulis pourrait être considérée comme une conséquence d'un mauvais contrôle de la trajectoire de la moto. La [figure 4.9](#) montre l'évolution de l'angle du guidon dans le segment S2 pour les classes ■ et ■. Nous pouvons clairement constater qu'il y a plus de variabilité dans la classe ■ comparée à celle ■. Cela signifie que les participants ayant effectué ces séquences ont essayé de corriger leurs trajectoires à travers le guidon. Par ailleurs, nous posons l'hypothèse que les participants dont les séquences se retrouvent dans la classe ■ n'ont pas eu besoin de corriger leurs trajectoires, car leurs approches du virage étaient déjà optimales.



**Figure 4.9 :** Évolution de l'angle du guidon dans le segment S2 pour les séquences constituant les deux classes : ■ et ■

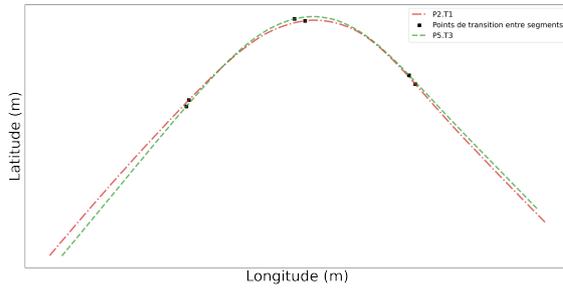
Pour vérifier cette hypothèse, nous prenons deux séquences P5.L3 et P2.L1 pour respectivement les classes ■ et ■ afin d'effectuer une analyse approfondie de leur comportement lors de la manœuvre de la prise de virage. Nous pouvons observer sur les [figure 4.10a](#) et [figure 4.10b](#) où sont représentées les deux trajectoires pendant les séquences P2.T1 et P5.T3 que les conducteurs ont adopté des manières différentes lors de l'approche d'un virage.

1. En franchissant le virage, le participant P2 se déplace du bord extérieur de la courbe vers le bord intérieur à l'approche de l'apex<sup>1</sup> « point de corde », voir [figure 4.10a](#). Après avoir dépassé l'apex, ce participant revient vers le bord extérieur en sortant de la courbe. Cette manœuvre a pour but de permettre à ce participant de progresser le plus rapidement possible dans un virage. Ce phénomène peut être constaté sur la [figure 4.10c](#) où sont représentées l'accélération longitudinale et la vitesse pratiquée par ce participant. En conséquence de ce comportement dynamique, nous pouvons observer sur la [figure 4.10e](#) que durant le segment S2 (qui est la première partie de la manœuvre de virage), la courbe de l'angle de roulis est caractérisée par des étapes qui correspondent à de fortes variations dynamiques autour de zéro pour la variation de la vitesse du roulis. En même temps, nous pouvons observer sur la [figure 4.10g](#) la variabilité de l'angle de direction du guidon. Ce qui signifie que ce participant essaie de maintenir la moto dynamiquement stable autour d'un certain angle de roulis.
2. Dans la seconde approche du virage adoptée par le participant P4, nous pouvons remarquer que ce dernier essaie d'adopter une position plus proche de la ligne centrale dans le segment S1, voir [figure 4.10b](#). Il a essayé de garder cette position tout au long du virage (S2 et S3). Ceci pourrait avoir été fait afin d'obtenir une meilleure visibilité du virage. En ce qui concerne le comportement dynamique, il y a moins de variabilité au niveau des dynamiques longitudinale et de roulis par rapport au cas décrit précédemment, voir [figure 4.10d](#), [figure 4.10f](#) et [figure 4.10h](#).

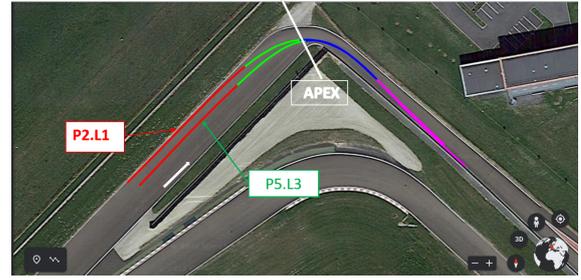
---

<sup>1</sup>Apex : C'est le point central-intérieur d'un virage, cependant ce point n'est pas nécessairement le milieu du virage. C'est à partir de ce point qu'un pilote commence à sortir du virage.

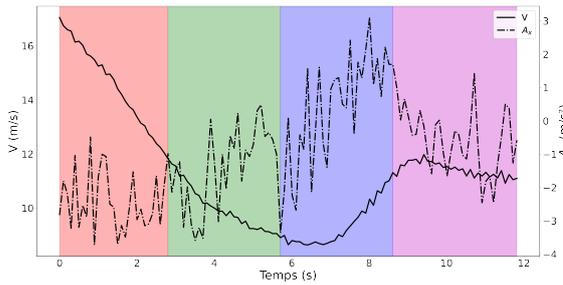
### 4.3. Application de la méthodologie proposée et résultats



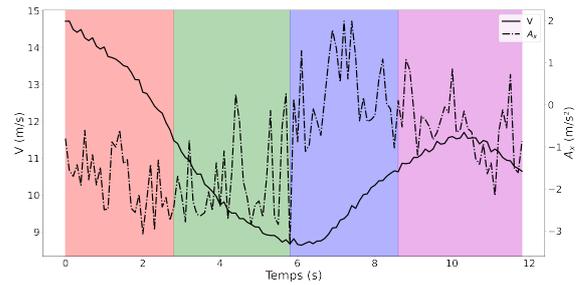
(a) Trajectoires des séquences P2.T1 et P5.T3.



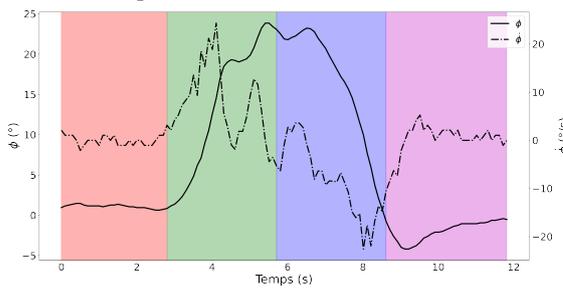
(b) Trajectoires projetées sur Google Earth des séquences P2.T1 et P5.T3.



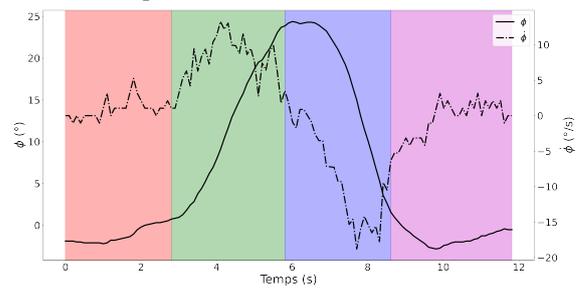
(c) Vitesse et accélération longitudinale pratiquées lors de la séquence P2.T1.



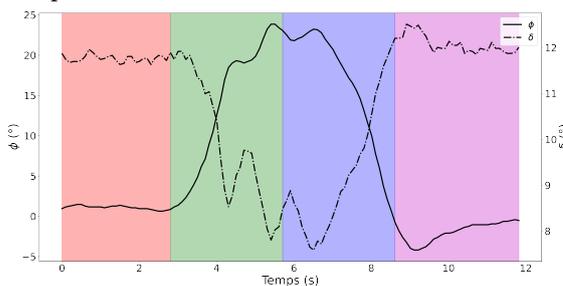
(d) Vitesse et accélération longitudinale pratiquées lors de la séquence P5.T3.



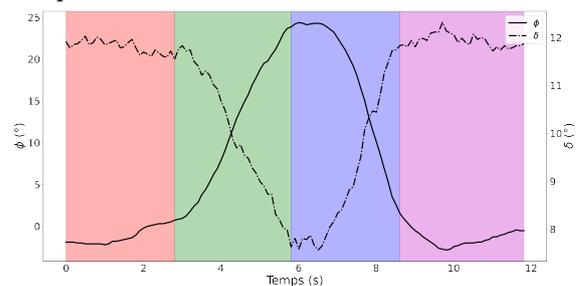
(e) Angle roulis ( $\phi$ ) et vitesse du roulis ( $\dot{\phi}$ ) lors de la séquence P2.T1.



(f) Angle roulis ( $\phi$ ) et vitesse du roulis ( $\dot{\phi}$ ) lors de la séquence P5.T3.



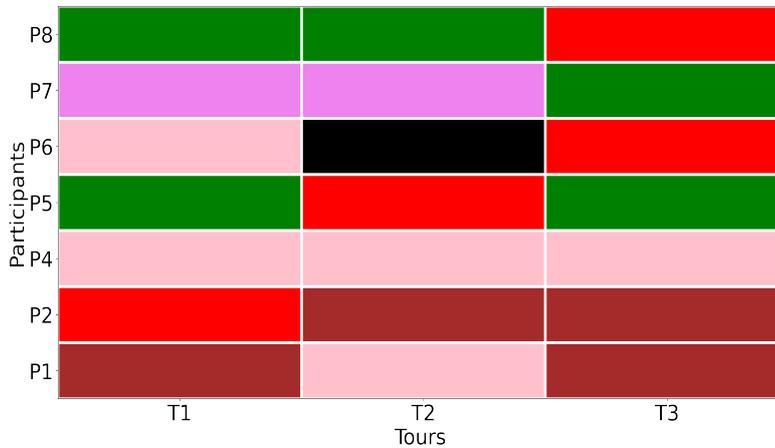
(g) Angle roulis ( $\phi$ ) et angle du guidon ( $\delta$ ) lors de la séquence P2.T1.



(h) Angle roulis ( $\phi$ ) et angle du guidon ( $\delta$ ) lors de la séquence P5.T3.

**Figure 4.10** : Comportement dynamique des deux séquences P2.T1 et P5.T3 pendant la manœuvre de prise de virage segmentée en différentes phases : de gauche à droite ( S1 : en rouge, S2 : en vert, S3 : en bleu, S4 : en violet).

La variabilité significative qui caractérise le comportement des conducteurs lors de la consigne de conduite CL mérite d'être soulignée, voir [figure 4.11](#).



**Figure 4.11** : Résultats de la classification des tours de conduite pour chacun des 7 participants.

Un tel résultat était prévisible car les participants essaient de se familiariser davantage avec la moto et la piste durant cette phase d'habituance. Nous constatons une variabilité d'un tour à l'autre pour un même participant excepté pour le P4. Ce dernier a conservé la même position pendant ses trois tours, voir [figure 4.12](#). De plus, sur la base de ses données déclaratives, ce participant pourrait être considéré, d'une part, comme un conducteur expérimenté, et d'autre part, il est le seul à avoir déclaré une préférence pour la consigne de conduite AC. Dans l'étude effectuée dans le [3](#), fondée sur une autre approche, nous avons constaté que ce participant était le seul à avoir respecté les consignes données (SC et AC). Nous avons également constaté que la consigne CL était conforme à celle AC ; ce qui est cohérent avec la préférence déclarée du participant P4.



**Figure 4.12** : Trajectoires représentées sur Google Earth des trois séquences de conduite réalisées par le participant P4.

## 4.4 Conclusion

Dans ce chapitre, le principal objectif consistait en l'étude du comportement des conducteurs de 2RMs lors de prise de virage. Ce comportement résulte de l'interaction entre les actions du conducteur, la dynamique du véhicule et l'infrastructure sur laquelle évolue la moto. Le comportement adopté peut différer d'un conducteur à un autre suivant différents facteurs : l'expérience du conducteur, l'interaction avec l'infrastructure, la perception visuelle du conducteur. Par conséquent, la détection de l'entrée du virage peut être une tâche très difficile à cause de l'ambiguïté existante au niveau de la définition d'une entrée de virage et de la diversité de pratiques chez les conducteurs de 2RMs. Par ailleurs, cette tâche constitue un challenge dans le cas d'un 2RM en raison de la forte implication du conducteur dans la tâche de conduite. Ainsi, son expérience pourrait jouer un rôle important dans la conduite, notamment lors la prise de virage.

Une méthodologie en deux étapes consistant en une segmentation automatique des mesures relatives à la dynamique du roulis suivie d'une classification a été mise en œuvre pour relever ce défi. Cette partie de la thèse a porté sur le développement d'outils de segmentation et de classification de séquences de conduite d'un 2RM lors de la prise de virage. Pour ce faire, l'approche proposée repose sur un modèle de régression logistique à processus cachés. Ce modèle sert à la segmentation non supervisée des séquences de conduite. Des caractéristiques statistiques extraites de chaque segment ont été utilisées comme entrées du classifieur CAH pour classifier les comportements de conduite à l'entrée du virage. Les résultats obtenus sur une base de données réelle, les données correspondantes à la consigne de conduite naturelle CL, ont permis de montrer l'efficacité de la méthodologie proposée à la fois pour la détection de l'entrée du virage et pour la détection de styles de conduite différents en termes de dynamique en approche de virage, et notamment, des participants ayant rencontré des difficultés de maîtrise de leur trajectoire sur cette partie du virage. Les résultats obtenus montrent la capacité de la méthodologie proposée à détecter des événements de conduite « critiques » des 2RMs lors de la prise de virage.



---

### Conclusion et Perspectives

---

Ces travaux de thèse ont pour objectif principal la mise au point d'outils d'analyse des comportements des conducteurs de 2RMs particulièrement lors de la prise d'un virage. Pour mieux appréhender les comportements des conducteurs de 2RMs lors de la prise d'un virage, il est nécessaire d'avoir des données réelles reproduisant au mieux une situation de conduite. Cela est d'autant plus nécessaire puisque dans la littérature, la plupart des travaux concernant le comportement des conducteurs de 2RMs se fondent en général sur des vidéos de conduites ou des données issues de simulateur. C'est ainsi que, dans ces travaux de thèse, nous avons présenté un chapitre relatif à la mise en place d'une expérimentation réelle de conduite en virage avec un 2RM instrumenté avec divers capteurs. Différentes mesures de conduite ont été recueillies sous trois consignes, notamment la conduite libre (CL), la conduite avec corps (AC) et la conduite sans corps (SC). Ces mesures, provenant des différents capteurs, ont été centralisées puis mises sous un format utilisable par un logiciel de traitement de données.

Dans le chapitre 3, nous avons utilisé ces données relatives au virage dans sa globalité pour effectuer une analyse des comportements des conducteurs de 2RMs lors de la prise d'un virage. Nous avons présenté une méthodologie fondée sur approche de fouille de données conduisant à une classification des comportements de conduite de 8 participants à l'expérimentation. Une interprétation des différentes classes obtenues a permis d'identifier la mesure de l'angle du guidon comme principal élément discriminant ces comportements. En outre, une confrontation avec les données déclaratives a permis d'identifier les participants ayant respecté ou non les consignes de conduite sans corps (SC) et avec corps (AC). Le participant P4 a été le seul à avoir respecté ces deux consignes citées précédemment. Par ailleurs, il ressort de ces analyses que le participant P3 qui se trouve être un stagiaire peut être considéré comme un conducteur atypique.

Le chapitre 4 s'inscrit dans la continuité du 3 en proposant une analyse beaucoup plus fine des comportements des conducteurs de 2RMs lors de la prise d'un virage. En effet, dans ce chapitre, un travail relatif à la segmentation multidimensionnelle du virage en 4 segments a été effectué en amont. La segmentation a été réalisée à l'aide d'algorithmes reposant sur une approche de régression, notamment le MRHLP et MHMMR. Les mesures de la dynamique du roulis (angle roulis et vitesse roulis) de chacune des séquences de conduite des participants ont été utilisées pour l'application des algorithmes de segmentation. À l'issue de cette segmentation, une analyse de la prise de virage à l'aide d'une caractérisation statistique de chacun des segments S1 et S2 suivie d'une classification des différents participants ont été effectuées. Il ressort de cette classification une nette séparation entre les séquences de conduites du segment S1 et celles du segment S2. Une analyse des mesures relatives de la consigne conduite libre (CL) des sous-classes du segment S2 a permis d'identifier des styles de conduite différents en termes de dynamique en approche de virage, et notamment, des participants ayant rencontré des difficultés de maîtrise de leur trajectoire sur cette partie du virage. Les résultats obtenus montrent la capacité de la méthodologie proposée à détecter des événements de conduite « critiques » des 2RMs lors de la prise de virage. Ces résultats permettront de proposer des outils de formation, apportant des éléments de connaissance et de compréhension nécessaires, pour une évaluation objective de la prise de virage « sécuritaire ».

### **Perspectives :**

Ce travail de thèse ouvre plusieurs perspectives. Tout d'abord, le processus mis en place pour l'acquisition de données relatives à la conduite d'un virage peut être amélioré. En effet, une des principales difficultés dans ce travail de thèse a été la délimitation exacte de l'entrée du virage. Sur un plan opérationnel, une amélioration de ce processus consiste à mettre des balises pour délimiter les différentes parties du virage. Ceci permettra d'avoir une délimitation commune à tous les participants à l'expérimentation, et d'avoir ainsi un cadre commun d'évaluation et de comparaison pour tous les participants. Ensuite, l'approche d'évaluation de la prise de virage proposée dans ce travail de thèse est fondée sur une première phase de segmentation de la trajectoire dans laquelle des algorithmes de segmentation sont utilisés suivie d'une phase de caractérisation statistique des différents segments. Cependant, cette caractérisation statistique pourrait causer une perte de l'aspect séquentiel caractérisant la conduite d'un véhicule, et rendant ainsi, la détection des phases de transitions entre les différents segments moins précise. Une des solutions envisageaient à ce problème pourrait être d'utiliser des algorithmes non supervisés adaptés aux données séquentielles, et ce en prenant en compte la notion de voisinage entre les mesures - et par la même occasion, permettant de réaliser en même temps une segmentation et une classification [26].

La méthodologie et les outils proposés sont donc généralisables à des manœuvres de conduite variées telles que la prise de virage et le changement de voie, que ces manœuvres soient réalisées

---

sur route ou sur piste par des véhicules routiers quelconques. Il est à noter que notre méthodologie, comme tout autre approche guidée par les données, est tributaire de la qualité des données en entrées de l'étape de segmentation et à leur quantité. Pour notre étude nous avons utilisé des capteurs « haut de gamme » précis mais coûteux et de ce fait difficilement utilisables à grande échelle. Toutefois et du fait des progrès technologiques constants on peut imaginer utiliser les capteurs présents dans les Smartphones ou dans les drones « grand public » pour effectuer des recueils de données (roulis, tangage, lacet, et GPS) à « bas coûts » et à grande échelle. La question posée par une telle instrumentation à grande échelle étant le contrôle du positionnement des capteurs sur le véhicule.

Enfin, toutes ces perspectives citées ci-dessus, nous envisageons de les mettre en œuvre dans le cadre du projet ANR eMC2 : « Motocyclettes électriques et sécurité routière » (2022-2026). Le projet eMC2 propose de combler les lacunes sur la connaissance des pratiques réelles des conducteurs de 2RMs électriques en s'appuyant sur le développement d'outils et méthodes permettant de mesurer les comportements de ces derniers et de les comparer à ceux des usagers de 2RMs thermiques. Les questions de recherche que nous proposons d'investiguer dans le cadre de ce projet sont : (1) comprendre les difficultés auxquelles font face les conducteurs qui passent des 2RM thermiques aux électriques, ceci en observant un panel de conducteurs, et ainsi identifier les difficultés et les situations à risque rencontrés par ces derniers. (2) étudier les pratiques de ce panel après une phase d'apprentissage de la conduite des 2RM électriques. (3) étudier l'impact de la motorisation électrique versus thermique sur la prise d'information visuelle en virage, et par conséquent, sur les trajectoires réalisées.



---

## Bibliographie

---

- [1] Seraina Agramunt, *Naturalistic driving behaviour and self-regulation practices among older drivers with bilateral cataract: a prospective cohort study*, Ph.D. Thesis, 2018.
- [2] Ferhat Attal, *Classification de situations de conduite et détection des événements critiques d'un deux roues motorisé*, Theses, 2015.
- [3] Samuel Aupetit, *Etude ergonomique de l'apprentissage de la conduite moto dans une perspective de prévention*, Ph.D. Thesis, 2010. Thèse de doctorat dirigée par Buttelli, OlivierEspié, Stéphane et Riff, Jacques Sciences et techniques des activités physiques et sportives Orléans 2010.
- [4] James C Bezdek, *Numerical taxonomy with fuzzy sets*, Journal of mathematical biology **1** (1974), no. 1, 57–71.
- [5] Robin A Blanchard, Anita M Myers, and Michelle M Porter, *Correspondence between self-reported and objective measures of driving exposure and patterns in older drivers*, Accident Analysis & Prevention **42** (2010), no. 2, 523–529.
- [6] M Boccione, Federico Cheli, Elisabetta Leo, and M Pezzola, *Experimental identification of kinematic coupled effects between driver–motorcycle*, Imac–xxv: a conference & exposition on structural dynamics, 2007, pp. 19–22.
- [7] Vladimir V. Bochkarev and Eduard Yu. Lerner, *Zipf and non-zipf laws for homogeneous markov chain*, 2013.
- [8] Rodrigo C Brasileiro, Victor LF Souza, and Adriano LI Oliveira, *Automatic trading method based on piecewise aggregate approximation and multi-swarm of improved self-adaptive particle swarm optimization with validation*, Decision Support Systems **104** (2017), 79–91.
- [9] Salvatore Cafiso, Alessandro Di Graziano, Oliver Giudice, and Giuseppina Pappalardo, *Using gps data to detect critical events in motorcycle rider behaviour*, International Journal of Mobile Network Design and Innovation **5** (2014), no. 4, 195–204.
- [10] F. Chamroukhi, S. Mohammed, D. Trabelsi, L. Oukhellou, and Y. Amirat, *Joint segmentation of multivariate time series with hidden process regression for human activity recognition*, Neurocomputing **120** (2013), 633–644. Image Feature Detection and Description.
- [11] ———, *Joint segmentation of multivariate time series with hidden process regression for human activity recognition*, Neurocomputing **120** (2013Nov), 633–644.
- [12] F. Chamroukhi, A. Samé, G. Govaert, and P. Aknin, *Time series modeling by a regression approach based on a latent process*, Neural Networks **22** (2009), no. 5-6, 593–602.
- [13] Pushpa Choudhary and Nagendra R Velaga, *Mobile phone use during driving: Effects on speed and effectiveness of driver compensatory behaviour*, Accident Analysis & Prevention **106** (2017), 370–378.

- 
- [14] Shivam Singh Chouhan, Ankit Kathuria, and Chalumuri Ravi Sekhar, *Examining risky riding behavior in india using motorcycle rider behavior questionnaire*, Accident Analysis & Prevention **160** (2021), 106312.
- [15] Vittore Cossalter, *Motorcycle dynamics*, Lulu.com, 2006.
- [16] Vittore Cossalter, Alberto Doria, and Roberto Lot, *Steady turning of two-wheeled vehicles*, Vehicle system dynamics **31** (1999), no. 3, 157–181.
- [17] Vittore Cossalter, Roberto Lot, and Stefano Rota, *Objective and subjective evaluation of an advanced motorcycle riding simulator*, European transport research review **2** (2010), no. 4, 223–233.
- [18] Mark A Elliott, Christopher J Baughan, and Barry F Sexton, *Errors and violations in relation to motorcyclists' crash risk*, Accident Analysis & Prevention **39** (2007), no. 3, 491–499.
- [19] Sonja Engmann and Denis Cousineau, *Comparing distributions: the two-sample anderson–darling test as an alternative to the kolmogorov–smirnov test*, Journal of Applied Quantitative Methods **6** (201109), 1–17.
- [20] Martin Ester, Hans-Peter Kriegel, Jörg Sander, Xiaowei Xu, et al., *A density-based algorithm for discovering clusters in large spatial databases with noise.*, kdd, 1996, pp. 226–231.
- [21] David Freedman, Robert Pisani, and Roger Purves, *Statistics (international student edition)*, Pisani, R. Purves, 4th edn. WW Norton & Company, New York (2007).
- [22] Moshe Fridman, *Hidden markov model regression*, Ph.D. Thesis, 1993.
- [23] ———, *Hidden markov model regression*, 1993.
- [24] Karl Pearson F.R.S., *Liii. on lines and planes of closest fit to systems of points in space*, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science **2** (1901), no. 11, 559–572.
- [25] David Hallac, Peter Nystrop, and Stephen Boyd, *Greedy gaussian segmentation of multivariate time series*, 2018.
- [26] David Hallac, Sagar Vare, Stephen Boyd, and Jure Leskovec, *Toeplitz inverse covariance-based clustering of multivariate time series data*, Proceedings of the 23rd acm sigkdd international conference on knowledge discovery and data mining, 2017, pp. 215–223.
- [27] Zhilin He and Chun-Hsing Ho, *An improved clustering algorithm based on finite gaussian mixture model*, Multimedia Tools and Applications **78** (2019), no. 17, 24285–24299.
- [28] B. Higgs and M. Abbas, *A two-step segmentation algorithm for behavioral clustering of naturalistic driving styles*, 16th international ieee conference on intelligent transportation systems (itsc 2013), 2013, pp. 857–862.
- [29] Frank J. Massey Jr., *The kolmogorov-smirnov test for goodness of fit*, Journal of the American Statistical Association **46** (1951), no. 253, 68–78.
- [30] R. Kalsoom and Z. Halim, *Clustering the driving features based on data streams*, Inmic, 2013, pp. 89–94.
- [31] Eamonn Keogh, Kaushik Chakrabarti, Michael Pazzani, and Sharad Mehrotra, *Dimensionality reduction for fast similarity search in large time series databases*, Knowledge and Information Systems **3** (200201).
- [32] Sheila G Klauer, Thomas A Dingus, Vicki L Neale, Jeremy D Sudweeks, David J Ramsey, et al., *The impact of driver inattention on near-crash/crash risk: An analysis using the 100-car naturalistic driving study data* (2006).
- [33] Teuvo Kohonen, *Self-organized formation of topologically correct feature maps*, Biological cybernetics **43** (1982), no. 1, 59–69.
- [34] Natália Kováčsová, Marco Grottoli, Francesco Celiberti, Yves Lemmens, Riender Happee, Marjan P Hagenzieker, and Joost CF de Winter, *Emergency braking at intersections: A motion-base motorcycle simulator study*, Applied ergonomics **82** (2020), 102970.
- [35] Burcu Kulahcioglu, Serhan Ozdemir, and Bora Kumova, *Application of symbolic piecewise aggregate approximation (paa) analysis to ecg signals*, 17th iasted international conference on applied simulation and modelling, 2008.

## Bibliographie

---

- [36] G. N. Lance and W. T. Williams, *A General Theory of Classificatory Sorting Strategies: 1. Hierarchical Systems*, The Computer Journal **9** (196702), no. 4, 373–380, available at <https://academic.oup.com/comjnl/article-pdf/9/4/373/1101470/9-4-373.pdf>.
- [37] P. V. Lemonakis, N. E. Eliou, T. Karakasidis, and G. Botzoris, *A new methodology for approaching motorcycle riders' behavior at curved road sections*, European Transport Research Review **6** (2014Sep), no. 3, 303–314.
- [38] Paul Loiseau, Chaouki Nacer Eddine Boulouf, Philippe Chevrel, Fabien Claveau, Stéphane Espie, and Franck Mars, *Rider model identification: neural networks and quasi-LPV models*, IET Intelligent Transport Systems **14** (2020), no. 10, 1259–1264.
- [39] ———, *Rider model identification using dynamic neural networks*, IFAC World Congress 2020, 2020July.
- [40] J MacQueen, *Classification and analysis of multivariate observations*, 5th berkeley symp. math. statist. probability, 1967, pp. 281–297.
- [41] J. Macqueen, *Some methods for classification and analysis of multivariate observations*, In 5-th berkeley symposium on mathematical statistics and probability, 1967, pp. 281–297.
- [42] Nils Magiera, Herbert Janßen, Martin Heckmann, and Hermann Winner, *Rider skill identification by probabilistic segmentation into motorcycle maneuver primitives*, 2016 iee 19th international conference on intelligent transportation systems (itsc), 2016, pp. 379–386.
- [43] H. B. Mann and D. R. Whitney, *On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other*, The Annals of Mathematical Statistics **18** (1947), no. 1, 50–60.
- [44] Seyed Abbas Motevalian, Mohsen Asadi-Lari, Habibollah Rahimi, and Mehrdad Eftekhari, *Validation of a persian version of motorcycle rider behavior questionnaire*, Annals of advances in automotive medicine/annual scientific conference, 2011, pp. 91.
- [45] In Jae Myung, *Tutorial on maximum likelihood estimation*, J. Math. Psychol. **47** (2003feb), no. 1, 90–100.
- [46] Abhishek Nayak, S. Murigendrappa, and Ashwin Suresh, *Stability enhancement of a powered two wheeler vehicle under curve negotiation*, Vehicle System Dynamics **7** (201607), 85–95.
- [47] Marilee Nugent, Giovanni Savino, Christine Mulvihill, Mike Lenné, and Michael Fitzharris, *Evaluating rider steering responses to an unexpected collision hazard using a motorcycle riding simulator*, Transportation research part F: traffic psychology and behaviour **66** (2019), 292–309.
- [48] ONISR, *La sécurité routière en france : Bilan de l'accidentalité de l'année 2019*, 2020.
- [49] Nishant Mukund Pawar, Rashmeet Kaur Khanuja, Pushpa Choudhary, and Nagendra R Velaga, *Modelling braking behaviour and accident probability of drivers under increasing time pressure conditions*, Accident Analysis & Prevention **136** (2020), 105401.
- [50] Eleni Petridou and Maria Moustaki, *Human factors in the causation of road traffic crashes*, European journal of epidemiology **16** (2000), no. 9, 819–826.
- [51] A. Popov, Stuart Rowell, and J. Meijaard, *A review on motorcycle and rider modelling for steering control*, VEH SYST DYN **48** (201006), 775–792.
- [52] Isabelle Ragot-Court and Pierre Van Eslande, *Projet compar : Les comportements de conduite et leurs déterminants dans l'accidentalité des deux-roues motorisés*, IFSTTAR/DSCR, 2016.
- [53] Huorong Ren, Xiujuan Liao, Zhiwu Li, and Abdulrahman Ai-Ahmari, *Anomaly detection using piecewise aggregate approximation in the amplitude domain*, Applied Intelligence **48** (2018), no. 5, 1097–1110.
- [54] Roy S. Rice, *Rider skill influences on motorcycle maneuvering*, Sae technical paper, 197802.
- [55] Richard Rowe, Gabriela D Roman, Frank P McKenna, Edward Barker, and Damian Poulter, *Measuring errors and violations on the road: A bifactor modeling approach to the driver behavior questionnaire*, Accident Analysis & Prevention **74** (2015), 118–125.

- [56] Jorge Santos, Natasha Merat, Sandra Mouta, Karel Brookhuis, and Dick De Waard, *The interaction between driving and in-vehicle information systems: Comparison of results from laboratory, simulator and real-world studies*, Transportation Research Part F: Traffic Psychology and Behaviour **8** (2005), no. 2, 135–146.
- [57] F. W. Scholz and M. A. Stephens, *K-sample anderson-darling tests*, Journal of the American Statistical Association **82** (1987), no. 399, 918–924.
- [58] ———, *K-sample anderson-darling tests*, Journal of the American Statistical Association **82** (1987), no. 399, 918–924.
- [59] Pavel Senin, *Dynamic time warping algorithm review*, Information and Computer Science Department University of Hawaii at Manoa Honolulu, USA **855** (2008), no. 1-23, 40.
- [60] Harpreet Singh and Ankit Kathuria, *Analyzing driver behaviour under naturalistic driving conditions: A review*, Accident Analysis & Prevention **150** (2021), 105908.
- [61] Sarra Smaiah, *Système de reconstruction de trajectoire pour véhicules deux roues motorisés*, Ph.D. Thesis, 2020. Thèse de doctorat dirigée par Bouaziz, Samir et Sadoun, Rabah Robotique université Paris-Saclay 2020.
- [62] D. Trabelsi, S. Mohammed, F. Chamroukhi, L. Oukhellou, and Y. Amirat, *Supervised and unsupervised classification approaches for human activity recognition using body-mounted sensors*, Proceedings of the xxth european symposium on artificial neural networks, computational intelligence and machine learning (esann), 2012April, pp. 417–422.
- [63] Charles Truong, Laurent Oudre, and Nicolas Vayatis, *Selective review of offline change point detection methods*, Signal Processing **167** (2020Feb), 107299.
- [64] Ingrid van Schagen and Fridulv Sagberg, *The potential benefits of naturalistic driving for road safety research: Theoretical and empirical considerations and challenges for the future*, Procedia-social and behavioral sciences **48** (2012), 692–701.
- [65] Ingrid van Schagen, Ruth Welsh, Agatha Backer-Grondahl, Marika Hoedemaeker, Tsippy Lotan, Andrew Morris, Fridulv Sagberg, and Martin Winkelbauer, *Towards a large scale european naturalistic driving study: final report of prologue: deliverable d4. 2* (2011).
- [66] EI Vlahogianni, George Yannis, John C Golias, N Eliou, and P Lemonakis, *Identifying riding profiles parameters from high resolution naturalistic riding data*, Proceedings of the 3rd international conference on road safety and simulation (rss2011), september, 2011, pp. 14–16.
- [67] Chao Wang, Mohammed Quddus, and Stephen Ison, *The effects of area-wide road speed and curvature on traffic casualties in england*, Journal of transport geography **17** (2009), no. 5, 385–395.
- [68] D. L. Weakliem, *A critique of the bayesian information criterion for model selection*, Sociological Methods & Research **27** (1999), no. 3, 359–397.
- [69] DAVID H. WEIR, *Motorcycle dynamics and rider control*, VSD **6** (1977), no. 2-3, 187–190.
- [70] Antoine Zambelli, *A data-driven approach to estimating the number of clusters in hierarchical clustering*, F1000Research **5** (201608).
- [71] S Zamfir, R Drosescu, and R Gaiginschi, *Practical method for estimating road curvatures using onboard gps and imu equipment*, Iop conference series: Materials science and engineering, 2016, pp. 012114.
- [72] Zhongheng Zhang, Fionn Murtagh, Sven Van Poucke, Su Lin, and Peng Lan, *Hierarchical cluster analysis in clinical research with heterogeneous study population: Highlighting its visualization with r*, Annals of Translational Medicine **5** (201701), 75.
- [73] Shaopeng Zhu, Shintaroh Murakami, and Hidekazu Nishimura, *Motion analysis of a motorcycle taking into account the rider's effects*, Vehicle System Dynamics **50** (2012), no. 8, 1225–1245.

## Bibliographie

---

- [74] Ilka Zöllner, Bettina Abendroth, and Ralph Bruder, *Driver behaviour validity in driving simulators—analysis of the moment of initiation of braking at urban intersections*, Transportation research part F: traffic psychology and behaviour **61** (2019), 120–130.