Biosignals for driver's stress level assessment : functional variable selection and fractal characterization

Neska El Haouij

NNT : 2018SACLS191

Biosignals for Driver's Stress Level Assessment: Functional Variable Selection and Fractal Characterization

Thèse de doctorat de l'Ecole Nationale d'Ingénieurs de Tunis,
Université de Tunis El Manar
et de l'Université Paris-Saclay,
préparée à l'Université Paris-Sud

École Doctorale Sciences et Techniques de l'Ingénieur (ED-STI)
Spécialité de doctorat : Sciences et Techniques de l'Information et de la Communication
École Doctorale n°574 : Mathématiques Hadamard (EDMH)
Spécialité de doctorat : Mathématiques appliquées

Thèse présentée et soutenue à Orsay, le 04 Juillet 2018, par

Neska El Haouij

Composition du Jury :

Philippe Besse Pr., Univ. Toulouse, (– IMT)	Rapporteur
Nabil Gmati Pr., Univ. El Manar, ENIT, (– LAMSIN)	Rapporteur
Anne Gégout-Petit Pr., Univ. Lorraine (– Institut Elie Cartan)	Examinatrice
Mériem Jaïdane Pr., Univ. El Manar, ENIT (– U2S)	Co-Directrice de thèse
Pascal Massart Pr., Univ. Paris-Sud, (– LMO)	Président du Jury
Mathilde Mougeot Pr., ENSIIE et Univ. Paris Diderot, (– LPMA)	Examinatrice
Jean-Michel Poggi Pr., Univ. Paris Descartes & Univ. Paris-Sud (– LMO)	Directeur de thèse
Raja Ghozi MCF, Univ. El Manar, ENIT (– U2S)	Invitée
Sylvie Sevestre-Ghalila MCF, CEA (– CEA-LinkLab)	Invitée

Thèse de doctorat

Dedicated to my mother

Acknowledgments

When I look back at the last three years, I sometimes wonder if I did a work or I just tried to channel the energy, knowledge, and support of every person around me into a work resulting in this thesis project.

My four supervisors are amazing. They patiently assisted me and were available when I asked for any help. If I respect the timeline of the thesis project, I will go back to the first idea that was shared between Sylvie and Mrs Jaïdane, just after the I'CityForAll project. They gave me the opportunity to write a draft of the thesis proposal and they proposed the team to work within the thesis project.
Jean-Michel powered me with his encouragements and his generosity with time and ideas. I suspect that most of the good initiatives I had were first secretly planted by him. He always knows how to let me find my own way back to productivity in research.
Mrs Jaïdane balanced her care and skepticism of every word, graph, and thought. She is always encouraging the multidisciplinary and applied studies. Her teasing and provocation improved me as a researcher and aroused my curiosity. Thanks.
Mrs Ghozi was from the beginning enthusiast and supportive. It was always a pleasure to discuss with her about everything, write emails, and correct the submitted papers. She gave me a boost of morale when I needed it.
From the beginning of postgraduate years, Sylvie was a mentor to me, and let me choose a project in the Affective Computing field. She believed in my skills and offered me funding to make my Ph.D. experience productive and stimulating. She was very protective and supportive.

I would like to acknowledge with gratitude all the members of my thesis committee, especially the two examiners who read carefully the manuscript and gave me precious feedbacks.

Special greetings go for all the participants in the real-world driving experiments: the drivers, the engineers who developed the environmental prototype, and all my colleagues for sharing their previous experiences.

It was a pleasure to work in the CEA-LinkLab and to have Amira, Yosra, Rabâa, and Nader just next to me. We used to share everything from the most difficult moments to the happiest ones. We shared especially food and recipes of dessert. A special thanks goes to Amira for her huge help in signal processing and interpretation. Thanks also to Safa, Sana, Emna, Seddik and Soufiane. I would like also to thank all my friends in the LMO and U2S. Alternating my stays between the two labs allowed me to meet new people and to share our experiences.

If I survived all the paper work that was needed during my thesis, it was thanks to the staff in the U-PSUD doctoral school that facilitated all the operations. Many thanks to Rakia, Raja and Dhouha, from the ENIT doctoral school, for their help.

I am indebted to my family for their continuous support and love. Thanks mum, Lyna, Narjes, Habib, Idriss, Souhir and Nadhir.
A special acknowledgment goes to Chokri Ezzine who always encouraged me especially in the most difficult moments. And finally, my deep gratitude goes to Rafik who encouraged me and made me stronger.

Publications

Articles in peer-reviewed journal

- N. El Haouij, J.-M. Poggi, R. Ghozi, S. Sevestre-Ghalila, and M. Jaïdane, "Random forest-based approach for physiological functional variable selection for driver's stress level classification", Stat Methods Appl, 2018. https://doi.org/10.1007/s10260-018-0423-5
- N. El Haouij, R. Ghozi, J.-M. Poggi, S. Sevestre-Ghalila, and M. Jaïdane, "Self-similarity Analysis of Electrodermal Activity for Driver's Stress Level Characterization", submitted.

Article in proceedings of an international conference

- N. El Haouij, J.-M. Poggi, S. Sevestre-Ghalila, R. Ghozi, and M. Jaïdane. 2018. "AffectiveROAD System and Database to Assess Driver's Arousal State". In SAC 2018: SAC 2018: Symposium on Applied Computing, Track IoT, April 9-13, 2018, Pau, France. ACM, New York, NY, USA.

Conference presentations

- N. El Haouij, *et al.*, "Random Forest-Based Approach for Physiological Functional Variable Selection for Driver's Stress Level Classification", 16^{th} Annual Conference of the European Network for Business and Industrial Statistics (ENBIS), September 11-15, 2016, Sheffield, United Kingdom.
- N. El Haouij, *et al.*, "City versus highway driver arousal state analysis for stress level assessment", 48^{th} days of Statistics of Société Française de Statistique (SFdS), May 30-June 3^{rd}, 2015, Montpellier, France.
- N. El Haouij, *et al.*, "Driver Arousal State Complexity Assessment in Urban Spaces", International Congress of Cognition, Emotion and Motivation (CEM), November 4-7, 2015, Hammamet, Tunisia.
- N. El Haouij, *et al.*, "Feature extraction and selection of electrodermal reaction towards stress level recognition : two real-world driving experiences", 47^{th} days of Statistics of Société Française de Statistique (SFdS), June 1-5, 2015, Lille, France.

Open database

AffectiveROAD: "Affective Recognition Of the Arousal of the Driver" available immediately under request on elhaouij.nsk@gmail.com. The database will be soon available for a direct public download, on the CEA-LinkLab [1] website and on the bank of PhysioNet [2].

[1] In October 2017, a declaration acts the opening of the CEA Regional Office for Africa and the Middle East which will continue and carry the same research fields and projects of the CEA-LinkLab.

[2] https://physionet.org/physiobank/

Contents

1 Introduction : Contexte, motivation et problématique **9**

1.1 Évaluation de l'état émotionnel du conducteur dans un contexte de conduite en situations réelles . . . 10

1.2 Biosignaux pour la reconnaissance du niveau de stress du conducteur . . . 11

1.3 Description des bases de données : *Drivedb* et *AffectiveROAD* . . . 14

1.4 Sélection des variables fonctionnelles basée sur les forêts aléatoires . . . 15

1.5 Analyse de l'auto-similarité dans les signaux AED . . . 16

1.6 Principales contributions et plan du rapport . . . 17

1.6.1 Chapitre 3 : Approche basée sur les forêts aléatoires pour la sélection de variables fonctionnelles physiologiques pour la classification du niveau de stress du conducteur . . . 17

1.6.2 Chapitre 4 : Analyse de l'auto-similarité de l'activité électrodermale pour la caractérisation du niveau de stress du conducteur . . . 19

1.6.3 Chapitre 5 : *AffectiveROAD* : un système et une base de données pour évaluer l'état d'éveil du conducteur . . . 20

1.6.4 Chapitre 6 : Sélection des biosignaux pour l'évaluation du stress du conducteur et caractérisation de l'auto-similairité de l'activité électrodermale . . . 22

2 Introduction: Context, motivation, and problem statement **25**

2.1 Driver's affect assessment in urban spaces . . . 25

2.2 Biosignals for the driver's stress level recognition . . . 26

2.3 Databases description: *Drivedb* and *AffectiveROAD* . . . 29

2.4 Random-Forests for functional variable selection . . . 30

2.5 Self-similarity analysis of EDA signals . . . 31

2.6 Main contributions and report outline . . . 32

2.6.1 Chapter 3: Random Forest-Based Approach for Physiological Functional Variable Selection for Driver's Stress Level Classification . . . 32

2.6.2 Chapter 4: Self-similarity Analysis of Electrodermal Activity for Driver's Stress Level Characterization . . . 34

2.6.3 Chapter 5: *AffectiveROAD* System and Database to Assess Driver's Arousal State . . . 35

2.6.4 Chapter 6: Biosignals Selection and self-similar Characterization of the Electrodermal Activity for the Assessment of the Driver's Stress . . . 37

3 Random Forest-Based Approach for Physiological Functional Variable Selec-

tion for Driver's Stress Level Classification **39**
3.1 Introduction . . . 40
3.2 Experimental protocol and data collection . . . 42
3.2.1 Real-world driving protocol . . . 42
3.2.2 Cohort description . . . 43
3.2.3 Data Construction . . . 43
3.3 Random Forests and variables selection using variable importance . . . 45
3.3.1 Random Forests and Variable Importance measure . . . 45
3.3.2 Variable Selection using Random Forest-based Recursive Feature Elimination 46
3.4 The 3 steps of the variable selection approach . . . 47
3.4.1 Step 1. Wavelet decomposition of the physiological functional variables . . 48
3.4.2 Step 2. Physiological functional variable removal . . . 49
3.4.3 Step 3. Wavelet levels selection . . . 50
3.5 Variables selection on the *drivedb* database . . . 51
3.5.1 Step 1. Wavelet decomposition of the physiological functional variables . . 51
3.5.2 Step 2. Physiological functional variables removal . . . 51
3.5.3 Step 3. Wavelet levels selection from the three retained physiological variables 54
3.5.4 Performances of the proposed "blind" approach . . . 57
3.6 Discussion . . . 59
3.7 Conclusion . . . 60

4 Self-similarity Analysis of Electrodermal Activity for Driver's Stress Level Characterization **62**
4.1 Introduction . . . 63
4.2 Data and Methods . . . 64
4.2.1 Real-world driving protocol and data description . . . 64
4.2.2 Fractional Brownian Motion . . . 65
4.2.3 Estimation of Hurst exponent . . . 66
4.2.4 The actual procedure of Hurst exponent estimation . . . 67
4.3 Results and discussion . . . 68
4.3.1 EDA and scale invariant processes . . . 68
4.3.2 EDA-based "stress" level characterization . . . 69
4.3.3 Individual driver-based analysis . . . 70
4.3.4 Discussion . . . 71
4.4 Conclusion and perspectives . . . 72

5 AffectiveROAD System and Database to Assess Driver's Arousal State **74**
5.1 Introduction . . . 75
5.2 Overview of real-world driving experiences systems . . . 76
5.2.1 MIT system for driver's stress detection . . . 76
5.2.2 hciLab system for driver's workload assessment . . . 76
5.2.3 Warwick-JLR system for driver monitoring research . . . 76

5.3 AffectiveROAD sensors network . . . 77
5.3.1 Physiological Wearable Sensors . . . 77
5.3.2 Ambient Environment Sensors of AffectiveROAD . . . 78
5.4 Real-world driving protocol . . . 78
5.4.1 Apparatus . . . 79
5.4.2 Cohort Description . . . 80
5.4.3 Route Description . . . 81
5.5 Data Collection and Construction . . . 82
5.6 Related works and discussion . . . 83
5.7 Conclusion and future work . . . 84

6 Biosignals Selection and Self-similar Characterization of the Electrodermal Activity for the Assessment of the Driver's State of Stress 85
6.1 Introduction . . . 86
6.1.1 Problem Statement . . . 86
6.1.2 Related Work . . . 87
6.1.3 Chapter contributions and outline . . . 87
6.2 *AffectiveROAD* database description . . . 88
6.2.1 Real-world driving protocol . . . 88
6.2.2 Data collection . . . 89
6.2.3 Data preprocessing . . . 91
6.2.4 Subjective stress metric construction . . . 92
6.2.5 Questionnaire Analysis . . . 93
6.3 Methods . . . 95
6.3.1 The 3 step-based approach for variable selection . . . 95
6.3.2 Self-similarity analysis . . . 96
6.4 Driver's arousal state recognition based on variable selection using Random Forests 97
6.4.1 Data description . . . 97
6.4.2 Physiological variable selection and arousal level classification . . . 98
6.4.3 Bio-signals selection and arousal level classification . . . 103
6.4.4 A regression model of the subjective stress metric on the biosignals . . . 105
6.5 EDA self-similarity analysis . . . 107
6.5.1 EDA and scale invariant processes . . . 108
6.5.2 Driving experiment-based analysis . . . 108
6.5.3 Individual driver-based analysis: right vs left EDA . . . 109
6.6 Discussion . . . 110
6.7 Conclusion . . . 112

7 Conclusion and Perspectives 113
7.1 Summary and conclusions . . . 113
7.2 Perspectives and future works . . . 114

References **122**

Chapter

1

Introduction : Contexte, motivation et problématique

Cette étude vise à évaluer le niveau de stress du conducteur dans un contexte de conduite réelle dans des espaces urbains, en utilisant ses mesures physiologiques. En effet, il est important d'étudier l'état émotionnel de l'être humain et de déterminer sa relation avec la performance de la tâche de conduite et le bien-être du conducteur. Ceci peut servir pour les concepteurs d'espaces urbains, les constructeurs de véhicules et les services de transport. Ce chapitre introductif vise à présenter le contexte général de l'application et la motivation théorique de ce travail. On commence tout d'abord par expliquer les raisons pour lesquelles nous utilisons une base de données ouverte et nous créons également une nouvelle base de données, qui fournissent toutes les deux un ensemble de biosignaux, mesurés lors des expériences de conduite in-vivo sur différents conducteurs. Pour les deux bases de données, l'importance des signaux physiologiques a été estimée à l'aide de modèles basés sur les forêts aléatoires. Une méthode de sélection de variables fonctionnelles a été utilisée pour fournir un ordre d'importance et sélectionner les biosignaux les plus pertinents pour la classification du niveau de stress des conducteurs. L'activité électrodermale (AED) du conducteur est apparue comme l'une des variables fonctionnelles les plus importantes et "endurantes" à la procédure de sélection. Une étude préliminaire de l'auto-similarité de l'AED est développée, offrant une analyse de complexité de cet indicateur physiologique. Ce chapitre se termine par un résumé des principaux résultats et contributions de la thèse.

1.1 Évaluation de l'état émotionnel du conducteur dans un contexte de conduite en situations réelles 10

1.2 Biosignaux pour la reconnaissance du niveau de stress du conducteur 11

1.3 Description des bases de données : *Drivedb* et *AffectiveROAD* 14

1.4 Sélection des variables fonctionnelles basée sur les forêts aléatoires 15

1.5 Analyse de l'auto-similarité dans les signaux AED 16

1.6 Principales contributions et plan du rapport 17

1.1 Évaluation de l'état émotionnel du conducteur dans un contexte de conduite en situations réelles

Le monde connaît une urbanisation croissante puisque ses deux tiers vivront dans les villes d'ici 2030 [114]. Ainsi, des solutions appropriées doivent donc être développées en tenant compte de différents acteurs clés, tels que concepteurs des espaces urbains, constructeurs de véhicules et services de transport, afin d'améliorer la sécurité et le bien-être des conducteurs.
Le département américain des transports (USDOT) par exemple, a lancé un programme de systèmes de transport intelligents (ITS) pour résoudre les problèmes et les défis de l'intégration efficace et équitable de l'automatisation dans le système de transport USDOT-ITS. L'ITS et les technologies associées permettent l'intégration facile des nouvelles fonctionnalités automatisées dans les véhicules. Ces fonctionnalités permettent de réduire le nombre d'accidents provoqués par des erreurs humaines, d'augmenter l'autonomie des personnes handicapées et âgées pendant les voyages, de réduire la consommation d'énergie en réduisant à la fois la conduite agressive et le temps de trajet.
La Commission Européenne a également créé des programmes de financement tels que Horizon 2020 afin de favoriser et de soutenir la recherche en Europe. Les principaux objectifs de l'appel à projet "Smart, Green and Integrated Transport", proposé dans le cadre d'Horizon 2020, concernaient l'amélioration de la mobilité, de la sûreté et de la sécurité [75]. Ce programme encourage le développement d'approches centrées sur l'utilisateur afin d'accroître l'acceptation des nouvelles technologies et d'améliorer le bien-être de l'utilisateur du véhicule.
Rappelons que la conduite d'un véhicule est une tâche complexe qui nécessite une attention visuelle, des capacités physiques et un contrôle émotionnel. Avec une urbanisation croissante, le temps passé par le conducteur dans le véhicule augmente ; son bien-être doit donc être évalué [35]. En fait, l'administration nationale chargée de la sécurité routière aux États-Unis (NHTSA) a rapporté dans une enquête que la principale raison des accidents était attribuée au conducteur, dans 94% des cas. En effet, les distractions internes et externes, la surveillance inadéquate et l'inattention étaient les causes d'accidents les plus fréquemment signalées [58].
Les sources internes de distraction dans un véhicule augmentent avec la montée en puissance des systèmes embarqués et peuvent perturber la tâche principale de conduite. Ainsi, le domaine d'étude d'Interaction Homme-Machine (IHM) et l'Affective Computing ont cherché des moyens pour réduire la charge mentale afin d'optimiser le processus nécessaire pour utiliser une interface donnée ou même une machine [63, 123, 124]. Les sources externes de distraction sont les scènes auditives et visuelles dynamiques de la ville perçues par le conducteur.
Dans un contexte de conduite urbaine, il faut considérer trois acteurs principaux : la ville, le véhicule et le conducteur, qui sont en interaction et échange d'informations constants. Ces interactions ont une influence directe sur l'état émotionnel du conducteur. Par conséquent, il est important non seulement d'évaluer l'état du véhicule et de son environnement externe, mais aussi de suivre l'état du conducteur et les paramètres de conduite du véhicule.
Il est donc crucial d'identifier et d'utiliser des technologies non intrusives et portables afin d'évaluer l'état de vigilance, la charge mentale, le niveau de stress, d'attention et d'éveil. Ces éléments sont importants dans l'évaluation et la conception de solutions qui s'adaptent aux besoins du conducteur.
Dans cette thèse, nous abordons les changements de niveau de stress du conducteur car des niveaux extrêmes de stress nuisent à la performance de la conduite et peuvent provoquer un inconfort ainsi qu'un danger potentiel. Divers capteurs physiologiques mesurant la fréquence cardiaque (HR), la fréquence respiratoire (BR), l'électromyogramme (EMG) et l'activité électrodermale (AED) sont utilisés, ainsi que des évaluations subjectives pré-expérience et post-expérience de conduite.

1.2 Biosignaux pour la reconnaissance du niveau de stress du conducteur

Il existe un lien inhérent entre le niveau d'éveil, la vigilance, le "stress" et la charge mentale en exécutant une tâche donnée. En général, un état de "stress" survient lorsque la capacité de réponse n'est pas suffisante pour maîtriser les demandes attendues dans une situation donnée [106, 116]. Pendant la conduite d'un véhicule, de multiples facteurs cognitifs et émotionnels jouent un rôle clé dans la prise de décision relative à la sécurité et au confort du conducteur [53].
La conduite provoque un stress qui peut être dû à divers facteurs, tels que la sensibilité du conducteur et sa réaction aux stimuli environnementaux, comme les conditions routières, la circulation, les conditions météorologiques, le bruit ambiant, etc. Ceci s'ajoute à un niveau potentiel d'appréhension lié aux tâches nécessaires estimées pour atteindre une destination prédéfinie.
Au cours de la dernière décennie, la technologie mobile est devenue un nouveau facteur qui rend la tâche de conduite plus complexe et peut être considérée comme une nouvelle source de stress pour les conducteurs de véhicules [142]. On rappelle que les tâches de conduite primaires comprennent généralement la direction des roues, le freinage et l'accélération. Les tâches de conduite secondaires sont souvent liées à la sécurité et comprennent des actions sur la boîte de vitesses, les essuie-glaces et les clignotants. Par contre, le confort du conducteur relève de ce que l'on considère comme des tâches de conduite tertiaire : régler la climatisation, activer la radio et les appareils mobiles tels que les téléphones portables, les tablettes, le GPS [4], etc.
Au cours des dernières années, plusieurs outils et méthodes ont été proposés pour caractériser les niveaux de vigilance et de stress du conducteur. On notera en particulier les dispositifs mesurant le mouvement oculaire, les détecteurs de visage et de regard [64, 88]. Cependant, l'évolution actuelle des technologies sans fil et non-invasives a donné de l'importance aux signaux physiologiques pour une mesure in vivo - en situation en temps réel - de l'état émotionnel d'un conducteur. Différents descripteurs dérivés de divers signaux physiologiques ont été utilisés pour atteindre cet objectif, et cela dans différents contextes tels que la simulation de vols pour les pilotes [146], les trajets quotidiens alternant une conduite dans la ville et en autoroute [81] ou dans le contexte de conduite sur routes rurales [83].
Il est essentiel d'identifier les changements au niveau des biosignaux qui reflètent le niveau de stress du conducteur. De plus, la sélection des variables les plus pertinentes dans la détermination de l'état du conducteur est très importante. En effet, ce choix offre un moyen d'identifier la liste des capteurs physiologiques à prendre en compte. Il devient alors possible de fournir une évaluation objective du niveau de stress du conducteur. En plus du suivi des biosignaux, il est important d'inclure et d'étudier la corrélation entre les évaluations subjectives des changements d'état émotionnel du conducteur et les mesures objectives recueillies (à savoir les biosignaux). Ceci est généralement réalisé en utilisant des enquêtes, des évaluations vidéo et des études basées sur des mesures de performance de la tâche de conduite (telles que l'inclinaison de la voie de circulation et le changement de voie).
Dans le contexte d'expériences de conduite réelle, l'approche classique de reconnaissance des niveaux de stress repose sur la sélection des méthodes appropriées pour le prétraitement des données, l'extraction et enfin la fusion des descripteurs qui expliquent le mieux l'annotation subjective du stress. Dans la littérature, il est intéressant de noter que les descripteurs extraits sont basés sur les recommandations des experts (physiologistes) et sont donc appelés des *descripteurs experts*.
Dans cette thèse, nous profitons de la nature des données, en considérant ces biosignaux comme des variables fonctionnelles, puis en adaptant une approche existante de sélection des variables fonctionnelles pour finalement offrir une liste des variables physiologiques ordonnées selon leur importance ainsi qu'un modèle basé sur les forêts aléatoires permettant de classifier le niveau de stress. Une telle approche permet d'offrir une procédure automatique pour prédire l'état du conducteur. Il convient de noter que la méthode utilisée dans notre travail peut être appliquée à

tout type d'état du conducteur (comme la fatigue, la somnolence ou le mal du transport).
On s'est également intéressé dans cette thèse à la caractérisation des biosignaux par mesures fractales, souvent utilisée afin d'étudier la complexité de ces signaux pour décrire l'état de santé ou de maladie d'une personne [21, 48, 139]. Cette caractérisation est généralement basée sur des considérations physiques et physiologiques. Cependant, les propriétés fractales de l'AED n'ont pas été complètement explorées jusqu'à présent. Notre travail établira l'importance de l'AED dans l'analyse et le suivi du niveau de stress du conducteur et offrira une caractérisation préliminaire de l'auto-similarité des signaux AED.
Ci-dessous, nous donnons un aperçu des signaux physiologiques les plus utilisés dans l'analyse du niveau de stress des conducteurs.

Le rythme cardiaque (HR)

La rythme cardiaque (HR) est l'un des indicateurs de santé et d'activité physique, généralement égal ou proche du pouls palpable à n'importe quel endroit. Il mesure habituellement le nombre de battements cardiaques par minute. Il reflète l'activité cardiaque et est affecté par de nombreux facteurs. Le HR est contrôlé par le système nerveux autonome (ANS) [95, 135]. Comme l'activité de l'ANS est altérée pendant les périodes de stress et de fatigue, le HR a été à l'origine de plusieurs études sur les émotions dans divers domaines comme la médecine [29, 107] et le sport [20, 92, 112, 127].
Le HR est en général estimé à partir d'un capteur de détection du volume sanguin (BVP) ou déduit d'un électrocardiographe (ECG). Le capteur BVP, également connu sous le nom de photopléthysmographie-PPG, utilise une technologie basée sur l'émission de la lumière pour détecter le taux de sang. C'est un capteur non intrusif, généralement placé sur la peau et ne nécessitant pas de gel ou d'adhésif. L'ECG mesure l'activité cardiaque en utilisant des électrodes placées sur la peau qui détectent les tensions résultantes des battements cardiaques. Avant de placer les électrodes, la peau est préparée et nettoyée. Habituellement, les caractéristiques HR utilisées sont : les bandes d'énergie à basse fréquence (LF) et à haute fréquence (HF) de la variabilité HR. La bande LF correspond généralement à des plages de fréquence dans l'intervalle $(0-0.15Hz)$, tandis que la bande haute fréquence (HF) correspond à $(0.15-0.4Hz)$, ce qui est considéré d'origine parasympathique dans l'analyse classique du HR.

La fréquence respiratoire (BR)

La fréquence respiratoire (BR) fournit le moyen le plus simple et le plus facile à capturer pour estimer les variables respiratoires. Il est utilisé comme un indicateur des états émotionnels tels que le stress, l'éveil et la charge mentale [136]. Ces enregistrements précis sont basés sur des estimations d'échange de gaz à travers les poumons. Cette méthode est intrusive et gêne ainsi les activités de la personne telles que parler et conduire. Par conséquent, une mesure alternative est généralement utilisée en capturant l'expansion de la cavité thoracique. Cette expansion est mesurée par une ceinture thoracique : l'élastique est étiré par l'inhalation et renvoyé à l'état de base par l'expiration.
La fréquence de respiration a été utilisée dans différentes études comme une mesure de l'état émotionnel des conducteurs des voitures et des pilotes d'avion [15, 82, 100, 136]. Une respiration plus rapide est habituellement signalée comme étant causée par une activité émotionnelle et/ou physiologique inhabituelle. Une irrégularité dans les modèles de respiration peut être causée par des émotions négatives [59].
Dans l'étude de référence de Healey et Picard [81], six descripteurs ont été extraits du signal de respiration, à la fois dans le domaine temporel et fréquentiel. Pour le domaine temporel, la moyenne et la variance ont été calculées sur les signaux de fréquence respiratoire normalisés. L'étude de Healey et Picard [81] a rapporté que ces deux descripteurs reflètent le volume pul-

monaire au repos et la quantité globale de variation du signal de la fréquence respiratoire. Pour le domaine fréquentiel, les descripteurs de puissance spectrale représentant l'énergie dans chacune des quatre bandes ($0-0.1Hz$, $0.1-0.2Hz$, $0.2-03Hz$ et $0.3-0.4Hz$) ont été jugés importants dans la reconnaissance des émotions [125].

L'électromyogramme (EMG)

L'électromyogramme (EMG) mesure les signaux électriques musculaires produits pendant la contraction musculaire, appelés potentiels d'action musculaire. Ces signaux sont transmis via le tissu musculaire de la même façon que les signaux des nerfs traversent le corps. L'EMG surfacique est une méthode non intrusive utilisée pour capturer l'information contenue dans les potentiels d'action musculaire.
Habituellement, on utilise trois électrodes : deux d'entre elles sont placées dans le même axe du muscle d'intérêt et la troisième est placée hors de cet axe [33, 55].
L'EMG a été utilisé dans plusieurs études pour analyser les émotions traduites sur le visage [49, 87], la valence émotionnelle [79, 150] et le stress [80, 102, 103]. Dans cette dernière étude, seule la moyenne de l'EMG brute a été extraite. Cependant, l'EMG semble relativement peu important selon le classement empirique des mesures primaires pour mesurer le stress dans l'étude [142].

L'activité électrodermale (AED)

L'activité électrodermale (AED), également connue sous le nom de réponse galvanique cutanée ou de conductance cutanée, est une mesure de l'activité du système nerveux autonome (ANS). Elle reflète les changements survenant dans la résistance de la peau causée lorsque les glandes de la peau produisent une sueur ionique [23]. Les changements se produisant dans l'activité de la glande sudoripare produisant des changements dans la résistance de la peau peuvent être capturés en faisant passer un faible courant électrique à travers deux électrodes positionnées sur la surface de la peau.
L'AED est basée sur des processus physiologiques qui changent relativement lentement, en les comparant aux signaux physiologiques de fréquence plus élevée comme l'ECG, l'EMG ou l'électroencéphalogramme (EEG). L'AED a deux composantes principales : un niveau de base tonique (long terme) et des réponses phasiques (court terme) superposées au niveau tonique. L'utilisation croissante de l'AED dans l'évaluation de diverses émotions humaines [23, 54, 69], en tant qu'indicateur fiable de l'état émotionnel humain, nécessite une analyse spécifique de ce signal physiologique.
La réponse phasique est appelée aussi réponse électrodermale (RED). Le terme "startle" est généralement utilisé pour décrire une réponse à un stimulus spécifique. Il se produit lorsque le sujet est confronté à des changements dans l'environnement tels qu'un bruit soudain, un changement de luminosité, etc ou en raison d'un changement émotionnel interne, tel que l'attente ou l'anticipation d'un événement [23, 38]. La RED est généralement caractérisée par l'ensemble de descripteurs suivant :

- La latence (RED Lat.) quantifiant la quantité de temps entre le stimulus et le début de la réponse.
- La magnitude (RED Mag.) mesurant l'amplitude de la RED.
- Le temps de montée (RED Dur.) caractérisant la durée entre le début de la réponse et le pic de la réponse.
- L'aire contenue sous la RED approximée en général par un modèle triangulaire $area = \frac{1}{2} \times RED\ Mag. \times RED\ Lat.$

Des travaux antérieurs (par exemple [157]) ont étudié la dépendance des signaux AED à l'emplacement du capteur de mesure. Ils ont signalé une différence des mesures AED entre 16 emplacements choisis sur la peau humaine.

1.3 Description des bases de données : *Drivedb* et *AffectiveROAD*

Dans notre étude, deux bases de données ont été considérées : *drivedb* proposée par MIT et *AffectiveROAD* construite dans le cadre de notre travail. Ces deux bases de données sont décrites ci-dessous :

Drivedb [1] : Cette base de données a d'abord été proposée par une étude pionnière en 2005 par Healey et Picard [81]. Elle fournit un ensemble de plusieurs signaux physiologiques mesurés dans un contexte de conduite réelle dans la région de Boston. Les ensembles de données physiologiques mis à disposition sur le site web PhysioNet [66], sont ceux de 17 expériences de conduite. Plusieurs études ont utilisé cette base de données ouverte et bien annotée [8, 30, 39]. Les signaux physiologiques mis à disposition sont : l'activité électrodermale (AED) mesurée sur le pied et la main de chaque conducteur, l'électromyogramme (EMG), la respiration qui est aussi la fréquence respiratoire (BR) et le rythme cardiaque (HR). L'étude [81] a utilisé les mesures physiologiques afin de prédire les niveaux de stress des conducteurs lorsqu'ils passent de la conduite en ville à la conduite sur autoroute, avec des périodes de repos.
La métrique de stress considérée dans le cadre de *drivedb* est basée sur cette hypothèse : on suppose que la conduite en ville induit un niveau de stress élevé, que la conduite sur autoroute induit un niveau de stress moyen et que les périodes de repos correspondent aux faibles niveaux de stress. Cette hypothèse a été validée par une étude subjective basée sur des enregistrements vidéo de chaque expérience de conduite et une annotation faite par des experts, en plus des réponses subjectives des conducteurs recueillies via le questionnaire pré et post-expérience [81].
Il convient de noter que le terme stress, utilisé à l'origine pour *drivedb* et dans une partie de notre travail (à savoir les chapitres 2 et 3), fait référence au stress négatif. Une notion plus globale du stress est proposée plus tard dans ce travail.

AffectiveROAD : Cette base de données a été développée dans le cadre de ce travail de thèse. Inspiré par les expériences de conduite de *drivedb*, un protocole similaire a été défini, dans le contexte de la conduite dans les routes de la région du Grand Tunis. Les principales différences entre *drivedb* et *AffectiveROAD* résident dans le type de paysage urbain des routes parcourues et dans la métrique subjective de stress. En fait, les expériences ont été menées pendant l'été dans la ville de l'Ariana qui fait partie de la région du Grand Tunis. Ce protocole expérimental a été conçu pour considérer le triplet environnement-conducteur-véhicule. Pour cela, un prototype de plateforme de capteurs a été conçu. Il comprend plusieurs capteurs qui mesurent la température intérieure du véhicule, l'humidité, la pression et le niveau sonore ambiant. Deux Gopro [2] qui enregistrent les scènes visuelles et acoustiques internes et externes sont utilisées. Deux types de réseaux de capteurs physiologiques sont placés sur le conducteur : Empatica E4 [3] et la ceinture thoracique Zephyr [4] Bioharness 3.0. Ces appareils capturent l'AED sur les poignets droit et gauche, le HR (extrait à partir de BVP et de l'ECG), le BR, la température de la peau, la posture du conducteur et l'accélération du véhicule. Afin de garder la même configuration pour les deux bases de données, on a considéré une alternance entre conduite en ville et en autoroute

[1] http://physionet.org/physiobank/database/drivedb/
[2] https://shop.gopro.com/cameras
[3] https://www.empatica.com/research/e4/
[4] https://www.zephyranywhere.com/

avec deux périodes de repos de 15 minutes.
Une métrique de stress construite en temps réel est proposée par l'expérimentateur, qui annotait subjectivement le stress global perçu dans l'intervalle $[0, 1]$, 0 pour désigner un niveau faible de stress et 1 pour annoter un stress élevé. La notion de stress considérée dans cette base de données inclut le stress négatif, la charge mentale, l'attention, la complexité de l'environnement et le niveau d'éveil.
Un protocole de validation de cette métrique subjective consistait à présenter à chaque participant l'évolution de la métrique du stress, avec deux enregistrements vidéo (scènes internes et externes de l'expérience de conduite). Un curseur qui balaye la courbe de la métrique du stress est synchronisé avec les deux scènes visuelles. Cet ensemble est montré au conducteur à qui l'on demande de valider et de corriger les valeurs de la métrique, si nécessaire, en fonction de sa mémoire du stress perçu. Pour accélérer la validation du score subjectif, la vidéo avance rapidement et le conducteur est libre de l'arrêter lorsque cela est nécessaire pour vérifier un événement spécifique.

La base de données *AffectiveROAD* a un contexte particulier par rapport aux bases de données existantes. En effet, les expériences de conduite ont été menées au Grand Tunis, qui est situé dans un pays du sud de la Méditerranée où les températures sont connues pour être plus élevées que celles des pays situés dans la partie nord, pendant l'été. En outre, la manière de conduire dépend de la culture du conducteur. Ainsi, les notions de stress, de charge mentale, d'attention, de complexité de l'environnement et d'éveil ne peuvent pas être dissociées des habitudes liées à la culture locale. Par exemple, un piéton peut apparaître et traverser la route pendant que les feux rouges sont allumés. Cet événement peut être très stressant dans une autre culture, tout en étant perçu comme un événement assez commun pour les conducteurs tunisiens.

1.4 Sélection des variables fonctionnelles basée sur les forêts aléatoires

Cette thèse adapte des approches statistiques pour analyser les variables physiologiques. Elle vise à fournir une méthode qui permet de reconnaître l'état d'un conducteur véhicule, et ce sans connaissance a priori sur les données ni un accès aux informations qui caractérisent chaque signal. La variable d'interêt considérée dans notre étude est l'état du conducteur, plus précisement son niveau de stress. Cette variable est généralement créée en se basant sur une métrique subjective construite suite à une session d'annotation des vidéos collectées durant les expériences, déduite des questionnaires après l'expérience ou même des annotations fournies par des experts. Pour cela, nous adaptons ici des techniques d'apprentissage supervisé. Classiquement, des descripteurs basés sur une construction experte sont extraits à partir des signaux physiologiques. Par la suite, des méthodes de fusion de données sont utilisées afin de prédire le niveau de stress. Comme les signaux physiologiques sont échantillonés en terme de temps, des méthodes d'analyse de données fonctionnelles peuvent être considérées comme un outil naturel d'analyse pour ce type de données. Les données fonctionnelles sont souvent représentées en utilisant des fonctions de base. Plusieurs types de fonctions de bases existent tels que l'ensemble des polynômes, séries de Fourier, ondelettes, etc. [130]. Dans cette thèse, les fonctions de base d'ondelettes sont utilisées.

Les méthodes d'apprentissage supervisé visent à estimer le lien entre les variables explicatives $\mathbf{X} = (X^1, \ldots, X^p)$ (dans notre cas, $\mathbf{X}$ désigne les signaux physiologiques) et une variable d'intérêt Y (le niveau de stress). Il s'agit donc d'estimer une fonction mesurable $f : \mathbb{X} \to \mathcal{Y}$. L'estimation de cette fonction est basée sur l'échantillon $D_n = \{(\mathbf{X}_1, Y_1), \ldots, (\mathbf{X}_n, Y_n)\}$ des n i.i.d observations. Dans cette thèse, deux situations sont considérées : une classification lorsque l'on considère le niveau de stress ayant 3 classes : faible, moyen, élevé, donc $\mathcal{Y} = \{L, M, H\}$ et un problème de régression qui va être présenté dans le chapitre 5. Une métrique subjective continue de stress où $Y \in \mathbb{R}$ est proposée.

Nous basons notre étude sur une approche non paramétrique puisqu'il n'est pas simple de

décrire la distribution de probabilité $P_{(\mathbf{X},Y)}$ de $(\mathbf{X}, Y)$ en utilisant des méthodes paramétriques. L'algorithme des forêts aléatoires (RF) est sélectionné car il est utilisé pour résoudre les problèmes de classification et de régression. De plus, cette méthode non-paramétrique a montré de bonnes performances [161]. En outre, les RF offrent un sous-produit intéressant qui est l'importance des variables (VI) permettant de quantifier l'importance de chaque variable physiologique dans l'explication du niveau de stress du conducteur.

L'objectif principal de l'analyse des signaux physiologiques dans le contexte de conduite dans le monde réel est de déterminer l'ensemble parcimonieux de variables physiologiques offrant une reconnaissance adéquate du niveau de stress du conducteur. Par conséquent, la méthode utilisée doit fournir un moyen de sélectionner un modèle offrant une bonne performance de reconnaissance avec un nombre minimal de variables.
Nous avons opté pour une méthode d'élimination récursive des descripteurs basée sur les forêts aléatoires (RF-RFE) proposée dans [70]. L'approche bénéficie de la mesure de l'importance des variables (VI). Une variable est considérée comme importante, si en brisant le lien entre cette variable et la variable d'intérêt, l'erreur de prédiction augmente. Cette notion a été étendue dans [71] pour les groupes de variables et propose ainsi la VI groupée.
L'approche RF-RFE peut être résumée dans les étapes principales suivantes : une RF est construite, la VI (ou la VI groupée) et l'erreur de prédiction sont calculées. La variable la moins importante est supprimée. Ces étapes sont répétées jusqu'à ce que toutes les variables soient supprimées. Les variables retenues correspondent à celles contenues dans le modèle minimisant l'erreur de prédiction.

1.5 Analyse de l'auto-similarité dans les signaux AED

La physiologie fractale se base sur l'utilisation de modèles et de mesures qui caractérisent les propriétés d'invariance d'échelle, lors de la description et de la compréhension de la physiologie humaine. Elle est considérée comme un domaine bien établi dans différentes sciences théoriques et appliquées. La complexité du corps humain et sa dynamique ont été ainsi décrites en utilisant plusieurs mesures fractales : les structures anatomiques du corps humain, telles que les réseaux de neurones du cerveau ainsi que diverses dynamiques d'activités physiologiques, se sont révélées en effet être fractales [163]. Les séries temporelles correspondant aux intervalles inter-battements cardiaques, inter-respiratoires et inter-pas, ont été par exemple étudiées, révélant un comportement fractal clair [78, 118, 119]. Les mesures fractales sont considérées par ailleurs comme des indicateurs fiables de la santé physique humaine. Par exemple, les caractéristiques fractales ont été extraites du signal électromyogramme (EMG) [132] et de la série temporelle de fréquence cardiaque (HR) [21] pour caractériser et reconnaître les états de fatigue. L'analyse fractale des signaux d'électroencéphalogramme (EEG) a été utilisée pour distinguer différents niveaux de vigilance humaine [117]. Ces indicateurs fractals ont été utilisés pour caractériser l'état émotionnel humain [139].

Il convient de noter que malgré le fait que de nombreux signaux physiologiques humains ont été explorés en utilisant des méthodes fractales, l'analyse fractale de l'activité électrodermale (AED) reste lacunaire et assez récente. Par exemple, des indicateurs non linéaires tels que l'indice de Lyapunov, l'exposant de Hurst parmi d'autres mesures de "complexité", ont été estimés sur des signaux AED recueillis lors de séances de cinéma afin de détecter l'état émotionnel incluant bonheur, tristesse et peur [31]. Même si certaines caractéristiques fractales ont été calculées à partir de l'AED, aucun détail sur l'étude des propriétés d'auto-similarité de ce signal physiologique n'a été fourni.

Dans cette thèse, nous proposons une analyse fractale préliminaire du signal AED basée sur un processus auto-similaire afin de mieux comprendre les différents niveaux de stress du conducteur concernant le contexte de conduite réelle : la ville par rapport aux zones autoroutières. Pour cela,

nous avons adopté l'approche de la modélisation des signaux physiologiques proposée par Eke *et al.* [48], où une étape de classification des signaux est réalisée, avant de procéder à l'analyse fractale. Elle consiste à identifier d'abord si la série temporelle correspond à un modèle de Bruit Gaussien Fractionnaire (FGN) ou de Mouvement Brownien Fractionnaire (FBM).

Nous rappelons qu'un Mouvement Brownien Fractionnaire (FBM) est un processus gaussien auto-similaire ayant des incréments stationnaires, qui est caractérisé par un exposant H, où $0 < H < 1$. Il est largement utilisé dans plusieurs domaines dont la physique, la biologie, les télécommunications et la finance, entre autres. Sa large utilisation découle du fait qu'un seul paramètre, l'exposant de Hurst H, est nécessaire et capture les dépendances à court et long terme, la régularité du processus et ses propriétés fractales. En fait, un FBM devient moins irrégulier lorsque l'indice H passe de 0 à 1. Plus précisément, un FBM est décrit comme anti-persistant si ($0 < H < \frac{1}{2}$), chaotique quand ($H = \frac{1}{2}$) et persistant pour ($\frac{1}{2} < H < 1$) [5]. Nous notons que le processus formé par les incréments d'un FBM est un FGN.

Il convient de noter que les mesures AED correspondent visuellement aux trajectoires d'un processus FBM. Cela a été confirmé dans notre étude. Pour cela, l'exposant de Hurst H a été estimé. Plusieurs techniques d'estimation de l'exposant de Hurst existent : des méthodes ont été dvéloppées dans le domaine temporel (par exemple, l'analyse rescaled range), le domaine fréquentiel (le log-périodogramme, par exemple) ou dans le domaine temps-échelle (estimation basée sur les ondelettes) [44].

Dans ce travail, nous adoptons une méthode basée sur les ondelettes pour estimer l'indice de Hurst sur les signaux AED, puisque les ondelettes fournissent un cadre naturel pour représenter les processus invariants en échelles. En fait, l'analyse par ondelettes a été largement utilisée pour détecter des caractéristiques à différentes échelles. De plus, cette méthode présente des avantages statistiques et informatiques décisifs. Elle offre des estimations non biaisées même pour les données de petite taille [6].

1.6 Principales contributions et plan du rapport

Le reste du rapport de thèse est organisé en quatre chapitres principaux. Il convient de noter que chaque chapitre peut être lu séparément et indépendamment et donc certaines informations peuvent être redondantes. Dans ce qui suit, nous présentons brièvement le contenu de chaque chapitre, en présentant ses principaux objectifs, le cadre de collecte de données, la méthode et les outils utilisés dans l'analyse des données. Les résultats, leur interprétation et les extensions sont également présentés.

1.6.1 Chapitre 3 : Approche basée sur les forêts aléatoires pour la sélection de variables fonctionnelles physiologiques pour la classification du niveau de stress du conducteur

Ce chapitre adapte et utilise une approche basée sur les forêts aléatoires pour la sélection des variables fonctionnelles physiologiques afin de classifier le niveau de stress du conducteur dans le contexte des expériences de conduite dans des situations réelles. Ce chapitre a fait l'objet d'un article, publié dans la revue "Statistical Methods and Applications" [51].

La première analyse a été effectuée sur des données issues d'une base de données *drivedb* ouverte disponible sur le site PhysioNet, proposée par l'étude de Healey et Picard [81]. Les mesures physiologiques d'intérêt sont : l'activité électrodermale mesurée au niveau de la main et du pied gauche du conducteur, l'électromyogramme, la fréquence respiratoire et le rythme cardiaque, extraites des jeux de données complets relatifs aux dix expériences de conduite effectuées sur deux

types de routes (ville et autoroute).
La plupart des études qui utilisent les données de *drivedb* reposent sur l'extraction de descripteurs avant de procéder à la classification du niveau de stress des conducteurs dans des expériences de conduite dans des situations réelles. Cependant, notre étude où une approche d'analyse des variables fonctionnelles, couplée à une décomposition en ondelettes multi-résolution, est considérée comme l'entrée d'une procédure de classification basée sur les forêts aléatoires.
L'approche proposée consiste, dans un premier temps, à décomposer les variables fonctionnelles en utilisant la base d'ondelette de Haar. L'élimination des variables fonctionnelles non pertinentes a par la suite été effectuée en se basant sur un score d'endurance proposé, qui agrège les informations relatives à la fréquence de sélection et au rang des variables physiologiques. En fait, le score reflète la capacité d'une variable à persister (dans le sens d'être sélectionnée et bien classée) lors de l'application de la procédure de sélection de variables répétée 10 fois en utilisant l'approche RF-RFE et l'importance groupée des variables. Les niveaux d'ondelettes correspondants aux variables fonctionnelles retenues ont été classés en fonction du même score. Enfin, les niveaux d'ondelettes les plus "persistants" à la procédure de sélection ont été retenus.
Pour le cas étudié, les résultats de l'analyse suggèrent que les signaux d'électromyogramme (EMG) et de rythme cardiaque (HR) sont moins pertinents que les signaux relatifs à l'activité électrodermale et à la fréquence respiratoire. De plus, l'activité électrodermale (AED) mesurée sur le pied du conducteur s'est avérée plus pertinente que celle mesurée sur la main.

Le fait que l'AED apparaisse comme importante dans la classification du niveau de stress du conducteur est un résultat attendu puisque l'étude de Healey et Picard [81] a rapporté que les descripteurs liés à l'AED et au HR étaient les plus corrélés à la métrique du stress qui résulte des évaluations subjectives basées sur les enregistrements vidéo. De plus, les résultats de notre étude sont en accord avec des études plus récentes ([23, 122]), qui ont révélé l'importance de l'emplacement du capteur AED dans une expérience donnée. À cet égard, notre étude, qui comprenait à la fois l'AED main et pied, a montré que l'AED du pied est aussi importante, et même plus pertinente que celle de la main pour détecter les niveaux de stress des conducteurs dans des situations de conduite réelle.
La méthode utilisée fournit une procédure "aveugle" (puisqu'elle ne nécessite aucune information préalable) de classification du niveau de stress du conducteur. Le modèle considéré suite à cette analyse a permis d'obtenir une erreur proche de la performance de l'approche basée sur l'expert en terme d'erreur. De plus, l'approche utilisée propose de nouveaux descripteurs physiologiques basés sur les niveaux d'ondelettes sélectionnés.
Plusieurs indices de performance ont été évalués en utilisant des taux d'erreur fournis par les modèles sélectionnés. Pour cela, nous avons évalué les performances du modèle, en termes de taux de mauvaise classification, en ne considérant que les variables retenues. Nous avons ensuite effectué une sorte de validation croisée pour l'estimation de l'erreur du modèle, en tenant compte de la particularité de la cohorte : des conducteurs qui ont refait l'expérience de conduite plusieurs fois. Enfin, nous avons examiné les performances du modèle résultantes de l'approche "aveugle" proposé par ce travail, en les comparant aux résultats de l'approche basée sur des connaissances "expertes". Les résultats montrent que les performances des deux approches sont comparables.

Dans ce chapitre, la variable à expliquer a été construite en fonction du type d'itinéraire parcouru par le conducteur, elle dépend donc principalement de l'hypothèse en supposant que le niveau de stress augmente en ville et diminue lorsque le participant est au repos. Une telle hypothèse reste simplifiée car elle ne prend pas en compte plusieurs facteurs tels que l'état cognitif du conducteur, la charge mentale et l'anticipation de certaines situations. L'effet de la complexité de l'environnement visuel et auditif externe pourrait aussi affecter la performance de conduite ([84]), donc le niveau de stress du conducteur. Des informations supplémentaires peuvent être collectées en utilisant des mesures caractérisant l'état émotionnel du conducteur, la complexité de l'environnement de conduite et les informations du véhicule (voir [52]).

Résumé du chapitre

Méthodes :
- Approche pour la sélection de variables fonctionnelles basée sur RF-RFE et l'importance des variables groupées.

Base de données : Base de données *drivedb*.
Acquis du chapitre :
- Variables physiologiques sélectionnées.
- Descripteurs basés sur les ondelettes sélectionnés.
- Performances du modèle retenu qui sont proches des performances du modèle expert.

Perspectives du chapitre :
- Analyse plus approfondie des signaux AED (chapitre 3).
- Développement d'un protocole de conduite dans des situations réelles et acquisition des données à l'aide de plusieurs plateformes de réseaux de capteurs (chapitre 4).
- Adaptation et application des approches utilisées dans les chapitres 2 et 3 à la nouvelle base de données introduite au chapitre 4.

1.6.2 Chapitre 4 : Analyse de l'auto-similarité de l'activité électrodermale pour la caractérisation du niveau de stress du conducteur

Ce chapitre propose une analyse de la complexité de l'activité électrodermale (AED) qui est l'un des signaux physiologiques mesurés sur le conducteur. Ce signal a été retenu dans le modèle final qui classifie le mieux le niveau de stress du conducteur. L'AED a été trouvée dans la littérature comme étant fortement corrélée avec le stress, comparée à d'autres indicateurs physiologiques [23]. Notre propre analyse au chapitre 2 a confirmé cette conclusion. En effet, l'AED a la caractéristique unique d'être contrôlée uniquement par l'activité de la branche sympathique du Système Nerveux Autonome (SNA) connue pour être prépondérante dans les états de stress [23].

Au cours des dernières décennies, plusieurs études ont confirmé les propriétés fractales dans divers signaux physiologiques et ont démontré que cette caractérisation était étroitement liée à la santé physique et mentale de l'homme.

En effet, la physiologie fractale est un domaine bien établi [47, 48, 139]. Nous notons en particulier que certaines des premières études qui ont examiné les signaux de l'électroencéphalogramme (EEG) pendant le sommeil, ont révélé que les comportements chaotiques existent pour l'EEG [13]. Plusieurs méthodes ont été utilisées pour analyser le comportement non-linéaire de l'EEG [7, 120]. Les indicateurs de complexité résultants, tels que l'exposant de Hurst, l'exposant de Lyapunov et la dimension fractale, présentent de bons indicateurs des personnes normales et épileptiques [90]. De plus, ces indicateurs ont permis de reconnaître différents états mentaux [110]. Plus récemment, des études ont utilisé la théorie des fractales et du chaos pour étudier les activités cérébrales et diverses déficiences mentales [148, 167]. Plusieurs études supplémentaires ont révélé un comportement fractal inhérent dans les séries chronologiques EMG, HR et BR.

Ce chapitre examine comment le stress peut être capturé via une analyse fractale en se basant uniquement sur les enregistrements de l'activité électrodermale (AED), dans la conduite en situations réelles. Plus précisément, nous caractérisons l'AED via un processus auto-similaire. Pour cela, le Mouvement Brownien Fractionnaire (FBM), paramétré via l'exposant de Hurst H, est choisi pour modéliser les changements de l'AED dans un contexte de conduite réelle. Pour caractériser l'invariance d'échelle présente dans l'AED, le processus FBM et son exposant H, estimés via une approche basée sur les ondelettes, sont utilisés. Plus précisément, une sélection automatique de la gamme d'échelle est proposée afin de détecter la linéarité dans le diagramme en échelle logarithmique.

La procédure est appliquée aux signaux AED, extraits de la base de données publique *drivedb*,

capturés à l'origine sur le pied et la main des conducteurs lors d'une expérience de conduite réelle conçue pour évoquer différents niveaux d'éveil et de stress. L'exposant de Hurst estimé H offre une distinction dans les niveaux de stress lors de la conduite sur l'autoroute par rapport à la ville, avec une référence pour l'état de repos produisant un niveau de stress minimal. Plus précisément, les estimations de H diminuent lorsque la complexité environnementale augmente. De plus, presque toutes les valeurs estimées sont supérieures à 0.5, ce qui suggère que le signal AED a une dépendance à long terme. L'exposant de Hurst H estimé sur les signaux AED mesurés au niveau du pied permet une meilleure caractérisation de la tâche de conduite que l'AED mesuré au niveau de la main. Cette analyse d'auto-similarité capture la complexité du signal AED. Une telle analyse a été appliquée à divers signaux physiologiques dans la littérature mais pas à l'AED ; un signal qui s'est avéré être le plus corrélé avec l'état émotionnel de l'être humain.

Cette étude a suggéré que l'AED mesuré au niveau du pied est fiable dans le suivi du niveau de stress pendant la conduite. Cette découverte est cohérente avec les études qui recommandent un examen attentif de l'emplacement des capteurs AED et de la tâche en question [23, 122]. Nous notons que plusieurs autres méthodes peuvent être utilisées pour estimer l'indice H. L'effet de la taille de fenêtre des segments AED mérite d'être mieux étudié. Bénéficiant du fait que les ondelettes sont faciles à mettre en oeuvre avec une faible complexité de calcul, l'approche proposée peut être facilement intégrée dans le véhicule et pourrait servir à faire le suivi en temps réel de l'état d'éveil du conducteur.

Résumé du chapitre

Méthodes :
- Processus auto-similaire pour caractériser le signal physiologique.
- Approche basée sur les ondelettes pour estimer l'exposant de Hurst.
- Sélection automatique de la gamme d'échelles d'ondelettes où l'invariance d'échelle de l'AED a été observée.

Base de données : Base de données *drivedb*.

Acquis du chapitre :
- L'AED peut être caractérisée en utilisant le processus de Mouvement Brownien Fractionnaire et l'exposant de Hurst correspondant (H).
- H permet une distinction entre les niveaux de stress de conduite en ville, en autoroute et au repos.
- Identification des gammes d'échelles d'ondelettes où l'invariance d'échelle de l'AED a été observée.

Perspectives du chapitre :
- Vérification de l'analyse de l'auto-similarité de l'AED sur un nouveau jeu de données, introduit et décrit dans le chapitre 4 et recueilli dans un contexte de conduite réelle similaire. Cette analyse sera détaillée dans le chapitre 5.

1.6.3 Chapitre 5 : *AffectiveROAD* : un système et une base de données pour évaluer l'état d'éveil du conducteur

Ce chapitre présente une nouvelle base de données, appelée *AffectiveROAD*, inspirée de *drivedb*, qui résulte des expériences de conduite réelle menées dans la région du Grand Tunis. Le travail de ce chapitre a été publié dans la section "Internet of Things" (IoT) du 33^{eme} Symposium ACM/SIGAPP on Applied Computing [52].

Plusieurs méthodes ont été utilisées dans le domaine de la conduite automobile pour évaluer l'état du conducteur. Ces méthodes peuvent être soit subjectives, basées sur le retour de l'utilisateur, soit objectives en utilisant les mesures physiologiques de l'utilisateur. Évaluer l'état mental et émotionnel du conducteur dans un contexte réel n'est pas une tâche facile. Pour pallier

à cela, plusieurs solutions basées sur l'utilisation du simulateur existent. Ainsi, les simulateurs offrent un cadre sûr et peu coûteux pour l'entraînement des pilotes d'avion et des pilotes automobiles. Cependant, ce contexte reste biaisé dans l'évaluation de l'état émotionnel d'un conducteur de véhicule dans un environnement complexe (ville et autoroute).

Avec le développement des technologies de capteurs portables, l'évaluation de l'état d'un conducteur dans les situations réelles est devenue plus facile. Malgré le fait que plusieurs études ont exploré des expériences de conduite dans le monde réel, l'accès aux données collectées n'est pas facile.
Dans le contexte des véhicules intelligents, la surveillance de l'état émotionnel humain doit considérer le contexte et le cadre d'étude afin de prendre en compte les interactions entre le conducteur, son véhicule et son environnement ambiant. Dans ce chapitre, nous concevons et développons une plateforme multi-capteurs qui mesure les changements physiologiques humains, les indicateurs environnementaux ambiants à l'intérieur du véhicule et la vitesse du véhicule.
La plateforme proposée *AffectiveROAD* fournit non seulement un suivi physiologique en temps réel, mais enrichit également les outils actuellement disponibles pour le suivi des états émotionnels et cognitifs humains. Grâce à cette plateforme, plusieurs indicateurs d'état de conduite tels que le stress et l'éveil peuvent être développés et validés. Deux types de capteurs physiologiques sans fil sont utilisés pour mesurer l'activité électrodermale, le rythme cardiaque, la température de la peau, la respiration et le mouvement du conducteur. De plus, nous avons développé un réseau de capteurs permettant de mesurer la température ambiante, l'humidité, la pression et la luminosité. La vitesse du véhicule est extraite des données GPS (Global Position System) capturées à l'aide d'un smartphone. Deux appareils GoPro sont utilisés pour capturer les scènes internes et externes.
Le but de ce chapitre est de décrire un protocole de conduite dans des situations réelles pour collecter des données en utilisant les outils à base de dispositifs de l'Internet des Objets et d'annoncer la publication d'une base de données pour le suivi de l'état du conducteur. Nous proposons 13 jeux de données liés aux expériences de conduite effectuées dans différents types de routes : ville et autoroute. De plus, une partie importante de la base de données couvrant les données physiologiques et environnementales est mise à la disposition du public. L'utilisation de cette plateforme peut être utile dans l'évaluation de l'état du conducteur, en particulier pour les situations qui peuvent nécessiter plus d'éveil et d'attention. Cela peut aider essentiellement au développement d'interfaces axées sur le besoin de l'être humain et à l'amélioration des systèmes d'interaction homme-machine (HMI) existants. De plus, nous fournissons aux constructeurs de véhicules une telle plateforme qui peut être utilisée pour réaliser de nouvelles expériences dans un contexte de conduite réelle, fournissant ainsi aux scientifiques une base de données ouverte qui peut être facilement explorée et utilisée.
La plateforme proposée peut être étendue en utilisant les protocoles de connectivité pour centraliser la collecte de données à partir de capteurs physiologiques (via Bluetooth), de capteurs environnementaux (intégrés aujourd'hui dans des véhicules intelligents) et d'informations du véhicule issues du bus CAN. En outre, un module GPS peut être branché dans la plateforme afin de remplacer le smartphone.
Le trajet, proposé dans le contexte d'une conduite réelle, est conçu en alternant différents types de routes, supposés provoquer différents niveaux de stress. Il est donc important d'étudier l'effet de type des routes sur le niveau de stress du conducteur et de vérifier si l'hypothèse, concernant les différents niveaux de stress provoqués par les types de route, reste vérifiée pour une expérience de conduite différente. Une telle analyse peut être effectuée en utilisant la métrique continue du stress basée sur l'annotation des enregistrements vidéo.

Résumé du chapitre

Méthodes :
- Conception et conduite du protocole expérimental réalisé dans un contexte de conduite réelle dans la région du Grand Tunis.
Base de données : Nouvel ensemble de mesures physiques et physiologiques comme une nouvelle base de données *AffectiveROAD*.
Acquis du chapitre :
- Base de données accessible au public, y compris les biosignaux, les données environnementales et la métrique subjective du stress.
- Prototype de plateforme de réseau de capteurs qui mesure les paramètres environnementaux à l'intérieur du véhicule.
Perspectives du chapitre :
- Application des méthodes utilisées (au chapitre 2 et au chapitre 3) sur les ensembles de données nouvellement collectés d'*AffectiveROAD*.

1.6.4 Chapitre 6 : Sélection des biosignaux pour l'évaluation du stress du conducteur et caractérisation de l'auto-similairité de l'activité électrodermale

Ce dernier chapitre de la thèse présente l'analyse et les résultats de la reconnaissance du niveau de stress du conducteur en utilisant les données nouvellement collectées de la base de données *AffectiveROAD*. Pour cela, nous adoptons la sélection des variables fonctionnelles et les approches d'analyse fractale précédemment présentées et appliquées à la base de données *drivedb* de PhysioNet.
Des études empiriques ont établi la relation U inversée, connue sous le nom de Yerkes-Dodson law [166], reliant le niveau d'attention (ou d'éveil) à l'exécution d'une tâche donnée. En effet, la performance augmente avec l'éveil, atteignant un point optimal puis décline lorsque le niveau d'éveil devient trop élevé. Ce point de performance optimale dépend de facteurs individuels, de la capacité fonctionnelle humaine et de la complexité de l'environnement de travail. Lors de l'exécution d'une tâche de conduite de véhicule, les niveaux d'éveil sous-optimaux peuvent entraîner un risque plus élevé d'accidents. Par conséquent, l'état émotionnel du conducteur devrait être évalué. Des protocoles pour obtenir un tel état ont été développés en utilisant plusieurs biosignaux dans deux configurations : expériences de conduite réelle [81, 98, 133] et dans des configurations en utilisant un simulateur [28, 129, 169]. Dans le contexte de conduite réelle, il est crucial de déterminer la liste des biosignaux du conducteur et l'information physique de l'environnement de conduite (interne et externe au véhicule). Cela aidera dans la sélection de capteurs physiques et physiologiques afin d'évaluer et de comprendre l'état émotionnel du conducteur, en termes de plusieurs variables de contrôle. La conception, la collecte et la construction d'une variable d'intérêt, reflétant l'état émotionnel du conducteur, est une étape critique dans toute étude qui vise la reconnaissance cognitive et émotionnelle humaine. Habituellement, des hypothèses sur différents niveaux de stress au cours d'une expérience de conduite donnée sont mises en avant, puis vérifiées en utilisant une évaluation subjective, une annotation d'expert ou des enregistrements vidéo.
Par exemple, l'étude de Healey et Picard [81] propose une métrique de stress négatif basée sur le type de route en supposant que la conduite en ville est plus stressante que la conduite sur autoroute (comparée à une période de repos).
La base de données *AffectiveROAD* construite dans ce travail, fournit une nouvelle approche pour construire la métrique subjective de stress basée sur une évaluation en temps réel de la complexité de l'environnement et la perception de l'expérimentateur de l'éveil et du stress négatif des conducteurs [52].

Dans ce chapitre, nous présentons deux analyses principales des biosignaux recueillis en vue de mieux comprendre le stress du conducteur, à partir de la base de données *AffectiveROAD*. Pour cela, nous considérons l'activité électrodermale mesurée sur les poignets droit (AED droit) et gauche (AED gauche) du conducteur, la fréquence respiratoire (BR), le rythme cardiaque (HR), la posture (Post) et la métrique subjective du stress. Une première analyse visant à sélectionner les principaux prédicteurs des niveaux de stress du conducteur est réalisée en utilisant l'approche de sélection de variables. Ceci a été réalisé en utilisant une décomposition en ondelettes des biosignaux indiqués, qui sont ensuite traités par l'algorithme d'élimination récursive des descripteurs basée sur les forêts aléatoires. Pendant la phase de classification du niveau de stress, le modèle sélectionné conserve trois niveaux d'ondelettes de l'AED et de la posture du corps du conducteur (Post). De plus, en utilisant la métrique de stress décrite au chapitre 4, un modèle de régression sélectionne également 3 niveaux d'ondelettes de AED droit du conducteur et de la posture corporelle. La deuxième analyse principale de ce chapitre a examiné comment le stress, dans la conduite réelle dans le contexte d'*AffectiveROAD*, est capturé via une analyse de l'auto-similarité de l'activité électrodermale (AED). Une invariance d'échelle des signaux correspondants aux segments AED du poignet gauche et droit a été observée. En fait, en utilisant l'approche par ondelettes pour estimer l'exposant de Hurst, une linéarité a été observée dans tous les diagrammes à échelle logarithmique pour les niveaux d'ondelettes allant de 1 à 5. Ceci conforte le résultat dans [50] puisque les niveaux trouvés ici présentant la linéarité correspondent aux mêmes gammes de fréquence trouvées dans leur travail. Les résultats suggèrent que l'AED présente le comportement d'un processus persistant auto-similaire avec tendance. En effet, les valeurs estimées de l'exposant de Hurst pour l'AED mesurée au poignet droit et gauche sont supérieures à 0.5.
La distinction du niveau de stress entre la conduite dans la ville et en autoroute n'a pas été clairement vérifiée en gardant le même comportement de H comme dans le cas de l'AED de *drivedb*. Lorsque l'on considère l'AED mesurée au poignet droit, les résultats obtenus peuvent s'expliquer par des événements inhabituels (comme répondre à un appel téléphonique pendant la période de repos où le conducteur devait se reposer et se détendre ou rouler le long d'une autoroute encombrée). L'exposant de Hurst estimé peut être ainsi considéré comme un bon indicateur de l'état global qui combine le niveau d'éveil, le stress négatif, la charge mentale et la perception de la complexité de l'environnement.
Nous notons que l'exposant de Hurst peut être considéré comme un descripteur pertinent pour le suivi de l'état émotionnel global du conducteur, qui comprend l'éveil, le stress négatif et même la charge mentale, parmi autres facteurs.

Résumé du chapitre

Méthodes :
- Approche pour la sélection des variables fonctionnelles basée sur RF-RFE et l'importance des variables groupées.
- Processus auto-similaire pour caractériser le signal de l'AED mesuré sur les poignets gauche et droit.

Base de données : Base de données *AffectiveROAD*.
Acquis du chapitre :
- Biosignaux sélectionnés.
- Descripteurs basés sur les ondelettes sélectionnés.
- Plage raisonnable d'échelles où l'invariance d'échelle est observée lors de la caractérisation de l'AED avec un processus auto-similaire.

La dernière partie du manuscrit fournit un résumé des travaux de la thèse en mettant en avant certaines contributions concernant l'analyse des données et les bases de données utilisées. Les

recherches en cours et les perspectives de l'acquisition et de l'analyse des données sont également décrites dans le dernier chapitre.

Chapter

2

Introduction: Context, motivation, and problem statement

The framework of this study focuses on vehicle driver's stress level assessment in urban space, using physiological measurements. Indeed, there is an inherent need to study the human affective state in this context, for urban space designers, vehicle manufacturers and transportation departments and to analyze its relationship with task performance and overall well-being. This introductory chapter aims to present the general context of the application and the theoretical motivation behind this work. The choice to use an open database and the creation of a new one, both providing a collection of biosignals, is explained. For both databases, the order of importance of the physiological signals was estimated using Random Forest-based models. A functional variable selection method was used to provide an order and to select the most reliable biosignals in the stress level recognition. The driver's Electrodermal activity (EDA) emerged as one of the most enduring functional variable. A preliminary EDA self-similarity study is developed, offering a complexity analysis of this physiological indicator. This chapter concludes with a summary of the thesis main results and contributions.

2.1 Driver's affect assessment in urban spaces 25
2.2 Biosignals for the driver's stress level recognition 26
2.3 Databases description: *Drivedb* and *AffectiveROAD* 29
2.4 Random-Forests for functional variable selection 30
2.5 Self-similarity analysis of EDA signals 31
2.6 Main contributions and report outline 32

2.1 Driver's affect assessment in urban spaces

According to the United Nations, two thirds of world will live in cities by 2030 [114]. In fact, the world is witnessing an increasing urbanization and thus appropriate solutions must be developed, taking into account various factors and key players, such as urban space designers, vehicle manufacturers and transportation departments, in order to enhance drivers' safety and well-being.
For instance, the U.S. Department of Transportation (USDOT) launched an Intelligent Transportation Systems (ITS) program to address problems and challenges towards safe, efficient, and

equitable integration of automation into the transportation system [115]. The ITS and related technologies allow the easy and smooth incorporation of the new automated features into the vehicles. Such features enable to reduce the number of crashes caused by human error, to increase the autonomy of impaired and elderly people during travels, to reduce the energy consumption when mitigating the aggressive driving, and to reduce the travel time [43]. The European Commission created also funding programs such as the Horizon 2020 in order to foster and to support the research in Europe. The main objectives of the European "Smart, Green and Integrated Transport" challenge, proposed as a part of the Horizon 2020, concerned the mobility, safety and security improvement [75]. Such program encourages to develop user-centric approaches in order to increase the new technologies acceptance and to enhance the well-being of the vehicle user.

We recall that vehicle driving is a complex task that requires visual attention, physical abilities, emotional control and cognition. With the increasing urbanization trend, the time spent by the driver in the vehicle is increasing. Thus, his/her well-being should be assessed and addressed [35]. Even though safety has been known as a key element in a driving performance evaluation, the driver's affective state remains more fundamental not only in safety considerations but also in the overall driving experience quality. In fact, the National Highway Traffic Safety Administration (NHTSA) reported in a survey, that the main reason of accidents was assigned to the driver, in 94% of the total crashes. Indeed, the internal and external distractions, the inadequate surveillance, and the inattention were found to be the most frequently reported causes of crashes [58].
The vehicle internal sources of distraction, that may impair the driving primary task, increase with the rise of today's vehicle embedded systems. In fact, the Human-Computer Interaction (HCI) and Affective Computing fields of study have been working towards ways of reducing the human cognitive workload in order to optimize the process needed to use a given interface or a machine [63, 123, 124]. When driving, the external distractions refer to the city dynamic auditory and visual scenes perceived by the driver.

In an urban driving context, one should consider the three main actors which are the city, the vehicle and the driver, which are in constant interactions and information exchange. These interactions have a direct influence on the driver's affective state. Therefore, it is important to assess, via a network of sensors, not only the inside and outside vehicle conditions, but also to track the driver's state, in addition to the vehicle driving parameters.
It is thus crucial to identify and use non-intrusive and wearable technologies to assess the driver's state of alertness, mental workload, stress level, attention, and affect arousal, which are important in the assessment and design of user-centric solutions.

In this thesis, we address the driver's stress level changes since extreme levels of stress impair the driving task performance and can induce not only discomfort, but also can lead to unsafe driving. Various physiological sensors capturing the Heart Rate (HR), Breathing Rate (BR), Electromyogram (EMG), and Electrodermal Activity (EDA) are used, along with pre-experience and post-experience subjective evaluations.

2.2 Biosignals for the driver's stress level recognition

There is an inherent link between the level of arousal, alertness, "stress", and mental workload during the performance of any task. In general, a state of "stress" arises when the responses capacity are not sufficient to master the expected demands in a given situation [106, 116]. During vehicle driving, multiple cognitive and emotional factors play key role in decision making in relation with the safety and comfort of the driver [53].
Driving induces stress that can be attributed to various factors, such as the driver's own inherent susceptibility, his/her reaction to the environment stimuli, such as road conditions, traffic, weather, timing of the day, ambient noise, etc. This is an addition to a potential level of apprehension related to foreseen or estimated tasks needed to reach a preset destination. Over the

past decade, the mobile technology was found to be a new factor that complicates oftentimes the driving task and could be a new source of stress for vehicle drivers [142].
It worth recalling here that primary driving tasks typically include wheel steering, braking, accelerating. While secondary driving tasks are often related to safety, and include actions on gearbox, windscreen wipers, and turn signals. The driver's comfort falls into what is considered as tertiary driving tasks, for example: regulating the air conditioner, activating radio and mobile devices such as cell phones, tablets, GPS [4].

In recent years, a multitude of tools and methods were proposed for capturing and characterizing driver's alertness and stress levels. We note in particular eye movement monitors, facial and gaze detectors [64, 88]. Audio and visual smart human-vehicle reactions have been also used as assessment tools [40]. However, current evolution of wireless and portable technologies have brought back the importance of physiological signals in capturing, in-vivo and in real-time configuration, the affective state of a driver, for instance. Metrics derived from diverse physiological signals have been used to achieve this goal, and this in different contexts such as simulation of flights for pilots [146], normal daily commutes in the city and highway [81], and real-world driving in rural roads [83].
The ability to identify the biosignals changes that reflect the driver's stress level is of essence. Moreover, selecting the most relevant variables is more important; since this choice offers a way to identify the list of physiological sensors to be considered. This would provide an objective assessment to the driver's stress level. In addition to tracking the biosignals, it is important to include and to study the correlation between the subjective evaluations of the driver's affective state changes and the collected objective measures (namely the biosignals). This is usually achieved using surveys, video ratings, and task performance-based studies (such as road lane inclination and lane changing).

In the context of real-world driving experiments, the classical stress level recognition approach relies on the selection of the appropriate tools for data preprocessing and features extraction and fusion, that best match the subjective annotation. Throughout the literature, it is worth noting that the extracted features are based on the field expert (physiologists) recommendations, and are thus called expert-based features.
In contrast, in this thesis, we benefit from the nature of data, by considering these biosignals as functional variables, and then adapt an existing functional variable selection approach to ultimately provide us with an importance ranking, a list of the selected most relevant physiological variables, and a Random Forest-based model providing driving stress level recognition. Ultimately, this will offer an automated procedure to predict the driver's state. It should be noted that the used method in our work can be applicable to all types of driver's affective state (for instance: fatigue, drowsiness, motion sickness), thus not limited only to stress.
We recall that the fractal measures have been widely used to characterize the biosignals in order to study the complexity of such signals and to describe the healthy state and the illness [21, 48, 139]. This characterization is based usually on physical and physiological consideration. However, the fractal properties of the EDA have not been fully explored so far. Our work will establish the importance of EDA in driver's stress analysis and will offer a preliminary self-similarity characterization of the EDA recordings.
Below, we give an overview of the most used physiological signals in drivers' stress level analysis.

Heart Rate (HR)

The heart rate (HR) is one of the most basic health and physical activities indicator, usually equal or close to the pulse that can palpated at any peripheral place. It usually measures the number of heart beats per minute. It reflects the heart activity and is affected by many factors. The HR is controlled by the Autonomic Nervous System (ANS) [95, 135]. Since the ANS activity is altered during stress, and fatigue periods, the HR, as a main feature revealing the ANS activity,

was the basis in several emotion studies [49] in various fields (for instance medicine [29, 107] and sport [20, 92, 112, 127]).
The HR is usually estimated from a Blood Volume Pulse (BVP) sensor or deduced from an Electrocardiograph (ECG). The BVP sensor, also known as a photo-plethysmography-PPG, uses a light-based technology to sense the rate of blood. It is non-intrusive sensor, usually placed on the skin and does not require any gel or adhesives. The ECG measures heart activity using electrodes placed on the skin that detect voltages resulting from the heart beats. Before placing the electrodes, the skin is prepared and cleansed.
Usually, the used HR features are: the energy at low frequency (LF) and high frequency (HF) bands of HR variability. The LF band corresponds generally to ranges $(0 - 0.15Hz)$, while the High-frequency (HF) band corresponds to $(0.15 - 0.4Hz)$, which is considered of parasympathetic origin in classical HRV analysis.

Breathing Rate (BR)

The Breathing Rate (BR) provides the simplest way to estimate the respiratory variables since it is the easiest one to capture. It is used as an indicator of emotional states such as stress, arousal and mental workload [136]. Its accurate recordings is based on estimates of gas exchange through the lungs. This method is intrusive and thus inhibits real-world activities such as talking and driving. Therefore, an alternate measure is generally used by capturing the chest cavity expansion. This expansion is captured via a chest belt: the elastic is stretched by the inhalation and returned to the baseline state by the exhalation.
The breathing rate was used in different studies as a measure of the driver's and pilot's affective state [15, 82, 100, 136]. Faster respiration is usually reported as caused by emotional or/and physiological activity and irregularity in respiration patterns can be caused by negative emotions [59].
In the reference study of Healey and Picard [81], six features were extracted from the respiration signal in both time and frequency domain. For the time domain, the mean and the variance computed on the normalized breathing rate signals. The study of Healey and Picard [81] reported that those two features reflect the resting lung volume and the overall amount of variation in the breathing rate signal. For the frequency domain, spectral power features representing the energy in each of the four bands $(0-0.1Hz$, $0.1-0.2Hz$, $0.2-03Hz$, and $0.3-0.4Hz)$ were found important in the emotion discrimination [125].

Electromyogram (EMG)

The electromyogram (EMG) measures the muscle electrical signals produced during muscle contraction called muscle action potentials. These signals are transmitted via muscle tissue similar to the way nerves signals travel through the body. Surface EMG is a non-invasive method used to capture the information contained in the muscle action potentials.
Usually, we use 3 electrodes, two of them are placed in the same axis of the muscle of interest and the third one is placed out of this axis [33, 55].
The EMG has been used in several studies to analyze the emotions from the face [49, 87], the emotional valence [79, 150] and the stress [80, 102, 103]. In this last study, only the mean of the raw EMG was extracted. However, the EMG seems to be relatively not important according to the empirical ranking of primary measures for measuring stress in the review study [142].

Electrodermal Activity (EDA)

The Electrodermal Activity (EDA) also known as the galvanic skin response (GSR) or skin conductance (SC) is a measure of the Autonomic Nervous System (ANS) activity. It reflects changes

occurring in the skin resistance caused when skin's glands produce ionic sweat [23]. Changes occurring in the sweat gland activity producing changes in skin resistance can be captured by passing a low electrical current across two electrodes positioned on the surface of the skin.
The EDA is based on relatively slow changing physiological processes, compared to higher frequency signals like ECG, EMG or Electroencephalogram (EEG). The EDA has two main components: a tonic (long-term) baseline level and short term phasic responses superimposed on the tonic level. The increasing use of the EDA in assessing various human emotions [23, 54, 69], as a reliable indicator of human affect arousal in various applications, calls for a specific analysis of this physiological signal.
The phasic response is called also Electrodermal response (EDR) or skin conductance startle response. The term startle response is usually used to describe a response to a specific stimulus. It occurs when the subject is facing changes in surrounding environment (eg. a sudden sound, change in lighting, etc.) or due to an internal thought process, such expectation or anticipation [23, 38]. The EDR has been represented by the following set of features:

- Latency (EDR Lat.) quantifying the amount of time between the stimulus and the onset of the response.
- Magnitude (EDR Mag.) measuring the amplitude of the EDR.
- Rise time (EDR Dur.) characterizing the duration between the onset of the response and the peak of the response rise.
- EDR area contained under the response approximated here by a triangular model $area = \frac{1}{2} \times EDR\ Mag. \times EDR\ Lat.$

Previous works (for instance [157]) studied the dependence of the EDA recordings and the location of measurement. They reported a difference of the EDA measurements between the 16 locations across the body.

2.3 Databases description: *Drivedb* and *AffectiveROAD*

In our study, two databases have been considered: *drivedb* proposed by MIT and *AffectiveROAD* constructed as a part of this study, which are presented below:

Drivedb[1] **database**: This database was firstly proposed by a 2005 pioneer study of [81], made available in PhysioNet [66] website. It provides a collection of different physiological signals gathered a real-world driving context in Boston area. The physiological sets of data made available is that of 17 driving experiences (called drives). Several studies were based on this open and well-annotated database [8, 30, 39]. The physiological signals made available are: the Electrodermal activity (EDA) measured on the foot and the hand of each driver, the electromyogram (EMG), the respiration (referred to RESP in next chapter) which is also the breathing rate (BR), and the Heart Rate (HR). The study [81] used the physiological measures in order to classify the drivers stress levels as they switched between city and highway driving, with rest periods.
The stress metric included in *drivedb* database is based on the following: it is assumed that city driving induces high stress level, highway driving induces medium stress level and rest periods (no driving) correspond to the low stress levels. This assumption was validated by a subjective study based on video recordings of each drive and expert annotation, in addition to drivers feedback collected via pre and post experience questionnaire [81].
It should be noted that the term stress, originally used for *drivedb*, and in part of our work (namely Chapters 2 and 3), refers to the distress (negative stress). A more comprehensive use of

[1] http://physionet.org/physiobank/database/drivedb/

stress is proposed later in this work.

***AffectiveROAD* database**: This database was developed as a part of this thesis work. Inspired by the *drivedb* experiments, a similar protocol was defined, in the context of real-world city versus highway driving in Grand Tunis area. The main differences between the *drivedb* and *AffectiveROAD* lie in the type of urban landscape of the routes traveled, and the subjective stress metric. In fact, the experiments were conducted during summer time in the city of Ariana which is part of Grand Tunis area. Ultimately, this experiment protocol was designed to sense the triplet environment-driver-vehicle. For that, a prototype of sensors-based platform was designed and implemented. It includes several sensors that capture the inside-vehicle temperature, humidity, pressure and ambient sound level. Two Gopro that record the internal and external visual and acoustic scenes are used. Two types of physiological sensor networks are placed on the driver namely: Empatica E4 and Zephyr Bioharness 3.0 chest belt. These devices capture the EDA on the right and the left wrists, the HR (from BVP and ECG), the BR, the skin temperature, the driver posture and the vehicle acceleration. In order to be able to compare the results of our developed driving experiments on the two databases, an alternation between city and highway driving and two 15 minutes rest periods, were introduced.

A real-time stress metric was attributed by the driver's companion (experimenter), who annotates subjectively the perceived overall stress in the continuum $[0, 1]$, 0 for no stress, and 1 for extreme stress. We consider here that the annotated "stress" includes distress, workload, attention, perceived environment complexity and affect arousal. A validation protocol of this subjective metric consists on presenting to each participant the evolution of the stress metric with two video recordings (internal and external scenes simultaneously playing). A sweeping of the stress metric and the corresponding two visual scenes are shown side by side to the driver who is asked to validate and to correct the rating, if necessary, based on his/her memory recollection of the perceived overall stress. To speed up the validation of the subjective score, the video is fast forward and the driver is free to stop when necessary to check for specific event.

The constructed *AffectiveROAD* database has a particular context compared to the existing databases. In fact, the driving experiments were conducted in Grand Tunis, which is located in a south Mediterranean country where temperatures are known to be higher than those located in the northern side, during summer. In addition, the driving culture is different from a Tunisian driver to a foreigner. Indeed, the concepts of distress, workload, attention, environment complexity and affect arousal can not be dissociated when driving, and especially in the context of Tunisian driving. This is mainly due to the acquired habits related to the local culture. For instance, a pedestrian can surprisingly appear crossing the road while the red lights are on. This event can be very stressful in another culture, while perceived as a normal event for Tunisian drivers.

2.4 Random-Forests for functional variable selection

This thesis adapts statistical approaches for physiological variables analysis. It seeks ultimately to provide a method allowing to recognize the driver's state, without prior information or knowledge about the data.
The variable of interest considered in our study is the driver's state, precisely, his/her stress level. This variable is usually created based on subjective metric such as video rating sessions, post experience questionnaires and sometimes expert-based annotations. Thus, supervised statistical learning techniques can be considered as appropriate tools here.
Classically, expert-based features are extracted from different physiological signals and methods of data fusion are used in order to predict the driver's stress level. Since physiological signals are time sampled data, functional data analysis methods can be considered in our study as a natural tool for data analysis. Functional data are usually represented using basis functions. Several

types of basis function exist such as the collection of polynomials, Fourier series, wavelets, etc. [130]. In this thesis, the wavelet basis functions are used.

Supervised learning methods aim to estimate the link between explanatory variables $\mathbf{X} = (X^1, \ldots, X^p)$ (in our case, $\mathbf{X}$ denotes the physiological signals) and a variable of interest Y (the stress level). It consists thus to estimate a measurable function $f : \mathbb{X} \rightarrow \mathcal{Y}$. The estimation of such function is based on sample $D_n = \{(\mathbf{X}_1, Y_1), \ldots, (\mathbf{X}_n, Y_n)\}$ of n i.i.d observations. In this thesis, we are faced to two problems: classification when considering the stress level having 3 classes: Low, medium, high, thus $\mathcal{Y} = \{L, M, H\}$ and regression problem as will be presented in Chapter 5. A continuous subjective stress metric where $Y \in \mathbb{R}$.

We base our study on a non-parametric approach since it is not simple to describe the probability distribution $P_{(\mathbf{X},Y)}$ of $(\mathbf{X}, Y)$ using parametric methods. Random-Forests (RF) algorithm is selected since it is used in resolving both: classification and regression problems. Moreover, this non-parametric method was shown to offer great performances [161]. In addition, the RF offers an interesting byproduct which is the Variable Importance (VI) allowing to quantify the importance of each physiological variable in the explanation of the driver's stress level.

One main objective, in physiological signals analysis in the context of real-world driving experience, is to determine the parsimonious set of physiological variables offering an adequate recognition of the driver's stress level. Therefore, the used method should provide a way to select model offering a good recognition performance with a minimum number of variables. We opted for the use of Recursive Feature Elimination based on Random Forests (RF-RFE) proposed in [70]. It benefits from a main important feature offered by the RF, which is the Variable Importance (VI), proposed firstly in [25]. A variable is considered as important, if when breaking the link between this variable and the variable of interest, the error of prediction increases. This notion was extended in the work of [71] for the groups of variables, and called the grouped VI.
The RF-RFE approach can be summarized in the following main steps: a RF is built, the VI (or the grouped VI) and the prediction error are computed. The less important variable is removed. These steps are repeated until all variables are removed. The retained variables correspond to those contained in the model minimizing the prediction error.

2.5 Self-similarity analysis of EDA signals

Fractal physiology focuses on the use of patterns and measures that characterize the scale invariance properties, when describing and understanding human physiology. It is considered as a well-established field in different theoretical and applied sciences. The complexity of the human body and its dynamics were described using a wide range of fractal measures. In fact, the human body anatomical structures such as the brain neural networks as well as various dynamics of physiological activities were proved to be fractals [163].

For instance, the time series corresponding to the heart inter-beats, inter-breaths, and inter-strides intervals, have been extensively studied revealing clear fractal behavior [78, 118, 119]. In addition, the fractal measures are considered as a reliable indicators of human physical health. These fractal indicators were used to characterize the human affective state [139]. For instance, fractal-based features were extracted from Electromyogram (EMG) signal [132] and Heart Rate (HR) time series [21] to characterize and recognize states of fatigue. Fractal analysis of Electroencephalogram (EEG) signals was used to distinguish different human vigilance levels [117].

It is worth noting that despite the fact that wide range of human physiological signals were explored using fractal methods, the Electrodermal activity (EDA) remains of timid focus. For instance, nonlinear indicators such as Lyapunov, Hurst exponent among other "complexity" measures, have been estimated on EDA signals collected during movie watching sessions in order to detect the affective state including happiness, sadness and fear [31]. Even though some fractal

features were computed from the EDA, no details on the study of the self-similarity proprieties of this physiological signal were provided.

In this thesis, we propose a preliminary fractal analysis of the EDA signal based on a self-similar process in order to further understand the different stress levels of the driver regarding the real-world driving context: city versus highway areas. For that, we have adopted the approach of physiological signal modeling proposed by Eke *et al.* [48], where a step of signal classification is performed, before proceeding to the fractal analysis. It consists on identifying first if the time series fits a Fractional Gaussian Noise (FGN) or Fractional Brownian Motion (FBM) models.

We recall that a Fractional Brownian Motion (FBM) is a Gaussian self-similar process with stationary increments, which is characterized by an exponent H, where $0 < H < 1$. It is widely used in several fields which have included physics, biology, telecommunications and finance, among others. Its wide use rises from the fact that only one parameter, the Hurst exponent H, is needed and captures the short and long range dependencies, the process regularity, and its fractal properties. In fact, a FBM gets less irregular as the H index increases from 0 to 1. More precisely, a FBM is labeled as anti-persistent if $(0 < H < \frac{1}{2})$, chaotic when $(H = \frac{1}{2})$ and persistent for $(\frac{1}{2} < H < 1)$ [5]. We note that the incremental process of a FBM is a FGN.

It should be noted that the EDA measurements correspond visually to FBM sample paths. This has been formally confirmed in our study. For that, the Hurst exponent H should be estimated. Several techniques of Hurst exponent estimation exists, methods have been performed in time domain (for instance, the rescaled range analysis), frequency domain (log-periodogram, for example) or in a combined time-scale domain (wavelet-based estimation) [44].

In this work, we adopt a wavelet-based method to estimate the Hurst index on EDA signals, since wavelets provide a natural framework for representing scale invariant processes. In fact, wavelet analysis has been widely used to detect features at different scales. In addition, this method presents optimal statistical and computational advantages. It offers unbiased estimates even for data of small length [6].

2.6 Main contributions and report outline

The remainder of the thesis report is organized into four main chapters. It should be noted that each chapter can be read separately and independently, and thus some repetition maybe encountered. In what follows, we provide a brief presentation of the content of each chapter, by introducing its main objectives, the data collection framework and the method and tools used in the data analysis. The results, their interpretation and extensions are presented as well.

2.6.1 Chapter 3: Random Forest-Based Approach for Physiological Functional Variable Selection for Driver's Stress Level Classification

This chapter adapts and uses a Random-Forest-based approach of functional variable selection, to provide physiological variable selection for real-world driver's stress level classification. It was the focus of an original paper, published in the Statistical Methods and Applications Journal [51].

Our analysis is performed on experimental data extracted from *drivedb*: an open database available on PhysioNet website, which is a subset of the original data constructed and used in Healey and Picard study [81]. The physiological measurements of interest are: electrodermal activity captured on the driver's left hand and foot, electromyogram, breathing rate, and heart rate, extracted from complete data sets related to ten driving experiments carried out in two types of routes (city and highway) and rest area.

Most studies using *drivedb* data relied on features extraction for drivers stress level classifi-

cation in real-world driving experiments. This is in contrast with our study where a functional variable approach, coupled with a multi-resolution wavelet decomposition, are considered as the input of a Random Forest-based classification procedure.
The developed approach consisted, at a first stage, in functional variables decomposition using the Haar wavelet. The irrelevant functional variables were then removed based on a proposed "endurance" score, which aggregates the information related to the selection frequency and the rank of the physiological variables. In fact, it reflects the variable ability to persist (in the sense of being selected and well ranked) when applying the variable selection procedure 10 times, based on RF-RFE and grouped variable importance. The corresponding wavelet levels of the retained functional variables were ranked based on the same score. Finally, the most "enduring" wavelet levels to the selection procedure were retained.
For the considered case of study, the analysis results suggest that the Electromyogram (EMG) and the heart rate (HR) signals are less relevant when compared to the Electrodermal and the breathing rate signals. Furthermore, the Electrodermal (EDA) activity measured on the driver's foot emerged as more relevant than that captured on the hand.

The fact that the EDA emerges as important in driver's stress level recognition is expected since Healey and Picard original study reported that features related to EDA and HR were the most correlated to the stress metric extracted from video-based subjective evaluations [81]. Furthermore, our findings are in alignment with more recent studies ([23, 122]), which revealed the importance of the EDA sensor placement in a given experiment. In that regard, our study, which included both Hand and Foot EDA, has indicated that the Foot EDA is as important and may be even more relevant in detecting stress levels of the drivers in real-world driving situations. The used method provides a "blind" procedure (since it does not require any prior information) of driver's stress level classification that resulted in a model error close to the expert-based approach performance. In addition, the used approach suggested new physiological features based on the wavelet levels. Moreover, the proposed approach offered a ranking of physiological variables according to their importance in driver's stress level classification in city versus highway driving, with restful period as a baseline reference.

Several performance indexes were evaluated using different error rates. For that, we assessed the model performances, in terms of misclassification rate, with regards to the retained variables in the classification of the driver's stress level. We performed then a cross-validation like-procedure for model error estimation, taking into account the particularity of the cohort used in the experiments. Finally, we examined the performances of our model using the proposed "blind" approach, as it compares with the results from the "expert-base" approach. The results show that our blind approach performances were comparable (i.e. within the same range) to the expert based ones.

In this chapter, the variable of interest was built according to the type of route traveled by the driver, thus it depended mainly on the hypothesis assuming that the stress level increases when driving in the city and decreases when the participant is at the rest. Such hypothesis remains simplified since it does not take into account several factors such as the driver's cognitive state, mental workload, and anticipation of some situations. The effect of the external visual and auditory environmental complexity could as well affect the driving performance ([84]) thus the driver's stress level. Additional information can be gathered using measurements characterizing the affective state of the driver, the driving environment and the inside vehicle complexity (see [52]).

Chapter summary

Method:
- Approach for functional variable selection based on RF-RFE and grouped Variable Importance.
Database: *drivedb* database.
Main outcomes:
- Selected physiological variables.
- Selected wavelet-based features.
- Performances of the retained model close to the performances of the expert-based model.
Follow-up tasks:
- Deeper analysis of EDA signals (Chapter 3).
- Development of real-world driving protocol and an experimental setup using a dedicated sensor networks-based platform for data acquisition (Chapter 4).
- Adapt and apply the approaches used in Chapters 2 and 3 to the new database introduced in Chapter 4 (focus of Chapter 5).

2.6.2 Chapter 4: Self-similarity Analysis of Electrodermal Activity for Driver's Stress Level Characterization

This chapter focuses on a "complexity" analysis of the Electrodermal Activity (EDA) which is one of the physiological signals measured on the driver in the open database *drivedb*, and was retained in the final model that best classify the driver's stress level. The EDA was found in the literature to be highly correlated with stress, compared with other physiological indicator [23]. Our own analysis in Chapter 2, has confirmed this finding with specific EDA preferential ranking. In fact, EDA has the unique characteristic of being controlled solely by the activity of the sympathetic branch of the Autonomic Nervous System (ANS) known to be preponderant in stress states [23].

During the past decades, several studies have confirmed the fractal properties in a variety of physiological signals and such characterization was shown to be in close link with human physical and mental health. Indeed, fractal analysis has been a well-established field in biomedical sciences [47, 48, 139]. We note in particular that some of the first studies which examined the Electroencephalogram (EEG) recordings, during sleep, revealed that chaotic behaviors exist for the EEG [13]. Several methods were used to analyze the nonlinear behavior of human EEG [7, 120]. The resulted complexity indicators, such as Hurst exponent, Lyapunov exponent, and fractal dimension presented good discriminators of normal and epileptic persons [90]. In addition, these indicators allowed to recognize various mental states [110]. More recently, studies used fractal and chaos theory to investigate brain activities and various mental impairments [148, 167]. Several additional studies revealed an inherent fractal behavior in EMG, HR, BR time series.

This chapter investigates how the stress in real-world driving can be captured via a fractal analysis, relying only on the Electrodermal activity (EDA) recordings. Specifically, we characterize the EDA via a self-similarity process. For that, the Fractional Brownian Motion (FBM), parameterized via the Hurst exponent H, is evoked to model the EDA changes in a real-world driving context. To characterize the EDA scale invariance, the FBM process and its corresponding exponent H, estimated via a wavelet-based approach, are used. Specifically, an automatic scale range selection is proposed in order to detect the linearity in the log-scale diagram.

The procedure is applied to the EDA signals, extracted from the public *drivedb* database, captured originally on the foot and the hand of the drivers during a real-world driving experiment designed to evoke different levels of arousal and stress. The estimated Hurst exponent H offers a distinction in stress levels when driving in highway versus city, with a reference to restful

state of minimal stress level. Specifically, the estimated H decreases when the environmental complexity increases. In addition, almost all the estimated values are greater than 0.5 suggesting that the EDA signal has a long-range dependence. Furthermore, the H estimated on the Foot EDA signals allows a better characterization of the driving task than the Hand EDA. This self-similarity analysis captures the complexity of the EDA signal. Such analysis was applied to various physiological signals in literature but not to the EDA, a signal which was found to correlate most with human affect.

This study suggested that the Foot EDA is reliable in monitoring the stress level while driving. This finding is consistent with the studies that recommend a close examination of EDA sensor placement and the subject task under exploration [23, 122]. We note that several other methods can be used to estimate H index. It is worth re-examining the effect of EDA window size processing. Benefiting from the fact that the wavelets are easy to implement with low computational complexity, the proposed approach can be easily integrated in vehicle and could serve for "real-time" monitoring of a driver's arousal state.

Chapter summary

Method:
- Self-similar process to characterize physiological signal.
- Wavelet-based approach for the Hurst exponent.
- Automatic wavelet scale range selection where the scale invariance of the EDA was found.

Database: *drivedb* database.
Main outcomes:
- The EDA can be characterized using Fractional Brownian Motion process and its corresponding Hurst exponent (H).
- H distinguished city from highway driving stress levels, with rest periods.
- Identification of the scale invariance ranges of the EDA.

Follow-up tasks:
- EDA self-similarity analysis verification on new sets of EDA measurements, introduced and described in Chapter 4, and were collected in a similar real-world driving context. This analysis will be detailed in Chapter 5.

2.6.3 Chapter 5: *AffectiveROAD* System and Database to Assess Driver's Arousal State

This chapter introduces a newly constructed database, labeled *AffectiveROAD*, inspired by *drivedb*, but set up in real-world driving in Grand Tunis area. The work of this chapter has been published in the track "Internet of Things" of 33^{rd} ACM/SIGAPP Symposium On Applied Computing [52].

Several methods have been explored in the automobile driving field to assess the driver's state. These methods can be either subjective, based on the user's reporting or objective via the user's physiological measurements. Assessing the driver's mental and affective state in a real-world context is not an easy task. To mitigate this, several solutions relying on the use of simulator have been used. In fact, simulators provide a low-cost and safe framework for the training of pilots and car racers. However, it remains very limited in assessing the "true" affective state of a vehicle driver in a complex environment (City and Highway), as is the focus of this work.

With the expansion and affordable cost of wearable sensors technologies, assessing a driver's state in real-world became easier. Despite the fact that several studies explored real-world driving experiments, the access to the collected data is not readily available.
In the context of smart vehicles, human affective state monitoring should be based on a context-

aware system in order to consider the interactions between the driver, his vehicle and his ambient environment. In this chapter, we design and develop a multi-sensors platform, that captures the human physiological changes, the ambient environmental indicators inside the vehicle, and the vehicle speed. Such platform is labeled *AffectiveROAD*.

The proposed *AffectiveROAD* platform does not only provide real-time physiological monitoring, but also enriches the currently available tools for human affective and cognitive states tracking. Thanks to this platform, several driver's state indicators such as stress and arousal may be developed and validated. Two types of wireless physiological sensors are used to monitor the Electrodermal activity, the heart rate, the skin temperature, the respiration, and the motion of the driver. Moreover, we developed a sensor network allowing to capture the ambient temperature, humidity, pressure, and luminosity. The vehicle speed is extracted from the Global Position System (GPS) data captured using a smartphone. Two GoPro devices are used to capture the internal and external scenes.
The purpose of this chapter is to describe a real-world driving protocol to collect data using the proposed IoT-based materials and to announce the publication of a database for driver's state monitoring research. We propose 13 datasets related to drives in different road types: city and highway. Moreover, the important part of database covering the physiological and the environmental data is released for public use.

The use of such platform may be useful in driver's state assessment in particular for the situations that may need more awareness and attention. Especially, this may help in the development of Human-centric IoT-based interfaces and in the enhancement of the existing Human-computer interaction (HCI) systems. Furthermore, we provide vehicle manufacturers with such a platform that may be used to conduct more real-world experiments, thus providing scientists with an open database that can be easily explored and used.
The proposed platform can be extended by using the connectivity protocols in order to centralize the data collection from physiological sensors (via Bluetooth connections), environmental sensors (integrated nowadays in smart vehicles or use the proposed environmental sensors) and vehicle CAN information. In addition, a GPS module can be plugged into the device in order to pass away from smartphone.
The route, proposed in the context of a real-world driving, is designed by alternating different types of roads, assumed to induce different stress levels. Therefore, it is important to study the effect of road types on the driver's stress level and to check if the assumption, concerning the different stress levels induced by the route types, still verified for a different driving experience. Such analysis can be performed using the developed video-based continuous stress metric in *AffectiveROAD* database.

Chapter summary

Method:
- Design and field test of experimental protocol conducted in real-world driving context in Grand Tunis Area.
Database: New set of physical and physiological measurements in *AffectiveROAD* database
Main outcomes:
- Publicly available database including biosignals measurements, environmental data and subjective stress metric.
- A prototype of sensors network-based platform that captures the environmental parameters inside the vehicle.
Follow-up tasks:
- Apply the used methods (in Chapter 2 and Chapter 3) on the newly collected datasets of *AffectiveROAD*.

2.6.4 Chapter 6: Biosignals Selection and self-similar Characterization of the Electrodermal Activity for the Assessment of the Driver's Stress

This last chapter of the thesis presents the analysis and results of driver's stress level recognition using the newly collected data of *AffectiveROAD* database. For that we adopt the functional variable selection and the fractal analysis approaches previously presented and applied to *drivedb* database of PhysioNet.

Empirical studies have established the inverted-U-relationship, known as Yerkes-Dodson law [166], relating the level of attention (or arousal) to the performance of a given task. In fact, performance increases with arousal, reaching an optimal point and then declines when the level of arousal become too high. This point of optimal performance depends on individual factors, human functional capacity, and the task environment complexity. When performing a task of vehicle driving, suboptimal arousal levels may result in a higher risk for accidents. Therefore, the affective state of the driver should be assessed. Protocols to elicit such state have been developed using several biosignals in two configurations: real-world driving experiences [81, 98, 133] and simulator-based settings [28, 129, 169]. In real-world driving contexts, it is crucial to determine the list of biosignals of the driver and physical information of the driving environment (internal and external to the vehicle). This leads to guidelines for the selection of physical and physiological sensors in order to assess and understand the driver's affective state, in terms of several control variables.

The design, collection and construction of a variable of interest, reflecting the driver's affective state, is a critical step in any study on human cognitive and emotion recognition. Usually, assumptions on different stress levels during a given driving experience are put forth, then verified using subjective evaluation, expert annotation and video recordings.
For instance, the study of Healey and Picard [81] proposed a distress metric based on the road type assuming that city driving is more stressful than highway driving (compared with restful period created especially to establish lowest stress level).
The constructed *AffectiveROAD* database in this work, provides a unique approach to a subjective stress metric based on a real-time evaluation of the environment complexity, and the experimenter perception of the evoked drivers' arousal and distress [52].

In this chapter, we present two main analyses of the collected biosignals toward a broader understanding of driver's stress, based on the publicly made "AffectiveROAD" database. For that, we consider the Electrodermal activity measured on the right (RightEDA) and the left (LeftEDA) wrists of the driver, the breathing rate (BR), the heart rate (HR), the posture (Post) and the developed subjective stress metric.

A first main analysis aiming to select the main predictors of the stress levels of the driver, is performed, using the variable selection approach. This was achieved using a wavelet decomposition of the indicated biosignals, which are then processed through the Random-Forests recursive feature elimination algorithm. During the stress level classification stage, the selected model retains three wavelet levels of the Right EDA and Driver's body posture (Post), with a 31% error rate of misclassification averaged over 100 iterations. Moreover, using the described stress metric (in Chapter 4), a regression model selects also 3 wavelet levels of the driver's Right EDA and body Posture with a root mean square error of 20%, average over 100 iterations.

The second main analysis of this chapter investigated how the stress, in real-world driving in the context used in *AffectiveROAD*, is captured via a self-similar analysis of Electrodermal activity (EDA). A scale invariance of both left and right wrist EDA segments was observed. In fact, when using the wavelet-based approach to estimate the Hurst exponent, a linearity was observed in all the log-scale diagrams for the wavelet levels ranging from 1 to 5. This supports the finding in [50] since the levels found here exhibiting the linearity corresponds to the same frequency ranges found in their work.

The results suggest that the EDA exhibits the behavior of a self-similar persistent process with trend. Specifically, the estimated values of the Hurst exponent for both RightEDA and LeftEDA are greater than 0.5.
The stress level distinction between City and Highway driving was not clearly verified by a lower H index in city versus that of highway, as was the case in the *drivedb* EDA sets. When considering the RightEDA, the results obtained can be explained by the unusual events (such as answering a phone call during the rest period when he/she was supposed to rest and relax or driving along a congested highway), that occurred during each drive and the estimated Hurst exponent can be considered thus as a good indicator of the overall stress that combines the affect arousal, the distress, the workload and the environment complexity perception.
We note that the Hurst exponent can be considered as a relevant feature for the monitoring of the driver's affective state at large, which includes the arousal, the distress and even the workload, among other factors.

Chapter summary

Method:
- Approach for functional variable selection.
- Self-similar process to characterize physiological signal of EDA on left and Right wrists.

Database: AffectiveROAD database.
Main outcomes:
- Selected biosignals.
- Selected wavelet-based features.
- Wavelet ranges of the scale invariance when characterizing the EDA with self-similar process

The last part of the report provides a summary of the work undertaken, including highlights on some contributions on the data analysis, and the used databases. On going researches and emerging perspectives on the data acquisition and analysis are outlined.

Chapter

3

Random Forest-Based Approach for Physiological Functional Variable Selection for Driver's Stress Level Classification

Abstract This chapter deals with physiological functional variables selection for driver's stress level classification using random forests. Our analysis is performed on experimental data extracted from the *drivedb* open database available on PhysioNet website. The physiological measurements of interest are: electrodermal activity captured on the driver's left hand and foot, electromyogram, respiration, and heart rate, collected from ten driving experiments carried out in three types of routes (rest area, city, and highway). The contributions of this work touch on the method as well as the application aspects. From a methodological viewpoint, the physiological signals are considered as functional variables, decomposed on a wavelet basis and then analyzed in search of most relevant variables. On the application side, the proposed approach provides a "blind" procedure for driver's stress level classification, giving close performances to those resulting from the expert-based approach, when applied to the *drivedb* database. It also suggests new physiological features based on the wavelet levels corresponding to the functional variables wavelet decomposition. Finally, the proposed approach provides a ranking of physiological variables according to their importance in stress level classification. For the case under study, results suggest that the electromyogram and the heart rate signals are less relevant compared to the electrodermal and the respiration signals. Furthermore, the electrodermal activity measured on the driver's foot was found more relevant than the one captured on the hand. Finally, the proposed approach also provided an order of relevance of the wavelet features.

3.1 Introduction 40
3.2 Experimental protocol and data collection 42
3.3 Random Forests and variables selection using variable importance 45
3.4 The 3 steps of the variable selection approach 47
3.5 Variables selection on the *drivedb* database 51
3.6 Discussion 59
3.7 Conclusion 60

This chapter written in collaboration with J.-M. Poggi, R. Ghozi, S. Sevestre-Ghalila and M. Jaïdane, is published [51].

3.1 Introduction

The goal of this study is to classify stress levels of drivers in real-world driving situations. To this end, a random forest-based method is used for the selection of physiological functional variables captured during driving sessions. First, we present the "affective computing" context of our work. Then, we give an overview of previous studies that used the *drivedb* database of physiological measurements. Subsequently, we describe our method for functional data analysis and variable selection. Finally, we discuss the importance of grouped variables.

Stress recognition during driving

In recent years, the field of affective computing emerged as an interdisciplinary field combining computer science and engineering with cognitive science, physiology, and psychology. This field is expected to be beneficial for improving products in the automotive industry in order to enhance the quality of the driving experience and the driver's comfort ([22]). Actually, according to the American Highway Traffic Safety Administration, high stress levels negatively impact drivers reactions, especially in critical situations ([147]), which is one of the lead causes in vehicle accidents. Hence, the ability to detect driver's stress in a timely fashion may provide key indicators for valuable and timely decision making. In addition, a driver's stress level monitoring is important in order to avoid traffic accidents and hence promote safe and comfortable driving. For instance, researches have used sensors for data collection in order to build models for better human affective state understanding and to design smart human-centric interfaces ([151]).

Physiological data such as Electrocardiogram (ECG), Electromyogram (EMG), Electrodermal Activity (EDA) and Respiration (Resp) have been proven to reflect stress levels during driving tasks ([39, 134, 144, 145]). Usually, various types of features are extracted from these physiological measurements. It is thus crucial to select only features that are relevant in the recognition of the different stress levels during a given task (such as driving in different routes). In this context, there is a dearth of literature that focused on features selection and the study of the correlation between the selected features and the driving route complexity ([144]).

In practice, given that real-world driving data collection is very challenging, having access to an open database with physiological measurements captured while driving, offers great benefits in carrying studies such as the one at hand. Unfortunately, very few open databases are available. We have used in our study the open database "*drivedb*", available on the PhysioNet website ([66]). The data were captured during real-world driving experiments that were originally collected by [80]. The database includes various physiological measurements namely: ECG, EMG, EDA (measured on hand and foot), and respiration of 17 drives performed along a fixed itinerary, which switches between city and highway driving conditions. The original study analyzed twenty two features, derived from non-overlapping segments of physiological signals, extracted from the rest, highway, and city periods of all drives ([81]). A correlation analysis was performed between the features extracted from the physiological signals and the driver's affective state assessed by a stress metric created from the experiment video recording. We recall that the original study of [81] found that the driver's stress level was correlated the most with the EDA and the Heart Rate (HR) extracted from the ECG. However, that study did not formally address the selection of physiological features. Moreover, it did not consider any use of multi-resolution analysis of any physiological signal.

It is worth noting that since then, several studies have relied on the *drivedb* database. We note in

particular [30, 91, 143] studies that used ECG signals while [86] used ECG and Respiration data. Several studies on stress level recognition relied on the complete set of the physiological variables: we note in particular the work of [168] that proposed an approach based on Bayesian Network for the fusion of features extracted from the whole *drivedb* database. [8] presented an evaluation of the mean and standard deviation computed on the various *drivedb* physiological signals, covering all rest and driving periods of the experiments. In [39] study, the importance of features selection on stress level recognition was explored using the whole data of *drivedb*.
We note that the bibliographical study of [142] found that the stress modeling techniques did not include Random Forests (RF). Recently, [12, 68, 73] used RF as a classification technique for emotion recognition. But there is no study that benefited from the variable importance offered by RF in order to select or rank the physiological variables allowing stress level recognition.

Variable Selection and functional data
We recall that the random forests algorithm, introduced by [25], is based on aggregating a large collection of tree-based estimators. These methods have relatively good predictive performances in practice and work well in high dimensional problems. The RF power was demonstrated in several studies, summarized in [161]. Moreover, random forest methods provide several measures of the importance of the variables with respect to the prediction of the outcome variable. It has been shown that the permutation importance measure introduced by Breiman, is an efficient tool for variables selection ([41, 61, 71]). One of the main issues of variable selection methods in high dimensional data (data with small number of observations compared to the number of variables) is their instability. Indeed, the set of selected variables may change when slight perturbation is introduced to the training set. One solution to solve this instability consists in using bootstrap samples and a stable solution can then be obtained by aggregating selections performed on several bootstrap subsets of the training data. A similar solution is in fact proposed in this work by introducing a selection through repetitions of random forests based methods, intensively based on bootstrap resampling.
Functional Data Analysis (FDA) is a field that analyzes data for which explanatory variables are functions (of time in our case). One possible approach in FDA consists in considering the projections of the observations into finite dimensional spaces such as splines, wavelets, Fourier (see for instance [56, 130]). Many studies ([131, 156]) propose classification or regression methods for functional data in one of two possible situations: with one functional predictor and recently for several functional variables. Classification based on several possibly functional variables has also been considered using the CART algorithm for similar driving experiences in the study of [126], and using SVM in [165]. Variable selection using random forests was recently performed in the study of [60]. In our study, we will adopt multivariate functional data using random forests and the grouped variable importance measure proposed by [70].
The main contributions of this study are twofold: on the methodological side, it takes advantage of the functional nature of the physiological data and offers a procedure of data processing and variable selection. For that, the physiological signals are decomposed on a common wavelet basis and then analyzed in search of important variables using grouped variable importance. This analysis is applied on two levels of data selection strategies, based on a proposed "endurance" score. On the applied side, the proposed method provides a blind (i.e. without prior information) procedure of driver's stress level classification that does not depend on the extraction of expert-based features of physiological signals.

Chapter Outline
This chapter is organized as follows. After this introductory section, Section 2 is dedicated to the description of the protocol and the database used in this study. Section 3 recalls the random forest model and the variable selection procedure based on the variable importance measure. Section 4 presents the three main steps of variables selection. Results of the variables selection applied to

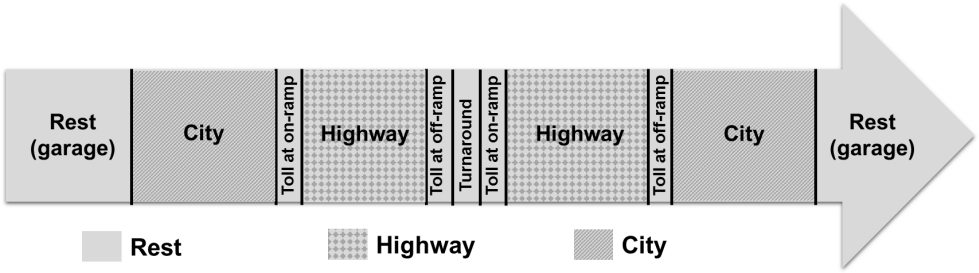


Fig. 3.1 Description of the driving setting. For simplification, the toll and the turnaround segments are considered as city driving.

the "drivedb" database are presented in Section 5. Finally, Section 6 and 7 present the discussion and the conclusions of the work.

3.2 Experimental protocol and data collection

The *drivedb* database was selected in our study for two main reasons: it is an open access database and its content has been explored in several studies due to its clear annotation. In fact, each dataset related to a drive includes a marker signal providing annotation of each driving segment, namely: city, highway driving, and the rest periods. In this section, we present the real-world driving protocol, originally proposed by [80]. Specifically, we will describe the cohort and detail the data construction process.

3.2.1 Real-world driving protocol

All the driving experiments were based on a drive path which extended over 20 miles of open roads in the greater Boston area ([81]). The driver was asked to follow a set of instructions in order to keep the drive regular. She/he was requested to respect the speed limits and to shut down the radio. The choice of the driving path was considered to imitate a typical daily commute and thus would induce reactions revealing stress levels usually evoked in normal daily driving trip. To ensure that all drivers made the same route, they were shown a map of the driving path before the driving session. An observer accompanied each participant on his/her drive. The observer sat in the rear seat in order not to interfere with the driver's normal behavior and affective state. This observer has also to monitor the experiment using a laptop displaying the recorded physiological signals and videos recording the inside and the outside scenes of the vehicle. Depending on the traffic density, the drive duration varies between 50 and 90 minutes. All driving sessions began with a 15 minutes rest period for the driver, after which the driver drove the car out of the garage, onto a ramp until he/she reached a congested city street. The driver then went into a main street, which has several traffic signals. It is assumed that driving on this road produces high stress-levels. The path continued from the city into a highway, where medium stress levels were usually evoked. After exiting the highway, the driver turned around, re-entered the highway in the opposite direction, followed the same route, and got back to the garage. After a complete driving session, the driver was requested to rest again. For each of the rest periods, the driver is asked to sit in the garage inside the vehicle (with an idle engine) with closed eyes. Thereby, this setting inducing the lowest stress level along the experiment was introduced to establish baseline measurements. Each driving session includes periods of rest, highway, and city driving (see Fig. 3.1), meant to produce respectively low, medium, and high stress levels. These assumptions were validated by two methods developed in [81]: the first method included a survey and the second consisted on a score derived from observable events and actions coded from the video tapes recorded during each drive.

Table 3.1 Description of the different 10 drives. Note that each participant is labeled by a sequence composed of gender (M or F) followed by the number of years of the driving experience. No information was available on Ind 4.

Drive	Participant label	Date (mm-dd-yy)	Duration (hh:mm:ss)
1	M-3	07-28-99	1:24:15
2		08-04-99	1:20:46
3	M-4	07-15-99	1:28:38
4		08-05-99	1:21:11
5		08-13-99	1:10:52
6	F-8	08-02-99	1:21:16
7		08-05-99	1:21:13
8		08-06-99	1:23:04
9		08-09-99	1:17:38
10	Ind 4	07-16-99	1:04:57

Table 3.2 Details of the missing data for the 7 drives excluded from the analysis due to the reported incomplete data in the *drivedb* database.

Drive Number	**Missing data**
01	Marker
02	Hand EDA and EMG
03	HR, EMG and Marker
04	EMG
13	Hand EDA
14	HR
17a & 17b (two parts of one experiment)	Marker is not complete

3.2.2 Cohort description

The set of data used in our work refers to 10 drives carried out by 4 participants denoted by M-3, M-4, F-8 and Ind 4, whose individual driving sequence details are shown in Table 3.1. According to the original study of [80], participant M-3, contributed in the database by 2 drives and was an undergraduate male. He had three years of driving experience though he was not a regular driver. Participant M-4 is an undergraduate student with over four years of driving experience. He had not driven a month prior to the experiment. He contributed to the database by 3 driving experiments. Participant F-8 was a female undergraduate having eight years of driving experience. Four driving experiments in the database are related to this driver. Finally, Ind 4 participated by just one driving experience. No gender or driving experience information of Ind 4 were provided.

3.2.3 Data Construction

The open database *drivedb*, available in PhysioNet website, contains only 17 out of the 24 total driving experiments (Drive01, Drive02, ..., Drive16, Drive17a, Drive17b).
In our study, only 10 out of those 17 driving experiences are considered (namely: Drive05, Drive06, ..., Drive12, Drive15, Drive16), since not all data were complete for the listed drives. In fact, seven out of the 17 drives report incomplete data including missing information, as indicated in Table 3.2.
For each of the 10 considered drives, seven segments, of 5 minutes each, were extracted for analysis

in our work, as shown in Fig. 3.2. However, two 5-minute segments (correspond to the second rest period of Drive 5 and Drive 10) were not considered since the markers that annotate the end of the second rest period are missing. Thus only 68 segments were subject of our analysis. In order to avoid driver's both memory recall and anticipatory effects, only the middle segments of highway and city driving are retained as representative samples of respectively medium and high stress levels. This choice was proposed by [81] where also the last 5 minutes of the two rest periods are considered, and this was chosen in order to give the participant enough time to relax from the previous tasks, thus inducing a low stress baseline.

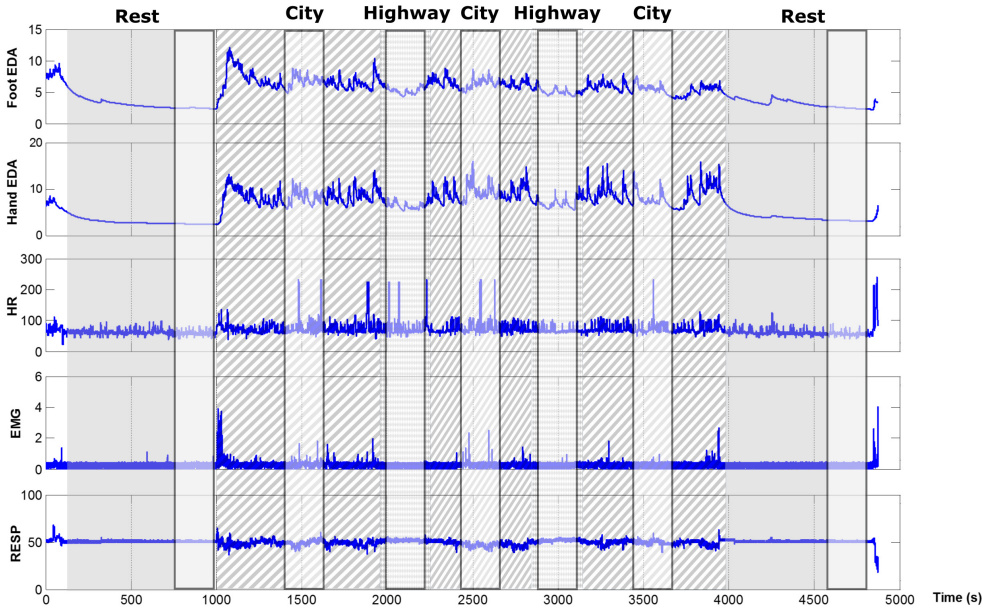


Fig. 3.2 Illustration of segment extraction of different physiological signals of Drive 7. Note that the physiological data were stored based on the same sampling frequency $F_s = 15.5\ Hz$.

Since *drivedb* data is in raw form, a standardization task is performed. For that, we adopt the same preprocessing proposed by [81], were the data were normalized in order to ensure the comparability between individuals.
For instance, EDA standardization, proposed by [104], is achieved by centering the EDA values using the minimum value (during the first rest period), then dividing by the EDA overall range (i.e. the difference between the maximum of the signal during the entire experiment and the minimum value during the first rest period). The HR, EMG and RESP signals are normalized by subtracting from each sample the mean value of the each signal over the first rest period. The physiological data were stored based on one same sampling frequency $F_s = 15.5\ Hz$.
Let $\boldsymbol{S}$ be the vector of physiological signals used here as explanatory variables. We extract five minutes segments. The choice of this window size allows the extracted segments to be informative and comparable. Five minutes corresponds to 4650 samples based on the used sampling period $\Delta t = \frac{1}{F_s} = \frac{1}{15.5} = 0.065s$.
In order to project these segments into a wavelet basis (see Section 3.4.1), we keep the first 2^{12} samples (since a power of two simplifies the discrete wavelet decomposition) which corresponds to 4096 samples. Functions $\boldsymbol{S} = (S^1, \ldots, S^p) \in \mathcal{S}^p$ where $S^j = \{S^j(t) \in \mathbb{R}, t \in [0,T]\} \in \mathcal{S}$ can be considered as functions of finite energy over $[0,T]$ where $T = 264.26s$ and $j = 1, \ldots, p$ where $p = 5$.
Let i be the index of the extracted segment, $i = 1, \ldots, N$ where $N = 68$.
For a given i, $S_i(t) = (S_i^1(t), \ldots, S_i^p(t))$ presents 5 physiological signals used as explanatory variables

corresponding to the stress level y_i,

$$y_i = \left\{ \begin{array}{c} \text{H=High stress level,} \\ \text{M=Medium stress level,} \\ \text{L=Low stress level.} \end{array} \right\}$$

3.3 Random Forests and variables selection using variable importance

3.3.1 Random Forests and Variable Importance measure

Let $\boldsymbol{X} = (X^1, \ldots, X^p)$, $\boldsymbol{X} \in \mathbb{R}^p$ denote the explanatory variables and Y the response variable that takes numerical values in the case of regression context, and a class label in the case of classification.
Let $L = \{(\boldsymbol{X}_1, Y_1), \ldots, (\boldsymbol{X}_n, Y_n)\}$ be a learning set, consisting in n independent observations of the vector $(\boldsymbol{X}, Y)$.
Random Forests (RF) is a non-parametric statistical method, originally introduced by [25] as an extension and improvement of decision trees [26]. RF provides estimators of either a Bayes classifier ($f : \mathbb{R} \mapsto \mathcal{Y}$ minimizing the classification error $P(Y \neq f(\boldsymbol{X}))$), in the classification case, and of the regression function g that verifies $Y = g(\boldsymbol{X}) + \varepsilon$ with $E[\varepsilon|\boldsymbol{X}] = 0$, in the regression case ([76]). A random forest is basically a set of decision trees, constructed over n_{tree} bootstrap samples $L^1, \ldots, L^{n_{tree}}$ of the training set L.
At each node, a fixed number of variables is randomly picked to determine the splitting rule among them. The trees are not pruned so all the trees of the forest are maximal trees. The resulting learning rule is the aggregation of all the estimators resulting from those trees, denoted by $\hat{f}_1, \ldots, \hat{f}_{n_{tree}}$. The response value at a new point x is hence obtained by the aggregation which consists in building the following

$$\hat{f}(x) = \left\{ \begin{array}{ll} \frac{1}{n_{tree}} \sum_{k=1}^{n_{tree}} \hat{f}_k(x) & \text{in regression,} \\ \\ arg \underset{1 \leq c \leq C}{max} \sum_{k=1}^{n_{tree}} 1_{\hat{f}_k(x)=c} & \text{in classification.} \end{array} \right. \tag{3.1}$$

We now introduce the Out-Of-Bag (OOB) sample that will be used to define the variable importance measure, provided by the RF. For each tree t, the OOB sample, denoted by OOB_t, is the set of observations that are not included in the bootstrap sample used to construct t. This OOB sample leads to a "smart" estimate of the error by calculating the error of tree t on the observations of OOB_t which can be used as test sample. This estimate is called the OOB error of a tree t. In order to assess the contribution of a variable to explain the response variable of interest in the RF model, we can use the permutation importance measure originally proposed by [25], and denoted henceforth as VI. It is defined as the mean increase, over all the trees in the forest, of the OOB error of a tree obtained when randomly permuting the variables in the OOB samples. The OOB error of a tree is measured by the Mean Square Error (MSE) for the regression case and by the misclassification rate for the classification on the OOB sample.
More formally, a variable X^j is considered important if when breaking the link between X^j and Y, the prediction error increases. The prediction error of each tree $\hat{f}$ is evaluated among its OOB sample with the empirical estimator

$$\hat{R}(\hat{f}, \bar{L}) = \left\{ \begin{array}{ll} \frac{1}{|\bar{L}|} \sum_{i:(\boldsymbol{X}_i, Y_i) \in \bar{L}} (Y_i - \hat{f}(\boldsymbol{X}_i))^2 & \text{in regression,} \\ \\ \frac{1}{|\bar{L}|} \sum_{i:(\boldsymbol{X}_i, Y_i) \in \bar{L}} 1_{\hat{f}(\boldsymbol{X}_i) \neq Y_i} & \text{in classification.} \end{array} \right. \tag{3.2}$$

Let $\left\{\bar{L}_k^j,\ k = 1, \ldots, n_{tree}\right\}$ refers to the set of OOB permuted samples resulting from random permutations of the values of the j-th variable in each out-of-bag sample OOB_{t_k}.

The VI is defined as the mean increase in the prediction error (estimated thanks to the OOB error) over all trees, and estimated by

$$\hat{I}(X^j) = \frac{1}{n_{tree}} \sum_{k=1}^{n_{tree}} \left[\hat{R}(\hat{f}_k, \bar{L}_k^j) - \hat{R}(\hat{f}_k, \bar{L}_k) \right]. \tag{3.3}$$

Several authors were interested in the numerical study of this VI (see [11, 111, 149]). Some theoretical results are also available in this regard (see [101], [70]). For instance, [170] proved that $\hat{I}(X^j)$ converges to $I(X^j)$, where

$$I(X^j) = E[(Y - f(\boldsymbol{X}^{(j)}))^2] - E[(Y - f(\boldsymbol{X}))^2]. \tag{3.4}$$

where $\boldsymbol{X}^{(j)} = (X^1, ..., X'^j, ..., X^p)$ is a random vector such that X'^j is an independent replication of X^j, independent of Y and of $X^k, k \neq j$.
The selection of groups of variables is a relevant topic in itself, see for example the Group Lasso allowing to select groups of variables in the linear model using the lasso selection strategy (see [14]). In our case, obvious interesting groups of variables are, for a given functional variable, the coefficients of the function considered as a whole or the wavelet coefficients at a given level, for example. In order to generalize the random forests based selection strategies, [70] extend the permutation importance definition for a group of variables. To estimate the permutation importance of a group of variables denoted by $\mathbf{X^J}$, let us consider for each $k \in \{1, ..., n_{tree}\}$, $\bar{\mathbf{L}}_k^J$ the permuted set of $\bar{\mathbf{L}}_k$ resulting by randomly permuting the group $\mathbf{X^J}$ in each OOB sample $\bar{\mathbf{L}}_k^j$. The permutation importance of $\mathbf{X^J}$ is estimated by

$$\hat{I}(\mathbf{X^J}) = \frac{1}{n_{tree}} \sum_{k=1}^{n_{tree}} \left[\hat{R}(\hat{f}_k, \bar{\mathbf{L}}_k^J) - \hat{R}(\hat{f}_k, \bar{\mathbf{L}}_k) \right]. \tag{3.5}$$

When the number of variables considered in a group increases, the grouped VI may increase ([70]), which is particularly true in the case of an additive model and independent variables. In order to take group size into account, we shall use the normalized version of the VI ([70]):

$$I_{nor}(\mathbf{X^J}) = \frac{1}{|J|} I(\mathbf{X^J}). \tag{3.6}$$

3.3.2 Variable Selection using Random Forest-based Recursive Feature Elimination

Random Forests-based Recursive Feature Elimination (RF-RFE) algorithm, summarized in Table 3.3, was originally proposed by [70], and inspired from [74] that introduced the Recursive Feature Elimination algorithm for SVM (SVM-RFE).

Table 3.3 Summary of the selection algorithm based on RF-RFE.

1. Split the whole data L into a training set L_T containing $\frac{2}{3}$ of the data and a validation set L_V containing the remaining $\frac{1}{3}$. Set the subset of the variables $\mathcal{V}$ to the whole explanatory variables.
2. Fit a random forest model using L_T and considering the set of variables $\mathcal{V}$
3. Compute the VI measure (respectively the grouped VI measure)
4. Compute the error using the validation sample L_V
5. Eliminate the least important variable (resp. group of variables) and update $\mathcal{V}$.
6. Repeat 2-5 until no further variables (resp. groups of variables) remain
7. Select the variables (resp. the groups of variables) involved in the model minimizing the prediction error

Let us first sketch the statistical model used in the Random Forest framework. Let $L = \{(\boldsymbol{S}_1, Y_1), \ldots, (\boldsymbol{S}_n, Y_n)\}$ be a learning set, consisting in n independent observations of the random vector $(\boldsymbol{S}, Y)$. The $(S_i(t), y_i)$, introduced in Section 3.2.3, are then considered as realization of unknown distribution of this sample. We aim to build an estimator of the Bayes classifier $f : \mathbb{R} \mapsto \mathcal{Y}$ that minimizes the classification error $P(Y \neq f(\boldsymbol{S}))$. We denote by $\hat{f}$ an estimator belonging to the fully non parametric set of models given by the random forests framework introduced in Section 3.3.1.
The proposed algorithm in Table 3.3 will be briefly explained in the case of scalar variables since the case of functional variables is similar. In fact, the variable groups are predefined depending on the type of data, and in our case this definition depends on the functional reconstruction procedure. At the first step, the dataset is randomly split into a training set containing two thirds of the data and a validation set containing the remaining one third. Steps 2-5 of the procedure are a loop. This loop starts by fitting the model using all the explanatory variables using Random Forests. Then, the variables are ranked from the most to the least important based on their VI values, computed on the training set. Then, the least important predictor is eliminated, the model is re-trained and the prediction error computed on the validation set. The variable ranking and elimination steps are repeated until no variable remains. The final model is chosen by minimizing the prediction error. It should be noted that at each iteration, the predictors importance is recomputed according to the model composed of the reduced set of explanatory variables. Indeed, variations to this approach are conceivable, namely using an ascending strategy where the variables are sequentially invoked according to a given ranking of their respective importance obtained once and for all at the beginning. These variants are suboptimal but less time consuming, which can be especially interesting in the high dimensional case. This exact same algorithm detailed in Table 3.3 can be applied for the selection of functionnal variables, as well as wavelet coefficients grouped by wavelet level.

3.4 The 3 steps of the variable selection approach

One goal of the affective computing research in the automotive field is to recognize the driver's stress level. Although using many sensors to capture physiological signals is likely to increase accuracy of the results, but integrating many sensors in the vehicle may be a complex task. In fact, there is a trade-off between the adequate number of sensors and stress level recognition accuracy.

Consequently, we propose an approach that not only removes irrelevant physiological variables from the classification model, but also selects the most relevant wavelet features among the wavelet levels resulting from the decomposition of the retained variables; allowing an optimal stress-level recognition.
A summary of the 3-step procedure is listed in Table 3.4. This procedure repeatedly applies the RF-RFE algorithm described in Table 3.3. A detailed description of our approach is given below.

Table 3.4 Summary of the proposed 3-step approach.

Step 1. Wavelet decomposition of the physiological functional variables.

Step 2. Physiological functional variable removal:

1. Repeat 10 times: RF-RFE $(G(1), ..., G(p))$.
2. Compute an "endurance" score for each group $G(\ell)$ as following:

$$score(G(\ell)) = \sum_{m=1}^{p} nbocc(G(\ell), Q_m) \times [(p - m) + 1]. \tag{3.7}$$

where $1 \leq \ell \leq p$, Q_m is the list of the selected variables, all over the 10 executions, at the rank m, $m = 1, ..., p$, and $nbocc(G(\ell), Q_m)$ corresponds to the number of occurrences of $G(\ell)$ in the list Q_m.

3. Remove the less "enduring" functional variables (those of a score below a threshold δ).

Step 3. Selection of wavelet levels among the R retained functional variables:

1. Repeat 10 times:
RF-RFE $(\{G(1, S_1), ..., G(J, S_1), ..., G(1, S_R), ..., G(J, S_R)\})$.
2. Compute a selection score for each group $G(j, S_r)$ as following:

$$score(G(j, S_r)) = \sum_{m=1}^{p} nbocc\ (G(j, S_r), Q_m) \times [((J \times R) - m) + 1]. \tag{3.8}$$

where $1 \leq j \leq J$, Q_m is the list of the selected wavelet levels, all over the 10 executions, at the rank m, $m = 1, ..., p$, $p = J \times R$, and $nbocc\ (G(j, S_r), Q_m)$ corresponds to the number of occurrences of $G(j, S_r)$ in the list Q_m.

3. Select the most "enduring" wavelet levels (those having a score higher than a threshold δ').

3.4.1 Step 1. Wavelet decomposition of the physiological functional variables

For the sake of clarity, we shall first recall few basic concepts related to wavelets and their relationship to functional variables. Given a space of functions $\mathcal{F}$ (typically, $L^2([0, 1])$which is the space of square integrable functions defined on $[0, 1]$) and a probabilistic space Ω, a functional random variable is a measurable application $S : \Omega \to \mathcal{F}$. A function of $\mathcal{F}$ is defined on $[0, 1]$ ($[0, T]$ equivalently in our specific situation) with values in $\mathbb{R}$.
Starting from compactly supported scaling function ϕ and mother wavelet ψ, the sequence of functions $\phi_{j,k}(t) = 2^{j/2}\phi(2^j t - k)$ and $\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k)$ are obtained by dyadic translations and dilations of ϕ and ψ. This allows us to build several orthonormal wavelet bases, for any $j_0 \geq 0$

$$\mathcal{B} = \left\{\phi_{j_0,k}, k = 0, \ldots, 2^{j_0} - 1\right\} \bigcup \left\{\psi_{j,k}, j \geq j_0, k = 0, \ldots, 2^j - 1\right\}. \tag{3.9}$$

Therefore, a wavelet transform of a function s of $L^2([0,1])$ can be written as an expansion on $\mathcal{B}$:

$$s(t) = \sum_{k=0}^{2^{j_0}-1} c_{j_0,k}\phi_{j_0,k}(t) + \sum_{j=j_0}^{\infty}\sum_{k=0}^{2^{j_0}-1} d_{j,k}\psi_{j,k}(t), \tag{3.10}$$

with $c_{j,k} = \left\langle s, \phi_{j,k} \right\rangle_{L^2}$ and $d_{j,k} = \left\langle s, \psi_{j,k} \right\rangle_{L^2}$.
The first term in Equation (3.10) is the smooth approximation of s at level j_0 while the second term is the detail part of the wavelet representation, then the $c_{j,k}$ are the scaling coefficients and $d_{j,k}$ are the wavelet coefficients at level j and position k. Thanks to the formalism of multi-resolution analysis of L^2, introduced by [105], the whole space is the limit of a sequence of nested subspaces V_j of approximation associated to the scale levels $j \in \mathbb{Z}$ and of increasing resolution. It allows thus to define the Discrete Wavelet Transform (DWT) as a simple algorithm to perform for an input signal of length $N = 2^J$ and for the maximum level of wavelet given by the size N of the sampling grid. The wavelet decomposition of s given in a similar form as Equation (3.10) is thus as follows,

$$\tilde{s}_J(t) = c_0\phi_{0,0}(t) + \sum_{j=0}^{J-1}\sum_{k=0}^{2^j-1} d_{j,k}\psi_{j,k}(t), \tag{3.11}$$

where c_0 is the single scaling coefficient and $d_{j,k}$ denotes the empirical wavelet coefficients derived from applying the DWT to the sampled values.
For a given wavelet function, the physiological functional variables are decomposed into a common wavelet basis at the first step of the approach.

3.4.2 Step 2. Physiological functional variable removal

The wavelet coefficients provided by the first step of the proposed approach are considered as inputs, which are then grouped by physiological variable. Let $G(\ell)$ denote the group of the wavelet coefficients resulting from the wavelet decomposition of the physiological variable S_ℓ. The RF-RFE algorithm (see Table 3.3), is applied then on these groups of wavelet coefficients $G(\ell)$ where $1 \leq \ell \leq p$, p is the number of the physiological variables considered in the initial model. As a result, the list of the physiological variables ordered using the VI measure and the number of the selected variables are provided.
It should be mentioned that the values of the results vary for different execution repetitions. Therefore, a procedure is proposed to improve the stability of the method. This procedure consists in executing the RF-RFE 10 times and computing a selection score that helps, at this step, to remove the less "enduring" physiological variables. The "endurance" of a variable measures the ability for it to be selected in the first ranks. The "endurance" score is built based on the number of occurrences of the variable in the lists of selected variables, and on the weights attributed according to the selection rank (see Table 3.5 for more details on the weights). This score aggregates information concerning the results of 10 executions of the RF-RFE algorithm over the entire set of variables.
Table 3.5 illustrates this result. For each execution e, the algorithm offers a list of the ranked variables $G(.)_e^1, ..., G(.)_e^p$ and a number of the selected variables $nbsel_e$. Let us define Q_m as a list that aggregates the selected variables, all over the 10 executions, at the rank m where $1 \leq m \leq p$. For instance, if $nbsel_e \geq 1$ for an execution e $(1 \leq e \leq 10)$ then Q_1 contains the list of the 10 selected variables at the first rank. Q_m may be empty in the case that no variable was selected at the rank m and it may contain several occurrence of the same variable. Let $nbocc(G(.), Q_m)$ denote the number of occurrences of a variable $G(.)$ in the list Q_m.
The "endurance" score of a variable is the weighted sum of the number of occurrences of the variable all over the lists Q_m, $1 \leq m \leq p$. The weight of a variable in a list of rank m is equal

to $(p-m)+1$. The computation of this score on the physiological variables is detailed in Step 2 of Table 3.4. The obtained selection scores allow to rank physiological variables according to their "endurance" in the selection procedure. Those having a score less than a threshold δ are removed. Thus, R physiological variables are retained and to be considered in the next step of the procedure.

Table 3.5 Summary of results of the 10 executions of the RF-RFE.

Execution Num	**Weights**				**Number selected var**
	p	$p-1$	...	1	
Execution 1	$G(.)_1^1$	$G(.)_1^2$	...	$G(.)_1^p$	$nbsel_1$
⋮	⋮			⋮	
Execution 10	$G(.)_{10}^1$	$G(.)_{10}^2$	...	$G(.)_{10}^p$	$nbsel_{10}$

3.4.3 Step 3. Wavelet levels selection

This step consists in applying the same procedure detailed in Section 3.4.2, except that for this step the wavelet coefficients are grouped by wavelet levels. In addition, the coefficients considered here concern only the R retained physiological functional variables after the removal process is applied. Let S_r, $1 \leq r \leq R$, be the r-th physiological variable retained after removal based on Step 1. Let $G(j, S_r)$ denote the group of wavelet coefficients resulting from the decomposition of the variable S_r on the wavelet scale level j, $1 \leq j \leq J$ and J is the common maximum scale level that physiological variables S_r, $r = 1, ..., R$ can be decomposed on.
Let $nbocc(G(j, S_r), Q_m)$ denote the number of occurrences of the group of wavelet coefficients $G(j, S_r)$ in the list Q_m, where $1 \leq m \leq p$. In this case, $p = R \times J$. The "endurance" score in this step is computed in the same way as in Step 1. It is the weighted sum of the number of occurrences of the group of wavelet coefficients $G(j, S_r)$ all over the lists Q_m, $1 \leq m \leq p$. This score provides then a ranking of the wavelet levels according to their "endurance" in the selection procedure. This step aims mainly to select the most "enduring" wavelet levels. Thus, wavelet levels having an endurance score higher than δ' are selected.
Let us denote the procedure detailed at the subsection 3.3.2 as
RF-RFE $(Grp_1, ..., Grp_p)$ then:

- the procedure described in Subsection 4.2 is indicated with $Grp_\ell = G(\ell)$, corresponding to wavelet coefficients grouped by physiological variables ℓ, $\ell = 1, ..., p$.

- the procedure described in Subsection 4.3 is indicated with $Grp_\ell = G(j, S_r)$, corresponding to wavelet coefficients grouped by wavelet scale j, $j = 1, ..., J$ of the selected functional variable S_r, $r = 1, ..., R$.

Remark: It must be noted that the removal of the functional variables (respectively wavelet levels) depends on the choice of the threshold δ (respectively δ'). A classical strategy is to apply a rule similar to the elbow criterion that allows to determine the number of factors in the PCA, see [89]. We propose here a classical trick: we look at the consecutive differences of the selection score. When starting from the highest values of the endurance scores, we select variables just before the biggest gap or when starting from the lowest values of endurance scores, remove variables just after the big gap. The threshold can then be selected based on a visual inspection, if the number of variables is small.

3.5 Variables selection on the *drivedb* database

In this section, we present the different variable selection strategies and their applications on the *drivedb* database. The objective of variable selection is not only to remove irrelevant functional variables, but also to retain the most relevant wavelet-based features in stress level recognition.

3.5.1 Step 1. Wavelet decomposition of the physiological functional variables

Recall that five physiological signals are considered, namely *Hand EDA*, *Foot EDA*, *HR*, *RESP* and *EMG*. In a first stage, we perform a functional variable decomposition using the Haar wavelet, which is considered as the simplest one. This choice is sufficient for our analysis since the main goal at this stage is to benefit from a basic scale information without the need for choosing a suitable wavelet type, which is required for instance, in the detection of some specific events or patterns. We recall that the Haar wavelet's mother wavelet function $\psi(t)$ is given by:

$$\psi(t) = \begin{cases} 1 & 0 \leq t < \frac{1}{2}, \\ -1 & \frac{1}{2} \leq t < 1, \\ 0 & \text{otherwise.} \end{cases} \tag{3.12}$$

and the corresponding scaling function $\phi(t)$ is given by:

$$\phi(t) = \begin{cases} 1 & 0 \leq t < 1, \\ 0 & \text{otherwise.} \end{cases} \tag{3.13}$$

We have opted for a full wavelet decomposition corresponding to 12 decomposition levels, which correspond to the maximum levels compatible with $2^{12} = 4096$ samples.
For all computational tasks, we used the `R` software ([155]), with the `randomForest` package proposed by [27] and `RFgroove` package developed by [72].

3.5.2 Step 2. Physiological functional variables removal

In this section, the importance of physiological functional variables considered in our study will be presented. Since not all of these variables are important, we will proceed to the elimination of the less important physiological signals, which do not contribute significantly in the identification of stress level.

Importance of physiological variables

Data resulted from the physiological functional variable decomposition provide 4096×5 scalar variables corresponding to the wavelet coefficients. Taking advantage of the grouped VI, the scalar variables were grouped by physiological variable and the VI of each considered group was computed. Thereby, the grouped VI allowed to consider only 5 informative VI instead of 20480 VI of the scalar variables.
In order to examine the VI of each group, we present in Fig. 3.3 the boxplots of the grouped VI values, computed for 100 executions.

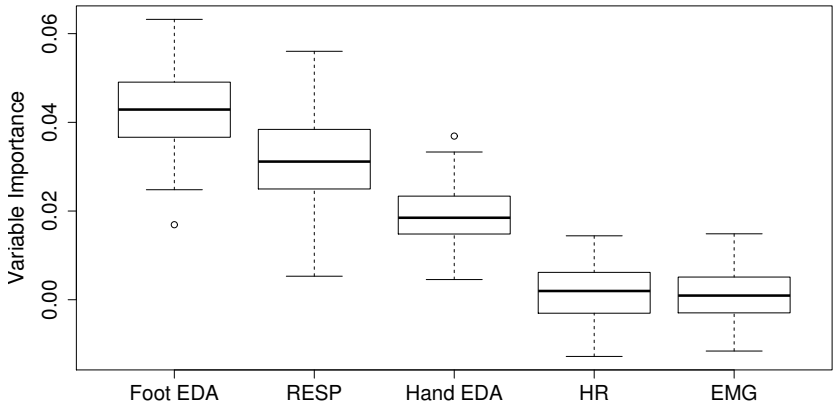


Fig. 3.3 Boxplots of grouped VI by physiological signals for 100 executions.

It can be noticed that *Foot EDA* is the most important physiological variable with a VI distribution around 4%. *RESP* is the second most important variable since the median value of the distribution of 100 VI is about 3%. *Hand EDA* comes out important when compared to the *EMG* and *HR*. The distribution of its 100 VI is around 2%. *Foot EDA* and *Hand EDA* are found to be among the most important variables. This result is in agreement with [81] findings that EDA is among the most correlated variable with the driver's stress level. *EMG* and *HR* variables are found to have a distribution of values corresponding to 100 VI around 0.

RF-RFE on the grouped physiological functional variables

In order to eliminate the less relevant physiological variables, we applied the RF-RFE algorithm ten times to the wavelet coefficients grouped by functional variables. Each execution offers a ranking of the 5 variables and provides the list of variables retained by the selected procedure. Table 3.6 shows the result of 10 executions of the RF-RFE algorithm on the five physiological variables. The shaded cells correspond to the final retained variables. For instance, *Foot EDA*, *RESP*, and *Hand EDA* are selected in the first execution.
Based on these 10 executions, *Foot EDA* is always included in the selected model. Additionally, *EMG* and *HR* (except for one execution) are excluded from the list of the variables selected in the model. It can be noticed that the number of selected variables and the order of importance of these selected variables vary from an iteration to another. Thus, we propose to use the procedure detailed in Table 3.4 in order to aggregate the information contained in these 10 executions and to select the most "enduring" physiological variables.

Ex. Num	Weights					Num sel var
	5	4	3	2	1	
1	Foot EDA	RESP	Hand EDA	HR	EMG	3
2	HR	RESP	Hand EDA	Foot EDA	EMG	4
3	Foot EDA	Hand EDA	HR	EMG	RESP	1
4	Foot EDA	RESP	Hand EDA	EMG	HR	1
5	RESP	Foot EDA	Hand EDA	HR	EMG	2
6	Foot EDA	RESP	Hand EDA	EMG	HR	2
7	Foot EDA	Hand EDA	RESP	HR	EMG	1
8	Foot EDA	RESP	Hand EDA	HR	EMG	1
9	Foot EDA	Hand EDA	RESP	EMG	HR	1
10	Foot EDA	RESP	Hand EDA	HR	EMG	3

Table 3.6 Selected model for 10 executions of the RF-RFE algorithm. The shaded cells corresponds to the retained variables.

Fig. 3.4 displays the "endurance" score of the five physiological variables computed based on equation (3.7). The results confirm the order found by the VI measure (see Fig. 3.3) where both *EMG* and *HR* had low endurance score. Moreover, the endurance score reveals that *Foot EDA* is the most important variable in driver's stress level classification, followed by *RESP* and *Hand EDA*.

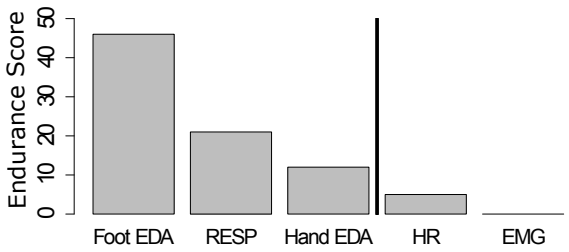


Fig. 3.4 Endurance score of the five physiological variables for 100 runs. The two last physiological variables are removed.

Thus it is confirmed again that the EDA is an important variable in driver's stress level recognition. These findings corroborate with the findings of [81], who reported that features extracted from the *EDA* and *HR* have the highest correlation with the stress metric built from the driving video recordings. However, it should be noted that the *Foot EDA* was not considered in the original study. The original study ([81]) determined the most correlated physiological signals to driver's stress using features extracted from four physiological variables, namely the mean and the standard deviation of the *Hand EDA*, *EMG*, *RESP*, the mean and 4 metrics extracted from the *HR*.
In contrast with the results of [81], *RESP* is selected in our analysis. This can be explained by the fact that the wavelet-based features used here may be more informative than the mean and the standard deviation used in the original study. In addition, the *RESP* wavelet features are proved sensitive in detecting the transient changes in respiration signal ([10]). *HR* was found to be of less relevance in our proposed model. This may be due to additional information contained

in the expert-based features that are not present in the proposed wavelet features.
Our analysis having included both Hand and Foot EDA, has revealed that the Foot EDA is as important as Hand EDA and may be even more relevant for the experienced stress levels of the drivers. Usually, EDA is measured on the fingers or on the palms. However, these two placements are not preferred in real-world task performance because they may hinder hand functionality. Measuring EDA on foot or shoulders is more practical, while having no significant difference from the signal captured from the fingers ([157]). Several other studies have also shown the importance of the placement of the EDA sensor ([23, 122, 157]). Hence, one can suggest to consider Foot EDA in future studies that aim to recognize stress levels experienced in real-world driving.
Based on the "endurance" score, the final model retains the three physiological variables: *Foot EDA*, *RESP*, and *Hand EDA* as most relevant in stress level classification. In the next subsection, the wavelet levels of these selected three variables will be considered. We will then seek the model based on their most "enduring" wavelet levels to the selection procedure.

3.5.3 Step 3. Wavelet levels selection from the three retained physiological variables

The goal of this step is to identify, for the retained physiological functional variables, the corresponding wavelet levels most able to predict the driver's stress class. To achieve this goal, the wavelet coefficients are now grouped by levels.
Let us denote the approximation of the signal s as $A12$ and the detail at level j by Dj then the wavelet decomposition can be written as follows

$$\tilde{s} = A12 + \sum_{j=1}^{12} Dj, \tag{3.14}$$

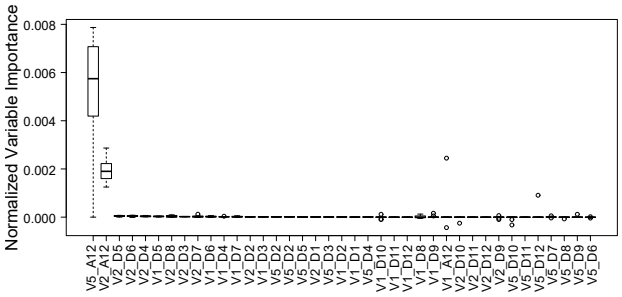
where $A12(t)$ is reconstructed from the single scaling coefficient at level 12 and details of increasing resolution Dj are reconstructed from the DWT wavelet coefficients ($d_{j,k}$) (see Equation (3.11)).

Wavelet levels importance

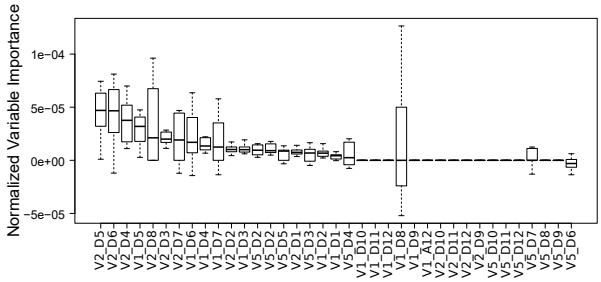
Let V1 stand for *Hand EDA*, V2 for *Foot EDA*, and V5 for *RESP*. We simply denote by Vx_A12 or Vx_Dj the wavelet decomposition components of the functional variable Vx. Fig. 3.5 (a) depicts the distribution of the VI computed over 10 executions of the different wavelet levels of the three physiological variables. The distributions show that the approximation level ($A12$) of both *Foot EDA* and *RESP* are the most important wavelet levels in the stress level classification. The variation of the distribution of the other wavelet levels is not clearly visible in Fig. 3.5 (a). In order to better investigate the relative variation of small VI, we present a graph obtained by removing the dominant variables $V2\_A12$ and $V5\_A12$ and the outliers in Fig. 3.5 (b). We find that for the *Hand EDA* (V1), levels from 1 to 8 are the most important. Wavelet levels from 3 to 8 are important for the *Foot EDA* (V2). For the *RESP* (V5), levels 1 and 2 are the most important.
In the next paragraph, we compute the "endurance" score for all wavelet levels that allows to select the final model.

Wavelet levels selection

When applying the RF-RFE algorithm to the groups composed by the wavelet levels of the three retained variables several times, the selected model varies in terms of the number of the selected levels, the validation error and even the ranking of the different wavelet levels. In order to reduce this variability, an "endurance" score, introduced earlier by Equation (3.8), is proposed. It allows



(a) Grouped VI of the V1, V2 and V5 wavelet levels.



(b) Grouped VI of the V1, V2 and V5 wavelet levels without V2_A12 and V5_A12.

Fig. 3.5 Grouped VI of the wavelet levels for 10 iterations.

to rank the different wavelet levels according to their ability to persist in the selection procedure. The plot of the ranked variables using this score is displayed in Fig. 3.6. There are two ways to

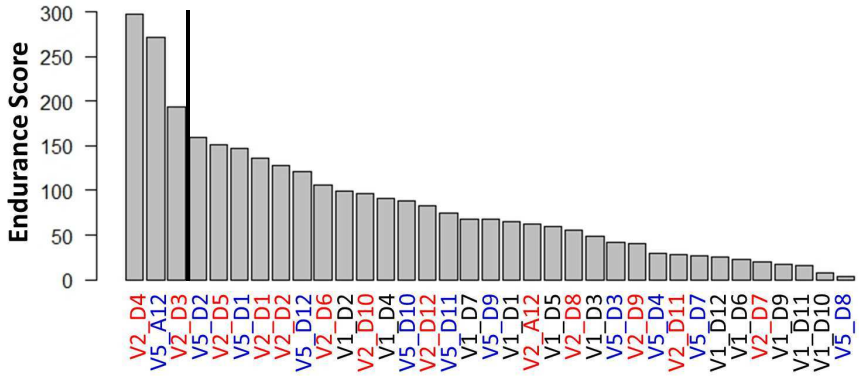


Fig. 3.6 Wavelet levels selection for 10 executions.

select variables of the final model, when examining the consecutive differences of the "endurance" scores: one can start from the highest value and cut off at the largest gap of this difference or start from the lowest score and cut off at the largest gap of this difference. When applying the first approach on the wavelet details corresponding to the three retained physiological variables, we select $V2\_D4$ and $V5\_A12$. If the second way of selection is chosen, $V2\_D4, V5\_A12$, and $V2\_D3$ will be selected. Three wavelet signals that correspond to levels 4 and 3 of details and the approximation at level 12, among $39 = 13 \times 3$ wavelet signals, are selected in our case. This can be justified by the fact that the third level ($V2\_D3$) is not too different from $V2\_D4$.
Based on the selected wavelet levels, we examine then the reconstructed physiological signals. For instance, Foot EDA wavelet details on levels 3 and 4 and RESP wavelet approximation at level 12 corresponding to Drive 07 are depicted in Fig. 3.7. Indeed, using only two selected levels for Foot EDA allows to distinguish between the three segments of city, highway, and rest. The reconstructed approximation level of the respiration (Fig. 3.7) highlights the fact that the 3

periods can be easily recognized. It can be also noticed that the value of RESP approximation level decreases when the stress level increases which is aligned with the conclusion of the study of [99], showing that stress is negatively correlated with the respiration rate.

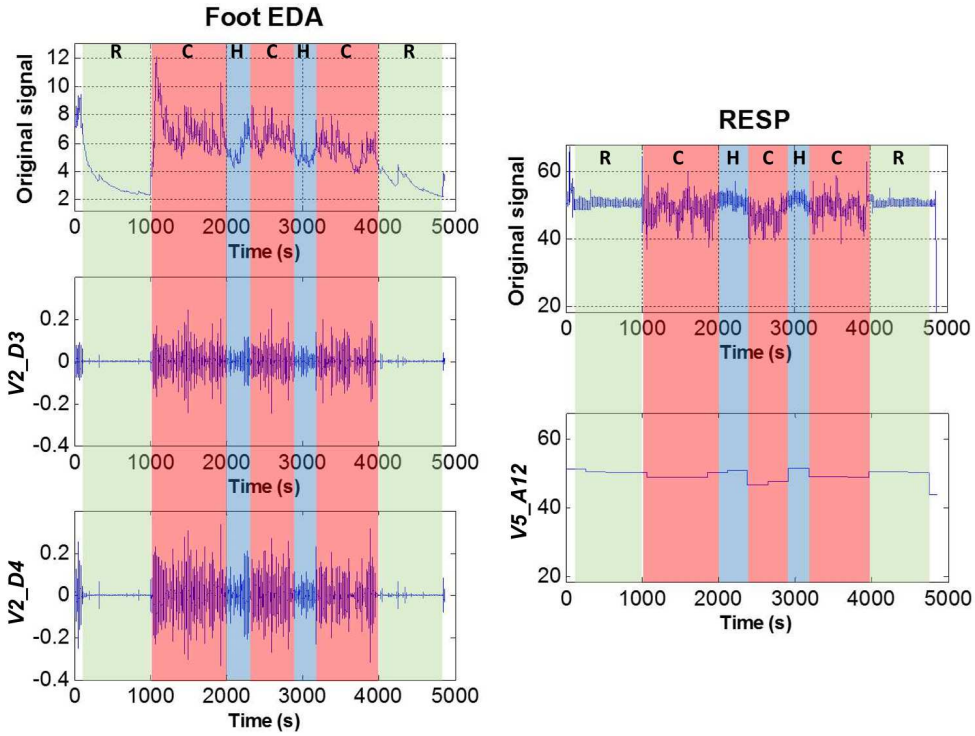


Fig. 3.7 Illustration of the reconstructed signals corresponding to Drive 07 for Foot EDA (left column) and RESP (right column), based on the three selected wavelet levels (see Fig. 3.6). The letter "R" corresponds to rest period, "C" to city driving and "H" to highway driving.

An additional "unidimensional" information

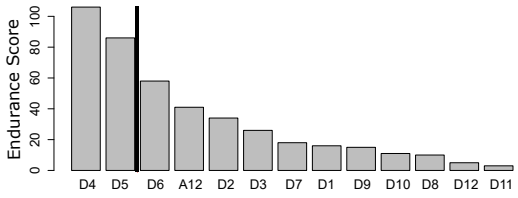
In this subsection, we propose to investigate, for each physiological functional variable, wavelet levels that are the most able to classify drivers stress level. Therefore, three models, aiming each at recognizing stress level based on a given physiological variable, are considered. The selection of wavelet levels is then performed independently, for model composed by *Hand EDA*, *Foot EDA* and *RESP* respectively. For that, Equation (3.8) is used first to compute the "endurance" score of the different wavelet levels of the *Foot EDA*. We note that two wavelet detail levels are selected, based on the "endurance" score shown in Fig 3.8 (a), which are namely levels 4 and 5. For this physiological predictor, the wavelet level 4 was selected in the final model that combines the different wavelet levels of the three physiological predictors.

Fig. 3.8 (b) reveals that the first and the second details of the wavelet levels and the approximation at level 12 are selected for the RESP predictor.

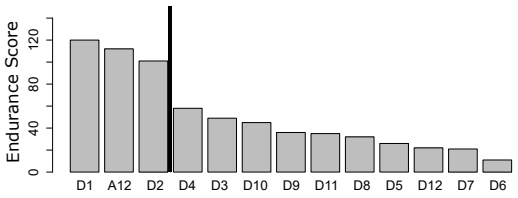
Fig. 3.8 (c) shows *Hand EDA* wavelet levels ordered by endurance score. When starting from the smallest value of the score and looking at the consecutive difference, levels 1 to 8 (except 7) are selected for *Hand EDA*. The same result was found based on VI analysis of *Hand EDA* levels (see Fig. 3.5 (b)). In addition, the "endurance" score proposes that $D4$, $D5$ then $D6$, are in the same order as for *Foot EDA* (see Fig. 3.8 (c) for comparison).

When considering separate models composed by each physiological variable, the proposed endurance score provides a similar order of the wavelet levels as given by VI of the global model (see Fig. 3.5). This additional separate analysis supports the findings based on the proposed endurance score.

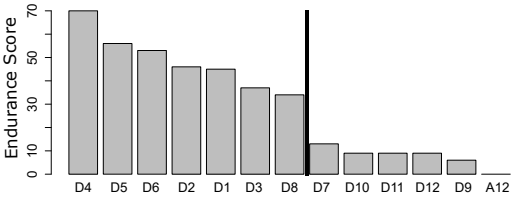
In the next subsection, the performance of the 3-step based approach will be presented and an assessment of the final misclassification rate will be performed.



(a) Foot EDA



(b) RESP



(c) Hand EDA

Fig. 3.8 Wavelet levels endurance score of the three retained physiological variables after 10 executions.

3.5.4 Performances of the proposed "blind" approach

Recall that the "blind" 3-step approach applied on the *drivedb* database resulted in the selection of Hand EDA, Foot EDA and RESP as the most relevant functional variables, in addition to their most important wavelet decomposition levels. Specifically, this resulted in a total of 3 wavelet signals corresponding to levels 3 and 4 of the Foot EDA wavelet details and level 12 of the RESP wavelet approximation.

In this section, we first assess the performances of the RF-based model when considering these retained variables in the classification of the driver's stress level, in terms of misclassification rate. Then we perform a cross-validation like-procedure for model error estimation taking into account the particularity of the cohort used in the experiments. Finally, we proceed with evaluation of the resulting performance of the proposed "blind" approach, when compared to those obtained by the "expert-based" approach.

Recall that a total of 669 corresponding wavelet coefficients were finally considered as input variables to the RF algorithm (for the *Foot EDA*, $512 = 2^{(12-3)}$ and $256 = 2^{(12-4)}$ coefficients and 1 coefficient corresponding to the level 12 of the *RESP* wavelet approximation).

In order to evaluate the performances of the proposed blind approach, two different procedures were considered: the first one consists in evaluating the error of the model, composed of the whole set of variables. The second procedure considers the error resulting from a wavelet level selection among the three signals corresponding to the retained wavelet levels using the RF-RFE method summarized in Table 3.3.

The model error corresponds to the average of the misclassification rate obtained in each iteration of the evaluation procedure, which is repeated 100 times in order to reduce the variability of the estimation. The misclassification rate is estimated in two different ways: the error obtained using a test set ($\frac{1}{3}$ of the data) and the OOB error embedded in the RF algorithm.

As can be seen in Table 3.7 (a), the model errors resulting from the wavelet selection are lower than those obtained when considering the complete set of coefficients. The model error assessed on the test set with wavelet level selection is 0.28 (sd = 0.09) while the OOB error is 0.27 (sd = 0.02). When the 669 coefficients are considered, the OOB error is 0.31 (sd = 0.03) and the error estimated

using the test set is 0.31 ($sd = 0.09$).

Table 3.7 Model error: misclassification rate averaged over 100 executions.

(a) Performances of the "blind" 3-step approach.

Input variables	Test error	OOB error
With selection	0.28 (sd=0.09)	0.27 (sd=0.02)
Without selection	0.31 (sd=0.09)	0.31 (sd=0.03)

(b) Performances of the "blind" 3-step and expert-based approaches for cross validation like procedure. Config1 (training set size = 54 and test set size = 14), Config2 (training set size = 48 and test set size = 20) and Config3 (training set size = 40 and test set size = 28.)

Input variables	Selection	Config1	Config2	Config3	Mean
Wavelet features	With selection	0.19 (sd=0.05)	0.22 (sd=0.04)	0.26 (sd=0.03)	0.22
	Without selection	0.31 (sd=0.04)	0.22 (sd=0.05)	0.27 (sd=0.04)	0.27
Expert-based features		0.33 (sd=0.01)	0.14 (sd=0.02)	0.25 (sd=0.04)	0.25

Recall that the *drivedb* datasets are based on a real-world driving of four participants, some of whom repeated the experiment. In order to compute the misclassification rate, we adopt a particular cross-validation procedure taking into account this particularity of the experiments (see Fig. 3.9).
In fact, the error assessment was evaluated for two models: one composed of the whole set of features (without selection) and the second procedure considered the resulting error based on a wavelet level selection, among the retained three levels using RF-RFE (with selection).
In Table 3.7 (b), we consider the three configurations of the cohort disposition defined in Fig. 3.9. The partitioning of the dataset into 3 configurations is done as follows: Config k consists in considering the data of the driver k, $k = 1, \ldots, 3$ as the test set, while the retained data related to the other drivers is considered as the training set as shown in Fig. 3.9. We note that the data related to driver 4 was considered in the training set of the 3 configurations since he performed only one test drive.

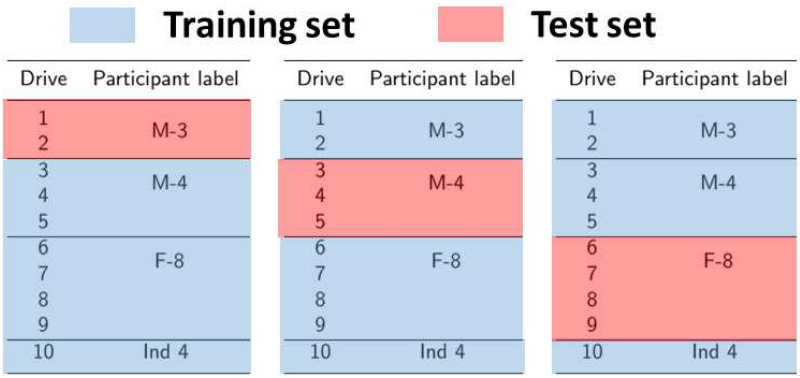


Fig. 3.9 Three configurations of training and test sets choice in the cross validation like procedure.

Table 3.7 (b) shows the mean error over the 3 configurations of the "blind" approach in both cases: with and without selection, in addition to the model error resulting from the RF using the

18 expert-based features shown in Table 3.8. The estimated model error considering the wavelet features (without selection) is 0.27 while it is 0.22 with selection. We note that the expert-based approach had a model error of 0.25.

Table 3.8 Features related to the EDA (Foot and Hand) and to the respiration signals proposed by [81].

Type of features	**Signals**	**Features**
6 statistical features	Hand EDA Foot EDA RESP	Mean Standard deviation
4 power spectral power features	RESP	Spectral power density in the bands 0-0.1 Hz, 0.1-0.2 Hz, 0.2-0.3 Hz, 0.3-0.4 Hz
8 startle features	Hand EDA Foot EDA	Total number of startle Sum of the magnitudes Sum of the duration Sum of the areas under the startles

The expert-based features, proposed by [81], was also used in the [39] study, which relied on the *drivedb* dataset as well. In such study, the PCA method was used for variable selection. The misclassification rate was estimated using 10 fold cross validation approach. The misclassification rate average based on five fusion algorithms (including LDF, C4.5, SVM, NB and KNN), applied to the PCA selected features, was found to be equal to 0.30 ($sd = 0.07$) while the misclassification rate of the expert based features (22 features) is 0.33 ($sd = 0.10$). The best misclassification rate, estimated using the PCA selected features and the SVM as fusion method, was found to be equal to 0.21 ($sd = 0.03$). The worst misclassification rate, estimated considering the whole features with KNN was equal to 0.38 ($sd = 0.07$). We note that when considering the LDF method, used as well in [81], the misclassification rate was found to be around 0.34 for the two sets of features (complete set of features and PCA selected features).
In summary, we note that our blind approach performances are comparable (within the same range) of the expert based ones.

3.6 Discussion

In this section, we summarize the rationale for our approach and discuss the obtained results. First, we recall that this study was based on the open PhysioNet *drivedb* database, which is a subset of the original data constructed and used in [81] study.
Most studies using *drivedb* data relied on features extraction as a basis for drivers stress level classification in real-world driving experiments. This is in contrast with our study where a functional variable approach, coupled with a multi-resolution wavelet decomposition, are considered as the input for Random Forest-based classification procedure.
The developed approach consisted, at a first stage, in functional variables decomposition using the Haar wavelet. The irrelevant functional variables were removed based on a proposed ?endurance? score. This score reflected the variable ability to persist (in the sense of being selected and well ranked) when applying the variable selection procedure 10 times, based on RF-RFE and grouped variable importance. The corresponding wavelet levels of the retained functional variables were ranked based on the same score. The most "enduring" wavelet levels to the selection procedure were finally retained.
The developed procedure, inspired from [70] work, was applied considering only models composed of $p, p-1, \ldots, 1$ variables, where p is the number of variables. Thus, not all variable combinations

were considered, which could lead to lower error rate. In fact, when p is large it is computationally costly to consider all variable combinations.
Recall that our "blind" 3-step approach allowing to classify driver's stress level using various physiological measurements, resulted in the ordered selection of the *Foot EDA*, *RESP* and *Hand EDA*. Moreover, a finer wavelet-based analysis suggested a total of 3 wavelet signals corresponding to the levels 3 and 4 of *Foot EDA* wavelet details and the approximation corresponding to level 12 of *RESP*. The fact that the EDA emerges as important in driver's stress level recognition is expected since [81] original study reported that features related to *EDA* and *HR* are the most correlated to the stress metric extracted from video-based subjective evaluations.
Furthermore, our findings are in alignment with more recent studies ([122], [23]), which revealed the importance of the EDA sensor placement in a given experiment. In that regard, our study, which included both Hand and Foot EDA, has indicated that the Foot EDA is as important and may be even more relevant in detecting stress levels of the drivers in real-world driving situations.
Several performance indexes were evaluated using different error rates. For that, we assessed the model performances, in terms of misclassification rate, with regards to the retained variables in the classification of the driver's stress level. We then performed a cross-validation like-procedure for model error estimation, taking into account the particularity of the cohort used in the experiments. Finally, we examined the performances of our model using the proposed "blind" approach, as it compares with the results from the "expert-based" approach. The results show that our blind approach performances were comparable (i.e. within the same range) to the expert based ones.

3.7 Conclusion

This study is based on physiological functional variable selection for driver's stress level classification using random forests. The physiological datasets used here consisted in ten real-world driving experiments extracted from the open *drivedb* database. Several physiological signals were collected using portable sensors during real-world driving experiments in three types of routes (rest area, city, and highway) in the Boston area. These signals are: the electrodermal activity measured on the driver's left hand and foot, electromyogram, respiration, and the heart rate.

The contributions of this work are twofold since they touched on the method as well as the application aspects of this problem. The proposed approach was based on RF Recursive Feature Elimination (RF-RFE) and grouped variable importance applied to two different levels of data selection strategies: physiological and wavelet-based variables. For that, all physiological signals were decomposed on a common Haar wavelet basis and then analyzed in search of important variables using a proposed "endurance" score.

The developed method provided a "blind" procedure (without prior information) of driver's stress level classification that resulted in a model error close to the expert-based approach performance. In addition, our approach suggested new physiological features based on the wavelet levels. Moreover, the proposed approach offered a ranking of physiological variables according to their importance in driver's stress level classification in city versus highway driving, with restful period for baseline reference. For the considered case of study, the analysis results suggest that the electromyogram and the heart rate signals are less relevant when compared to the electrodermal and the respiration signals. Furthermore, the electrodermal activity measured on the driver's foot emerged as more relevant than that captured on the hand.

The different changes of road type are the same for all the experiments, since all the drivers follow the same road-map, but occur at different instants. By using classical time warping techniques, one could synchronize these events across the experiments. This would enable the interpretation of each wavelet coefficient and could allow a deeper analysis of the signal by adding an other step to select the most important coefficients within the details selected in Step 3 of our procedure.

One way to further improve the methodology adopted in this study would be to consider costs that penalize the high versus low confusion. In fact, confusing a low stress level with a high one is more serious than predicting the high level as medium one, or the medium as high stress level. This can be easily achieved by introducing different misclassification costs to grow each tree of the forest. The second variant is a preprocessing of the segments extracted from the driving experience recordings. A pre-selection of the coefficients in order to reduce dimension can be performed using wavelet based techniques or Principal Components Analysis for instance, to ensure better stability of the variable selection process. Finally, we could also consider a preprocessing of each driving experience considered as a whole and redefine new original signals realigned on a single time interval by synchronizing typical events.

In this study, the variable of interest was built according to the type of route traveled by the driver, thus it depended mainly on the hypothesis assuming that the stress level increases when driving in the city and decreases when the participant is at the rest. Such hypothesis remains simplified since it does not take into account several factors such as the driver's cognitive state, mental workload, and anticipation of some situations. The effect of the external visual and sonic environmental complexity could as well affect the driving performance ([84]) thus the driver's stress level. Additional information can be gathered using measurements characterizing the affective state of the driver, the driving environment and the inside vehicle complexity (see [52]). Thereby, a finer analysis of the groupings can be proposed for the coefficients resulting from wavelet decomposition as well.

Chapter

4

Self-similarity Analysis of Electrodermal Activity for Driver's Stress Level Characterization

Abstract This chapter characterizes "stress" levels via a self-similarity analysis of the Electrodermal Activity (EDA). For that, the Fractional Brownian Motion (FBM), parameterized via the Hurst exponent H, is evoked to model the EDA changes in a real-world driving context. To characterize the EDA scale invariance, the FBM process and its corresponding exponent H, estimated via a wavelet-based approach, are used. Specifically, an automatic scale range selection is proposed in order to detect the linearity in the logscale diagram. The procedure is applied to the EDA signals, from the open database *drivedb*, captured originally on the foot and the hand of the drivers during a real-world driving experiment designed to evoke different levels of arousal and stress. The estimated Hurst exponent H offers a distinction in stress levels when driving in highway versus city, with a reference to restful state of minimal stress level. Specifically, the estimated H decreases when the environmental complexity increases. In addition, almost all the estimated values are greater than 0.5 suggesting that the EDA signal has a long-range dependence. Furthermore, the H estimated on the Foot EDA signals allows a better characterization of the driving task than the Hand EDA. This self-similarity analysis captures the complexity of the EDA signal. Such analysis was applied to various physiological signals in literature but not to the EDA, a signal which was found to correlate most with human affect. The proposed analysis could be useful in real-time monitoring of "stress" and arousal levels in urban driving spaces.

4.1 Introduction 63
4.2 Data and Methods 64
4.2.1 Real-world driving protocol and data description 64
4.2.2 Fractional Brownian Motion 65
4.2.3 Estimation of Hurst exponent 66
4.2.4 The actual procedure of Hurst exponent estimation 67
4.3 Results and discussion 68
4.3.1 EDA and scale invariant processes 68
4.3.2 EDA-based "stress" level characterization 69
4.3.3 Individual driver-based analysis 70

4.3.4 Discussion 71
4.4 Conclusion and perspectives 72

This chapter written in collaboration with R. Ghozi, J.-M. Poggi, S. Sevestre-Ghalila and M. Jaïdane, is submitted and under review in the Physiological Measurement journal.

4.1 Introduction

"Stress" is a concept that covers a large spectrum of human affect. It is associated with a multitude of emotional and cognitive states, ranging from frustration, anger, and joy to calm and easygoing states. The term stress covers two types of reactions: "distress" designating negative reaction and "eustress", first introduced by [141] referring to positive stress. Stress arises when the human responses capabilities to master stressors are not sufficient. Stressors are in fact challenges that exceed a critical level in terms of intensity, duration or both [106]. It can be defined as the unique set of responses (intellectual, emotional or physiological) to a stimulus [116]. It is thus evident that there is a close link between the level of arousal, attention, mental workload, and stress during the performance of a given task.
On the physiological level, stress is oftentimes associated with an elevated high blood pressure [32], increased heart rate (HR) [19], and increased Electrodermal activity (EDA) level [23]. In fact, EDA has been proved to be highly correlated with stress, especially as this signal is contorlled solely by the activity of the sympathetic branch of the Autonomic Nervous System (ANS) known to be preponderant in stress states [23].
Over the past decades, several studies have confirmed the intrinsic fractal properties in a variety of physiological signals and such characterization was shown to be in close link with human well-being. Indeed, the fractal analysis has been a well established field in biomedical sciences [47, 48, 139]. Some of the first studies which examined the Electroencephalogram (EEG) recordings, during sleep, revealed that chaotic behaviours exist for the EEG measured during sleep stages [13]. Several methods were used to analyze the nonlinear behaviour of human EEG [7, 120]. The resulted complexity indicators, such as Hurst exponent, Lyapunov exponent, and fractal dimension presented good discriminators of normal and epileptic persons [90]. In addition, these indicators allowed to recognize various mental states [110]. More recently, studies used fractal and chaos theory to investigate brain activities and various mental impairments [148, 167]. Several other studies on chaotic and nonlinear behavior focused on electromyogram (EMG) data analysis. In fact, nonlinear measures computed on the human EMG signals were able to characterize patterns of motor unit recruitment in both healthy and diseased patients [65]. The irregularity of heart beat activity was extensively studied (even during rest periods) as a health indicator using various methods to estimate fractal indicators [121]. For instance, complexity analysis of the HR was shown to be a key factor in the diagnosis of human emotion such as anxiety [42] and heart disease [96, 97]. More recent fractal studies of the HR have been used to examine different human activities. For instance, fractal analysis of HR was shown to be of highest sensitivity in distinguishing the heart rate dynamics during rest and exercise periods [162]. In [16] study, a locally fractional Gaussian noise was proposed, to fit marathon heartbeat series. The range of the low frequency fractal parameter allowed to detect heart problems which could help in the prevention of heart failures occurring during marathons. Fractal properties were also used in the analysis and characterization of human gait [77, 109]. In order to assist the diagnosis of breast cancer, multimodal image data (infrared thermograms) were analyzed using multi-fractal analysis [62].
In summary, extensive studies based on fractal analysis of the physiological signals have been carried out, with special focus on human physical and mental health. However, the EDA which has

strong ability to capture human affect, was not characterized thus far using the fractal measures. It is worth noting the relatively recent study of [31], that concerns the extraction of nonlinear EDA-based features in order to recognize the affective state of subjects, while watching different visual scenes.
In this chapter, a wavelet-based method is used to locally characterize an EDA signal using the Hurst exponent (H). The use of such method is motivated by the fact that wavelets are naturally suitable to analyze scaling processes, for they are easy to implement, and having fast and efficient algorithms [5]. In the context of a real-world driving experiment, we show that the estimated the Hurst exponent H, on different segments of EDA signals, offers a distinction between low, medium and high stress levels induced respectively by rest, highway and city driving. The fractal properties of the EDA signals (measured on foot and hand) are examined in the situation of repeated driving experiments, in order to characterize its ability to distinguish individual versus group tendencies as well.
The remainder of the chapter is organized as follows: Section 3.2 presents the EDA data and the method adopted in its self-similarity study. Section 3.3 details the results of the EDA self-similarity behavior, followed by a discussion. Section 3.4 gives a conclusion and perspectives of the study.

4.2 Data and Methods

In this section, we first give a description of the real-world driving protocol and we present the generated data. We then introduce the Fractional Brownian Motion (FBM) which is one of the classical self-similar processes, parametrized by the Hurst exponent. The theoretical framework of the estimation of Hurst exponent are briefly described with a special focus on the wavelet-based approach. Finally, we present the estimation procedure used in our study.

4.2.1 Real-world driving protocol and data description

The data used in this study is based on the public dataset[1] *drivedb* freely accessible on PhysioNet website [66]. This database resulted from the protocol, proposed by [80], that consisted of real-world driving over 20 miles in the Boston area. The driving route was chosen to mimic a daily trip that provoked different stress levels during a normal daily stress ranges [80]. The total duration of each drive depends on traffic congestion, and thus varied from 50 to 90 minutes. It included periods of rest, highway and city driving assumed to produce respectively low, medium and high levels of stress. Two subjective analyses, described in [81], allowed to validate these assumptions. They both consisted on an analysis of the survey results and the evolution of a score derived from video tapes recorded during each drive. Several physiological measurements were collected such as the electrodermal activity (EDA) captured on the hand and the foot (denoted by HandEDA and FootEDA resp.), the electromyogram (EMG), the electrocardiogram (ECG) and the driver respiration. The *drivedb* database contains signals related only to 18 driving experiments (hence drive). We propose here to consider only the 10 complete datasets related to drives performed by 4 participants. Two complete datasets resulting from 2 drives of Driver1 who has 3 years of driving experience. Driver2 contributed with 3 complete datasets related to 3 drives while 4 datasets from Driver3 who has 8 years of driving experience. Driver4 performed only one drive that provided the last complete dataset. Note that the EDA signals were low pass filtered in order to eliminate high frequency noise [81]. For the complete 10 drives, we select the HandEDA and the FootEDA measurements to characterize driver's stress level using the self-similar processes in this study.

[1] http://physionet.org/physiobank/database/drivedb/

4.2.2 Fractional Brownian Motion

A Fractional Brownian Motion (FBM) is a Gaussian self-similar process with stationary increments (sssi-) process with $0 < H < 1$. It is a popular self-similar (ss-) process, widely used in several fields ranging from physics or biology to telecommunication and finance. In fact, only one parameter, the Hurst exponent H, governs its properties such as short/long range dependence, regularity, and fractal properties explaining thus its versatility to inspect, scan and model various aspects of the data. In this section, we recall definitions of ss-processes, sssi-processes, FBM, and Long-range dependent (LRD) processes, mainly based on the paper of [152] which is a precious introductory text on such processes.
Self-similar processes are stochastic processes exhibiting some invariance properties in their distributions. A stochastic process $S = \{S(t)\}_{t\in\mathbb{R}}$ is a ss-process with $H > 0$ if for every $\lambda > 0$:

$$\{S_H(\lambda t)\}_{t\in\mathbb{R}} \overset{d}{=} \{\lambda^H S_H(t)\}_{t\in\mathbb{R}}. \tag{4.1}$$

where $\overset{d}{=}$ denotes the equality in terms of the finite dimensional distributions. This means that the distribution of a ss-process is essentially scale invariant. In fact, the statistical properties of the ss-process in one scale completely determine those of any other scale.
A process $S = \{S(t)\}_{t\in\mathbb{R}}$ has stationary increments if for all $u \in \mathbb{R}$,

$$\{S(t+u) - S(u)\}_{t\in\mathbb{R}} \overset{d}{=} \{S(t) - S(0)\}_{t\in\mathbb{R}}. \tag{4.2}$$

When a process is self-similar with an index H and has stationary increments, it is called H-sssi process. Such a process $S = \{S(t)\}_{t\in\mathbb{R}}$ with finite variance (for any $t \in \mathbb{R}, E[S^2(t)] < \infty$) has the following properties:

(i) $S(0) = 0$, and $S(-t) \overset{d}{=} -S(t)$.

(ii) For all $t \in \mathbb{R}$, if $H \neq 1$ then $E[S(t)] = 0$.

(iii) $E[S^2(t)] = |t|^{2H}\sigma^2$ where $\sigma^2 = E[S^2(1)]$.

(iv) The covariance function of H-sssi process S is defined for $t, t' \geq 0$ as

$$E[S(t)S(t')] = \frac{\sigma^2}{2}(|t|^{2H} + |t'|^{2H} - |t - t'|^{2H}). \tag{4.3}$$

Its popularity arises also from the proposition allowing to check whether a given process is FBM. In fact, if $S = \{S(t)\}_{t\in\mathbb{R}}$ is a centered Gaussian process with stationary increments such that $S(0) = 0$, and $E[S^2(t)] = |t|^{2H}\sigma^2$ for some $\sigma^2 > 0$ and $0 < H < 1$, then $S = \{S(t)\}_{t\in\mathbb{R}}$ is FBM.
Interpretation of a FBM path can be deduced from the value of its corresponding H. In fact, a FBM gets less irregular as H increases from 0 to 1. More precisely, a FBM can be divided into anti-persistent ($0 < H < \frac{1}{2}$), chaotic ($H = \frac{1}{2}$) and persistent ($\frac{1}{2} < H < 1$) processes (See Fig. 4.1).
Let us consider the increments of a H-sssi process $S = \{S(t)\}_{t\in\mathbb{R}}$, with $0 < H < 1$ defined by:

$$G_k = S(k+1) - S(k), \quad k \in \mathbb{Z}. \tag{4.4}$$

If $S = \{S(t)\}_{t\in\mathbb{R}}$ is a FBM, then its incremental process $\{G(k)\}_{k\in\mathbb{Z}}$ is called Fractional Gaussian Noise (FGN).
The increment sequence $\{G(k)\}_{k\in\mathbb{Z}}$ of a H-sssi process $S = \{S(t)\}_{t\in\mathbb{R}}$, with $0 < H < 1$ has the following properties:

(i) $\{G(k)\}_{k\in\mathbb{Z}}$ is stationary, $E[G_k] = 0$, and $E[G_k^2] = \sigma^2 = E[S(1)]^2$.

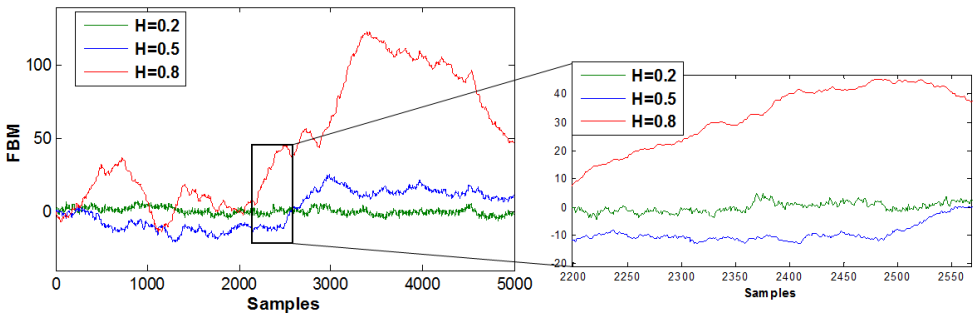


Fig. 4.1 Samples of simulated FBM for H=0.2 (green), H=0.5 (blue) and H=0.8 (red).

(ii) The auto-covariance function of the process $\{G_k\}_{k\in\mathbb{Z}}$ is defined by

$$\gamma(k) = \frac{\sigma^2}{2}(|k+1|^{2H} - 2|k|^{2H} - |k-1|^{2H}). \tag{4.5}$$

(iii) If $H = \frac{1}{2}$ then $\gamma(k) = 0$, if $0 < H < \frac{1}{2}$ then $\gamma(k) < 0$ and if $\frac{1}{2} < H < 1$ then $\gamma(k) > 0$, for $k \neq 0$.

(iv) If $H \neq \frac{1}{2}$ then

$$\gamma(k) \sim \sigma^2 H(2H-1)|k|^{2H-2}, \text{ as } k \to \infty. \tag{4.6}$$

Property (iv) implies that $\gamma(k)$ tends to 0 as $k \to \infty$ like a power function. If $\frac{1}{2} < H < 1$ then it tends to 0 so slowly that $\sum_{i=-\infty}^{+\infty} \gamma(k)$ diverges. The process exhibits thus a long-range dependence and is called LRD process. However, when $0 < H < \frac{1}{2}$, $\sum_{i=-\infty}^{+\infty} |\gamma(k)| < \infty$ and $\sum_{i=-\infty}^{+\infty} \gamma(k) = 0$, the process shows no long-range dependence.
Remark: Mention the link between the Hurst exponent and the fractal dimension since the self-similar process is often characterized by the fractal dimension D_f directly related to H by $D_f = 2 - H$.

4.2.3 Estimation of Hurst exponent

The most popular techniques of Hurst exponent estimation can be performed in time domain (for instance the rescaled range analysis), frequency domain (e.g. log-periodogram for example) or in a combined time-scale domain (wavelet-based estimation). One can refer to the book edited by [44] for an overview, theoretical and applied aspects of the different estimation methods.
In this work, we adopt a wavelet-based method since wavelets are considered as a natural tool in the study of scale invariant processes. Wavelet analysis has been widely used to detect features at different scales of the signal.
Let ψ be a mother wavelet having M first vanishing moments, thus there exists $M \in N^*$ such that:

$$\int_{\mathbb{R}} t^m \psi(t)dt = 0, \text{ for all } m \in \{0, 1, ..., M\}. \tag{4.7}$$

We define the family of functions $\psi_{j,k}$ by $\psi_{j,k}(t) = \frac{1}{2^{j/2}}\psi(2^{-j}t - k)$ with $j \in \mathbb{Z}$ and $k \in \mathbb{Z}$, 2^j is the *scale* and k is the *position*.
The discrete wavelet transform of a process $\{S(t), t \in \mathbb{R}\}$ is defined as:

$$d_{j,k} = \int_{\mathbb{R}} S(t)\psi_{j,k}(t)dt, \text{ where } j, k \in \mathbb{Z}. \tag{4.8}$$

The wavelet-based method to estimate the Hurst exponent H has been introduced first by [57] for self-similar processes and developed then by [5, 17, 108] for long range dependent (LRD)

processes. This study relies on [159] work and their associated routines available online on [160]. In fact, the proposed method to estimate H for a ss-process S (typically a FBM) is based on $(d_{j,k})_{k\in\mathbb{Z}}$ which is a zero-mean process for a fixed wavelet level $j \in \mathbb{Z}$ and

$$\mathbb{E}(d_{j,k}^2) = Var(d_{j,k}) \sim C(\psi, H)2^{j(2H+1)} \quad \text{for all} \quad 2^j > 0. \tag{4.9}$$

It was shown that for a stationary LRD process S (typically FGN) , $(d_{j,k})_{k\in\mathbb{Z}}$ is a zero-mean process for a fixed $j \in \mathbb{Z}$ and

$$\mathbb{E}(d_{j,k}^2) = Var(d_{j,k}) \sim D(\psi, H)2^{j(2H-1)} \quad \text{for} \quad 2^j \to \infty, \tag{4.10}$$

where $C(\psi, H)$ and $D(\psi, H)$ are positive constants which depend on ψ and H. Thus, for the two types of processes, the variance of the wavelet coefficients is a power law of 2^j of the form $2^{j\alpha}$ where $\alpha = 2H + 1$ in the case of ss-processes and $\alpha = 2H - 1$ for LRD processes. Therefore, the slope coefficient of a log-log regression of this variance on j provides an estimator of α and thus of H.
In practice, only a discrete sample path of a process is available. For a time series $(S_1, ..., S_N)$ of the random process S, the fast Mallat's algorithm [105] provides approximation of wavelet coefficients defined by:

$$e_{j,k} = \sum_{n=1}^{N} \psi_{j,k}(n)S(n). \tag{4.11}$$

Let us consider $\mu_j = \frac{1}{N/2^j}\sum_{i=1}^{N/2^j} e_{j,i}^2$ and $\sigma_j^2 = Var(log_2(\mu_j))$. An asymptotically unbiased and efficient estimator $\hat{\alpha}$ of the power law exponent α, proposed by [159], and was defined using the weighted regression on the logscale diagram that presents $log_2(\mu_j)$ versus $log_2(2^j) = j$. More precisely,

$$\hat{\alpha} = \sum_{j=j_1}^{j_2} w_j \times y_j = \sum_{j=j_1}^{j_2} w_j \times (log_2(\mu_j) - g_j), \tag{4.12}$$

where j_1 and j_2 are the wavelet scale limits of the range where the linearity of the logscale is observed, where $w_j = \frac{(Q\times j - Q_j)/\sigma_j^2}{Q\times Q_{jj} - Q_j^2}$, $Q = \sum \frac{1}{\sigma_j^2}$, $Q_j = \sum \frac{j}{\sigma_j^2}$, $Q_{jj} = \sum \frac{j^2}{\sigma_j^2}$,
and where g_j are constants used for the fact that $log_2(\mathbb{E}(d_{j,k}^2)) \neq \mathbb{E}(log_2(d_{j,k}^2))$.

4.2.4 The actual procedure of Hurst exponent estimation

The routines proposed in [160] allows to generate the logscale diagram, to estimate the Hurst exponent for different scaling processes (for instance ss-processes, LRD processes) and to automatically select the lower index j_1 of the scale range $[j_1, j_2]$ over which the logscale is linear for LRD processes. In this work, first of all, each signal segment is fully decomposed using the Haar wavelet. The logscale diagram is then built. If a global linearity is verified, the power law parameter $\hat{\alpha}$ is estimated and the exponent $\hat{H}$ can be deduced depending on the range of $\hat{\alpha}$. If $1 < \hat{\alpha} < 3$ then $\hat{H} = \frac{(\hat{\alpha}-1)}{2}$ and if $0 < \hat{\alpha} < 1$ then $\hat{H} = \frac{(\hat{\alpha}+1)}{2}$.
When the linearity is not verified for all the scales, an automatic selection of the wavelet scale range $[j_1, j_2]$ over which the logscale is linear, using the procedure given by Table 4.1. The proposed selection procedure assumes that a region of linearity is observed in a range of wavelet scales (see Fig. 4.2 commented in the next section).
Note that the choice of δ depends on the degree of the linearity in the logscale diagram. Indeed, the value of δ should be small when the linearity is strong. For instance, we pick $\delta = 0.01$ since the linearity in our case is strong.
We recall that this study aims mainly to characterize EDA signals using self-similar processes in order to assess the driver's stress level during a real-world driving experience.

Table 4.1 Automatic selection of the wavelet scales over which the logscale is linear.

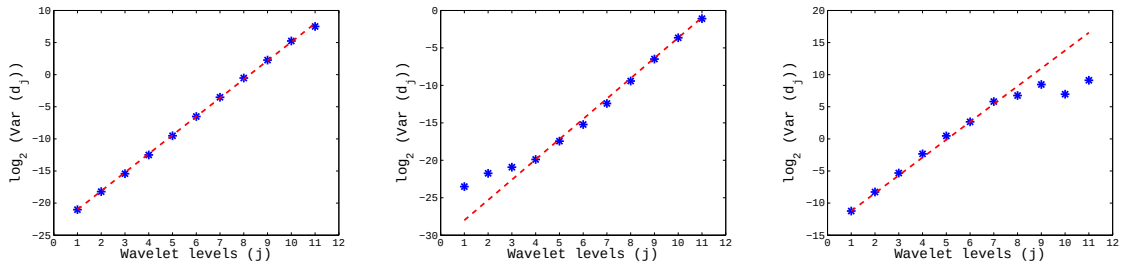
Inputs: $\delta > 0$, J and y_j for $j = 1...J$
Outputs: j_1 and j_2
1. Compute the slope of each line connecting two consecutive points of the logscale diagram, for all the wavelet scales, using the first difference
2. Compute the difference between the consecutive slopes
3. If the index that corresponds to the largest absolute value of the difference is greater than $J/2 + 1$ then
- $j_1 = 1$
- For $l \in [3...J]$, compute $R_a^2(l)$: the adjusted R^2 of the model that fits the data on the range $[j_1, l]$
- $j_2 = \text{argmax}_{3 \leq l \leq J}(M - R_a^2(l) < \delta)$ where $M = max\ (R_a^2(l))$
else
- $j_2 = J$
- For $m \in [1...J-3]$, compute $R_a^2(m)$: the adjusted R^2 of the model that fits the data on the range $[m, j_2]$
- $j_1 = \text{argmin}_{1 \leq m \leq J-3}(M' - R_a^2(m) < \delta)$ where $M' = max\ (R_a^2(m))$

4.3 Results and discussion

We present the self-similarity analysis performed on the EDA signals. We recall that we applied a wavelet-based approach to estimate the Hurst index H on the EDA captured on the hand and on the foot, extracted from the open *drivedb* database. Ultimately, we seek to check if the Hurst exponent, estimated on EDA signals, can be used to characterize the driver stress level evolution during real-world driving environment.
We recall that each EDA signal corresponding to a drive is subdivided into segments. A proper time window was selected to be long enough to capture specific information related to the human affective reactions. We recall that multiple physiological studies have confirmed that EDA has relatively long latencies (between 1 and 5 seconds in cases of skin cooling) compared to other physiological signals [46]. In addition, the time analysis window depends on the type of the experiment, mainly the type of the stimuli evoking EDA changes in the driver. The smallest time window for the statistical evaluation of the EDA ranges from 5 seconds to 1 minute [23]. In the *drivedb* data, the largest window size that can be considered is 5.20 minutes to make sure that every EDA segment corresponds to one and only one driving condition. This time duration corresponds to the second city driving of the drive 5. Therefore, we have opted for a time window analysis of 3 minutes.

4.3.1 EDA and scale invariant processes

Before estimating a value of H, the logscale diagram allows a visual inspection and a search for wavelet scales for which the ss-process properties hold. When inspecting Fig. 4.2 that depicts the logscale diagrams for the different EDA segments, a linearity trend is observed. The diagram can be linear for all the wavelet scales (cf. Fig. 4.2 (a)) or for some scale ranges (cf. Fig. 4.2 (b) and Fig. 4.2 (c)). The most encountered case consisted on the existing linearity for almost all scales expect for the highest scales (cf. Fig. 2 (c)). This case corresponds usually to city or highway driving while the case of the linearity for all wavelet scales or for almost all scales expect for the lowest ones is observed for rest periods.



(a) **(b)** **(c)**

Fig. 4.2 Plots of the logscale diagram: illustrations of the linear tendencies in the logscale graphs based on 3-minutes segments extracted from HandEDA.

Beyond the visual inspection, the logscale diagram offers a byproduct of such estimation procedure that illustrates the scale invariance character of EDA signal and that provides a realistic range of scales for which the phenomenon is observable.
Once the slope estimation is performed, one needs to destinguish a FBM ($1 < \alpha < 3$) from a FGN ($0 < \alpha < 1$). In fact in our case, all the values of the estimated slopes $\hat{\alpha}$ for the different EDA segments ranged between 1 and 3 with a median value for HandEDA of 2.51 and for FootEDA of 2.47, leading to $\hat{H} = \frac{(\hat{\alpha}-1)}{2}$.

4.3.2 EDA-based "stress" level characterization

We develop in this section the analysis of self-similarity of HandEDA and FootEDA per drive. The temporal evolution of the the Hurst exponent H is first inspected. For that, a 3-minutes segment is extracted, the corresponding H is estimated. One minute is skipped and the next 3-minutes are then considered and the corresponding Hurst exponent is estimated again. These two steps are repeated until the complete EDA signal is swept (cf. Fig. 4.3 for illustration).
The distributions of $\hat{H}$ estimated on segments from respectively HandEDA and FootEDA are shown in Fig. 4.4 and Fig. 4.5, for the different driving conditions (city, highway and rest) and this for each drive. It is worth noting that segments that fall in two different driving conditions are excluded from our analysis, for example a segment of 3 minutes that covers rest and city driving is not considered in our study. Fig. 4.4 and Fig. 4.5 indicate that the values of the estimated H are greater than 0.5 except for the FootEDA drive8. Values greater than 0.5 reflect that the driver's EDA, captured on both hand and foot, is a persistent positively correlated times series having a trend. This finding confirms the conclusions of [139] proposing that physiological sequences are persistent motion or noise characterized by a Hurst exponent greater than 0.5.

For Driver1, the values of $\hat{H}$ on both HandEDA and FootEDA corresponding to Drive1 are greater than those corresponding to Drive2. This indicates that the stress level for Driver1 increases with driving sequence. However, the boxplots corresponding to the $\hat{H}$ distributions computed on both HandEDA and FootEDA, for Driver2, are lower for this driver's first drive than his second one. This may be explained by the fact that he had not driven a month previous to the experiment and he is not used to drive in Boston area. One also notices that the boxplots corresponding to the values of the estimated H on FootEDA for Drive8 are lower than the other boxplots related to the distributions of $\hat{H}$ for Driver3. This would inform about more complexity encountered during the third drive. Significant interpretations could not been deduced on the rest period of Drive5 and Drive10 since these two conditions are missing the second rest period data.

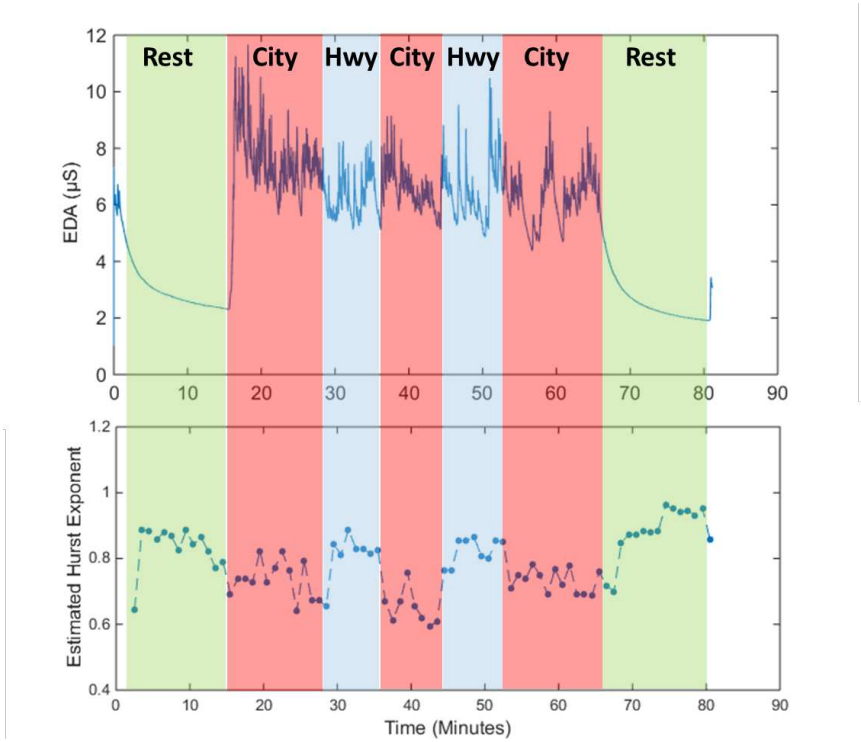


Fig. 4.3 FootEDA corresponding to the Drive 4 of *drivedb* database is depicted in the top and the corresponding estimated Hurst exponent is presented in the bottom of the Figure. The estimation is done on segments of 3 minutes with 1 minute of overlapping. The term "Hwy" is used to designate highway driving.

4.3.3 Individual driver-based analysis

We consider and compare the distributions of $\hat{H}$ for the periods: rest, city and highway, grouped by driver. The question to be answered here is: how does the $\hat{H}$ characterizes the city-highway-rest for the different drivers?
Fig. 4.6 shows that the values of $\hat{H}$ estimated on both HandEDA and FootEDA increase when the driving environment complexity decreases. In order to further understand the ability of EDA to characterize the stress level, a comparison of HandEDA versus FootEDA is of interest. The Wilcoxon test is performed in order to compare $\hat{H}$ distributions of the different driving conditions for the four drivers and for both HandEDA and FootEDA. For HandEDA, the boxplots corresponding to $\hat{H}$ estimated on city segments are significantly ($p < 0.01$) lower than those computed on highway segments, except for Driver4 ($p = 0.05$). The boxplots representing distributions of $\hat{H}$ estimated on highway driving segments are lower than rest periods ($p < 0.1$ for Driver3 and Driver4) and the difference is not significant for Driver1 and Driver2. For FootEDA, all the differences between City and Highway driving are significant ($p < 0.01$) and the boxplots of $\hat{H}$ distributions are significantly lower than rest ($p < 0.01$ for Driver4 and $p < 0.05$ for all the other drivers). This suggests that the FootEDA characterizes better the driving periods thus the stress level for each driver. In summary, when considering the different drivers, the difference between city and highway driving is significant for both HandEDA and FootEDA. However, this difference is not signficant for all drivers in the HandEDA.
Now we consider the estimated H for all drives as a whole then check the difference between the triplet city-highway-rest. Fig. 4.7 shows that the difference is important especially for HandEDA. This suggests that for the driving experiment, one should consider each drive separetly in order to better assess the self-similarity of the EDA signal.

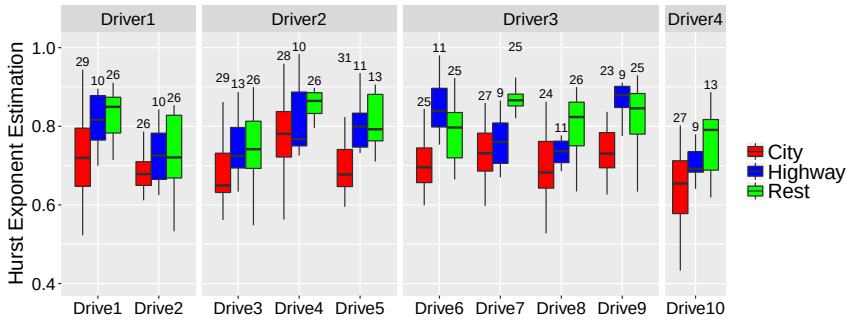


Fig. 4.4 The boxplots of estimated H on HandEDA per drive for the 4 drivers. For each drive, the boxplot is presented according to the different driving periods. The sample size is designated on the top of each boxplot.

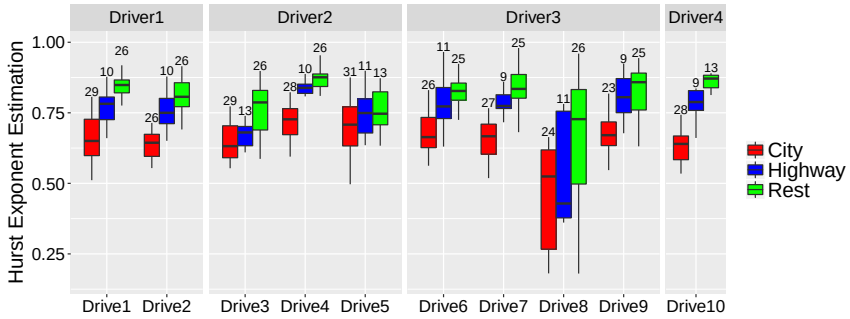


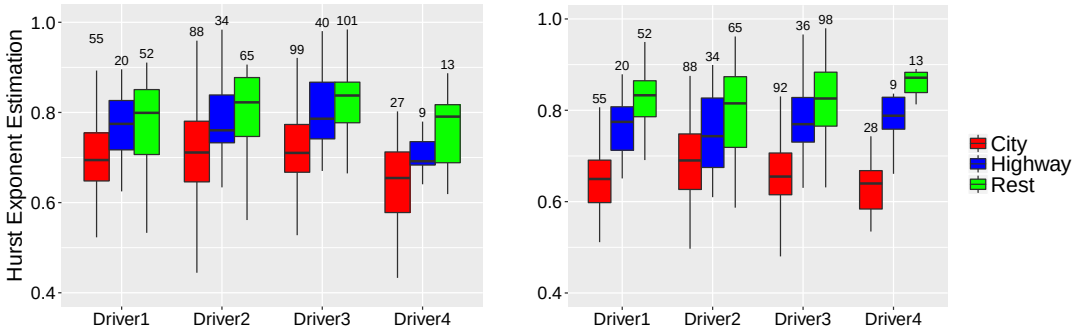
Fig. 4.5 The boxplots of estimated H on FootEDA per drive for the 4 drivers. For each drive, the distribution is presented according to the different driving periods. The sample size is designated on the top of each boxplot.

4.3.4 Discussion

This study investigated how the stress in real-world driving is captured via a fractal analysis of Electrodermal activity (EDA). EDA fluctuations were captured using the Hurst exponent estimated on rest, highway and city driving conditions. It should be mentioned that "stress" was used here in alignment with the original study of real-world driving [81]. In fact, the term "arousal" is more appropriate physiologically when using the EDA signal. The stress level here is assumed to be higher in city when compared to highway.

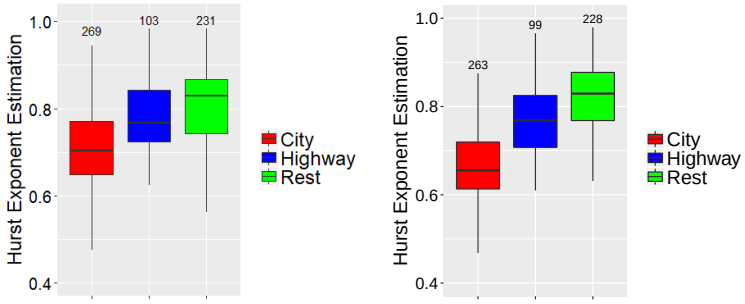
On the methodological side, segments extracted from EDA signals, were characterized as Fractional Brownian Motion (FBM) using Hurst exponent H. We used a wavelet-based analysis to estimate the Hurst exponent. Such approach appears natural since the non-stationarity and the self-similarity of FBM require respectively time and scale dependent analysis. Several other methods can be used to estimate H. Moreover, the window size of the extracted EDA segments is of 3 minutes. This choice was motivated by the fact that a minimum duration is between 5 seconds and 1 minute (in order to see a stimuli reaction in EDA segment) and the maximum is 5.20 minutes (due to the database limitation). This parameter deserves to be more investigated and a more elaborated study could be of interest, like a cross-validation type procedure. In addition, this study used Haar wavelet decomposition thus Daubechies wavelet with one vanishing moment. This number of vanishing moment could be changed and to inspect its effect on the characterization of the EDA as a ss-process and thus the resulting H index values.

In this study, we suggested that FootEDA is more appropriate to be used and allows to monitor the stress level while driving. This finding may be approved by the study [157] which proposed that EDA measurements captured on foot are similar to that acquired from the fingers. In fact, EDA is usually measured on fingers or on palms which are the best placements to obtain the



(a) **(b)**

Fig. 4.6 The boxplots of $\hat{H}$ estimated on (a) HandEDA and (b) FootEDA per driver. The number on the top of each boxplot designates the sample size.



(a) **(b)**

Fig. 4.7 The boxplots of $\hat{H}$ estimated on (a) HandEDA and (b) FootEDA per driving condition. The sample size is indicated on the top of each boxplot.

precise EDA measurements. However, these two placements are not preferred in real-world task performance because they may be intrusive and disturb the task to be performed such as driving. The estimated H of EDA signals proposed a novel feature that can inform about the driver's state. This feature should be compared to startle features, proposed by [81], in terms of stress level classification performances.

4.4 Conclusion and perspectives

This analysis evoked the characterization of the Electrodermal activity (EDA) using self-similar processes. For that, a wavelet-based method of Hurst exponent estimation was used to characterize the EDA using a Fractional Brownian Motion.
When applying this method to EDA data derived from real-world driving experience, a problem of nonlinearity in the logscale diagram was encountered. Thus, an automatic scale selection was proposed in search of linearity, which reflects the self similarity behavior. The estimated Hurst exponent offered a distinction in stress levels among the divers, as they proceed from highway to city, with reference to rest (established for minimal stress level). The boxplots of the estimated Hurst exponents are almost greater than 0.5 suggesting that EDA paths have long range dependence. Moreover, those boxplots were higher when the stress level decreases; this is most likely due to the difference in the scene complexity between the city and the highway. Benefiting from the fact that the wavelets are easy to program and fast, the proposed approach can be easily integrated in vehicle and could serve for "real-time" monitoring of a driver's arousal state.

Future works will examine a multi-fractal analysis in the case of the EDA signals. More non-linear indexes can be used to investigate the fractal properties of the EDA. We propose also to adapt [160] routines allowing to select the scale onset to the case of ss-process. Moreover, and since the *drivedb* database contains more physiological data, such as EMG, HR, and respiration, a study of the relation between the stress levels and such physiological signals may be used to better recognize the stress level. An on-line version of the proposed approach will be implemented and adapted for the case of the EDA signals since a real-time version of the wavelet-based estimator proposed by [137], already exists for the long range dependency processes.

Chapter

5

AffectiveROAD System and Database to Assess Driver's Arousal State

Abstract Thanks to the rise of new wearable and non-intrusive sensor technology, Internet of Things (IoT) contributes in human daily life improvement. In the context of smart vehicles, human affective monitoring should be based on a context-aware system in order to consider the interactions between the driver, his vehicle and his ambient environment. In this chapter, we propose *AffectiveROAD* platform, that sense the human physiological changes, the ambient environment inside the vehicle, and the vehicle speed. The proposed sensor-based solutions are not only providing real-time physiological monitoring, but also enriching the tools for human affective and cognitive states tracking. Thanks to this platform, several driver's state indicators such as stress and arousal may be developed and validated. Two types of wireless physiological sensors are used to monitor the electrodermal activity, the heart rate, the skin temperature, the respiration, and the motion of the driver. Moreover, we developed a sensor network allowing to capture the ambient temperature, humidity, pressure, and luminosity. The vehicle speed is extracted from the Global Position System (GPS) data captured using a smartphone. Two GoPro devices are used to capture the internal and external scenes. The purpose of this chapter is to describe a real-world driving protocol to collect data using the proposed IoT-based materials and to announce the publication of a database for driver's state monitoring research. We propose 13 datasets related to drives in different road types: city and highway. A part of the database concerning the physiological and the environmental data is released for public use.

5.1 Introduction 75
5.2 Overview of real-world driving experiences systems 76
5.2.1 MIT system for driver's stress detection 76
5.2.2 hciLab system for driver's workload assessment 76
5.2.3 Warwick-JLR system for driver monitoring research 76
5.3 AffectiveROAD sensors network 77
5.3.1 Physiological Wearable Sensors 77
5.3.2 Ambient Environment Sensors of AffectiveROAD 78
5.4 Real-world driving protocol 78

5.4.1 Apparatus 79
5.4.2 Cohort Description 80
5.4.3 Route Description 81
5.5 Data Collection and Construction 82
5.6 Related works and discussion 83
5.7 Conclusion and future work 84

This chapter in based on a paper written in collaboration with J.-M. Poggi, R. Ghozi, S. Sevestre-Ghalila and M. Jaïdane, which is published [52].

5.1 Introduction

The expansion of the technology of electronic devices and sensors leads the growth and the rise of demand in the Internet of Things (IoT) field. These sensors are connected to the Internet and to each other. Cisco IBSG [1] [34] predicted that by 2020, there will be 50 billion devices connected to Internet, which is equivalent to 6.6 connected devices per person. IoT technology can be used in several areas such as industrial systems, health care and medical systems, smart cities, transportation, and many others.
In smart vehicle systems, the main challenge is the design of human-centric systems that do consider all type of users: driver's with health problems, elderly with age related hearing loss that affect the driving performance [18], persons whom driving is their job such as taxi and bus drivers and young drivers without experience. Recently, the small size of sensors and devices, their good accuracy and non intrusiveness make them easier to be connected and integrated in vehicles.
Studies that aim to recognize driver's state are based on subjective and objective methods: subjectively, drivers report their state when responding to questionnaire [81] or when rating post-experience videos [81, 140]. Objective measurements are captured via sensors such as physiological sensors or ambient environment sensors. Such methods can be used when conducting experiments in car simulator [21, 94] or even in real-world driving context [134, 140, 153].
In the context of real-world driving experience, the used objective measurements are based on the collection of physiological data like in [81, 133], based on information gathered from the vehicle's Controller Area Network (CAN) [153] or from smartphones [140]. It should be mentioned that the used physiological sensors contains electrodes that are external to the sensors, snapped to the connectors.
In this Chapter, we propose a platform *AffectiveROAD* for affective recognition of the driver's arousal state. It is based on sensor network that measures physiological signals and ambient environment information. The considered physiological sensors in our system are based on wristband non-bulky offered by Empatica and Zephyr BioHarness 3.0 chest belt. We used such sensors in order to have the electrodes integrated in the sensor to avoid the intrusiveness and the discomfort caused by wearing external electrodes. Data related to the electrodermal activity, heart rate, skin temperature, hand motion and respiration are gathered using *AffectiveROAD* physiological sensors. The proposed system includes also ambient environment sensors that capture the state of the in-car ambiance. We developed a prototype of sensor network based on temperature-humidity-pressure, luminance, and air quality sensors. A Bluetooth connection allows the Empatica E4 to turn on the environmental sensors network. The novelty of *AffectiveROAD* environmental platform lies in the inclusion of sensors that inform about the in-vehicle thermal comfort which is an important factor in the driving task well performance [37].

[1] Cisco IBSG stands for Internet Business Solutions Group. IBSG is the premier leadership group within Cisco which is a leader in telecommunication and networking hardwares

One main focus of the chapter is the description of the physiological and environmental devices used to collect data during a real-world driving experience. We have designed a protocol involving several drivers. All participants along a route supposed to mimic a daily-life metropolitan trip alternating between city and highway driving periods assuming to induce respectively high and medium stress levels. Several data are collected from sensors, added to information coming from questionnaire and video-based rating session.
The rest of this chapter is organized as follows. In Section 4.2, an overview of sensors network in real-world driving experiences is provided. The description of *AffectiveROAD* sensors is presented in Section 4.3, while the real-world driving protocol is described in Section 4.4. Then, we present the *AffectiveROAD* database in Section 4.5, related work in Section 4.6 and conclude the chapter with future directions of research in Section 4.7.

5.2 Overview of real-world driving experiences systems

Several methods have been explored in the automotive domain to assess the driver's state. These methods can be either subjective, based on the user's reporting or objective such as capturing the user's physiological properties.
Assessing the driver's state in a real-world context is hardly possible. To mitigate this, several solutions rely on the use of simulator. However, this solution is not encouraged as the driver is aware that he navigates through a virtual world. With the expansion of sensors technologies, assessing the driver's state become easier. Despite the fact that several studies are achieved to do this, few researches proposed an open dataset of their collected data in order to help researchers to validate their works.
In this section, we propose to describe three main studies, offering an open database based on physiological and contextual data in order to monitor and assess the driver's state.

5.2.1 MIT system for driver's stress detection

The MIT system [80] is used to present methods for collecting and analyzing physiological data during real world driving tasks to determine a driver's relative stress level. Electrocardiogram, electromyogram, electrodermal activity and respiration were recorded continuously while drivers followed a set route through open roads in the greater Boston area. These sensors were connected to a FlexComp unit which was connected to an embedded computer in a modified Volvo S70 series station wagon. A part of the data collected during the experiments [2], is published online as the *drivedb* database available on the website of PhysioNet [66].

5.2.2 hciLab system for driver's workload assessment

Human-Computer-Interaction laboratory (hciLab) of University of Stuttgart presented a real-world driving study [140] assessing different physiological and contextual information. Environmental and vehicle data such as brightness, acceleration and GPS data were recorded using smartphone. Nexus 4 physiological sensing system is used to record driver's physiological data. It contains mainly electrocardiogram, electrodermal activity and skin temperature sensors. The data are publicly available on *hcilab* driving dataset [1].

5.2.3 Warwick-JLR system for driver monitoring research

The study [153] that was conducted by the Department of Computer Science of Warwick University co-funded by Jaguar Land Rover Cars (JLR)- monitored driver's workload using sensors accessible via the vehicle's Controller Area Network (CAN). The driver's workload is evaluated

using features extracted from ECG and EDA signals. Those signals were recorded via a GTEC USB biosignal amplifier (USBamp), connected to the ECG electrodes and EDA wires. A part of the collected data is accessible on the *Warwick-JLR* database [3].

5.3 AffectiveROAD sensors network

In this section, we present our proposed platform *AffectiveROAD*. This platform contains mainly physiological wearable sensors and ambient environment sensors.

5.3.1 Physiological Wearable Sensors

The used physiological sensors are based on a smart wristband provided by the company Empatica[2] and a smart chest belt designed by the manufacturer Zephyr[3] which is now a part of Medtronic[4], a leader in healthcare technologies.
Empatica E4[5] offers an unobtrusive monitoring allowing to record signals in laboratory as well as in inside-home contexts. The recorded data are uploaded in Empatica's secured cloud platform. Hence, users can easily access the recordings. The device offers the possibility to personalize and develop a mobile API in order to access and stream the gathered data in real-time .
It contains mainly Photoplethysmography sensor that measures the Blood Volume Pulse (BVP), a 3-axis accelerometer, an Electrodermal Activity (EDA) sensor, an infrared thermopile, an internal real-time clock and an event mark button. The Heart Rate (HR) is derived from the BVP. The EDA sensor is used to measure sympathetic nervous system arousal which is highly correlated to stress [81] and excitement [23, 128].
The weight of the E4 device is 25 g and the dimensions of the case are $44 \times 40 \times 16$ mm. These characteristics make the wristband non-bulky and easy to use (cf. Fig. 5.1).

Fig. 5.1 A photo of the plugged E4 in the left hand of a driver. It is unobtrusive device allowing to perform the driving task.

Moreover, the battery characteristics (such as charging time less than 2 hours and autonomy[6]) makes this product very useful. Its flash memory is of more than 60 hours of data storage.
Several modes are proposed for the E4 use: a recording and a streaming mode. When recording, the wristband stores data in its internal memory. Data are later downloaded via USB through

[2]https://www.empatica.com/
[3]https://www.zephyranywhere.com/
[4]http://www.medtronic.com/us-en/index.html
[5]https://www.empatica.com/e4-wristband
[6]In the streaming mode, the autonomy of the battery can exceed 20 hours. When using the memory mode, we have at least 36 hours of battery autonomy

the Empatica Manager. The E4 wristband connects to a smartphone or desktop computer via Bluetooth. Thus, the streaming mode is offered when using the real-time application and mobile API for iOS and Android mobile devices and desktop integration for Windows and Mac.
The second wearable device used in our platform is Zephyr BioHarness 3.0 (see Fig. 5.2). This latter contains a BioHarness Module, BioHarness Chest Strap, BioHarness Charging and Configuration Cradle. It is equipped by a Bluetooth technology to ensure communication with smartphone and tablet. Its electrical power supply is a rechargeable Lithium Polymer with nominal voltage equal to 3.7. The breathing rate sensing is assured using a pressure sensor pad detecting expansion of the rib cage following to breathing action. In addition, it contains an ECG sensor and a thermistor which is integrated to measure the device internal temperature. The driver's activity level and the device orientation are captured using an internal 3-axis accelerometer integrated in the BioHarness module.

Fig. 5.2 A photo of the Zephyr BioHarness 3 sensor that should be fixed on the chest.

5.3.2 Ambient Environment Sensors of AffectiveROAD

The platform (c.f. Fig. 5.3) is based on Intel Edison developer kit[7]. This kit offers an embedded IoT platform with strong computing performance, as well as Wi-Fi and Bluetooth connectivity. As Intel Edison is a low-power device, it may run on USB power and an external power adapter from 7 to 15 V. Thanks to its computing performance, its reduced size and its ability to be connected directly to the car lighter port, this platform is adapted to the use in vehicle.
We propose to evaluate the in-vehicle ambient environment using this platform. Data that we aim to collect concern mainly ambient temperature, humidity, pressure luminance and ambient sound level. The set of environmental sensors are wired connected to the shield such that two Grove touch sensors are embedded in order to offer to the user the possibility to easy put on/off the environmental platform. The luminance is captured using Grove luminance sensor. The air quality is measured using Clik air quality. Adafruit BME 280 allows to capture the temperature, the humidity and the pressure. The data were stored directly in Micro SD card. We established a bluetooth connection between this platform and one of the E4. Thus, the driver can switch on the platform from the wristband E4 placed on his hand. With such functionality, the experiment is facilitated and can be repeated easily since the driver can control from his hand the platform and have synchronized data after each drive.

5.4 Real-world driving protocol

A real-world driving experiment is designed. We propose in this section to describe the apparatus in Section 5.4.1, the cohort in Section 5.4.2, and the route description in Section 5.4.3.

[7]https://software.intel.com/en-us/get-started-edison-windows

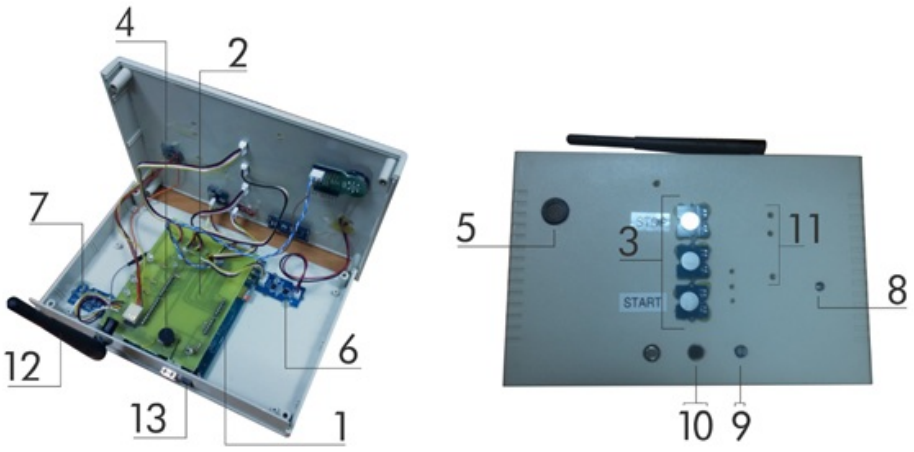


Fig. 5.3 An overview of the environmental *AffectiveROAD* platform: 1- Intel edison Arduino breakout 2- Shield in/out 3- Grove touch sensor 4- beeper 5- Clik air quality sensor 6- Grove Piezo vibration sensor 7- IMU grove MPU 9250 8- Adafruit BME 280 9- Grove luminance sensor 11- Leds 12- WIFI / Bluetooth antenna 13- Power supply.

5.4.1 Apparatus

All drivers are recruited from our institution in order to be covered by insurance. They all are used to drive every day since minimum 5 years. Each participant should possess his own car to make sure that he is used to drive it, and this choice is fixed in order to avoid the workload induced by the non-familiarity of the participant in driving a new car. The experimenter was the same person all over the whole drives. She did not interact with the driver unless she/he makes a mistake in the route or a problem of sensor. Each driver had to consent to participating voluntary to the experiment and having physiological signals, video and audio recorded during the whole experiment.

Table 5.1 Description of the different sensors used in the data acquisition

	Sensors	Signals	Fs
Body sensors	2 Empatica E4	EDA	4 Hz
		HR extracted from BVP	1 Hz
		Skin temperature	4 Hz
		Hand movement	36 Hz
	Zephyr Bioharness 3	Respiration Rate	1 Hz
		Skin temperature	
		HR	
Environmental sensors	Grove luminance sensor	Luminosity	4 Hz
	Adafruit BME 280	Temperature	
		Pressure	
		Humidity	
	Sound meter PCE-322A	Sound amplitude	10 Hz
Video recording	2 Go Pro	Videos	50 fps

Several sensors are used in order to capture the physiological state of the driver, the state of his

environment and the speed of the car. The list of the used sensors is detailed in Table 5.1. Fig. 5.4 shows the different used sensors in the experiment.

Fig. 5.4 An overview of the used sensors to capture the different signals

The Physiological sensors are plugged as shown in Fig. 5.5. The 2 Empatica E4 are placed on the left and the right arm of the driver. The Zephyr Bioharness 3 is placed, after wetting it, on his chest, with direct contact with his skin.

Fig. 5.5 Snapshot from the video capturing the inside of the car.

The environmental platform containing sensors that capture the humidity, temperature, pressure and luminance, the sound meter are placed in the rear seat, just behind the experimenter, to make sure that the data collection is ensured. Two GoPro, capturing the inside and the outside scenes are placed on the windshield of the car.

5.4.2 Cohort Description

As shown in Table **??**, 14 drives were performed by 10 participants: 5 female and 5 Male. Except for one participant (having 59 years old), the age varied between 24 and 34 years (Mean=29.9, sd=3.7). All of them held a valid driving license and are used to drive everyday. The driving

Table 5.2 Characteristic of drivers

Driver Ref.	Numb. of performed drives	Age	Gender	Driving experience	Driving is stressful?
NM	3	31	M	5	No
RY	2	34	F	11	No
MJ	1	59	F	37	Yes
BK	1	33	M	13	Yes
MT	1	33	F	12	No
EK	1	31	F	10	No
KSG	1	24	M	6	No
AD	1	29	M	11	No
GM	2	24	F	6	No
SJ	1	30	M	11	No

experience ranged from 5 to 37 years (Mean=11, sd=8.37). As mentioned in the Table resuming the 14 drives, three drivers repeated the experiments, NM repeated 3 times the experience and found that the stress decreased with drives since he get used to drive in the same path. RY was more stressed in her second drive as she was overwhelmed and did not sleep very well the night before the second drive.

5.4.3 Route Description

The experiment, described in Table 5.3, begins with 15 minutes of rest in our uncovered institution parking. The driver is sitting in her/his car, closing the eyes and the car engine is off. After this first rest period, the participant exited the parking and pass different streets in city. The driving in this area is presumed to induce high stress level due to the harsh atmosphere such as narrow routes, two red lights, a lot of vehicles, pedestrians, motorcycles and bikes.
Following the exit of the avenue, the driver goes out to big route imitating a highway driving. She/he spends about 8 minutes in a smooth route which is designed to provide a continuous driving. Then, the driver arrives to a roundabout before beginning a city driving spending about 10 minutes. This route is characterized by a lot of parked vehicles in front of restaurants or shops and there are no traffic lights. The driver arrives to a roundabout where she/he takes the same route to return back to the starting point which is the parking.

Table 5.3 Roadmap: Description of the proposed periods to be encountered during the experiments and the corresponding assumed evoked stress level.

Events	Reference	Duration (Minutes)	Assumed evoked stress level
Parking	Rest1	15	Low
City Driving	City1	10	High
Highway driving	Highw1	8	Medium
City Driving	City2	20	High
Highway driving	Highw2	8	Medium
City Driving	City3	10	High
Parking	Rest2	15	Low

The driving protocol consists in a set of paths along 31 km. In daily normal traffic, the whole experiment takes about 1h26 with 30 minutes of rest.

Table 5.4 Summary of Driving Events. The mark "X" is used to designate that the data are complete. "Hwy" means highway driving and "Rest" designates rest period. The date of each drive is expressed in the second column of the table in format of DD/MM. All the drives were conducted at the same year, 2017.
The special events listed in the column "Comments" are described in Appendix A.

Driver	When	Data										Comments
		Platform	Gopro In	Gopro Out	Micro	Soundmeter	GPS	E4	Bioharness	Arousal metric	Survey	
NM	03/05		X	X	X		X	X	X	X	X	Sunbeam in Rest
	26/05	X	X	X	X	X	X	X	X	X	X	
	19/07	X	X	X	X	X	X	X	X	X	X	Loud noise in Rest
RY	08/05		X	X	X		X	X	X	X	X	Talk in Rest
	23/05	X	X	X	X	X	X	X	X	X	X	Trucks in Hwy
BK	17/05	X	X	X	X	X	X	X	X	X	X	Road Pb in Hwy
MT	18/05	X		X	X		X	X	X	X	X	Phone use during rest
EK	19/05	X	X	X	X	X	X	X	X	X	X	
KSG	25/05	X	X	X	X	X	X	X	X	X	X	Distressed in Rest
AD	31/05	X	X	X	X	X	X	X	X	X	X	
GM	05/07	X	X	X	X	X	X	X	X	X	X	
	15/08	X	X	X	X	X	X	X	X	X	X	Road Pb in Hwy
SJ	18/07	X	X	X	X	X	X	X	X	X	X	Congested Hwy and dry throat in Rest

5.5 Data Collection and Construction

The experiments presented in Section 4.4 lead us to a data set containing physiological data collected from the different sensors and environmental data. In addition, we added a smartphone to offer GPS data. The database includes also sound level and the contextual data such as the video ranking and marks of special events that have been occurred during the drive.
After conducting the experiments, we figured out that some data are missing. The status of the collected data is summarized in Table 5.4.

The data set are anonymized and available upon request to the corresponding author. A part of the dataset will be available soon on the web site of our institution and on PhysioNet bank [8]. We will exclude the video from the data set for privacy reasons.
It should be mentioned that the drive duration varies due to different traffic conditions and driving behaviors, this results in different samples number from a drive to other. Driver's physiological information were recorded using the two physiological sensors, described in Section 3.1. Two Empatica E4 devices were plugged on the left and the right hand of the driver. The Zephyr BioHarness 3 was placed in the driver's chest. In order to capture information related to the driver environment, *AffectiveROAD* environmental platform was placed in the rear place of the vehicle in order to not influence the driver's comfort.
Two Gopro were plugged to capture the internal and the external scenes. The sampling rate is 50fps. Since the file size is very high, a conversion was made to reduce the sampling frequency to 25 fps. During each drive, the experimenter was in the rear seat building a subjective stress metric by evaluating the complexity of the driving scene, the distress and the workload that may be perceived by the driver. This were achieved using a slider, implemented in a laptop. The resulted rating is a metric ranging from 0 (no stressful) to 1 (extremely stressful). Since this metric is subjective and depends on the perception of the experimenter, a protocol of validation was proposed. For each drive, the driver had to validate the preprocessed subjective annotation provided by the experimenter. A video superimposing in parallel the inside and the outside video

[8] https://www.physionet.org/physiobank/

recordings with the plot of the subjective score. A cursor that is synchronized with the processing of the two videos indicates the advancement of the subjective score (see Fig. 5.6). Each driver should indicate if the resulted rating corresponds exactly to the perceived stress and correct it if it does not correspond to his perception.

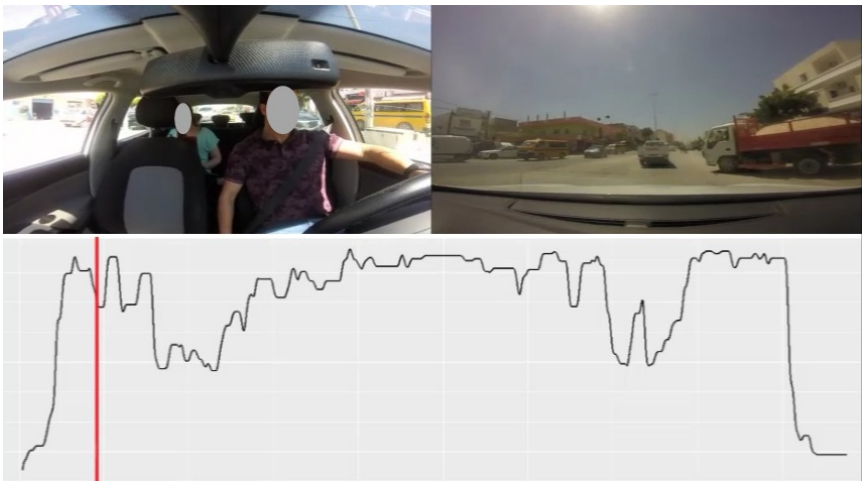


Fig. 5.6 Snapshot from the screen shown to each participant in order to validate the subjective stress metric.

5.6 Related works and discussion

Driver's state recognition is a key point in the design of intelligent systems that aim to ensure security and to improve driver's comfort. Several studies exist that offer system allowing to sense the human changes and the environment. Most of the studies include subjective measurements usually used to validate the objective information. However, when considering the objective measurements issued from sensors, not all studies are based on the same type of data. We consider here the three main studies described in Section 5.2, that offer open database. Table 5.5 summarizes the list of sensors used in the different studies. Despite the fact that MIT database is based

Table 5.5 Comparison between studies on driver's state recognition offering open database. The comparison is made in terms of used sensors and the provided data.

System	Physiological signals					Environmental data			GPS
	HR	EDA	Breathing Rate	Skin temperature	Hand motion	Luminance	Temperature, humidity and pressure	Sound level	
MIT [81]	extracted from ECG	On foot and hand	Yes	No	No	No	No	No	No
hciLab [140]	extracted from ECG	On left hand	No	Yes	Yes	Yes	No	No	Yes
Warwick-JLR [153]	extracted from ECG	On the non dominant hand	No	No	No	No	No	No	Yes
AffectiveROAD [52]	extracted from ECG and BVP	On left and right hand	Yes	Yes	Yes	Yes	Yes	Yes	Yes

on only physiological data and subjective measurement, numerous studies are achieved using the

available dataset.
The released datasets from the compared systems are not well annotated (except for MIT database) which makes difficult the use of the data. We propose to provide, with the dataset that will be soon available online, annotation files that allow to define the different route types and to mention if the driver encounter some event that may affect her/his state.

5.7 Conclusion and future work

In this chapter, we proposed *AffectiveROAD*: a sensor network system that captures physiological changes and environmental information. The proposed system includes two types of platforms, namely physiological platform and a prototype of an environmental platform. The physiological platform is composed of two unobtrusive Empatica E4 wristbands and a chest belt Zephyr Bioharness 3.0. Electrodermal activity, skin temperature, heart rate, respiration and hand motion are the main physiological signals collected using *AffectiveROAD* physiological platform. The prototype of the environmental platform that we developed offers the possibility to collect information concerning temperature-humidity-pressure, luminance, and air quality. *AffectiveROAD* system allows to conduct drives and to gather data with several drivers, in different types of route and during different weather conditions.

The use of such platform may be useful in driver's state assessment in particular for the situations that may need more awareness and attention. Especially, this may help in the development of Human-centric IoT-based interfaces, improve the existing Human-computer interaction (HCI) systems and monitor Heating, Ventilation and Air-Conditioning (HVAC) system.
By this work, we propose first to provide vehicle manufacturer with a platform that may be used to conduct many real-world experiments, then to provide scientists with an open database that can be easily explored and used. Collecting driving data in real-world context is really a challenge since it is costly.
Several studies on driver's state recognition can be conducted using *AffectiveROAD* database. Since the route of real-world driving protocol is designed by alternating different types of roads, assumed to induce different stress level, we will propose in future work to study the effect of road type on the driver's stress level. Such analysis can be based on the video-based continuous stress metric that we proposed. In addition, statistical methods will be used to extract from data relevant indicators to assess driver's performances and enhance his/her awareness. Moreover, features selection will be used in order to select sensors to be integrated in the vehicle. This will allow to efficiently predict the driver's state and reduce cost of equipments inside the vehicle.
Automated systems based on sensors and advanced decision algorithms allow to consider different driving functionalities. Nowadays, these technologies are precise so that can be used not only to assess users while driving but also to intervene in vehicle take over strategies.
As an extension to our system, we propose to use the connectivity protocols and benefit from the Intel Edison as a data acquisition device in order to centralize the data collection from physiological sensors (via Bluetooth connection), environmental sensors (integrated nowadays in smart vehicles or use the proposed environmental sensors) and vehicle CAN information. In addition, a GPS module can be plugged into the device in order to pass away from smartphone.

Chapter

6

Biosignals Selection and Self-similar Characterization of the Electrodermal Activity for the Assessment of the Driver's State of Stress

Abstract This chapter presents two analysis of biosignals toward driver's stress recognition. Data issued from the public database *AffectiveROAD* are considered. Specifically, we take the electrodermal activity measured on the right (RightEDA) and the left (LeftEDA) wrist of the driver, the breathing rate (BR), the heart rate (HR), the posture (Post) and the subjective stress metric. The first analysis aims to select predictors of the stress of the driver. An approach of variable selection based on wavelet decomposition and Random-Forests recursive feature elimination considering the grouped variable importance was applied. In the classification problem, the final model selects three wavelet levels of the Right EDA and Post and leads to a final error of 31%. Based on the stress metric, the regression model proposes to consider also 3 wavelet levels of Right EDA and Post with an error of the final model of 20%. The second analysis presents a characterization of the Electrodermal activity through self similar processes mainly the Fractional Brownian Motion and its corresponding Hurst exponent. The results suggest that the EDA exhibits the behavior of a self similar persistent process with trend. Moreover, the Hurst exponent can be considered as a relevant feature for the monitoring of the driver's affective state which includes the affect arousal, the distress, the mental workload, and the perception of the driver to his/her environment complexity.

6.1 Introduction . . . 86
6.1.1 Problem Statement . . . 86
6.1.2 Related Work . . . 87
6.1.3 Chapter contributions and outline . . . 87
6.2 *AffectiveROAD* database description . . . 88
6.2.1 Real-world driving protocol . . . 88
6.2.2 Data collection . . . 89
6.2.3 Data preprocessing . . . 91
6.2.4 Subjective stress metric construction . . . 92
6.2.5 Questionnaire Analysis . . . 93
6.3 Methods . . . 95

6.3.1 The 3 step-based approach for variable selection 95
6.3.2 Self-similarity analysis 96
6.4 Driver's arousal state recognition based on variable selection using Random Forests 97
6.4.1 Data description 97
6.4.2 Physiological variable selection and arousal level classification 98
6.4.3 Bio-signals selection and arousal level classification 103
6.4.4 A regression model of the subjective stress metric on the biosignals 105
6.5 EDA self-similarity analysis 107
6.5.1 EDA and scale invariant processes 108
6.5.2 Driving experiment-based analysis 108
6.5.3 Individual driver-based analysis: right vs left EDA 109
6.6 Discussion 110
6.7 Conclusion 112

This chapter is written in collaboration with R. Ghozi, J.-M. Poggi, S. Sevestre-Ghalila and M. Jaïdane. It is a draft of a paper to be submitted soon in the journal of IEEE Transactions in Affective Computing.

6.1 Introduction

The expansion of the amounts of information systems and the advanced technologies integrated inside the vehicle makes the assessment of the driver's mental workload and affective state an area of great importance. Indeed, the driver's state affects the driving performance thus his/her comfort, safety, and even other road users safety. According to Coughlin, the state of underload and fatigue due to a lack of sleep as well as situations of cognitive overload intervene significantly in the driver's ability to deal safely with the primary driving task [36]. An other serious issue encountered in human performance is under-arousal and over-arousal which is considered in several models of human performance (see an overview in [158]).
In human performance models, relationships between workload, task demands, performance are to be considered. In this study, we do consider different biosignals in order to characterize the driver's state of "stress" during a real-world driving performance. *It should be mentioned that the term "stress" is here used to designate the affect arousal of the driver, his/her mental workload, attention, distress, and his/her perception to the environment complexity as a whole.*
The affect arousal can be defined as the human state of being activated, physiologically or psychologically, as a response to emotional stimuli. It is associated with a physiological state of the body characterized by a relative intensified responsiveness and preparation for actions through the activation of many systems in neurons and body [113]. The arousal can be seen as a one dimension of the workload which is a multidimensional concept hard to quantify and define [67].

6.1.1 Problem Statement

Task performance and arousal have an empirical inverted U relationship, known as Yerkes-Dodson law [166]. In fact, performance increases with arousal, arriving to an optimal point and then declines when levels of arousal become too high. This optimal point depends on individual factors and capacities, and situation complexity. When driving, suboptimal arousal levels may result in a higher risk for accidents. Therefore, the affect arousal state should be assessed. Protocols to

elicit the driver's state have been developed including several biosignals acquisition in two main configurations: real-world driving experience [81, 98, 133] and simulator-based studies [28, 129, 169]. People when experimenting simulated tests, naturally tend to simulation or dissimulation [45]. Thus, it is preferred to conduct studies in a real-world context. In addition, the fact that the experiment is performed in a real traffic situations allows to consider the study findings and results more applicable to use in these real-world situations.
In the driving context, it is crucial to select the biosignals that allow a best recognition of the driver's state. In fact, this allows to first select the list of sensors to be considered when studying his/her state. Such list allows to assess in an off-line way and even on-line the evolution of the driver's state.
The definition, collection and construction of the variable of interest reflecting the driver's state is a critical task in the studies focusing on the driver's state recognition. specifically, the driver's stress level is usually defined based on assumptions, related to the road types, and supported then using subjective reports, expert annotation and video recording. For instance, the study of Healey and Picard [81] proposed a distress metric based on the road type assuming that city driving is more stressful than highway driving, while the rest period is supposed to evoke low stress level. This assumption was validated through expert annotation and questionnaire analysis. The protocols used to built this variable are complicated, time-consuming and costly.

6.1.2 Related Work

The relatively recent development of non-invasive wearable sensors makes the development of efficient solutions possible. New sensor technologies offer nowadays a non-invasive devices allowing to capture human changes during a daily-task performance. Several studies focus on the driver's state recognition such as the fatigue, drowsiness, inattention, stress, etc.

Table 6.1 reports 5 main studies based on biosignals, specifically the Electrodermal activity (EDA) which is considered as one of the most correlated physiological signals with human affective state [23]. Almost all the studies propose a public database that can be used by the scientific community. The particularity of the database AffectiveROAD is the subjective metric of stress based on a real-time evaluation of the environment complexity, and the experimenter perception of the evoked affect arousal and distress [52]. Moreover, the sensors used are wireless which allows a fast adaptation of the driver to the experiment.

6.1.3 Chapter contributions and outline

The used methods and approaches in this study are based on two main works: the first [51] offers an approach of functional variable selection while the second [50] provides a method to characterize the EDA through a self-similar process.
This chapter describes first the open *AffectiveROAD* database which offers a subjective stress metric based on a real-time annotation validated post experience by the driver that performed a video rating. The contributions of this work are two-folded. The first one aims to select predictors of the driver's stress level. One approach consists on a random-forest based classification model. The other benefits from the stress metric and proposes a regression model. The second analysis intended to characterize the EDA through self-similar processes in order to monitor the stress level.
The rest of the chapter is structured as follows: Section 5.2 describes the *AffectiveROAD* database and the data pre-precessing. A brief summary of the methods is provided in Section 5.3, namely the 3 step random-forest based approach for variables selection and the wavelet-based self-similar analysis. Section 5.4 lists the results of the driver's stress level recognition based on the classification model first then using the proposed stress metric regression. Section 5.5 allows to present the main result of the self-similarity characterization of the electrodermal activity. The discussion

Table 6.1 Studies achieved on driver's state recognition based on biosignals

Study	Variable of interest	Public database	Biosignals	Type of Physio. sensors	Other data
Healey and Picard (2005) [81]	Distress (road type with expert annotation and quest. analysis	Yes (drivedb[2])	EDA, ECG, EMG and Respiration	Wired	No
Rigas et al. (2011) [133]	Fatigue and stress (self-annotated)	No	EDA, ECG and Respiration	Wired	Weather, traffic and visibility
Taylor et al. (2013) [153]	Worload (test performance)	Yes (Warwick-JLR[3])	EDA and ECG	Wired	CAN bus data
Schneegass et al. (2013) [140]	Workload (video rating)	Yes (hciLab[1])	EDA, ECG and Body temperature	Wired	Location, brightness and acceleration
El Haouij et al. (2018) [52]	Stress (including distress, attention, affect arousal and workload)	Yes (AffectiveROAD[1])	EDA, HR, BR, and Posture	Wireless	In-vehicle temperature, humidity, pressure, and sound level

of the chapter is provided in Section 5.6. Finally, the conclusion along with the limitations of the current work and the future research work are presented in Section 5.7.

6.2 *AffectiveROAD* database description

6.2.1 Real-world driving protocol

The driving protocol consists in a set of paths along 31 km. In daily normal traffic, the whole experiment takes about 85 minutes, with 30 minutes of rest. The experience is inspired of the *drivedb* protocol described in [81]. In fact, it is designed to assess naturally occurring emotions while driving in real-world situations. In such conditions, the stressors are not controlled, but this protocol is designed in order to take the driver in situations that might typically be encountered in a daily trip. Two rest periods, at the beginning and the end of the drive, are set in order to provoke the least opportunity for stress. The driver alternates between city and highway driving assumed to provoke respectively high and medium stress levels. Incidental encounters such as tolls, road works, and pedestrian crossing may provide the highest stress level. The experiment consisted on following these steps: First of all, the driver should consent to participating voluntary to the experiment and having physiological signals, video and audio recorded during the whole experiment. She/he responds to a list of pre-experience questions. After that, the experimenter showed him how to wear the different physiological sensors. Then, they went together to the parking in order to prepare the driver's own car. The environmental sensors and two video

recording devices are plugged. This period allowed to the driver to get used to the protocol and to forget that he/she is putting physiological sensors. The experiment began with a period of 15 minutes allowing the driver to relax. Then, he/she began the driving period and finished with 15 minutes of rest. After returning to our institution, the driver answered a post experience questionnaire. The video annotation that allows the driver to correct the stress metric constructed by the experimenter is achieved an other day.
All drivers are recruited from the institutions CEA-LinkLab and Telnet in order to be covered by insurance. They all are used to drive every day since minimum 5 years. The experiments are achieved using each participant own car. This choice is motivated by the fact that she/he is used to drive it and this allows to avoid the workload induced by the non-familiarity of the participant when driving a new car. He/she should hold a valid driver license. The experimenter was the same person all over all drives. She did not interact with the driver unless the participant makes a mistake in the route or there is a problem of sensor. Several sensors are used in order to capture the physiological state of the driver, the state of his environment and the GPS data. Two Empatica E4 placed on the right and left wrists of the driver and a Bioharness Zephyr 3.0 are used. More details about the list of the used sensors and their placements can be found in [52] (Chapter 4 of the thesis report).
The road map of the path is depicted in Fig. 6.1: The experiment begins with 15 minutes of rest in the parking (see picture 1 in Fig 6.2). The driver is sitting in her/his car, closing her eyes and the car engine is off. After this first rest period, the participant exited the parking and drive along a zone that is a part of the Technopole. This zone is designated by "Z" in Fig.6.1 and corresponds to the picture 2 of Fig 6.2. After passing through a gateway, the participant

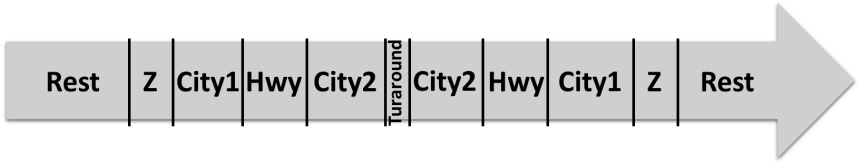


Fig. 6.1 Road map of the path drove by the different participants.

drive along different streets presumed to be stressful because of two red lights always crowded of vehicles since the road is narrow and the stream of passing vehicles is huge. This configuration is set in order to imitate city driving. The driving in this area provokes high stress level due to the harsh atmosphere such as narrow routes, a lot of vehicles, pedestrians, motorcycles and bikes. Following the exit of this avenue, the driver goes out to a route imitating a highway driving (see picture 4 in Fig 6.2). She/he spends about 8 minutes in a smooth route. Then, the driver arrives to a roundabout before beginning a city driving for about 10 minutes. This route is characterized by a lot of parked vehicles in front of restaurants or shops (see picture 5 in Fig 6.2). There are no traffic lights. The driver arrives to a big roundabout where she/he drives again along the same route to return back to the starting point which is the parking of our institution.

6.2.2 Data collection

Several physiological data were gathered from two main types of wireless device: Empatica E4 that was plugged on the left and the right wrist of the driver and Zephyr Bioharness 3.0 which is a chest belt. Specifically, the Electrodermal activity (EDA) measured on the right and the left wrist, the Heart Rate (HR), and the Breathing Rate (BR) are collected. Data are gathered post experience using the two different sensors softwares. The choice of such physiological signals is motivated by the experiments of [81, 83] conducted in order to assess and monitor drivers on rural roads. The chest belt provides in addition a biosignal that expresses the degrees of deviation from the vertical. An illustration of the raw 5 biosignals for a drive performed by a driver GM is depicted in Fig. 6.3. Each signal was sampled at a rate appropriate for capturing

Fig. 6.2 The different zones considered in the road map respectively Rest, Z, City1, Hwy and City2.

the information contained in the signal. The left and right EDA were sampled at 4 Hz, the HR extracted from the ECG, the BR and the Posture were sampled at 1 Hz. The experimenter that accompanies each participant during each drive, was charged to evaluate the ambiance in the vehicle and to analyze its complexity, the state of the driver and the road state. She had to built a real-time continuous metric which corresponds to the driver's state including his/her mental workload, attention, distress, affect arousal and his/her perception to environment complexity. For each drive, the driver had to be validate the constructed rate using the two video recordings corresponding to the driving experiment.
In total, 14 drives were completed, seven by drivers who completed the driving path only once and seven drives were accomplished by three drivers who repeated the experience on multiple days. One drive is not considered since the two rest periods were not achieved. The missing data related to the 13 drives were summarized in Table 5.4 in Chapter 4. The environmental data captured using *AffectiveROAD* platform were missing for two drives. One recording of the car inside scene using a Gopro is missing. The sound level is not captured for 3 drives.

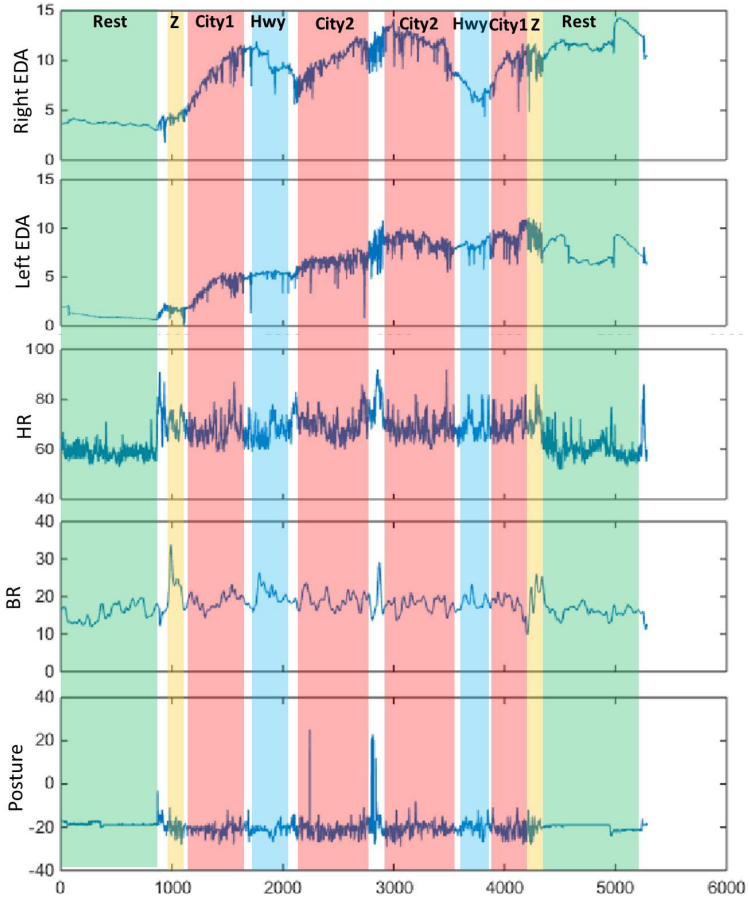


Fig. 6.3 Raw bio-signals for a drive. EDA signals captured on both right and left wrists, Heart Rate, Breathing Rate and Posture corresponding to the second drive of the driver GM.

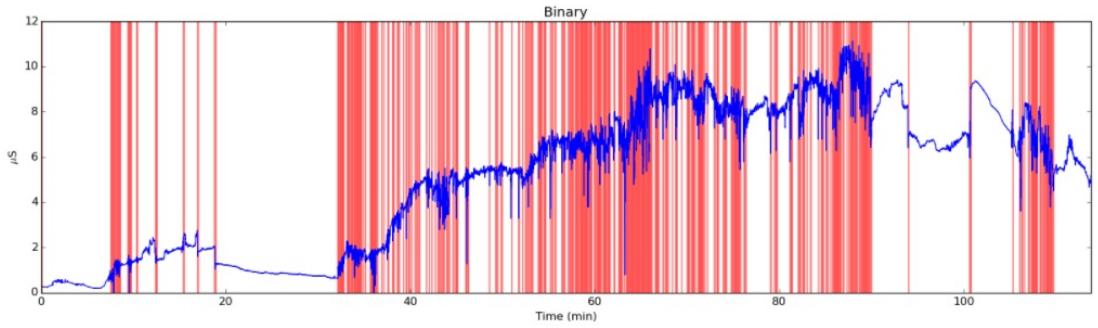
6.2.3 Data preprocessing

The data issued from Zephyr Bioharness 3.0 belt, namely the breathing rate (BR), the heart rate (HR), and the posture (Post) were filtered. However, EDA signals captured on the left and the right wrist of the driver are raw data.

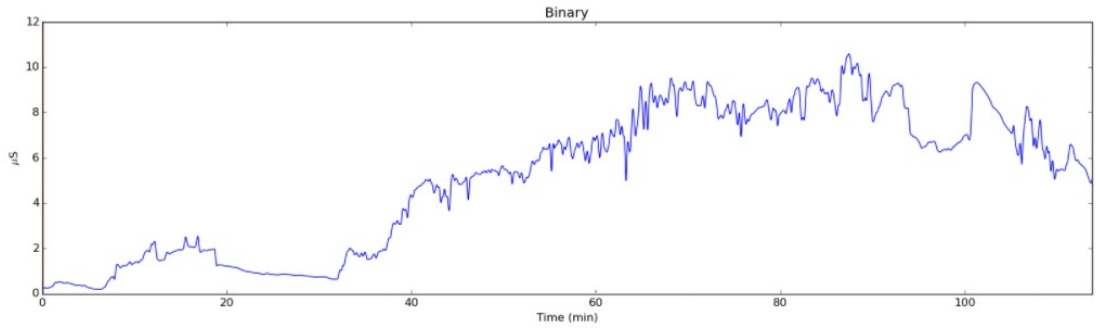
Signal preprocessing

A preprocessing step was encountered for the EDA signals. First of all, we detect artifacts using the web version of the software proposed by [154] on the website[2]. This explorer offers an adequate tool able to detect artifact (see Fig. 6.4 (a)). The advantage of the proposed software is the fact it accepts files issued from the sensor Empatica E4 and automatically label the epochs as artifacts or not. The signal is subdivided into 5 seconds non overlapping segments. We removed the 5 seconds segments corresponding to artifacts then we used a linear interpolation in order to replace the missing data.
The second step consisted on filtering the EDA signal using the Singular Spectrum Analysis (SSA) that was found to be the most appropriate for the EDA signal filtering (see [9] for details). Finally, a verification step was performed: Each signal was passed another time into the website to make sure that the signal does not contain anymore artifact (see Fig. 6.4 (b)).
Since the sampling frequency was different (4 Hz for the EDA and 1 Hz for the other signals), we performed an upsampling to 4Hz. In order to avoid the aliasing, a lowpass filter with a Kaiser window is applied during the resampling process.

[2]http://eda-explorer.media.mit.edu/dashboard/



(a)



(b)

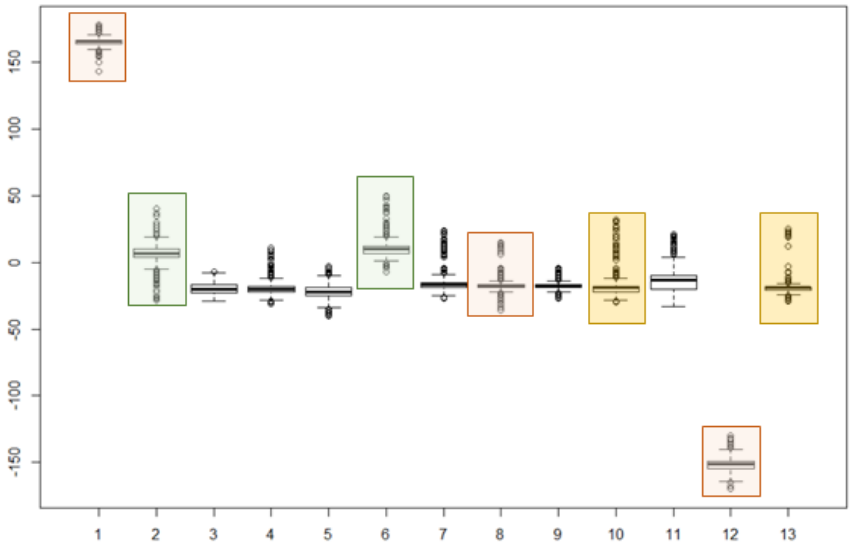
Fig. 6.4 Screenshot of the artifact detection from EDA Explorer applied to an EDA signal (a) before and (b) after preprocessing. This EDA measurement was captured on the left wrist of the driver GM during her second drive. Red vertical bars correspond to the placement of the different detected artifacts.

Statistical transformation

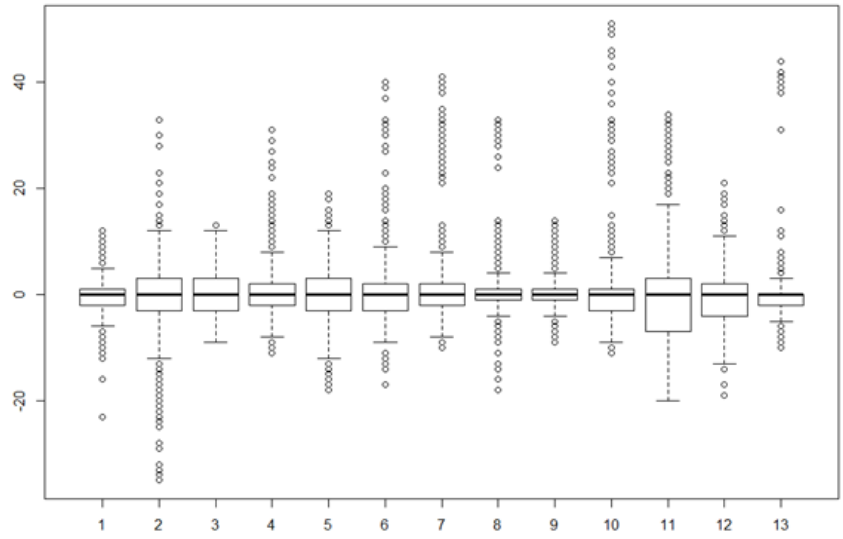
The posture variable depends on the body morphology of the participant. For instance, the boxplots 1, 8 and 12 of Fig. 6.5 (a) are relative to the same person. In order to eliminate this individual information, we propose to transform this variable by setting up the median value to 0 which corresponds to the vertical reference. This was achieved by subtracting the mean value of each drive from each sample (see Fig. 6.5 (b)). All the physiological signals were normalized using the min-max range correction.

6.2.4 Subjective stress metric construction

During each drive, the same experimenter was in the rear seat of the vehicle, observing the experiment, checking if all the sensors network work well and performing a real-time annotation of the overall perceived stress. Such metric is built based on a subjective evaluation of the complexity of the driving scene taking into account the outside vehicle unexpected events that can alter and induce a distress. It evaluates also the mental workload and the affect arousal that may be perceived by the driver. This evaluation was achieved using a laptop interface in format of a slider. The experimenter annotated the whole drive by attributing a score ranging from 0 (designating the no stress), to 1 which is used when extremely stressful situations were encountered.
A protocol was designed in order to validate and correct such subjective metric in a non-costly and fast way. In fact, we proposed to consider and to take into account the subjective assessment of the driver of the whole drive. The driver had to correct and validate this rating, which was smoothed using a Hanning filter with a window of 100 seconds. An interface which combines the



(a)



(b)

Fig. 6.5 Boxplots of the variable "Posture" per drive (a) before and (b) after the statistical processing.

video superimposing in parallel the inside and the outside video recordings, with the display of the subjective stress metric. A vertical cursor that is synchronized with the processing of the two videos, sweeps the subjective stress metric to indicate the corresponding score to each visual and acoustic scene (see Fig. 6.6). The validation protocol consists on presenting the annotation interface to the driver and to ask him/her to correct or validate the attributed metric. The video recordings were 4 times forward speed up, in order to spare the effort and the time to the driver and to avoid the fatigue that can deteriorate the quality of the provided response. The interface allows to modify the value of the stress metric which can be, after finishing the validation procedure, downloaded.

6.2.5 Questionnaire Analysis

We proposed a subjective study based on questions inspired from the questionnaire proposed by the study of Healey and Picard [81]. A French version is provided in Appendix B. The

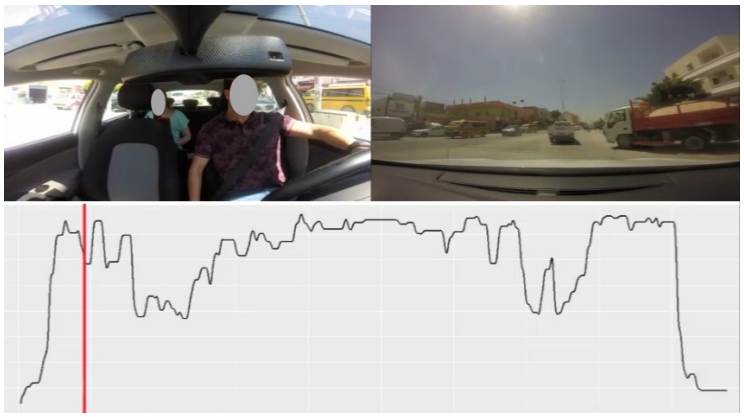


Fig. 6.6 Snapshot from the screen shown to each participant in order to validate the subjective score.

questionnaire contains three parts. The first one concerns the driver's information such as the name, age, and driving experience. The pre-experience section of the questionnaire concerns mainly the rating of the driving zones that will be encountered when conducting the experiments. The last section of the questionnaire contains the same questions of the pre-experience part as well as questions proposed in order to report if there is a difference between the stress perceived in the way going and back the trip. For each drive, the participant is asked to rate the driving zones on a scale of 1 to 5 where 1 is used to assess a feeling of no stress and 5 was used to represent a feeling of high stress.

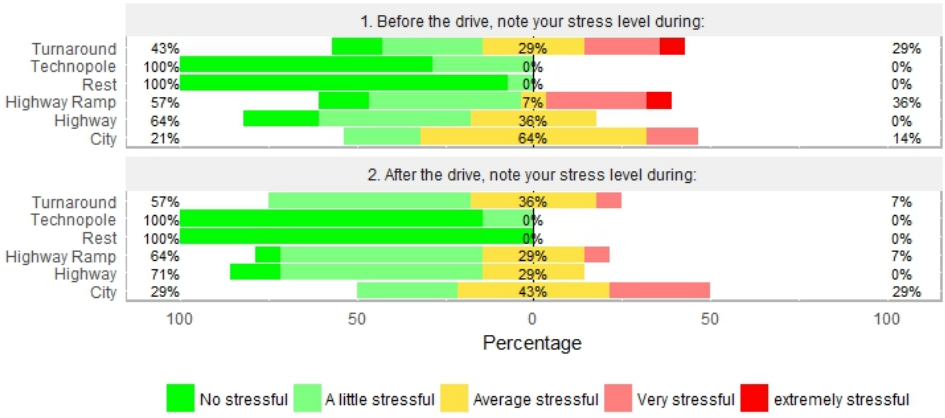


Fig. 6.7 Likert plot corresponding to the responses to the two questions on the perceived stress level. The driver is asked to rate the stress level before and after performing the drive.

As can be noted from the Likert plot presented in Fig. 6.7 and from Table 6.2, there is a coherence in the responses of the drivers performed before and after the drive when rating the three zones namely: city, highway and rest.
We focus on these three zones and we attribute a score to each level of the Likert scale. Thus, the scale is ranging from 1: no stress to 5: extremely stressful. Table 6.2 shows that drivers found the rest periods to be the least stressful, the highway driving to be more stressful than rest zone and the city driving to be the most stressful which confirms the finding of the subjective study of drivedb performed in [81].

Table 6.2 Rating questionnaire results before and after the experiment.

Periods	Pre-experience	Post experience
City driving	2.93 (sd=0.64)	3 (sd=0.76)
Highway driving	2.14 (sd=0.80)	2.14 (sd=0.69)
Rest period	1.07 (sd=0.28)	1 (sd=0)

6.3 Methods

The used methods were previously described in Chapter 2 and 3, thus we will briefly summarize the two main approaches namely: the 3 step-approach for variable selection and the fractal analysis.

6.3.1 The 3 step-based approach for variable selection

The approach proposed in [51] considers the wavelet decomposition of the functional variables, allows first to remove the irrelevant variables from the model, and than to select the most relevant wavelet features among the wavelet levels of the retained functional allowing an optimal stress level recognition. In this section, we present first the random forest framework. Then, we summarize the Recursive Feature Elimination based on the endurance score. Finally, we describe the variable selection based on the endurance score.

Random Forest and Variable Importance

Random Forest is a non-parametric method, introduced by [25], consisting on a set of decision trees built over a number of bootstrap samples of the training set. It provides estimators of the Bayes classifier that minimizes the classification error in the classification case and estimator of the regression function, in the case of regression problem. At each node, a fixed number of variables is randomly picked to determine the splitting rule among them. The trees are not pruned so all the trees of the forest are maximal trees. The resulting learning rule is the aggregation of all the estimators resulting from those trees.
The RF offers a measure which is the Variable Importance (VI) allowing to assess the importance of a variable in explaining the variable of interest. A measure of VI for a variable X_j is the difference in prediction accuracy before and after permuting X_j. In fact, by randomly permuting the variable X_j, its link with the response variable Y is broken. An extension of the VI allowing to estimate the VI for a group of variables is provided in [70]. It is based on the same definition of the VI, but considering for each tree the permuted set resulting by randomly permuting the group of variables. More details can be found in Chapter 3 Section 3.3.1.

Recursive Feature Elimination based on Random Forest (RF-RFE)

Random Forests-based Recursive Feature Elimination (denoted by RF-RFE) was proposed by [70]. The description of this procedure is based on the summary provided in [51]. The RF-RFE approach starts by splitting the dataset into a training and validation set (containing respectively $\frac{2}{3}$ and $\frac{1}{3}$ of data). Then, a random forest model is fitted using the training set and considering the whole set of variables. The VI or the grouped VI is computed and the error is evaluated using the validation set. The variable having the lowest VI is eliminated and the procedure is repeated until no further variables remain. The selected variables correspond thus to the variables involved in the model that minimizes the prediction error.

Variable selection based on the endurance score

The results obtained when applying the RF-RFE vary from an execution to an other in terms of the list and the total number of the selected variables. In order to stabilize the procedure, the endurance score is considered. Such score is a measure of the variable endurance to the selection procedure. The RF-RFE is repeated 10 times for the p variables $(V_1, ..., V_p)$. We determine the list of the selected variables all over the 10 runs for each rank r, $r = 1, ..., p$, denoted here by L_r and the total number of occurrences of a variable V_i, $i = 1, ..., p$ in the list L_r, (henceforth $noc(V_i, L_r)$). The endurance score of a variable i, $i = 1, ..., p$ (denoted by $score$) is computed using

$$score(V_i) = \sum_{j=1}^{p} noc(V_i, L_r) \times (p - r + 1). \tag{6.1}$$

To sum up, the 3-step based approach consists in considering the full wavelet decomposition of the signals, eliminating the less relevant variables and then selecting the most enduring wavelet levels. Precisely, the biggest gap detection of the ordered endurance score allows to first eliminate the less enduring functional variables. The wavelet coefficients of the kept variables are then grouped by wavelet levels. Thus, an endurance score is also computed for each level. The selection of the wavelet levels is then achieved based on the same technique of the inspection of the endurance score. The final model is mainly composed by the retained wavelet levels. The assessment of the final model is done by a cross-validation like procedure taking into account the particularity of the data.

6.3.2 Self-similarity analysis

Fractional Brownian Motion

Self similar processes (ss-processes) are stochastic processing exhibiting some invariant distributional properties over scales. Indeed, the statistical properties of the ss-process in any scale is completely determined by those of one scale. The Fractional Brownian Motion (FBM) is a Gaussian ss-process with stationary process with $0 < H < 1$, where H is the Hurst Exponent that governs the process properties such as the long/short range dependance, regularity and the fractal properties.
For a processes FBM denoted by S with a finite variance (for $t \in \mathbb{R}$, $E[S^2(t)] < \infty$), we have

$$E[S^2(t)] = |t|^{2H}\sigma^2, \tag{6.2}$$

where $\sigma^2 = E[S^2(1)]$. One can refer back to [152] for details about the FBM.

The wavelet-based estimator of the Hurst exponent

Several techniques to estimate the Hurst exponent are available such as the rescaled range analysis, the log-periodogram, the wavelet-based estimation, etc. (See [44] for an overview of the different estimation methods).
Our study uses the wavelet-based method since wavelets can be considered as a natural tool in the analysis of the scale invariant processes. The wavelet-based method to estimate the Hurst exponent H has been introduced first by [57] for self-similar processes and developed then by [5, 17, 108] for long range dependent (LRD) processes. This study relies on [159] work and their associated routines available online on [160].
In fact, this estimation method is based on the wavelet coefficient $(d_{j,k})_{k\in\mathbb{S}}$ of S for a given wavelet ψ. For a ss-process, the wavelet coefficients $(d_{j,k})_{k\in\mathbb{S}}$ is a zero-mean process for a fixed wavelet level $j \in \mathbb{Z}$ and

$$\mathbb{E}(d_{j,k}^2) = Var(d_{j,k}) \sim C(\psi, H)2^{j(2H+1)} \text{ for all } 2^j > 0, \tag{6.3}$$

where $C(\psi, H)$ is a positive constant which depends on ψ and H.
Thus, the variance of the wavelet coefficients is a power law of 2^j of the form $2^{j\alpha}$ where $\alpha = 2H+1$. Therefore, the slope coefficient of a log-log regression of this variance on j provides an estimator of α and thus of H. The graph depicting the log of the variance of the wavelet coefficients against log (2^j) is called the logscale diagram.

6.4 Driver's arousal state recognition based on variable selection using Random Forests

In this section, the used data are described, then two types of models are proposed. The first one is resulting after functional variables selection considering a Random Forest (RF) based classification model while the second is obtained when considering the stress metric and a regression RF based model.

6.4.1 Data description

The considered variable of interest is the stress of the driver. We recall that the term stress is used here to describe the state of the driver considering his/her affect arousal, distress and even his perception to the environment complexity. We annotate the data based on the same assumption introduced by Healey and Picard study [81] in the case of the distress. This annotation was used in several studies based on *drivedb* database. For city driving, the stress is assumed to be high. Highway driving is assumed to induce medium stress level and rest period evoke low stress since the participant is relaxing.
The explanatory data to be considered here are segments of 4.27 minutes duration extracted from the physiological signals corresponding to different driving zones. The segment extraction is inspired from Healey work [81]. The segments corresponding to city and highway driving zones are extracted from the middle driving zone in order to avoid the memory or the anticipation effects. The segments from the rest periods are extracted from the last minutes in order to obtain the optimal state of rest and relaxation. The number of extracted segments varies from a drive to another since the experiments are conducted in different days, thus the traffic density changes which results in different duration of drives. Table 6.3 lists the total number of segments per driver and the repartition of those segments by classes of stress. For instance, NM performed 3 drives, he contributes with 22 segments which 9 of them correspond to high stress level, 7 to medium stress level and 6 assuming to evoke low stress level. In order to consider the unexpected events that occurred during the rest periods and even when driving, we cross the subjective metric with the classes considered based on the assumption of [81]. Specifically, the mean of the subjective stress metric that corresponds to the extracted segments of 4.27 minutes of duration is considered.

Table 6.4 Contingency table between the mean subjective metric and the classes provided by the assumptions done in [81]

		Based on the subjective metric		
		High	Medium	Low
Based on the assumption in [81]	High	34	7	0
	Medium	0	23	0
	Low	0	4	22

Table 6.4 presents the contingency table between the subjective stress metric and the assumption based on the road type. It confirms the fact that for almost all the cases, the road type

Table 6.3 Repartition of the segments of 4.27 minutes duration per drives

Driver ref.	# drives	Segments	Number of segments per class		
			High	Medium	Low
NM	3	22	9	7	6
RY	2	14	6	4	4
BK	1	6	2	2	2
MT	1	8	3	5	0
EK	1	7	4	3	0
KSG	1	5	2	1	2
AD	1	6	2	2	2
GM	2	14	2	8	4
SJ	1	8	4	2	2

intervene in the stress assessment and the assumption is verified. Four segments from the rest period are classified as perceived medium stress level since some participants used their phone or speak with the experimenter. Seven segments extracted from city driving are classified as medium stress level.

6.4.2 Physiological variable selection and arousal level classification

In this section, the considered physiological variables are the electrodermal activity measured on the right wrist (RightEDA), on the left wrist (LeftEDA), the breathing rate (BR) and the heart rate (HR). The importance of physiological functional variables considered in our study will be presented. We will proceed to the elimination of the less important physiological signals, which do not contribute significantly in the identification of the stress level.

Importance of physiological variables

Data resulted from the physiological functional variables decomposition provide 1024×4 scalar variables corresponding to the wavelet coefficients. Taking advantage of the grouped VI, the scalar variables were grouped by physiological variable and the VI of each considered group was computed. Thereby, the grouped VI allowed to consider only 4 informative VI instead of 4096 VI of the scalar variables.
In order to examine the variable importance (VI) of each group, we present in Fig. 6.8 the boxplots of the grouped VI values, computed over 100 runs.
It can be noticed that *RightEDA* is the most important physiological variable with a VI distribution around 2%. The *HR* variable is found to have values corresponding to 100 VI around 0%.

Physiological variable elimination

In order to eliminate the less relevant physiological variables, we applied the RF-RFE algorithm ten times to the wavelet coefficients grouped by functional variables. Each run offers a ranking of the 4 variables and provides the list of variables retained by the selection procedure. Table 6.5 shows the result of 10 runs of the RF-RFE algorithm on the four physiological variables. The shaded cells correspond to the final retained variables. For instance, *Right EDA* is selected in the first run.

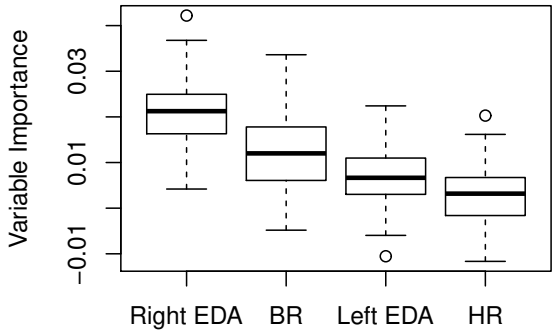


Fig. 6.8 Boxplots of grouped VI by physiological signals for 100 runs.

Based on these 10 runs, *Right EDA* is always included in the selected model (except one run). Additionally, *Left EDA* and *HR* (except for two runs) are excluded from the list of the variables selected in the model.
It can be noticed that the number of selected variables and the order of importance of these selected variables vary from an iteration to another. Thus, we propose to use the 3 step approach detailed in Section 6.3.1 in order to aggregate the information contained in these 10 runs and to select the most "enduring" physiological variables. Fig. 6.9 displays the "endurance" score of

Num	**Weights**				**Numb sel var**
	4	3	2	1	
1	Right EDA	Left EDA	BR	HR	1
2	BR	Right EDA	Left EDA	HR	4
3	BR	Right EDA	HR	Left EDA	3
4	Right EDA	BR	HR	Left EDA	2
5	Right EDA	BR	Left EDA	HR	1
6	BR	Right EDA	BR	Left EDA	2
7	BR	HR	Right EDA	Left EDA	3
8	BR	Left EDA	HR	Right EDA	1
9	Left EDA	Right EDA	BR	HR	2
10	BR	Right EDA	Left EDA	HR	3

Table 6.5 Selected model for 10 runs of the RF-RFE algorithm. The shaded cells corresponds to the retained variables.

the four physiological variables computed based on Equation 6.1. The results confirm the order found by the VI measure displayed in Fig. 6.8 where both *Left EDA* and *HR* had low endurance score. Moreover, the endurance score reveals that *Right EDA* is the most important variable in driver's stress level classification, followed by *BR*.

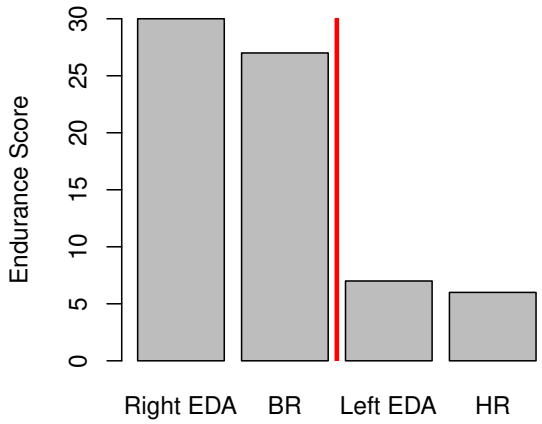


Fig. 6.9 Endurance score of the four physiological variables for 10 runs.

Wavelet levels selection from the retained physiological variables

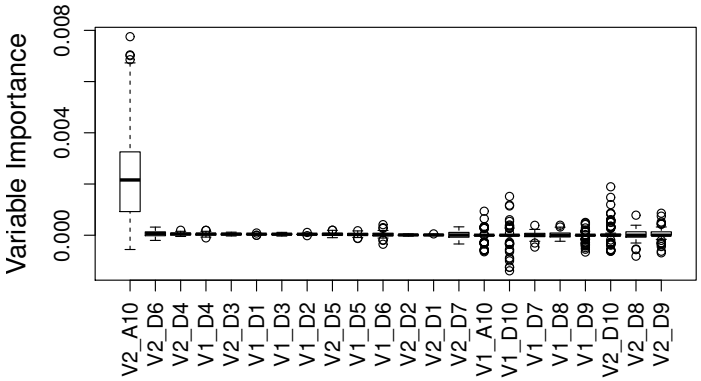
The goal of this step is to identify, for the retained physiological functional variables, the corresponding wavelet levels the most able to predict the driver's stress level. To achieve this goal, the wavelet coefficients are now grouped by levels.

Wavelet levels importance

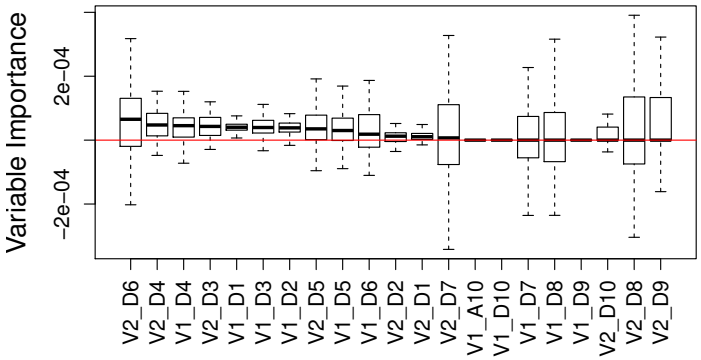
Let V1 stand for *RightEDA* and V2 for *BR*. We simply denote by Vx_A12 or Vx_Dj the wavelet decomposition components of the functional variable Vx. Fig. 6.10 (a) depicts the boxplots of the VI computed over 100 runs of the different wavelet levels of the two retained physiological variables. The boxplots show that the approximation level ($A10$) of the BR is the most important wavelet level in the stress level classification. The variation of the distribution of the other wavelet levels is not clearly visible in Fig.6.10 (a). In order to better investigate the relative variation of small VI, we present a graph obtained by removing the dominant variables $V2\_A10$ and the outliers in Fig. 6.10 (b). We find that for the *RightEDA* (V1), levels from 1 to 6 are the most important. Wavelet levels from 2 to 6 are important for the *BR* (V2).
In the next paragraph, we compute the "endurance" score for all wavelet levels that allows to select the final model.

Wavelet levels selection

When applying the RF-RFE algorithm to the groups composed by the wavelet levels of the three retained variables several times, the selected model varies in terms of the number of the selected levels, the validation error and even the ranking of the different wavelet levels. In order to reduce this variability, an "endurance" score, introduced earlier by Equation 6.1, is computed. It allows to rank the different wavelet levels according to their ability to persist in the selection procedure. The plot of the ranked variables using this score is displayed in Fig. 6.11.



(a) Grouped VI of the V1 and V2 wavelet levels.



(b) Grouped VI of the V1 and V2 wavelet levels without V2_A12.

Fig. 6.10 Grouped VI of the wavelet levels for 100 iterations for the V1 (RightEDA) and V2 (BR).

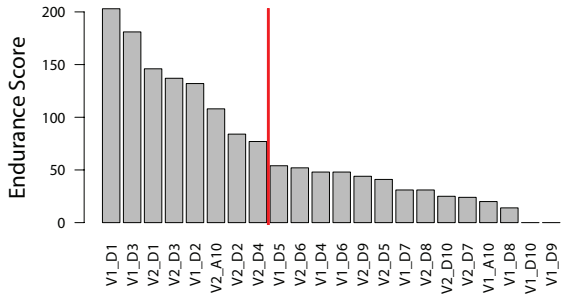


Fig. 6.11 Wavelet levels selection for 10 runs.

In order to select the wavelet levels, we examine the consecutive differences of the "endurance" scores starting from the lowest score and we cut off at the largest gap of this difference. When applying this approach on the wavelet details corresponding to the two retained physiological variables, we select the three first wavelet levels of details of $V1$ and the four first wavelet levels of details and the approximation at level 10 of $V2$. The three first wavelet details selected for the Right EDA correspond to 2-4Hz, 1-2Hz and 0.5-1Hz ranges. These features are relevant in characterizing the EDA since they are used in [154] work when detecting the EDA artifacts. These

ranges of frequencies were also selected in [51] work (which is Chapter 2 of the thesis report), that aims to classify the driver's stress level using the open database *drivedb*.

Assessing the final model

Recall that the 3-step approach applied on the "AffectiveROAD" database resulted in the selection of RightEDA and BR as the most relevant functional variables, in addition to their most important wavelet levels. Specifically, this resulted in a total of 8 wavelet signals corresponding to levels 1 to 3 of the Right EDA wavelet details and levels 1 to 4 of the BR and the level 10 of the BR wavelet approximation.
In this section, we first assess the performances of the RF-based model when considering these retained variables in the classification of the driver's arousal level, in terms of misclassification rate using a cross-validation like-procedure for model error estimation. This procedure takes into account the fact that some drivers repeated the experiment several times. Table 6.6 summarizes the four settings considered in the final model error computation. For instance, Setting 1 takes as test set the segments extracted from signals related to the driver NM and as training set the rest of segments related to the other drivers.

Table 6.6 Cross validation-like procedure

(a) Setting 1

Sets	Driver ref.	Segments	Number of segments per class		
			High	Medium	Low
Test 1	NM	22	9	7	6
Training 1	Others	68	25	27	16

(b) Setting 2

Sets	Driver ref.	Segments	Number of segments per class		
			High	Medium	Low
Test 2	RY and GM	28	8	12	8
Training 2	Others	62	26	22	14

(c) Setting 3

Sets	Driver ref.	Segments	Number of segments per class		
			High	Medium	Low
Test 3	BK, MT, and AD	20	7	9	4
Training 3	Others	70	27	25	18

(d) Setting 4

Sets	Driver ref.	Segments	Number of segments per class		
			High	Medium	Low
Test 4	EK, KSG, and SJ	20	10	6	4
Training 4	Others	70	24	28	18

Recall that a total of 2048 corresponding to the wavelet coefficients were considered as input variables to the RF algorithm. For each setting we compute the model error that corresponds to the average of the misclassification rate obtained over 100 runs. The final error corresponds to the average of the model errors for the four settings. This results of an error equal to 0.34. The same value of the error is obtained even when taking only the wavelet coefficients corresponding to the retained wavelet levels.

6.4.3 Bio-signals selection and arousal level classification

In this section, the posture (Post) is considered as a biosignal and is added to the list of the 4 physiological variables. First of all, the Variable Importance of these variables will be examined. Then, the endurance score will be computed in order to eliminate the less enduring variables. Finally, the variable selection will be achieved and the final model will be assessed.

Importance of the bio-signals

We consider here the scalar variables that correspond to the wavelet coefficients resulted from the functional variables decomposition providing thus 1024×5 variables.
In order to examine the variable importance (VI) of each group, we present in Fig. 6.12 the boxplots of the grouped VI values, computed over 100 runs. Fig. 6.12 shows that Post has the

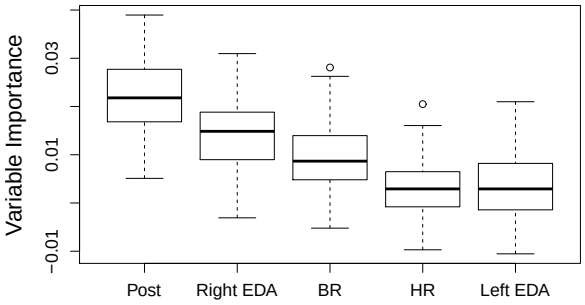


Fig. 6.12 Variable Importance of the 5 biosignals for 100 runs.

highest values of VI, than right EDA followed by BR. Left EDA and HR have 0% of VI. This suggests to eliminate those two variables if the VI was the criterion considered.

Bio-signals elimination

Based on the endurance score, we proceed to the variable elimination. We propose to eliminate BR, HR and Left EDA based on the graph depicted in Fig. 6.13. Left EDA and HR were eliminated in the first analysis considering only physiological signals.

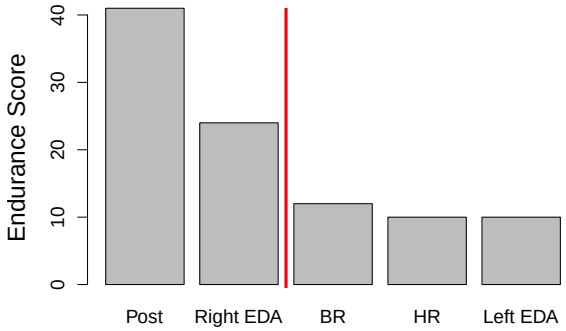


Fig. 6.13 Variable Importance of the 5 variables for 100 runs.

Wavelet levels selection

In this section, we propose to select the wavelet levels among the retained biosignals namely Post and Right EDA. Fig. 6.14 depicts the boxplots of the VI computed for 100 runs. It shows that the most important wavelet features are wavelet details corresponding to the levels 1 to 4 for RightEDA, and from 1 to 8 for Post.

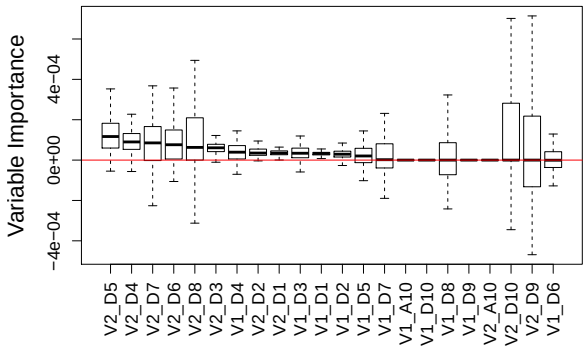


Fig. 6.14 Grouped VI of the wavelet levels for the V1 (RightEDA) and V2 (Post).

When computing the endurance score for the wavelet details over the different scales, we find that the first wavelet level of the RightEDA and the wavelet levels 1 and 4 of Post are the most enduring wavelet levels in the stress level classification (see Fig. 6.15).

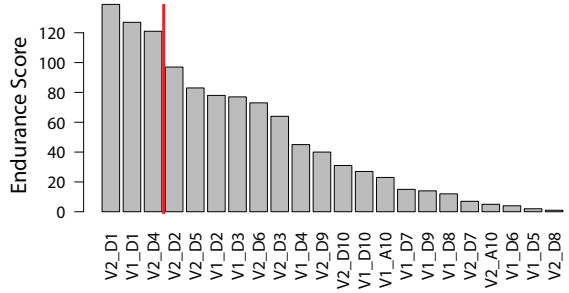


Fig. 6.15 Variable Importance of the 5 variables for 100 runs.

Assessing the final model

In order to assess the final model, the wavelet levels corresponding to the two retained biosignals, namely Right EDA and Post, are first considered. For each setting detailed in Table 6.6, the mean misclassification rate computed over 100 runs. Second, the selected wavelet levels are considered. In particular, the wavelet details corresponding to the first and the fourth levels of Post and the first level detail coefficients of Right EDA are selected. In order to assess the final model, we consider the mean of the misclassification rates computed over 100 runs for the settings 1, 2, 3 and 4. Table 6.7 lists the averages of the misclassification rates for the four settings. We notice then that the final error is the same, with or without selection.

Table 6.7 Average of the misclassification rates over 100 runs of the Random Forest models considering the retained Right EDA and Post. This assessment is computed for two configurations: when considering all the wavelet levels (first column) and for only the selected wavelet levels (second column).

	All the wavelet levels	Selected wavelet levels
Setting 1	0.43 (sd=0.09)	0.43 (sd=0.09)
Setting 2	0.39 (sd=0.06)	0.42 (sd=0.06)
Setting 3	0.41 (sd=0.09)	0.40 (sd=0.08)
Setting 4	0 (sd=0)	0 (sd=0)
Final error	0.31	0.31

6.4.4 A regression model of the subjective stress metric on the biosignals

The mean of the subjective stress metric corresponding to the extracted functional variables is considered in this section as the variable of interest. We aim to build a regression model that explain this variable of interest using a selected list of the five biosignals namely: Left EDA, Right EDA, BR, HR and Post. We want to check if the lists of the selected biosignals involved in the classification and the regression models are the same. Does the selected wavelet-based features remain the same for the regression model?

Variable importance and biosignals selection

Fig. 6.16 depicts the boxplots of the computed VI for the 5 variables Left EDA, Right EDA, BR, HR and Post. It shows that even for the regression, the Post is the most important variable followed by RightEDA. The same order is almost observed in the classification case except the Left EDA that became more important.

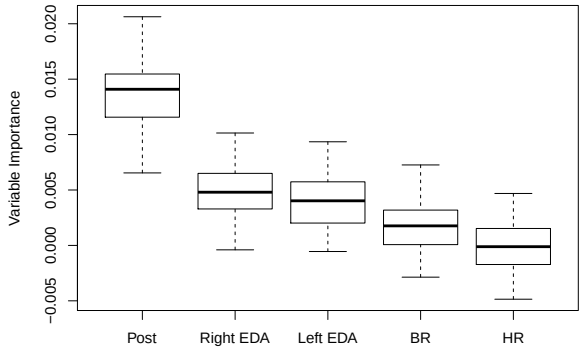


Fig. 6.16 Variable Importance of the 5 variables for 100 runs.

Variable elimination

In order to eliminate the less persistent biosignals, the same approach is repeated in the case of the regression model. Fig. 6.17 shows the endurance score of the 5 variables Left EDA, Right EDA, BR, HR and Post. Based on this score, Left EDA, BR and HR are eliminated. Thus, we perform the wavelet selection using RightEDA and Post.

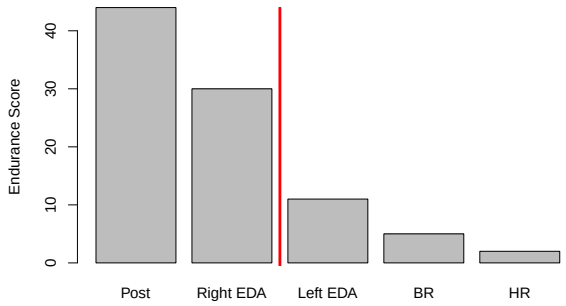


Fig. 6.17 Endurance of the 5 variables for 10 runs.

Wavelet levels selection

When selecting among the 22 wavelet levels corresponding to the two retained biosignals, we look at the ordered endurance score computed by wavelet level. The rule used to cut off the variables was to look at the difference of the consecutive scores and to choose the biggest value. Applying this trick, we do select three wavelet levels (c.f. Fig. 6.18) corresponding to the first and the fourth wavelet levels of RightEDA and the level 5 of the Post. Note that the first wavelet level is selected in the different models. The frequency ranges corresponding to this level are $2-4$ Hz since the sampling frequency is 4Hz. These ranges were also relevant and enduring to the selection approach in the work [51]. In fact, the selected wavelet detail signal, roughly corresponding to the ranges of $1.9-3.9$ Hz (with a sampling frequency Fs=15.5Hz), was included in the list of the selected wavelet-based features in the final model.

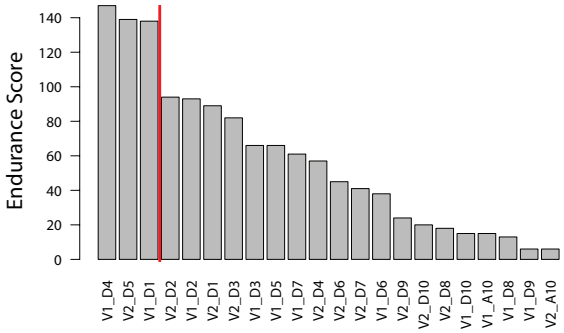


Fig. 6.18 Variable Importance of the wavelet levels of the two retained biosignals for 10 runs. V1 stands for RightEDA and V2 stands for Post.

Assessing the final model

The final model considers the two selected biosignals namely RightEDA and Post. These two signals are decomposed into a wavelet basis. The Random Forest model is constructed first considering all the wavelet levels. It is trained using the four training sets corresponding to the settings listed in Table 6.6. The root mean square error (rmse) is computed on the corresponding test sets. In order to increase the stability of the results, the error reported in Table 6.8 corresponds to the average of the obtained rmse over 100 runs of the Random Forest. The second column corresponds to the same error when considering only the three selected wavelet levels in Fig. 6.18. One can notice that the same errors are obtained when considering the 22 wavelet levels and when considering only the three wavelet levels. The final error of the regression model is 0.2.

Table 6.8 Average of the root mean square errors over 100 runs of the Random Forest models built using the retained Right EDA and Post. This assessment is computed for two configurations: when considering all the wavelet levels (first column) and for the selected wavelet levels (second column).

	All the wavelet levels	Selected wavelet levels
Setting 1	0.24 (sd=0.006)	0.23 (sd=0.006)
Setting 2	0.24 (sd=0.005)	0.25 (sd=0.006)
Setting 3	0.22 (sd=0.006)	0.23 (sd=0.008)
Setting 4	0.1 (sd=0.005)	0.1 (sd=0.005)
Final error	0.20	0.20

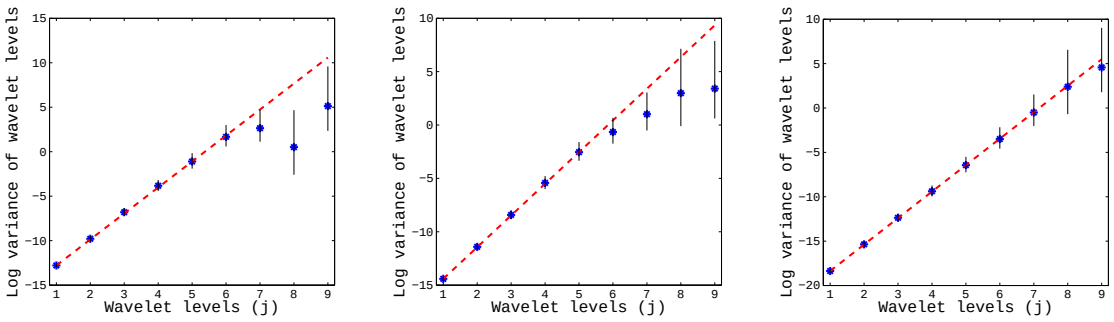
6.5 EDA self-similarity analysis

We present in this section the self-similarity analysis performed on the EDA signals. We apply the wavelet-based approach to estimate the Hurst index H on the EDA captured on the left and the right wrist, extracted from the open *AffectiveROAD* database. Ultimately, we seek to check if the Hurst exponent, estimated on EDA signals, can be used to characterize the driver stress level evolution during real-world driving environment.
The Hurst exponent is estimated for 3-minutes segments for each EDA signal corresponding to a drive. EDA segments of 3 minutes of duration extracted from the LeftEDA and RightEDA are

taken with an overlapping of 1 minute. In fact, the 3-minutes segment is taken, H is estimated then one minute is skipped and the next segment of 3 minutes of duration is considered and so on until the whole signal is swept. The segment duration is selected to be long enough to capture specific information related to the human affective reactions.

6.5.1 EDA and scale invariant processes

The logscale diagram allows, before even estimating the value of H, to inspect visually the linearity and the wavelet scales for which the ss-process properties hold. For instance, Fig. 6.19 (a) depicting the logscale diagram for a segment from the city driving of the second drive of GM, shows a clear linear pattern for the scales ranging from 1 to 6. The most encountered cases of the logscale diagrams are presented in Fig. 6.19. Such diagram can be linear for all the wavelet scales (cf. Fig. 6.19 (c)) or for the first scale ranges (cf. Fig. 6.19 (a) and 6.19 (b)).



(a) **(b)** **(c)**

Fig. 6.19 Illustration of linearity in logscale diagrams corresponding to the most encountered cases of linearity in the RightEDA of the second drive of GM. The vertical bars correspond to the 95% confidence interval estimated for each wavelet level.

The scale invariance character of EDA signal is provided by this visual inspection of the logscale diagram. It should be mentioned that for all the drives and their corresponding extracted segments, the linearity is checked for the wavelet levels ranging from 1 to 5. Thus, the estimation of the Hurst exponent is done based on the linear regression of the log variance of wavelet levels on wavelet levels for scales ranging from 1 to 5.

6.5.2 Driving experiment-based analysis

We analyze in this section the results of the self-similarity of RightEDA and LeftEDA per drive. One can notice that for the two signals all the values of the estimated H are greater than 0.5. This can reflect the fact that EDA is a persistent positively correlated time series with a trend.
In addition, we expect that more the task performance and the environment complexity increase more the boxplot of the estimated values of H is lower. The driving task complexity negatively varies with the estimated Hurst exponent. The ideal order is reflected in the boxplots depicted in Fig. 6.20 for the column AD for instance. This order is observed for the RightEDA for all the drives of NM, second drive of GM, EK and AD.
Let us comment the other boxplots where the order is not respected. The comments are listed in Chapter 4, Table 5.4. For RY, the boxplot of the estimated H in rest period is less than the other boxplots. This can be explained since the participant RY talked during the rest period of the first drive thus the arousal is higher and the task is switched from a simple relaxation to a talking task that may be a demanding task which produces physiological changes [164]. Her second drive exhibits a lower boxplot of estimated H corresponding to the highway segments,

which is not surprising since the highway was crowded especially with big trucks. The first drive corresponding to GM is characterized by values estimated from city segments comparable to the H values values estimated in the rest. In fact, the participant is used to drive in an other city where she encounters many vehicles and pedestrians. The road in her city is special thus she find that the city driving is not stressful. The boxplot of the highway is less than the other classes which means that the complexity is higher in highway. However, the highway for her is more stressful. It should be mentioned that during the highway segment the driver deviated from the expected road. For BK, The boxplot of the estimated H on highway suggests that the driving in this area is an over-arousal task. This can be explained by the fact that the driver deviated also from the expected path which stressed him.
The boxplot of estimated H on segments extracted from the rest period is comparable to the city and highway driving for the drive performed by MT. We believe that this is mainly due to her use of the smart-phone since she received a message and responded to it. KSG reported that he was distressed during the rest period since he was overthinking about an issue in his work. This was observed in his corresponding boxplot of rest. For SJ, a driving in a congested road characterizes his highway driving, thus the boxplot of the estimated H on segments extracted from highway is low than city boxplot. The rest boxplot is comparable to the city. This can be due to discomfort felt during rest caused by a dry throat and he wanted to cough.

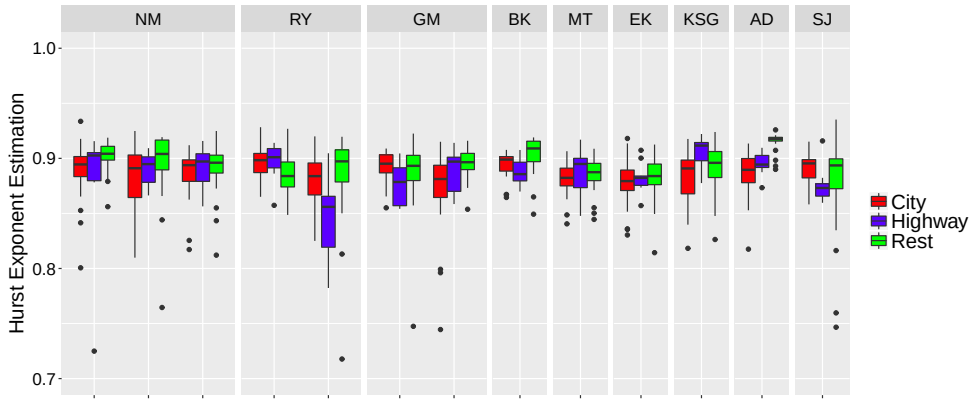


Fig. 6.20 The boxplot of the estimated H per drive for the 9 drivers for RightEDA. For each drive, the values are grouped by the different periods: city, highway and rest.

It is clear that for RightEDA, the boxplots corresponding to the values of the estimated H reflects the affect arousal and the perception of the driver of his/her environment, the mental workload and to the individual experience. However, this distinction is not clear for the leftEDA (see Fig. 6.21), the order of the boxplots is not the same as in RightEDA.

6.5.3 Individual driver-based analysis: right vs left EDA

We consider in this section the distribution of the estimated H for the periods city, highway, and rest grouped by driver. Fig. 6.22 depicts those boxplots for the case of the Right EDA. The same comments still true for the drivers performing the experiment only once namely BK, MT, ET, KSG, AD, and SJ.
For NM who repeated the experiment 3 times, the order assuming that city is more complex than highway which is more complex than rest is found. The participant RY is characterized by close median values of the three boxplots. This is mainly due to unexpected events like the congested highway and the fact that she talks in rest period. For GM, the individual-based presentation exhibits a comparable boxplots between the city and the highway thus characterizes the same perception of the two area driving. This was partly reported by the driver that found the highway driving is equivalent or even more stressful to her than the city driving since she was not used to.

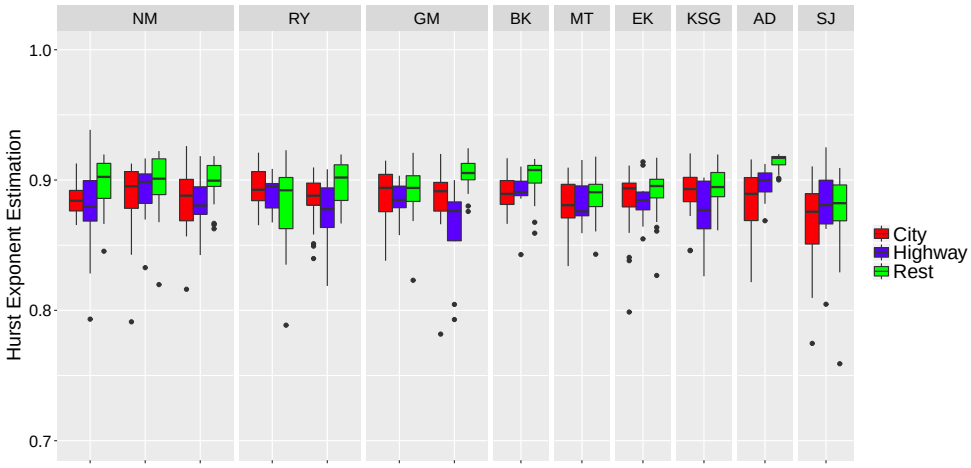


Fig. 6.21 The boxplot of the estimated H per drive for the 9 drivers for LeftEDA.

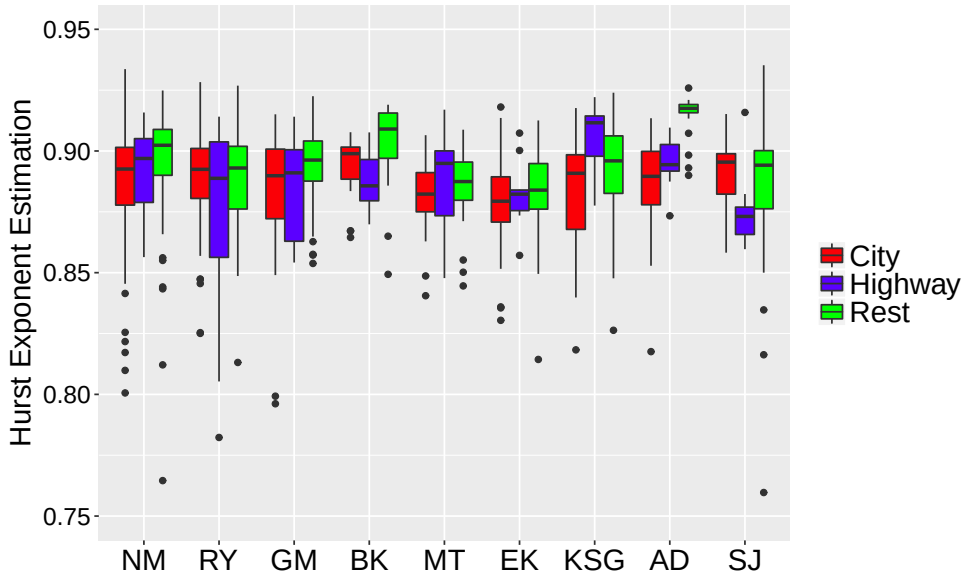


Fig. 6.22 Boxplots of the estimated H per participant for the Right EDA.

The same problem of interpretation is noticeable here for the the LeftEDA (see Fig. 6.23). No clear conclusions can be performed based on the different boxplots of the estimated Hurst exponent grouped by participant. This supports our finding of the functional variable selection study reporting that RightEDA is more important than LeftEDA.

6.6 Discussion

Vehicles are becoming more and more intelligent, responsive and communicating. Drivers expect that the car can sense him and feel his discomfort and well-being. To accomplish this goal, the affective state must be determined using objective measures, namely the biosignals. This study proposed two types of analysis for the evaluation of the driver's overall state of stress in a real-world driving context. Inspired from the protocol proposed by Healey and Picard [81], a driving protocol was designed and the data resulting from the use of such protocol is proposed publicly through *AffectiveROAD* database. The first data analysis aimed to recognize the driver's state

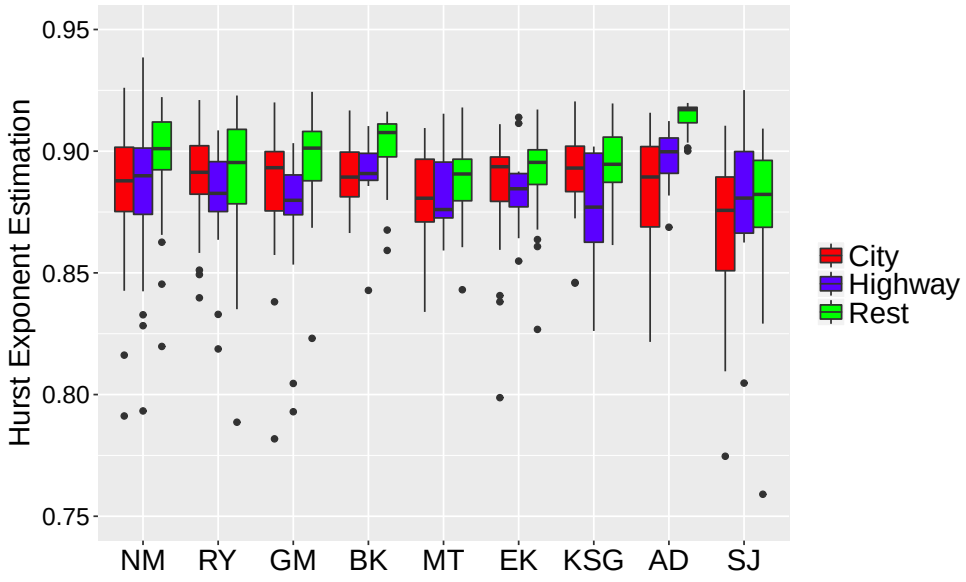


Fig. 6.23 Boxplots of the estimated H per participant for the Left EDA.

of stress. A Random-Forests-based classification approach is proposed. First, the physiological signals, namely RightEDA, LeftEDA, BR, and HR were considered. A selection of variables is performed in order to remove the less relevant physiological signals based on the endurance score that quantifies the persistence of a variable to the selection procedure. HR is found of null importance and the RightEDA and BR are retained. This first part of the analysis was achieved in order to compare with the results of the physiological variable selection achieved on *drivedb*, described in Chapter 2. Recall that FootEDA and BR were selected. The fact that the EDA and the BR are here retained confirms the findings.
After that, we include the Post which is a biosignal characterizing the degree of deviation from the vertical axis. It was found that is an important variable in the recognition of the state of stress. After applying the whole approach passing from biosignal selection to wavelet levels selection, the final proposed model considers three wavelet levels among 55 signals that correspond to the different wavelet levels offering an error of 0.31. This rate is comparable to those obtained in Chapter 2, based on the same approach applied on a the *drivedb* open database.
The second part of the first analysis takes advantage from the available subjective stress metric that was built based on real-time annotation proposed by the experimenter, than validated by a video rating session performed by the driver. This metric offers a gain of the time spent classically in the experiment annotation, a gain in the effort and even in the cost when resorting to experts to perform annotation. Moreover, this metric benefits from a detailed assessment of the environment perceived into the vehicle provided by the experimenter. The validation of such metric by the driver based on the video rating allows to consider the perception of the driver of the environment and his affect arousal during the drive performance. Based on this metric, we perform a RF-RFE to select the most enduring biosignals. The variables Post and RightEDA were an other time selected.
In the real-world driving experience, the non-dominant wrist/hand which usually corresponds to the left side, is used to capture the EDA. However, for both classification and regression Random-Forest based models performed in our study, RightEDA was selected and found more important than LeftEDA. This finding motivates the recommendation to consider both sides when measuring the EDA during the performance of real-world driving experience.
More signals present in the *AffectiveROAD* database such as the environment temperature, humidity, sound level, etc. can be considered in future work in order to increase the accuracy of the state of stress recognition.

The second analysis of this study investigated how the stress in real-world driving is captured via a self-similar analysis of Electrodermal activity (EDA). One conclusion concerns the scale invariance of the EDA segements checked for the RightEDA and LeftEDA. In fact, when using the wavelet-based approach to estimate the Hurst exponent, a linearity was observed in all the logscale diagrams for the wavelet levels ranging from 1 to 5. This supports the finding in Chapter 3 since the levels found here exhibiting the linearity corresponds to the same frequency ranges found in their work. The linearity of the logscale diagram is observed here for the levels 1 to 5, with a sampling frequency 4 Hz thus the scale invariance is verified for the ranges $0.13 - 4$ Hz. Those ranges correspond to the wavelet levels 3 to 7 when the sampling frequency is of 15.5 Hz. Moreover, the estimated values of the Hurst exponent for both RightEDA and LeftEDA are greater than 0.5. This suggests an other time that the EDA exhibits an other time a behavior of a persistent positive correlated time series with a trend (see Chapter 3 for details). In addition to the scale invariance, the EDA changes were captured using the Hurst exponent estimated by rest, highway and city driving conditions. When considering the RightEDA, the results obtained can be explained by the special events that occurred during each drive and the estimated Hurst exponent can be considered thus as a good indicator of the overall stress that combines the affect arousal, the distress, the mental workload and the environment complexity perception.

6.7 Conclusion

In summary, we have analyzed 5 biosignals that were collected during a real-world driving experience. The RightEDA, LeftEDA, BR, Post and HR are extracted from the open *AffectiveROAD* database. Two types of analysis were presented in this work. The first one aimed to recognize the stress of the driver using an approach combining wavelet decomposition, grouped variable importance, and Random-forests-based recursive feature elimination. This allows to select two variables from the five biosignals namely RightEDA and Post. Then, 3 wavelet signals that correspond to three different wavelet levels are selected in the final model. A cross validation like procedure is considered in the final error assessment. This procedure considers the fact that a driver may repeat the experiment several times. The final error (in terms of misclassification rate) results in 0.31. When considering the subjective stress metric as the variable of interest, a regression model is selected and offers also a selection of RightEDA and Post as the most enduring biosignals. Three wavelet levels are also selected and the final error is about 0.2.
The second analysis aimed to characterize the EDA through self-similar process mainly the Fractional Brownian Motion. A wavelet-based approach is used to estimate the Hurst exponent that characterizes the local irregularity and the long range dependence of such processes. When inspecting the logscale diagram, a linearity is observed for the wavelet scales ranging from 1 to 5, which defines thus a scale invariance of the EDA for this range. Moreover, all the estimated values of H are greater than 0.5 suggesting that the EDA can be considered as persistent with a trend. Even though the boxplots of the estimated H did not show a clear ordered distinction between the three periods, they still inform about the driver's state. Specifically, the boxplots corresponding to the RightEDA reflect the driver state and his/her stress level. This was validated by video recordings and the subjective report. This study suggests in summary that the EDA should be considered in the driver's state recognition. In addition, the Hurst exponent offers an important cue of the human affect state and can be considered toward a real-time monitoring of the driver's state.

Chapter

7 Conclusion and Perspectives

7.1 Summary and conclusions

With the rise of urbanization trend over the world and the new technological advances of wearable physiological sensors, the ability to track and evaluate the stress levels of vehicle's drivers is more feasible and its impact is of immediate interest to several stakeholders.

In this study, we have focused on the analysis of drivers stress level changes comparing different urban area: city versus highway situations. For that, we first studied physiological signals gathered from driving experiments conducted in the Boston area, described in a pioneering study [81] of MIT. These signals are grouped in the publicly available database *drivedb* accessible from PhysioNet website. We later developed our own database (*AffectiveROAD*) with local experiments in the Grand Tunis area, using different wearable biosensors. In addition, the database includes measurements related to physical factors inside and outside the vehicle. Subjective studies in both *drivedb* and *AffectiveROAD* were carried out to establish a stress metric.

The physiological signals were analyzed using the Random Forest Recursive Feature Elimination approach, in order to identify an order of relevance and endurance of such signals in the stress level recognition of drivers, considered in both databases. One of the main findings of this analysis concerned the particular importance of Electrodermal Activity (EDA), compared with Breathing Rate (BR), Electromyogram (EMG), and Heart Rate (HR).
Moreover, the used variable selection provided a finer order of features relevance, when these signals were decomposed on a wavelet basis. It is important to note that in both databases, the EDA was captured on driver in different placements. It emerged for *drivedb* database that the EDA captured on the foot was more revealing of the driver's stress, thus confirming the sensitivity of driving task to sensor placement choice.

Furthermore, it was interesting in this study to explore the complexity analysis of the EDA, from a fractal point of view, since this has not been applied to EDA, in contrast with other physiological indicators such as HR, EMG and BR. For that, a self-similarity characterization was performed based on Fractional Brownian Motion (FBM) characterized by the Hurst exponent (H). The values estimating H were computed in the cases of both individual based and drive-based analysis.
They revealed the presence of long-range dependency in EDA time series over a wide range of scales. Moreover, the individual-based fractal study of the *drivedb* EDA signals have shown more irregularities in city driving compared with that of highway. It is worth noting that the use of fractal physiology has been proven to be of wide use in medical community, where fractal measures captured different aspects of health and well-being. In this work, we have shown that the EDA carries fractal properties, and potentially can be more "reliable" in affective state recognition than other physiological signals.

The stress metric proposed in the *AffectiveROAD* database is more informative about the state of the driver. It reveals an overall state including the distress, workload, affect arousal and the perception of the driver to the complexity of the driver's environment.
This study focused on the stress level and can be applied to further investigate other states such

as drowsiness, fatigue, etc. using biosignals and the adapted statistical methods in this work.

7.2 Perspectives and future works

The current work is two-folded: on one hand the contributions touched upon the data analysis with its various aspects. On the other hand, the constructed *AffectiveROAD* database and the developed platform of sensors network, though inspired from the experience of Healey [81], opened up to further questions of interest and thus merit further investigations, as summarized below:

Method and result

- The used approach of Random Forests-Recursive Feature Elimination for variable selection was not always stable, and as remedy, an endurance score was introduced. A more exhaustive study of such solution should be examined for feature selection robustness.
  An extension of this method can be encountered by implementing the whole approach using the on-line Random-Forests [138]. Such implementation is of interest since the amount of data that can be gathered inside a vehicle is increasing. This is mainly due to the sensors embedded in the existing smart systems and human-machine interfaces in cars.

- Even though the fractal character of EDA time series was relatively validated, using the Fractional Brownian Motion for self-similarity analysis, in both driving databases, this property should be further examined in other EDA-based experiments, beyond vehicle driving tasks. Moreover, it is important to compare the ability of the fractal measures of the other physiological signals such as HR, BR and EMG in terms of their ability to characterize the stress level.
  Another extension consists on developing a mobile application providing EDA Hurst index estimation for a real-time drivers stress tracking.

- A correlation study between the resulting relevant objective stress level indicators (via the physiological features), and the subjective responses, collected via various techniques, needs to be further examined in order to better qualify and quantify the mapping of subjective versus objective stress recognition. This leads to a classical question that needs to be addressed in future study, which concerns the effect of the participant profile on his/her subjective reporting. Thus, it is important to examine the population profile (age, gender, experience, altered perceptions, etc.).

- Future studies would benefit from integrating the acquired vehicle internal and external ambient parameters. Namely, the *AffectiveROAD* database provides internal temperature, humidity, pressure and the sound level; in addition to vehicle motion information that can be extracted from GPS. It would be of interest to compare the importance of these parameters and see which of the environmental and the biosignals are more important in the stress recognition.

- The complexity of the auditory and visual scenes in city and highway driving should be part of future extensions of this work, for we know how important they can be on stress level alteration on humans. For instance, this can be achieved by detecting moving objects and analyzing the "video complexity" based on [24, 85, 93].

Data acquisition and collection

The developed *AffectiveROAD* database, made public, will allow researchers in Affective Computing to have access to a new multi-signals database with a continuous stress metric, thus opening

the way to further comparative studies and result verification and validation.

- The proposed *AffectiveROAD* platform is composed of an environmental sensors network and biosignals devices. The actual version of the platform allows to conduct different driving experiments. The elaboration of new protocols in order to conduct new driving experiments which aim for instance to assess the driver's stress level alteration while using current in-vehicle Human Computer Interfaces.

- An ongoing work consists on extending this data acquisition platform in order to develop a wireless version of *AffectiveROAD* environmental sensors network and to synchronize it with the biosignals devices.

- The developed subjective stress metric currently integrates different aspects of human affect such as workload, distress, affect arousal and perception of the driver's environment complexity. The purpose of such subjective metric is to facilitate the annotation of the perceived stress during each driving experiment. In fact, the classical protocol used to construct such metric are costly and involves the expert annotation. Our proposed protocol of the stress metric construction is conducted in order to spare the time and the effort of the annotation. However, it should be approved by in-lab studies where the environment is simpler and controlled. Such study is crucial to check if the proposed stress metric fits well with the real perceived stress by a participant in the context of a real-world experiment.

Appendices

Glossary

NHTSA National Highway Traffic Safety Administration

HCI Human-Computer Interaction

USDOT U.S. Department of Transportation

ITS Intelligent Transportation Systems

GPS Global Positioning System

CAN Controller Area Network

Physiological measurements

ANS Autonomic Nervous System

EDR Electrodermal Response

EDA Electrodermal Activity

BR Breathing Rate

HR Heart Rate

HRV Heart Rate Variability

ECG Electrocardiogram

EEG Electroencephalogram

BVP Blood Volume Pulse

PPG Photoplethysmography

EMG Electromygram

Post Posture

Units

BPM Beat Per Minute

μS Micro-Siemens

Hz Hertz

Appendix A

Description of the "*AffectiveROAD*" database: missing data and special events.

The data collected during driving experiments are detailed in Chapter 4 Table 5.4. The shaded cells correspond to missed data. Details about the most important special events occurring during each drive are listed in the column "Comments":

- *Sunbeam* is used to designate that the driver was disturbed by a ray of sunshine.
- *Loud noise* designates that a loud noise occurs.
- *Talk* designates that the driver talks with the experimenter.
- *Trucks* designates that too many trucks are encountered in the route.
- *Road Pb* is used to designate that the driver takes an other path thus deviating from the expected route.
- If the participant used his/her smartphone, this is reported in the table by *Phone use*.
- One driver reported that he was distressed during the rest period since he was thinking about an issue in his work. This is designated in the table by *Distressed*.
- *Congested* is used to designate a congested route where several vehicles are in the route and even vehicles are parked in the right side of the route.
- *Dry throat* designates that the participant encounters a discomfort because his throat feels dry and he wants to cough.

Appendix B

Questionnaire de conduite

Nom du sujet :
Age du sujet :
Numéro du sujet :
Numéro de la session :
Type de la voiture :
Date :

A travers ce questionnaire, vous serez amené à noter le niveau de stress de certains évenements ou zones de votre parcours ou à comparer certaines zones par rapport à d'autres avant et après votre expérience de conduite.

Questions générales :

1. Depuis quand vous avez votre permis de conduire ?

2. Vous conduisez :

   (a) Chaque jour

   (b) Quelques fois par semaine

   (c) Quelques fois par mois

   (d) Quelques fois par an

   (e) Rarement

3. Vous trouvez l'expérience de conduite toujours stressante ?

   (a) Oui

   (b) Non

4. En conduisant, vous vous sentez plus stressé que les autres personnes, moins ou au même niveau de stress ?

Questions pré-expérience :
Cocher la case qui peut correspondre au niveau de ton stress ressenti dans chaque zone.

Zone	Non stressante	Un peu stressante	Moyennement stressante	Très stressante	Extrêmement stressante
Parking					
Ville					
Autoroute					
Rondpoint					
Bretelle entrée/sortie autoroute					
Technopôle					

Pensez-vous que la chaleur a plus d'influence sur votre niveau de stress dans la voiture que n'en a le bruit ?

a Oui

b Non, le contraire

c La même influence

Questions post-expérience :

1. Votre niveau de stress ressenti pendant la dernière période de repos était-il plus ou moins élevé que lors de la première période de repos ? pourquoi ?
2. Avez-vous ressenti des niveaux de stress différents durant les périodes de conduite en ville (El Ghazela, El Mnihla, El Ghazela retour)
   Si oui, laquelle est plus stressante et pourquoi ?
3. Avez-vous ressenti des niveaux de stress différents lors de la conduite sur autoroute à l'aller et au retour?
   Si oui, laquelle est plus stressante et pourquoi ?
4. Avez-vous ressenti des niveaux de stress différents lors de la conduite dans les rond-points?
   Si oui, laquelle est plus stressante et pourquoi ?
5. D'après vous quel est le facteur qui a le plus influence votre niveau de stress :
6. Entre la chaleur et le bruit, qu'est ce qui a d'après vous le plus d'influence sur votre niveau de stress :
   (a) La chaleur
   (b) Le bruit
   (c) La même influence

Cocher la case qui correspond au niveau de stress ressenti dans chaque zone.

Zone	Non stressante	Un peu stressante	Moyennement stressante	Très stressante	Extrêmement stressante
Parking					
Ville					
Autoroute					
Rondpoint					
Bretelle entrée/sortie autoroute					
Technopôle					

Bibliography

[1] A Dataset of Real World Driving to Assess Driver Workload. https://www.hcilab.org/research/hcilab-driving-dataset/.

[2] Stress Recognition in Automobile Drivers. https://physionet.org/physiobank/database/drivedb/.

[3] The Warwick-JLR Driver Monitoring Dataset. http://www.dcs.warwick.ac.uk/dmd/.

[4] M. Ablaßmeier, T. Poitschke, S. Reifinger, and G. Rigoll. Context-aware Information Agents for the Automotive Domain Using Bayesian Networks. In *Proceedings of the 2007 Conference on Human Interface: Part I*, Beijing, China, 2007.

[5] P. Abry, P. Flandrin, M.-S. Taqqu, and D. Veitch. Self-similarity and long-range dependence through the wavelet lens. *Theory and applications of long-range dependence*, pages 527–556, 2003. Birkhäuser, Boston, MA.

[6] P. Abry, P. Flandrin, M.-S. Taqqu, D. Veitch, et al. Wavelets for the analysis, estimation and synthesis of scaling data. *Self-similar network traffic and performance evaluation*, pages 39–88, 2000.

[7] R. Acharya U., O. Faust, N. Kannathal, T. Chua, and S. Laxminarayan. Non-linear analysis of EEG signals at various sleep stages. *Computer Methods and Programs in Biomedicine*, 80(1):37–45, oct 2005.

[8] A. Akbas. Evaluation of the physiological data indicating the dynamic stress level of drivers. *Scientific research and essays*, 6(2):430–439, 2011.

[9] S. Aladağ, A. Güven, H. Özbek, and N. Dolu. A comparison of denoising methods for electrodermal activity signals. In *Medical Technologies National Conference (TIPTEKNO), 2015*, pages 1–4. IEEE, 2015.

[10] A.-H. Alkali, R. Saatchi, H. Elphick, and D. Burke. Short-time Fourier and wavelet transform analysis of respiration signal obtained by thermal imaging. In *2014 9th International Symposium on Communication Systems, Networks & Digital Sign (CSNDSP)*, pages 183–187. IEEE, Jul 2014.

[11] L. Auret and C. Aldrich. Empirical comparison of tree ensemble variable importance measures. *Chemometrics and Intelligent Laboratory Systems*, 105(2):157–170, 2011.

[12] D. Ayata, Y. Yaslan, and M. Kamasak. Emotion recognition via random forest and galvanic skin response: Comparison of time based feature sets, window sizes and wavelet approaches. In *2016 Medical Technologies National Congress (TIPTEKNO)*, pages 1–4. IEEE, oct 2016.

[13] A. Babloyantz, J.M. Salazar, and C. Nicolis. Evidence of chaotic dynamics of brain activity during the sleep cycle. *Physics Letters A*, 111(3):152–156, sep 1985.

[14] F.-R. Bach. Consistency of the group lasso and multiple kernel learning. *Journal of Machine Learning Research*, 9(Jun):1179–1225, 2008.

[15] H.-J. Baek, H.-B. Lee, J.-S. Kim, J.-M. Choi, K.-K. Kim, and K.-S. Park. Nonintrusive biological signal monitoring in a car to evaluate a driverâ€™s stress and health state. *Telemedicine and e-Health*, 15(2):182–189, 2009.

[16] J.-M. Bardet, I. Kammoun, and V. Billat. A new process for modeling heartbeat signals during exhaustive run with an adaptive estimator of its fractal parameters. *Journal of Applied Statistics*, 39(6):1331–1351, jun 2012.

[17] J.-M. Bardet, G. Lang, E. Moulines, and P. Soulier. Wavelet estimator of long-range dependent processes. *Statistical Inference for Stochastic Processes*, 3(1):85–99, 2000.

[18] A. Ben Jemaa, G. Irato, A. Zanela, A. Brescia, M. Turki, and M. Jaïdane. Congruent auditory display and confusion in sound localization: Case of elderly drivers. *Transportation Research Part F: Traffic Psychology and Behaviour*, 2017.

[19] G.-G Berntson and J.-T Cacioppo. Heart rate variability: Stress and psychiatric conditions. *Dynamic electrocardiography*, pages 57–64, 2004.

[20] A. Best and B. Braun. Using a novel data resource to explore heart rate during mountain and road running. *Physiological reports*, 5(8), 2017.

[21] R. Bittner, P. Smrčka, M. Pavelka, P. Vysokỳ, and L. Poušek. Fatigue indicators of drowsy drivers based on analysis of physiological signals. In *International Symposium on Medical Data Analysis*, pages 62–68. Springer, 2001.

[22] J. Bostrom. Emotion-sensing pcs could feel your stress. *PC World*, 2005.

[23] W. Boucsein. *Electrodermal Activity*. Springer US, 2012.

[24] C. Braillon, C. Pradalier, J.-L. Crowley, and C. Laugier. Real-time moving obstacle detection using optical flow models. In *Intelligent Vehicles Symposium, 2006 IEEE*, pages 466–471. IEEE, 2006.

[25] L. Breiman. Random forests. *Machine Learning*, 45(1):5–32, 2001.

[26] L. Breiman, J. Friedman, C.-J. Stone, and R.-A. Olshen. *Classification and Regression Trees*. The Wadsworth and Brooks-Cole statistics-probability series. Taylor & Francis, 1984.

[27] Breiman, L. and Cutler, A. randomforest: Breiman and cutler's random forests for classification and regression. *https://CRAN.R-project.org/package=randomForest, R package version 4.6-12*, 2015.

[28] K.-A. Brookhuis and D. de Waard. Monitoring drivers' mental workload in driving simulators using physiological measures. *Accident Analysis & Prevention*, 42(3):898–903, 2010.

[29] R.-M. Carney, K.-E. Freedland, P.-K. Stein, J.-A. Skala, P. Hoffman, and A.-S. Jaffe. Change in heart rate and heart rate variability during treatment for depression in patients with coronary heart disease. *Psychosomatic medicine*, 62(5):639–647, 2000.

[30] R. Chaudhary. Electrocardiogram comparison of stress recognition in automobile drivers on matlab. *Advance in Electronic and Electric Engineering*, 3(8):1007–1012, 2013.

[31] J. Cheng and G. Liu. Computing nonlinear features of skin conductance to build the affective detection model. In *2013 International Conference on Communications, Circuits and Systems (ICCCAS)*, pages 331–334. IEEE, Nov 2013.

[32] A.-V. Chobanian, G.-L. Bakris, H.-R. Black, W.-C. Cushman, L.-A. Green, J.-L. Izzo, D.-W. Jones, B.-J. Materson, S. Oparil, J.-T. Wright, et al. Seventh report of the joint national committee on prevention, detection, evaluation, and treatment of high blood pressure. *hypertension*, 42(6):1206–1252, 2003.

[33] R.-H Chowdhury, M.-B. Reaz, M.-BM Ali, A. Bakar, K. Chellappan, and T.-G. Chang. Surface electromyography signal processing and classification techniques. *Sensors*, 13(9):12431–12466, 2013.

[34] Cisco, IBSG. The internet of things: How the next evolution of the internet is changing everything, Apr 2011.

[35] G. Cookson. Inrix global traffic scorecard. 2018.

[36] J.-F. Coughlin, B. Reimer, and B. Mehler. Driver wellness, safety & the development of an awarecar. *AgeLab, Mass Inst. Technol., Cambridge, MA*, 2009.

[37] P. Danca, A. Vartires, and A. Dogeanu. An overview of current methods for thermal comfort assessment in vehicle cabin. *Energy Procedia*, 85(Supplement C):162 – 169, 2016. EENVIRO-YRC 2015 - Bucharest.

[38] M.-E. Dawson, A.-M. Schell, D.-L. Filion, and G.-G. Berntson. The electrodermal system. In John T. Cacioppo, Louis G. Tassinary, and Gary Berntson, editors, *Handbook of Psychophysiology:*, pages 157–181. Cambridge University Press, Cambridge, 001 2001.

[39] Y. Deng, Z. Wu, C.-H. Chu, and T. Yang. Evaluating feature selection for stress identification. In *Information Reuse and Integration (IRI), 2012 IEEE 13th International Conference on*, pages 584–591, Aug 2012.

[40] L.-S. Dhupati, S. Kar, A. Rajaguru, and A. Routray. A novel drowsiness detection scheme based on speech analysis with validation using simultaneous eeg recordings. In *2010 IEEE International Conference on Automation Science and Engineering*, pages 917–921, Aug 2010.

[41] R. Díaz-Uriarte and S.-A. De Andres. Gene selection and classification of microarray data using random forest. *BMC bioinformatics*, 7(1):3, 2006.

[42] D.-A. Dimitriev, E.-V. Saperova, and A.-D. Dimitriev. State anxiety and nonlinear dynamics of heart rate variability in students. *PLOS ONE*, 11(1):1–22, 01 2016.

[43] K. Dopart. Its research 2015-2019 automation. 2016.

[44] P. Doukhan, G. Oppenheim, and M.-S. Taqqu. *Theory and applications of long-range dependence*. 2003. Birkhäuser, Boston, MA.

[45] I.-E Dror, A.-E. Peron, S.-L. Hind, and D. Charlton. When emotions get the better of us: the effect of contextual top-down processing on matching fingerprints. *Applied Cognitive Psychology*, 19(6):799–809, 2005.

[46] R. Edelberg. Electrical properties of the skin. *Methods in psychophysiology*, pages 1–53, 1967. Baltimore: Williams & Wilkins.

[47] A. Eke, P. Hermán, J. Bassingthwaighte, G. Raymond, D. Percival, M. Cannon, I. Balla, and C. Ikrényi. Physiological time series: distinguishing fractal noises from motions. *Pflügers Archiv - European Journal of Physiology*, 439(4):403–415, Feb 2000.

[48] A. Eke, P. Herman, L. Kocsis, and L.-R Kozak. Fractal characterization of complexity in temporal physiological signals. *Physiological Measurement*, 23(1):R1–R38, Feb 2002.

[49] P. Ekman. Facial expression and emotion. *American psychologist*, 48(4):384, 1993.

[50] N. El Haouij, J.-M. Poggi, R. Ghozi, S. Sevestre-Ghalila, and M. Jaïdane. Self-similarity analysis of electrodermal activity for driver's stress level characterization.

[51] N. El Haouij, J.-M. Poggi, R. Ghozi, S. Sevestre-Ghalila, and M. Jaïdane. Random forest-based approach for physiological functional variable selection for driver's stress level classification. *Statistical Methods & Applications*, Feb 2018.

[52] N. El Haouij, J-M. Poggi, S. Sevestre-Ghalila, R. Ghozi, and M. Jaïdane. AffectiveROAD System and Database to Assess Driver's Arousal State. In *SAC 2018: Symposium on Applied Computing , April 9–13, 2018, Pau, France*, 2018.

[53] F. Eyben, M. Wöllmer, T. Poitschke, B. Schuller, C. Blaschke, B. Färber, and N. Nguyen-Thien. Emotion on the road: Necessity, acceptance, and feasibility of affective computing in the car. *Adv. in Hum.-Comp. Int.*, 2010:5:1–5:17, January 2010.

[54] R.-T. Faghih, P.-A. Stokes, M.-F. Marin, R.-G. Zsido, S. Zorowitz, B.-L. Rosenbaum, H. Song, M.-R. Milad, D.-D. Dougherty, E.-N. Eskandar, et al. Characterization of fear conditioning and fear extinction by analysis of electrodermal activity. In *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE*, pages 7814–7818. IEEE, 2015.

[55] D. Farina and F. Negro. Accessing the neural drive to muscle and translation to neurorehabilitation technologies. *IEEE Reviews in Biomedical Engineering*, 5:3–14, 2012.

[56] F. Ferraty and P. Vieu. *Nonparametric Functional Data Analysis: Theory and Practice (Springer Series in Statistics)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006.

[57] P. Flandrin. Wavelet analysis and synthesis of fractional brownian motion. *IEEE Trans. Inf. Theor.*, 38(2):910–917, September 1992.

[58] NHTSA's National Center for Statistics and Analysis. Critical reasons for crashes investigated in the national motor vehicle crash causation survey. 2015.

[59] N.-H Frijda. *The emotions*. Cambridge University Press, 1986.

[60] R. Genuer, J.-M. Poggi, and Tuleau-Malot. Vsurf: An r package for variable selection using random forests. *The R Journal*, 7(2):19–33, 2015.

[61] R. Genuer, J.-M. Poggi, and C. Tuleau-Malot. Variable selection using random forests. *Pattern Recognition Letters*, 31(14):2225 – 2236, 2010.

[62] E. Gerasimova, B. Audit, S.-G Roux, A. Khalil, O. Gileva, F. Argoul, O. Naimark, and A. Arneodo. Wavelet-based multifractal analysis of dynamic infrared thermograms to assist in early breast cancer diagnosis. *Frontiers in physiology*, 5:176, 2014.

[63] A. Gevins and M.-E. Smith. Neurophysiological measures of cognitive workload during human-computer interaction. *Theoretical Issues in Ergonomics Science*, 4(1-2):113–131, 2003.

[64] S. Ghosh, T. Nandy, and N. Manna. *Real Time Eye Detection and Tracking Method for Driver Assistance System*, pages 13–25. Springer India, New Delhi, 2015.

[65] J.-A. Gitter and M.-J. Czerniecki. Fractal analysis of the electromyographic interference pattern. *Journal of Neuroscience Methods*, 58(1-2):103–108, may 1995.

[66] A.-L. Goldberger, L. Amaral, L. Glass, J.-M. Hausdorff, P. Ivanov, R. Mark, J. Mietus, G.-B. Moody, C.-K. Peng, and H.-E. Stanley. Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals, http://physionet.org/. *Circulation 101(23):e215-e220; 2000 (June 13)*, 2000.

[67] D. Gopher and E. Donchin. Workload: An examination of the concept. 1986.

[68] A.-C. Granero, F. Fuentes-Hurtado, V. Naranjo Ornedo, J. Guixeres Provinciale, J.-M. Ausín, and M. Alcañiz Raya. A Comparison of Physiological Signal Analysis Techniques and Classifiers for Automatic Emotional Evaluation of Audiovisual Contents. *Frontiers in computational neuroscience*, 10:74, 2016.

[69] A. Greco, G. Valenza, L. Citi, and E.-P. Scilingo. Arousal and valence recognition of affective sounds based on electrodermal activity. *IEEE Sensors Journal*, 17(3):716–725, 2017.

[70] B. Gregorutti, B. Michel, and P. Saint-Pierre. Grouped variable importance with random forests and application to multiple functional data analysis. *Computational Statistics and Data Analysis*, 90:15–35, 2015.

[71] B. Gregorutti, B. Michel, and P. Saint-Pierre. Correlation and variable importance in random forests. *Statistics and Computing*, pages 1–20, 2016.

[72] Gregorutti, B. RFgroove: Importance Measure and Selection for Groups of Variables with Random Forests. *https://CRAN.R-project.org/package=RFgroove, R package version 1.1*, 2016.

[73] Z. Guendil, Z. Lachiri, C. Maaoui, and A. Pruski. Emotion recognition from physiological signals using fusion of wavelet based features. In *2015 7th International Conference on Modelling, Identification and Control (ICMIC)*, pages 1–6. IEEE, dec 2015.

[74] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik. Gene selection for cancer classification using support vector machines. *Machine Learning*, 46(1-3):389–422, March 2002.

[75] H2020. Smart, green and integrated transport. 2014.

[76] T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning*. Springer Series in Statistics. Springer New York Inc., New York, NY, USA, 2001.

[77] J.-M Hausdorff. Gait dynamics, fractals and falls: finding meaning in the stride-to-stride fluctuations of human walking. *Human movement science*, 26(4):555–589, aug 2007.

[78] J.-M. Hausdorff, C.-K. Peng, J.-Y. Wei, A.-L. Goldberger, and R. Jung. Fractal Analysis of Human Walking Rhythm. In *Biomechanics and Neural Control of Posture and Movement*, pages 253–264. Springer New York, New York, NY, 2000.

[79] R.-L Hazlett. Measuring emotional valence during interactive experiences: boys at video game play. In *Proceedings of the SIGCHI conference on Human Factors in computing systems*, pages 1023–1026. ACM, 2006.

[80] J. Healey. *Wearable and automotive systems for affect recognition from physiology*. PhD thesis, MIT Dept. of Electrical Engineering and Computer Science, 2000.

[81] J. Healey and R. Picard. Detecting stress during real-world driving tasks using physiological sensors. *IEEE Transactions on Intelligent Transportation Systems*, 6(2):156–166, June 2005.

[82] J. Healey, J. Seger, and R. Picard. Quantifying driver stress: Developing a system for collecting and processing bio-metric signals in natural situations. *Biomedical sciences instrumentation*, 35:193–198, 1999.

[83] M. Helander. Applicability of drivers' electrodermal response to the design of the traffic environment. *Journal of Applied Psychology*, 63(4):481–488, 1978.

[84] T. Horberry, J. Anderson, M.-A. Regan, T.-J. Triggs, and J. Brown. Driver distraction: The effects of concurrent in-vehicle tasks, road environment complexity and age on driving performance. *Accident Analysis & Prevention*, 38(1):185–191, 2006.

[85] M. Hu, S. Ali, and M. Shah. Detecting global motion patterns in complex videos. In *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*, pages 1–5. IEEE, 2008.

[86] M.-H. Imam, C.-K. Karmakar, A.-H. Khandoker, and M. Palaniswami. Effect of ECG-derived respiration (EDR) on modeling ventricular repolarization dynamics in different physiological and psychological conditions. *Medical & Biological Engineering & Computing*, 52(10):851–860, 2014.

[87] S. Jerritta, M. Murugappan, K. Wan, and S. Yaacob. Emotion recognition from facial emg signals using higher order statistics and principal component analysis. *Journal of the Chinese Institute of Engineers*, 37(3):385–394, 2014.

[88] Q. Ji and X. Yang. Real-time eye, gaze, and face pose tracking for monitoring driver vigilance. *Real-Time Imaging*, 8(5):357–377, October 2002.

[89] I. Jolliffe. *Principal Component Analysis*. Springer series in statistics, 2002.

[90] N. Kannathal, U. Rajendra Acharya, C.M. Lim, and P.K. Sadasivan. Characterization of EEG - A comparative study. *Computer Methods and Programs in Biomedicine*, 80(1):17–23, oct 2005.

[91] C. Karmakar, M.-H. Imam, A. Khandoker, and M. Palaniswami. Influence of psychological stress on qt interval. In *Computing in Cardiology 2014*, pages 1009–1012, Sept 2014.

[92] J. Karvonen and T. Vuorimaa. Heart rate and exercise intensity during sports activities. *Sports Medicine*, 5(5):303–311, 1988.

[93] D.-S. Kim and J. Kwon. Moving object detection on a vehicle mounted back-up camera. *Sensors*, 16(1):23, 2015.

[94] P. Konstantopoulos, P. Chapman, and D. Crundall. Driver's visual attention as a function of driving experience and visibility. Using a driving simulator to explore drivers' eye movements in day, night and rain driving. *Accident Analysis & Prevention*, 42(3):827–834, May 2010.

[95] A. Krstacić, G. Krstacić, and D. Gamberger. Control of heart rate by the autonomic nervous system in acute spinal cord injury. *Acta clinica Croatica*, 52(4):430–5, dec 2013.

[96] G. Krstacic, A. Krstacic, M. Martinis, E. Vargovic, A. Knezevic, A. Smalcelj, M. Jembrek-Gostovic, D. Gamberger, and T. Smuc. Non-linear analysis of heart rate variability in patients with coronary heart disease. In *Computers in Cardiology*, pages 673–675, 2002.

[97] L. Li, C. Liu, C. Liu, Q. Zhang, and B. Li. Physiological Signal Variability Analysis Based on the Largest Lyapunov Exponent. In *2009 2nd International Conference on Biomedical Engineering and Informatics*, pages 1–5, 2009.

[98] N. Li, J.-J. Jain, and C. Busso. Modeling of driver behavior in real world scenarios using multiple noninvasive sensors. *IEEE transactions on multimedia*, 15(5):1213–1225, 2013.

[99] H.-P. Lin, H.-Y. Lin, W.-L. Lin, and A.-CW Huang. Effects of stress, depression, and their interaction on heart rate, skin conductance, finger temperature, and respiratory rate: sympathetic-parasympathetic hypothesis of stress and depression. *Journal of Clinical Psychology*, 67(10):1080–1091, oct 2011.

[100] Y. Lin, H. Leng, G. Yang, and H. Cai. An intelligent noninvasive sensor for driver pulse wave measurement. *IEEE Sensors Journal*, 7(5):790–799, 2007.

[101] G. Louppe, L. Wehenkel, A. Sutera, and P. Geurts. Understanding variable importances in forests of randomized trees, 2013.

[102] R. Luijcks, H.-J. Hermens, L. Bodar, C.-J. Vossen, J. van Os, and R. Lousberg. Experimentally induced stress validated by emg activity. *PloS one*, 9(4):e95215, 2014.

[103] U. Lundberg, M. Forsman, G. Zachau, M. Eklöf, G. Palmerud, B. Melin, and R. Kadefors. Effects of experimentally induced mental and physical stress on motor unit recruitment in the trapezius muscle. *Work & Stress*, 16(2):166–178, 2002.

[104] D.-T. Lykken. Range correction applied to heart rate and to gsr data. *Psychophysiology*, 9(3):373–379, 1972.

[105] S.-G Mallat. Multiresolution approximations and wavelet orthonormal bases of $L^2(R)$. *Transactions of the American mathematical society*, 315(1):69–87, 1989.

[106] J.-E. Mc Grath. A conceptual formulation for research on stress. *Social and Psychological Factors in Stress*, 10:10–21, 1970.

[107] A. Morelli, M. Singer, V.-M. Ranieri, A. D'Egidio, L. Mascia, A. Orecchioni, F. Piscioneri, F. Guarracino, E. Greco, M. Peruzzi, et al. Heart rate reduction with esmolol is associated with improved arterial elastance in patients with septic shock: a prospective observational study. *Intensive care medicine*, 42(10):1528–1534, 2016.

[108] E. Moulines, F. Roueff, and M.-S. Taqqu. On the spectral density of the wavelet coefficients of long-memory time series with application to the log-regression estimation of the memory parameter. *Journal of Time Series Analysis*, 28(2):155–187, 2007.

[109] A. Munoz-Diosdado. Fractal and Multifractal Analysis of Human Gait. In *AIP Conference Proceedings*, volume 682, pages 243–250. AIP, 2003.

[110] K. Natarajan, R. Acharya U, F. Alias, T. Tiboleng, and S.-K. Puthusserypady. Nonlinear analysis of EEG signals at different mental states. *BioMedical Engineering OnLine*, 3(1):1–11, mar 2004.

[111] K.-K. Nicodemus, J.-D. Malley, C. Strobl, and A. Ziegler. The behaviour of random forest permutation-based variable importance measures under predictor correlation. *BMC Bioinformatics*, 11(1):1–13, 2010.

[112] J. Nims, M. Capozzi, M.-B. Hailey, and K. Crankson. Athletic performance monitoring system utilizing heart rate information, February 20 2018. US Patent 9,895,096.

[113] K. Niven and E. Miles. Affect arousal. In *Encyclopedia of Behavioral Medicine*, pages 50–52. Springer, 2013.

[114] Department of Economic and Social Affairs. World urbanization prospects: The 2014 revision. 2015.

[115] ITS Joint Program Office. Intelligent transportation systems (its) program overview.

[116] P. Orris, E.-D. Hartman, P. Strauss, J.-R. Anderson, J. Collins, C. Knopp, Y. Xu, and J. Melius. Stress among package truck drivers. *American Journal of Industrial Medicine*, 31:202–210, 1997.

[117] J. Pan, Q.-S. Ren, and H.-T. Lu. Vigilance analysis based on fractal features of eeg signals. In *Computer Communication Control and Automation (3CA), 2010 International Symposium on*, volume 1, pages 446–449. IEEE, 2010.

[118] C.-K. Peng, J. Mietus, J. M. Hausdorff, S. Havlin, H. E. Stanley, and A. L. Goldberger. Long-range anticorrelations and non-gaussian behavior of the heartbeat. *Phys. Rev. Lett.*, 70:1343–1346, Mar 1993.

[119] C.-K. Peng, J.-E Mietus, Y. Liu, C. Lee, J.-M. Hausdorff, H.-E. Stanley, A.-L. Goldberger, and L.-A. Lipsitz. Quantifying fractal dynamics of human respiration: age and gender effects. *Annals of biomedical engineering*, 30(5):683–692, 2002.

[120] E. Pereda, A. Gamundi, R. Rial, and J. González. Non-linear behaviour of human EEG: fractal exponent versus correlation dimension in awake and sleep stages. *Neuroscience Letters*, 250(2):91–94, jun 1998.

[121] J.-S. Perkiömäki, T.-H. Mäkikallio, and H.-V. Huikuri. Fractal and complexity measures of heart rate variability. *Clinical and experimental hypertension (New York, N.Y. : 1993)*, 27(2-3):149–158, 2005.

[122] P. Picard, S. Fedor, and Y. Ayzenberg. Multiple arousal theory and daily-life electrodermal activity asymmetry. *Emotion Review*, 8(1):62–75, 2016.

[123] R. Picard. *Affective Computing*. MIT Press, Cambridge, MA, USA, 1997.

[124] R. Picard. Affective computing for hci. In *HCI (1)*, pages 829–833, 1999.

[125] R. Picard, E. Vyzas, and J. Healey. Toward machine emotional intelligence: Analysis of affective physiological state. *IEEE transactions on pattern analysis and machine intelligence*, 23(10):1175–1191, 2001.

[126] J.-M. Poggi and C. Tuleau. Classification of objectivization data using cart and wavelets. *Proceedings of the IASC 07, Aveiro, Portugal*, pages 1–8, 2007.

[127] H. Portier, F. Louisy, D. Laude, M. Berthelot, and C.-Y. GuÉzennec. Intense endurance training on heart rate and blood pressure variability in runners. *Medicine and science in sports and exercise*, 33(7):1120–1125, 2001.

[128] D. Purves, G.-J. Augustine, D. Fitzpatrick, L.-C. Katz, A.-S. LaMantia, J.-O. McNamara, and S.-M. Williams. Physiological Changes Associated with Emotion. *Neuroscience. 2nd edition. Sunderland (MA): Sinauer Associates*, 2001.

[129] F. Putze, J.-P. Jarvis, and T. Schultz. Multimodal recognition of cognitive workload for multitasking in the car. In *Pattern Recognition (ICPR), 2010 20th International Conference on*, pages 3748–3751. IEEE, 2010.

[130] J. Ramsay and B.-W. Silverman. *Functional Data Analysis*. Springer Series in Statistics. Springer, 2005.

[131] J.-O. Ramsay and B.-W. Silverman. *Applied functional data analysis: methods and case studies*, volume 77. Springer New York, 2002.

[132] P. Ravier, O. Buttelli, R. Jennane, and P. Couratier. An emg fractal indicator having different sensitivities to changes in force and muscle fatigue during voluntary static muscle contractions. *Journal of Electromyography and Kinesiology*, 15(2):210–221, 2005.

[133] G. Rigas, Y. Goletsis, P. Bougia, and D.-I. Fotiadis. Towards driver's state recognition on real driving conditions. *International Journal of Vehicular Technology*, 2011, 2011.

[134] G. Rigas, C.-D. Katsis, P. Bougia, and D.-I. Fotiadis. A reasoning-based framework for car drivers stress prediction. In *Control and Automation, 2008 16th Mediterranean Conference on*, pages 627–632, June 2008.

[135] B.-F. Robinson, S.-E. Epstein, G.-D. Beiser, and E. Braunwald. Control of heart rate by the autonomic nervous system: studies in man on the interrelation between baroreceptor mechanisms and exercise. *Circulation Research*, 19(2):400–411, 1966.

[136] A.-H. Roscoe. Assessing pilot workload. why measure heart rate, hrv and respiration? *Biological Psychology*, 34(2):259 – 287, 1992. Special Issue Cardiorespiratory Measures and thier Role in Studies of Performance.

[137] M. Roughan, D. Veitch, and P. Abry. Real-time estimation of the parameters of long-range dependence. *IEEE/ACM Transactions on Networking (TON)*, 8(4):467–478, 2000.

[138] A. Saffari, C. Leistner, J. Santner, M. Godec, and H. Bischof. On-line random forests. In *Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on*, pages 1393–1400. IEEE, 2009.

[139] N. Scafetta, R.-E. Moon, and B.-J. West. Fractal response of physiological signals to stress conditions, environmental changes, and neurodegenerative diseases. *Complexity*, 12(5):12–17, may 2007.

[140] S. Schneegass, B. Pfleging, N. Broy, A. Schmidt, and Heinrich F. A data set of real world driving to assess driver workload. In *5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI'13). ACM, New York, NY, USA*, pages 150–157. IEEE, sep 2013.

[141] H. Selye. *Stress Without Distress*. McClelland and Stewart, 1974.

[142] A. Sharma, S. Velipasalar, S. Singh, D. Engel, and S. Gyawali. Effect of Freeway Level of Service and Driver Education on Truck Driver Stress - Phase 1. Technical report, Mid-America Transportation Center, 12 2012.

[143] K.-A. Sidek and I. Khalil. Automobile driver recognition under different physiological conditions using the electrocardiogram. *PC World*, 38:753–756, 2011.

[144] M. Singh and A.-B Queyam. Stress detection in automobile drivers using physiological parameters: A review. *International Journal of Electronics Engineering*, 5(2):1–5, 2013.

[145] R.-R. Singh, S. Conjeti, and R. Banerjee. Biosignal based on-road stress monitoring for automotive drivers. In *2012 National Conference on Communications (NCC)*, pages 1–5. IEEE, Feb 2012.

[146] M.-J. Skinner and P.-A. Simpson. Workload issues in military tactical airlift. *The International Journal of Aviation Psychology*, 12(1):79–93, 2002.

[147] R.-G. Smart, E. Cannon, A. Howard, P. Frise, and R.-E. Mann. Can we design cars to prevent road rage? *International Journal of Vehicle Information and Communication Systems*, 1(1-2):44–55, 2005.

[148] F.-M. Smits, C. Porcaro, C. Cottone, A. Cancelli, P.-M Rossini, and F. Tecchio. Electroencephalographic Fractal Dimension in Healthy Ageing and Alzheimer's Disease. *PloS one*, 11(2):e0149587, 2016.

[149] C. Strobl and A. Zeileis. Danger: High power! ? exploring the statistical properties of a test for random forest variable importance. *Proceedings of 18th International Conference on Computational Statistics*, 2008.

[150] J.-W. Tan, A.-O. Andrade, H. Li, S. Walter, D. Hrabal, S. Rukavina, K. Limbrecht-Ecklundt, H. Hoffman, and H.-C Traue. Recognition of intensive valence and arousal affective states via facial electromyographic activity in young and senior adults. *PloS one*, 11(1):e0146691, 2016.

[151] J. Tao and T. Tan. Affective computing: A review. In *International Conference on Affective computing and intelligent interaction*, pages 981–995. Springer, 2005.

[152] M.-S. Taqqu. Fractional brownian motion and long-range dependence. *Theory and applications of long-range dependence*, pages 5–38, 2003.

[153] P. Taylor, N. Griffiths, A. Bhalerao, D. Watson, X. Zhou, and T. Popham. Warwick-jlr driver monitoring dataset (dmd): A public dataset for driver monitoring research. In *Proceedings of the cognitive load and in-vehicle human-machine interaction workshop*, 2013.

[154] S. Taylor, N. Jaques, W. Chen, S. Fedor, A. Sano, and R. Picard. Automatic identification of artifacts in electrodermal activity data. In *Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE*, pages 1934–1937. IEEE, 2015.

[155] Team, R Core. R: A language and environment for statistical computing. r foundation for statistical computing, vienna, austria. *www. R-project. org*, 2016.

[156] S. Ullah and C.-F. Finch. Applications of functional data analysis: A systematic review. *BMC medical research methodology*, 13(1):43, 2013.

[157] M. Van Dooren, J.-J. De Vries, and J.-H. Janssen. Emotional sweating across the body: Comparing 16 different skin conductance measurement locations. *Physiology & Behavior*, 106(2):298 – 304, 2012.

[158] J.-BF. Van Erp, H.-JA. Veltman, and M. Grootjen. Brain-based indices for user system symbiosis. In *Brain-Computer Interfaces*, pages 201–219. Springer, 2010.

[159] D. Veitch and P. Abry. A wavelet-based joint estimator of the parameters of long-range dependence. *IEEE Transactions on Information Theory*, 45(3):878–897, 1999.

[160] D. Veitch and P. Abry. Matlab code for the wavelet based analysis of scaling processes: http://crin.eng.uts.edu.au/ darryl/index.html. 2002. Accessed: 2017-10-21.

[161] A. Verikas, A. Gelzinis, and M. Bacauskiene. Mining data with random forests: A survey and results of new tests. *Pattern Recognition*, 44(2):330–349, 2011.

[162] S. Wallot, R. Fusaroli, K. Tylén, and E.-M. Jegindø. Using complexity metrics with R-R intervals and BPM heart rate measures. *Frontiers in physiology*, 4:211, 2013.

[163] B.-J. West. Fractal physiology and the fractional calculus: a perspective. *Frontiers in physiology*, 1:12, 2010.

[164] G.-F. Wilson. Applied use of cardiac and respiration measures: Practical considerations and precautions. *Biological Psychology*, 34(2):163 – 178, 1992. Special Issue Cardiorespiratory Measures and thier Role in Studies of Performance.

[165] K. Yang, H. Yoon, and C. Shahabi. A supervised feature subset selection technique for multivariate time series. *Proceedings of the Workshop on Feature Selection for Data Mining: Interfacing Machine Learning with Statistics*, pages 92–101, 2005.

[166] R.-M Yerkes and J.-D. Dodson. The relation of strength of stimulus to rapidity of habit-formation. *Journal of comparative neurology*, 18(5):459–482, 1908.

[167] F. Zappasodi, L. Marzetti, E. Olejarczyk, F. Tecchio, and V. Pizzella. Age-Related Changes in Electroencephalographic Signal Complexity. *PloS one*, 10(11):e0141995, 2015.

[168] L. Zhang, T. Tamminedi, A. Ganguli, G. Yosiphon, and J. Yadegar. Hierarchical multiple sensor fusion using structurally learned Bayesian network. In *Wireless Health 2010 on - WH '10*, page 174, New York, New York, USA, 2010. ACM Press.

[169] R. Zheng, S. Yamabe, K. Nakano, and Y. Suda. Biosignal analysis to assess mental stress in automatic driving of trucks: Palmar perspiration and masseter electromyography. *Sensors*, 15(3):5136–5150, 2015.

[170] R. Zhu, D. Zeng, and M.-R. Kosorok. Reinforcement learning trees. Technical report, University of North Carolina, 2012.

List of Figures

3.1 Description of the driving setting. For simplification, the toll and the turnaround segments are considered as city driving. 42

3.2 Illustration of segment extraction of different physiological signals of Drive 7. Note that the physiological data were stored based on the same sampling frequency $F_s = 15.5\ Hz$. 44

3.3 Boxplots of grouped VI by physiological signals for 100 executions. 52

3.4 Endurance score of the five physiological variables for 100 runs. The two last physiological variables are removed. 53

3.5 Grouped VI of the wavelet levels for 10 iterations. 55

3.6 Wavelet levels selection for 10 executions. 55

3.7 Illustration of the reconstructed signals corresponding to Drive 07 for Foot EDA (left column) and RESP (right column), based on the three selected wavelet levels (see Fig. 3.6). The letter "R" corresponds to rest period, "C" to city driving and "H" to highway driving. 56

3.8 Wavelet levels endurance score of the three retained physiological variables after 10 executions. 57

3.9 Three configurations of training and test sets choice in the cross validation like procedure. 58

4.1 Samples of simulated FBM for H=0.2 (green), H=0.5 (blue) and H=0.8 (red). . . 66

4.2 Plots of the logscale diagram: illustrations of the linear tendencies in the logscale graphs based on 3-minutes segments extracted from HandEDA. 69

4.3 FootEDA corresponding to the Drive 4 of *drivedb* database is depicted in the top and the corresponding estimated Hurst exponent is presented in the bottom of the Figure. The estimation is done on segments of 3 minutes with 1 minute of overlapping. The term "Hwy" is used to designate highway driving. 70

4.4 The boxplots of estimated H on HandEDA per drive for the 4 drivers. For each drive, the boxplot is presented according to the different driving periods. The sample size is designated on the top of each boxplot. 71

4.5 The boxplots of estimated H on FootEDA per drive for the 4 drivers. For each drive, the distribution is presented according to the different driving periods. The sample size is designated on the top of each boxplot. 71

4.6 The boxplots of $\hat{H}$ estimated on (a) HandEDA and (b) FootEDA per driver. The number on the top of each boxplot designates the sample size. 72

4.7 The boxplots of $\hat{H}$ estimated on (a) HandEDA and (b) FootEDA per driving condition. The sample size is indicated on the top of each boxplot. 72

5.1 A photo of the plugged E4 in the left hand of a driver. It is unobtrusive device allowing to perform the driving task. . . . 77
5.2 A photo of the Zephyr BioHarness 3 sensor that should be fixed on the chest. . . . 78
5.3 An overview of the environmental *AffectiveROAD* platform: 1- Intel edison Arduino breakout 2- Shield in/out 3- Grove touch sensor 4- beeper 5- Clik air quality sensor 6- Grove Piezo vibration sensor 7- IMU grove MPU 9250 8- Adafruit BME 280 9- Grove luminance sensor 11- Leds 12- WIFI / Bluetooth antenna 13- Power supply. . . . 79
5.4 An overview of the used sensors to capture the different signals . . . 80
5.5 Snapshot from the video capturing the inside of the car. . . . 80
5.6 Snapshot from the screen shown to each participant in order to validate the subjective stress metric. . . . 83

6.1 Road map of the path drove by the different participants. . . . 89
6.2 The different zones considered in the road map respectively Rest, Z, City1, Hwy and City2. . . . 90
6.3 Raw bio-signals for a drive. EDA signals captured on both right and left wrists, Heart Rate, Breathing Rate and Posture corresponding to the second drive of the driver GM. . . . 91
6.4 Screenshot of the artifact detection from EDA Explorer applied to an EDA signal (a) before and (b) after preprocessing. This EDA measurement was captured on the left wrist of the driver GM during her second drive. Red vertical bars correspond to the placement of the different detected artifacts. . . . 92
6.5 Boxplots of the variable "Posture" per drive (a) before and (b) after the statistical processing. . . . 93
6.6 Snapshot from the screen shown to each participant in order to validate the subjective score. . . . 94
6.7 Likert plot corresponding to the responses to the two questions on the perceived stress level. The driver is asked to rate the stress level before and after performing the drive. . . . 94
6.8 Boxplots of grouped VI by physiological signals for 100 runs. . . . 99
6.9 Endurance score of the four physiological variables for 10 runs. . . . 100
6.10 Grouped VI of the wavelet levels for 100 iterations for the V1 (RightEDA) and V2 (BR). . . . 101
6.11 Wavelet levels selection for 10 runs. . . . 101
6.12 Variable Importance of the 5 biosignals for 100 runs. . . . 103
6.13 Variable Importance of the 5 variables for 100 runs. . . . 104
6.14 Grouped VI of the wavelet levels for the V1 (RightEDA) and V2 (Post). . . . 104
6.15 Variable Importance of the 5 variables for 100 runs. . . . 104
6.16 Variable Importance of the 5 variables for 100 runs. . . . 106
6.17 Endurance of the 5 variables for 10 runs. . . . 106
6.18 Variable Importance of the wavelet levels of the two retained biosignals for 10 runs. V1 stands for RightEDA and V2 stands for Post. . . . 107
6.19 Illustration of linearity in logscale diagrams corresponding to the most encountered cases of linearity in the RightEDA of the second drive of GM. The vertical bars correspond to the 95% confidence interval estimated for each wavelet level. . . . 108

6.20 The boxplot of the estimated H per drive for the 9 drivers for RightEDA. For each drive, the values are grouped by the different periods: city, highway and rest. . . . 109
6.21 The boxplot of the estimated H per drive for the 9 drivers for LeftEDA. 110
6.22 Boxplots of the estimated H per participant for the Right EDA. 110
6.23 Boxplots of the estimated H per participant for the Left EDA. 111

List of Tables

3.1 Description of the different 10 drives. Note that each participant is labeled by a sequence composed of gender (M or F) followed by the number of years of the driving experience. No information was available on Ind 4. 43

3.2 Details of the missing data for the 7 drives excluded from the analysis due to the reported incomplete data in the *drivedb* database. 43

3.3 Summary of the selection algorithm based on RF-RFE. 47

3.4 Summary of the proposed 3-step approach. 48

3.5 Summary of results of the 10 executions of the RF-RFE. 50

3.6 Selected model for 10 executions of the RF-RFE algorithm. The shaded cells corresponds to the retained variables. 53

3.7 Model error: misclassification rate averaged over 100 executions. 58

3.8 Features related to the EDA (Foot and Hand) and to the respiration signals proposed by [81]. 59

4.1 Automatic selection of the wavelet scales over which the logscale is linear. 68

5.1 Description of the different sensors used in the data acquisition 79

5.2 Characteristic of drivers 81

5.3 Roadmap: Description of the proposed periods to be encountered during the experiments and the corresponding assumed evoked stress level. 81

5.4 Summary of Driving Events. The mark “X” is used to designate that the data are complete. “Hwy” means highway driving and “Rest” designates rest period. The date of each drive is expressed in the second column of the table in format of DD/MM. All the drives were conducted at the same year, 2017. The special events listed in the column ”Comments” are described in Appendix A. 82

5.5 Comparison between studies on driver’s state recognition offering open database. The comparison is made in terms of used sensors and the provided data. 83

6.1 Studies achieved on driver’s state recognition based on biosignals 88

6.2 Rating questionnaire results before and after the experiment. 95

6.4 Contingency table between the mean subjective metric and the classes provided by the assumptions done in [81] 97

6.3 Repartition of the segments of 4.27 minutes duration per drives 98

6.5 Selected model for 10 runs of the RF-RFE algorithm. The shaded cells corresponds to the retained variables. 99

6.6 Cross validation-like procedure 102

6.7 Average of the misclassification rates over 100 runs of the Random Forest models considering the retained Right EDA and Post. This assessment is computed for two configurations: when considering all the wavelet levels (first column) and for only the selected wavelet levels (second column). 105

6.8 Average of the root mean square errors over 100 runs of the Random Forest models built using the retained Right EDA and Post. This assessment is computed for two configurations: when considering all the wavelet levels (first column) and for the selected wavelet levels (second column). 107

Titre : Biosignaux pour l'évaluation du niveau de stress du conducteur: sélection des variables fonctionnelles et caractérisation fractale de l'activité électrodermale.

Mots clés : Niveau de stress du conducteur, biosignaux, expérience de conduite réelle, sélection de variables fonctionnelles, ondelettes, analyse fractale.

Résumé : La sécurité et le confort dans une tâche de conduite automobile sont des facteurs clés qui intéressent plusieurs acteurs (constructeurs, urbanistes, départements de transport), en particulier dans le contexte actuel d'urbanisation croissante. Il devient dès lors crucial d'évaluer l'état affectif du conducteur lors de la conduite, en particulier son niveau de stress qui influe sur sa prise de décision et donc sur ses performances de conduite.

Dans cette thèse, nous nous concentrons sur l'étude des changements de niveau de stress ressenti durant une expérience de conduite réelle qui alterne ville, autoroute et repos.

Les méthodes classiques sont basées sur des descripteurs proposés par des experts, appliquées sur des signaux physiologiques. Ces signaux sont prétraités, les descripteurs ad-hoc sont extraits et sont fusionnés par la suite pour reconnaitre le niveau de stress.

Dans ce travail, nous avons adapté une méthode de sélection de variables fonctionnelles, basée sur les forêts aléatoires, avec élimination récursive des descripteurs (RF-RFE). En effet, les biosignaux, considérés comme variables fonctionnelles, sont tout d'abord projetés sur une base d'ondelettes. L'algorithme RF-RFE est ensuite utilisé pour sélectionner les groupes d'ondelettes, correspondant aux variables fonctionnelles, selon un score d'endurance. Le choix final de ces variables est basé sur ce score proposé afin de quantifier la capacité d'une variable à être sélectionnée et dans les premiers rangs.

Dans une première étape, nous avons analysé les signaux physiologiques tels que : fréquence cardiaque (HR), électromyogramme (EMG), fréquence respiratoire (BR) et activité électrodermale (EDA), issus de 10 expériences de conduite menées à Boston, de la base de données du MIT, drivedb. Dans une seconde étape, nous avons conduit 13 expériences in-vivo similaires, en alternant conduite dans la ville et sur autoroute dans la région de Grand Tunis. La base de données résultante, AffectiveROAD contient -comme dans drivedb- les biosignaux tels que le HR, BR, EDA mais aussi la posture du conducteur.

Le prototype de plateforme de réseau de capteurs développé, a permis de collecter des données environnementales à l'intérieur du véhicule (température, humidité, pression, niveau sonore et coordonnées géographiques) qui sont également inclues dans AffectiveROAD. Une métrique subjective de stress, basée sur l'annotation d'un observateur et validée a posteriori par le conducteur au vu des enregistrements vidéo acquis lors de l'expérience de conduite, complète cette base de données. Nous définissons ici la notion de stress par ce qui résume excitation, attention, charge mentale, perception de complexité de l'environnement par le conducteur.

La sélection de variables fonctionnelles dans le cas de drivedb a révélé que l'EDA mesurée au pied est l'indicateur le plus révélateur du niveau de stress du conducteur, suivi de la fréquence respiratoire. La méthode RF-RFE associée à des descripteurs non experts, conduit à des performances comparables à celles obtenues par la méthode basée sur les descripteurs sélectionnés par les experts. En analysant les données d'AffectiveROAD, la posture et l'EDA mesurée sur le poignet droit du conducteur ont émergé comme les variables les plus pertinentes. Pour les deux bases de données, l'emplacement du capteur EDA est apparu important dans la reconnaissance de ce niveau de stress.

Une analyse plus approfondie de l'EDA a par la suite été menée car ce signal a été retenu, pour les deux bases de données, parmi les variables fonctionnelles sélectionnées pour la reconnaissance du niveau de stress. Ceci est cohérent avec diverses études sur la physiologie humaine qui voient l'EDA comme un indicateur clé des émotions. Nous avons ainsi exploré le caractère fractal de ce biosignal à travers une analyse d'auto-similarité et une estimation de l'exposant de Hurst basée sur les ondelettes. L'analyse montre un comportement d'auto-similarité des enregistrements de l'EDA pour les deux bases de données, sur une large gamme d'échelles. Ceci en fait un indicateur potentiel temps réel du stress du conducteur durant une expérience de conduite réelle.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France

Title: Biosignals for Driver's Stress Level Assessment: Functional Variable Selection and Fractal Characterization.

Keywords: Driver's stress level, biosignals, real-world driving experience, functional variable selection, wavelets, fractal analysis.

Abstract: The safety and comfort in a driving task are key factors of interest to several actors (vehicle manufacturers, urban space designers, and transportation service providers), especially in a context of an increasing urbanization. It is thus crucial to assess the driver's affective state while driving, in particular his state of stress which impacts the decision making and thus driving task performance.

In this thesis, we focus on the study of stress level changes, during real-world driving, experienced in city versus highway areas.

Classical methods are based on features selected by experts, applied to physiological signals. These signals are preprocessed using specific tools for each signal, then ad-hoc features are extracted and finally a data fusion for stress-level recognition is performed.

In this work, we adapted a functional variable selection method, based on Random Forests Recursive Feature Elimination (RF-RFE). In fact, the biosignals considered as functional variables, are first decomposed using wavelet basis. The RF-RFE algorithm are then used to select groups of wavelets coefficients, corresponding to the functional variables, according to an endurance score. The final choice of the selected variables relies on this proposed score that allows to quantify the ability of a variable to be selected and this, in first ranges.

At a first stage, we analyzed physiological signals such as: Heart Rate (HR), Electromyogram (EMG), Breathing Rate (BR), and the Electrodermal Activity (EDA), related to 10 driving experiments, extracted from the open database of MIT: drivedb, carried out in Boston area. At a second stage, we have designed and conducted similar city and highway driving experiments in the greater Tunis area. The resulting database, AffectiveROAD, includes, as in drivedb, biosignals as HR, BR and EDA and additional measurement of the driver posture.

The developed prototype of the sensors network platform allowed also to gather data characterizing the vehicle internal environment (temperature, humidity, pressure, sound level, and geographical coordinates) which are included in AffectiveROAD database. A subjective stress metric, based on driver video-based validation of the observer's annotation, is included in AffectiveROAD database. We define here the term stress as the human affective state, including affect arousal, attention, mental workload, and the driver's perception of the environment complexity.

The functional variable selection, applied to drivedb, revealed that the EDA captured on foot followed by the BR, are relevant in the driver's stress level classification. The RF-RFE method along with non-expert based features offered comparable performances to those obtained by the classical method. When analyzing the AffectiveROAD data, the posture and the EDA captured on the driver's right wrist emerged as the most enduring variables. For both databases, the placement of the EDA sensor came out as an important consideration in the stress level assessment.

A deeper analysis of the EDA was carried out since its emergence as a key indicator in stress level recognition, for the two databases. This is is consistent with various human physiology studies reporting that the EDA is a key indicator of emotions. For that, we investigated the fractal properties of this biosignal using a self-similarity analysis of EDA measurements based on Hurst exponent (H) estimated using wavelet-based method. Such study shows that EDA recordings exhibits self-similar behavior for large scales, for the both databases. This proposes that it can be considered as a potential real-time indicator of stress in real-world driving experience.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France