UDRIVE

European Naturalistic Driving Study

Risk factors, crash causation and everyday driving

Deliverable 42.1

DOI: 10.26323/UDRIVE_D42.1

UDRIVE

European Naturalistic Driving Study

EUROPEAN COMMISSION

SEVENTH FRAMEWORK PROGRAMME

FP7-SST-2012.4.1-3

GA No. 314050

eUropean naturalistic Driving and Riding for Infrastructure and Vehicle safety and Environment

Deliverable No.	UDRIVE D42.1	
Deliverable Title	Risk factors, crash causation and everyday driving	
Dissemination level	Public	
Written By	Mandy Dotzauer, DLR Eric Stemmler, DLR Fabian Utesch, DLR Jonas Bärgman, SAFER Laurette Guyonvarch, LAB Jordanka Kovaceva, SAFER Hélène Tattegrain, IFSTTAR Meng Zhang, DLR Daryl Hibberd, University of Leeds Charles Fox, University of Leeds Oliver Carsten, University of Leeds	23-06-2017
Checked by	Jonas Bärgman, SAFER	23-06-2017
Approved by	Marika Hoedemaeker (TNO) *QA manager* Nicole van Nes (SWOV) *Project Coördinator*	25-06-2017 26-06-2017
Status	Final	30-06-2017

Please refer to this document as:

Dotzauer, M., Stemmler, E., Utesch, F., Bärgman, J., Guyonvarch, L., Kovaceva, J.,… Carsten, O. (2017). *Risk factors, crash causation and everyday driving.* UDRIVE Deliverable 42.1. EU FP7 Project UDRIVE Consortium. https://doi.org/10.26323/UDRIVE_D42.1

Acknowledgement:

The author(s) would like to thank Martin Winkelbauer (KFV) and Susanne Pröstl (KFV) for their valuable comments on previous drafts and Jonas Bärgman (SAFER) and Marika Hoedemaker (TNO) for performing the quality assurance on the final draft.

Disclaimer:

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 314050.

Executive Summary

The UDRIVE Naturalistic Driving Study (NDS) is the first large scale European project to observe driving behaviour directly in the field. Its goal was to identify risky driving behaviour, understand what drivers commonly do (everyday driving), why and when driver's attention is diverted from the road (distracted driving), focus on power two wheelers (PTW), pedestrians and cyclists as they are road users that are exceptionally exposed to crashes (vulnerable road users; VRU) and learn about the properties of sustainable driving behaviour (eco-driving). One of the advantages of this project is the unique opportunity to observe critical events while they occur, and with the help of the records, go back in time and investigate what may have caused them. Drivers from France (FR), Germany (GE), Poland (PL), Spain (SP), the Netherlands (NL), and the UK volunteered to participate in this study. Mechanics equipped their vehicles with cameras and sensors and thus created a fleet of 200 vehicles. This included 120 cars (FR, GE, PL, NL, and UK), 40 scooters (SP), and 40 trucks (NL). The data was collected in a box, referred to as data acquisition system (DAS), which was installed in the trunk of the vehicles. It recorded videos of seven to eight cameras, CAN, GPS and acceleration data. The data was collected between a time period of 12 to 21 months, accumulating 87 871 hours of data. This deliverable reports the results of normal and risky driving behaviour while the findings of the other research topics will be reported in respective deliverables.

The safety critical event definition section explains the procedure of creating safety critical event triggers (SCE). Actual crashes are very rare, even in a data collection of over 21 months and 200 vehicles. Thus, it is almost impossible to investigate crashes directly. Surrogate measures are used instead to identify and assess potential risk factors. Hard braking, sudden steering, and accelerations are used as surrogates for collisions. While it is reasonable to assume a connection between these surrogates and real crashes, researchers are still uncertain whether SCEs and crashes follow the same patterns. Nonetheless, SCEs are still the best option to investigate how crashes are caused. The process of finding these surrogates is the SCE trigger definition.

The objectives of the SCE trigger definition is to:

- Provide SCE-candidates that result in an unbiased selection of actual SCEs that are reasonably representative for actual crashes;
- Capture as many relevant SCEs as possible with minimum annotation resources and minimum selection bias.

Two approaches were used to define SCEs in the UDRIVE project. With the first approach, referred to as static kinematic triggers, data points exceeding the threshold values are searched for within a pre-defined temporal window. This approach relies on drivers' responses to unexpected traffic situations and extreme vehicle kinematics, such as hard braking at a certain time-to-collision (TTC). First, all situations are identified exceeding predefined thresholds. Then, a stratified random sample is drawn from all identified triggered situations. This sampled number of situations is reviewed by trained video analysts, so that only relevant SCE remain. The random sample and incorporation of sampling probabilities in estimation ensures that all results can be generalized to the total sample (i.e., to the set of events not reviewed and annotated).

The second approach is the definition of probabilistic triggers. Here, SCE triggers are defined based on the likelihood of certain events. The probability is estimated with a joint probability density distribution (JPDD) of the involved trigger parameters. Specific combinations of triggers are identified falling within a certain range of values. These ranges are chosen to reflect the probability level of critical events identifying rare and thus conspicuous events. Since this approach is new and not yet well established for NDS data, it will be evaluated. Based on the analysis, it can be determined whether one approach leads to better results compared to the other approach or whether the approaches complement each other. Therefore, the most efficient way of finding safety critical events can be determined.

Additionally to SCEs, episodes with a high relevance to road safety were investigated. On rural roads, more crashes occur than on highways and they are more severe than in cities. This makes them a highly relevant research area in respect to traffic safety. Results reported in this deliverable focus on overtaking on rural roads as an example for a complex driving manoeuvre leading to crashes. The influences of situational

factors, gender, country and type of overtaking on lateral and longitudinal acceleration were investigated. Drivers did not make the decision on whether, how and when to overtake another vehicle dependent on the presence of passengers, overtaking regulation, road curvature, vegetation, or secondary tasks. But drivers appeared to be conscious of the environment by avoiding overtaking in bends, in adverse weather conditions or when oncoming traffic was present. They also respected the overtaking regulations. Small differences between genders in longitudinal acceleration revealed that males are enthusiastic in leaving their lane while females are sporty in adjusting back to the travelling lane.

A further investigation of the link between the performance of overtaking manoeuvres and driver personality was conducted. All drivers in UDRIVE completed a pre-study suite of questionnaires that covered areas such as driver behaviour, attitudes, skill, perceived locus of control, and sensation-seeking tendencies. Drivers were categorised based on whether they ranked high or low on each subscale of these questionnaires, before comparing the tendency to overtake in each of these two groups. Drivers who performed a rural overtaking manoeuvre were more likely to self-report more speeding behaviours. However, these drivers also expressed stronger negative attitudes about both speeding and close following behaviours, suggesting a disconnection between driver attitudes and behaviours. Overtaking drivers had high sensation-seeking scores, reported more behavioural errors and violations, and more readily attributed the cause of accidents to other road users than the non-overtaking drivers. This combination of personality factors is a cause for concern because it suggests that the more risk-taking individuals may also be those who are least aware of the link between their own speeding behaviour and the likelihood of being involved in an accident.

UDRIVE also was an opportunity to collect information on driver experience insofar as it leaves untouched the real objectives of drivers and their economics of interests during a whole trip. Consequently, it presents valuable conditions for the analysis of secondary tasks (distraction), in addition to the analysis of critical driving situations whose high ecological validity is long-time claimed. 16 of the 30 French drivers were interviewed to clarify some critical sequences related to these two themes. These interviews were primed by video sequences acquired over the past 3 months (self-confrontation interviews) before complementary free-recall. The accuracy of the self-confrontation method in exposing causal chain (potentially) leading to critical events such as crashes is limited by its own standard framework: the delay between the actual and video-rendered events with its inevitable memory consequences. Such memory decays or confusions are particularly pernicious for reviving secondary task sequences. The present results suggest, however, that the method is fruitful in documenting risky situations. Indeed, risky situations have stronger emotional content than standard double task situations: they are thus more strongly stored and more easily recalled. Beyond methodological issues, the present study also provides data on three recent crashes (3) and driving situations experienced as uncomfortable. Typical revivals of risk-taking situations were mostly focused on difficulties due to infrastructural factors. Finally, the interviews contribute to better knowledge of participants' compensatory rules, whether efficient or not, when engaging in a secondary task or when taking risks.

ADAS use of cars in naturalistic conditions is addressed as well. The analysis is based on data recording as well as questionnaires filled in by participants. The dataset included data of all operation sites (OS) with unbalanced number of drivers (data available on March 15th, 2017). The context of cruise control (CC) and speed limiter (SL) use is described in terms of duration, total distance of trip as well as road type. In the studied dataset, 88% of the trips were driven without any ADAS activation. Only 2 % of the trips included speed limiter use. Cruise control and speed limiter were used in comparable conditions of road type and trip distance. 17% of the drivers don't know that their cars are equipped with CC/SL and more generally, 40% of the drivers confuse automatic high beam low beam function with automatic lighting function.

Concerning drivers' seatbelt use, the same data set was used (March 15th, 2017). Trees were used to track the most significant factors to explain wrong seatbelt behaviour (trip including driving without seat belt). The same method was used to study trip characteristic for full trips without belt. Driver country is the most significant parameter explaining the behaviour of driving for some distance without wearing the seatbelt.

The second contributing factor is gender. More male drivers do not wear a seatbelt. For whole trips without driver seatbelt use, the most significant variable is trip distance: this behaviour is more frequent for very short trips (<325m) and especially at night.

In the UDRIVE project, a comprehensive database of driver's speed choice was built under varying circumstances. One major contributing factor to traffic safety is speed choice/speeding. It was investigated how prevalent speeding behaviour is and when drivers decide to speed. The degree of speeding was determined by calculating the difference between the posted speed limit (from map data) and the driven speed (from CAN access). Two levels were distinguished, light speeding (11-15%) and heavy speeding (16-20%). The strongest factors related to speeding were severity of speeding, the posted speed limit, and time of day. Most of the speeding events occurred at low speed limits with an almost linear decline to higher levels. Light speeding was about 15-30% more frequent than heavy speeding for all speed limit levels except for 111-130 km/h. Speeding was more prevalent in France and the UK than in Poland, the Netherlands, or Germany. The most popular times of day for speeding were late night (26%), in the afternoon (28%) and in the morning (20%). Females had more light and males more strong speed violations and the speeding events in general lasted the longest in the Netherlands.

In terms of personality factors, drivers who committed at least 20 excessive speeding violations had a higher score on a composite negative driving personality traits scale. Self-reported speeding behaviour was also a good predictor of the likelihood of a driver showing a high number of excess speed occurrences. This shows that drivers are aware of their speeding behaviours to some extent, and are choosing to violate. This can possibly be explained by the high Fate subscale scores on the Traffic Locus of Control questionnaire for the frequent speeders, suggesting that drivers who speed with high frequency are not successfully linking excessive speed with increased accident likelihood. Drivers who frequently broke the speed limit also reported a high level of both aggressive and ordinary violations, and a wide range of driving 'bad behaviours' including mobile phone use, seatbelt misuse and red light violations. It would appear that speeding behaviours and other deviant driving behaviours co-occur, and that drivers 'collect' bad driving behaviours.

In addition to speeding events, close following events were identified. Here as well, several circumstances were investigated to understand when and why drivers follow too closely. Close following was defined as following another vehicle with a time headway (THW) under 1.5 s for longer than 1.5 s. Results show that drivers older than 50 years old comprised the largest group of close followers. The most close following events (~75%) were observed within posted speed limits of 31-50 km/h. Low speed limits, such as in urban areas appear to be predestined for observing close following events.

In order to analyse potential events that could create risky situations, we are interested in hard braking. To have a reference in terms of exposure, non-hard braking events were created and served as a baseline. Difference were observed for some factors (i.e., age, gender, speed limit, time of day, rain state, ADAS use, and infrastructure) being more often present in hard braking than in non-hard braking events between the operational sites (Table 6-56). German data show a very significant effect for gender and the categories of speed limit, a significant effect for age, time of day and type of infrastructure but show no effect of rainfall and ADAS use. French data show a significant effect for time of day and rainfall and a very significant effect on all other factors. Dutch data show a very significant effect on all factors except for the gender. The rain effect was not tested due to the lack of Dutch rain data. Polish data show no effect on gender, but a very significant effect for time of ADAS use and a very significant effect on all other factors. English data show no effect of the rain and a very significant effect on all other factors. The ANOVA tests on the times to reach the position of the previous vehicle (TIV) showed that they were significantly different between the brakes 1 and 3 for each operational site (Table 6-53) and for each type of infrastructure in which braking was performed (Figure 6-54).

Overall, the results presented in this deliverable give insight into natural driving behaviour on a level unprecedented in Europe. Many research questions were addressed within the project UDRIVE, other research questions could not be addressed during the project period. The data sets used for the analyses of the deliverable represent a subset of all data collected as at some point, data needed to be frozen in order to do the analyses. While analysing the data, more data was collected, pre-processed, and uploaded to the database. This means that 100% of the data will only be available after the project, inviting to replicate (and extend) the performed analysis with the whole dataset and investigating new research questions with it. Data will be made available after the project. Details about data after the project will be published at www.udrive.eu.

Table of contents

UDRIVE
European Naturalistic Driving Study

1 Introduction

1.1 The UDRIVE project

UDRIVE started 01.10.2012 and ended 30.06.2017 and had a total budget of 10.6 M of which 8 M were funded from the European Commission under Framework Program 7 (FP7) and the remainder was in-kind contributions of the partners. UDRIVE is a large naturalistic driving study in Europe. Nineteen partners across Europe have come together and defined research questions, developed data acquisition, collected and managed data, and finally, performed a first analysis on the UDRIVE dataset with respect to driver/rider behaviour related to traffic safety and the environment (i.e., with respect to emissions).

In UDRIVE, a data acquisition system (DAS) was developed collecting up to eight video streams and acquiring controller area network (CAN) data. The data was complemented by accelerometer and angular rate sensors, GPS, and MobilEye (www.mobileye.com) data. Mobileye is a collision warning system based on video image processing. For UDRIVE, most of the internal parameters were provided (e.g. to determine time headway, time to collision). This acquisition system was used across all operation sites (OS) and vehicle types. Car (i.e., Renault Megane and Clio) data was collected in the Netherlands, Poland, France, Germany, and UK, truck (i.e., Volvo) data in the Netherlands, and power two wheelers (i.e., Piaggio Liberty 125 Delivery) in Spain. Participant recruitment, organisation of the installation and de-installation of the DAS as well as all other everyday operations were handled by an operation site, except for Great Britain: two operation sites were installed.

Within UDRIVE, an online monitoring tool (OMT) was developed in order to monitor data collection, processing, and data management. Collected data was sent to one of the three local data centres (LDC, i.e., France, Germany, or Sweden). There, raw data was pre-processed and if necessary videos anonymized before being sent to the central data centre (CDC) in Sweden. At the CDC data was uploaded to a MySQL database and by means of SALSA (analysis tool developed in UDRIVE), researchers were able to develop, share, and apply Matlab algorithms for calculations and also to annotate videos. Further analysis of the data was done with SAS, SPSS, R, Phyton, MS Excel, or Matlab).

Four main research topics are addressed in UDRIVE:

1. everyday and risky driving
2. distraction and inattention
3. vulnerable road users
4. eco-driving

Those research areas were determined and prioritized together with stakeholders from governments, industry and academia. Within each research topic, relevant sub-topics were identified. In this report, results of everyday and risky driving are reported. Results of the other main topics are reported in other deliverables, respectively.

1.2 This work in the UDRIVE project

This report of the UDRIVE project covers the results of the analyses of everyday driving behaviour in terms of safe and unsafe behaviours. The report is part of work package 4.2 and presents results of the work done within the work package. It is the main deliverable as well as dissemination of WP4.2 of the UDRIVE project. Particularly, the following results are presented in the report:

1. Development and implementation of triggers for safety critical events (SCE) and a method for baseline selection
2. Results of everyday driving, overall and for different driver groups

3. Results of driving behaviour in specific situations (i.e. vehicle overtaking on rural roads) with respect to safe and unsafe behaviours

Study results using the self-confrontation technique

1.3 Background and aim/(research) questions of the work at hand

Data supporting a deep and comprehensive understanding of everyday and risky driving behaviour is difficult to gather. Driving simulator studies can be used to investigate certain aspects of driving such as investigating the effect of personality on driving performance and road traffic safety. This investigation is punctual and does not allow to conclude on everyday driving as everyday driving and driving performance is moderated by many different factors such as environmental factors (e.g., road type, traffic density)and situational factors (e.g. time of day, presence of passengers). Asking drivers about their everyday driving behaviour is also not the most reliable source as drivers may answer socially desirable. Naturalistic driving studies (NDS) provide the opportunity to address questions around everyday and risky driving that cannot be addressed any other way as the naturalistic driving setting is as natural as it can be achieved.

In order to address the research goal, different tasks were accomplished within the work package.

1.3.1 Additional data extraction and processing

In this task, additional data extraction processing were defined and performed. Focus laid on the definition of triggers of relevant SCEs and the selection of baseline events. In addition to using static trigger thresholds, a probabilistic approach has also been developed and applied. As visual inspection of potential safety-critical events is cost and time consuming, the aim of the task was to develop triggers minimizing false-positives as well as evaluating to what extent the methods complement each other. In addition, a method for selecting baseline events needed to be developed. A baseline represents an episode without a critical event. It is utilized to identify relevant factors that might have caused the critical event. Therefore, the groundwork for future risk calculations was done within the task.

1.3.2 Overtaking manoeuvres of motorized traffic on rural roads

Overtaking manoeuvres on rural roads are dangerous and one of the leading causes of fatal crashes. While motorway driving and urban driving have been investigated thoroughly in the past years, research on rural road driving has not been the focus of research. The NDS data collected gives us the opportunity to analyse overtaking manoeuvres to better understand the mechanisms behind initiating and executing overtaking manoeuvres. In addition to analysing environmental factors such as weather conditions, situational factors, such as overtaking when prohibited, and driving performance parameters such as lateral acceleration or time-to-collision at the time initiating a lane change, regional, age, and gender comparisons will also complement the analysis.

1.3.3 Self-confrontation

Self-confrontation is a method used to investigate risky events in more detail. It included showing videos of critical situations and risky situations (i.e., secondary tasks engagement) to drivers who experienced them and follow-up with interviews. This technique allows for acquiring drivers' recollection of events in certain near-crash situations and during secondary task engagements. The goal was to gain more insight into: a) how and when secondary tasks or risky sequences occur, b) what was the sequence of events having led to them, and c) what was the driver's role in this sequence (active or reactive). In addition, the applicability of this technique was assessed in the framework of NDS.

1.3.4 Descriptive analysis of everyday driving

Within this task various aspects of normal up to risky (but non-SCE-related) driving behaviour was analysed. This includes aspects, such as speed choice, gap acceptance, hard braking, ADAS use (specifically cruise control and speed limiter systems) and seatbelt use. The NDS administered within the EU project UDRIVE offers the opportunity to investigate the prevalence and risk of disregarding safety precautions. For example, speeding and close following are major contributing factors to crashes. In UDRIVE, we have the possibility to investigate such driving behaviour in more detail. Questions answered are, for example, whether speeding and close following is a deliberate act of driving (i.e. do some drivers engage in those behaviours more frequently than others). Along those lines, personality and cultural back ground may also influence speed choice and close following behaviour.

Hard braking is also investigated, since hard braking may be a driver's reaction to a safety critical situation. Such events may indicate a reaction to a suddenly developing situation. It can be a hazard to the surrounding traffic as well. Therefore, it can be investigated how drivers react in order to control the situation. Analyses of hard braking events include the examination of effects of, for example, age, gender, and cultural background on hard braking.

Wearing a seat belt is mandatory throughout Europe. Nonetheless, seatbelt usage is not thoroughly understood; therefore, another aim within the everyday driving analysis is to investigate and better understand seatbelt usage. In addition to gathering knowledge about how often drivers use their seatbelt, cultural differences are also of interest. Driver characteristics and environmental factors affecting seatbelt use are also investigated.

Throughout the last decades, more and more advanced driver assistant systems have been developed and implemented into vehicles. Those serve also the purpose to increase road traffic safety by either preventing crashes or minimizing the consequences (e.g., level of injury severity) of crashes. Aim of the research is to investigate the use of driver assistance systems; in particular, when do drivers use ADAS (e.g., urban, rural, motorway etc.)? In addition, to analysing questionnaires on what drivers know about their ADAS administered before data collection started the use of cruise control and speed limiter was analysed.

1.4 The contents and structure of the report

This report covers the main results of the tasks completed in WP42. In Chapter 2, the groundwork for future risk calculations is described. Two approaches of defining potential safety critical events are shown in detail. In addition, the method for selecting baselines (i.e., episodes without critical events) is shown. In Chapter 3, results of the investigation of overtaking manoeuvres on rural roads are described. In Chapter 4, the technique self-confrontation is introduced as means for analysing critical situations together with the driver experiencing the situation, its suitability in the context of NDS is assessed, and factors contributing to the critical situations are identified. In Chapter 4, the thoroughly analysis of everyday driving is presented. This includes results on speeding, close following, hard breaking, seatbelt and ADAS use. In the following chapters, findings are discussed and concluded and recommendations summarised.

2 SCE trigger definition

Note: based on strong advice from the UDRIVE Advisory Board, only very limited analyses of safety critical events (SCEs) were made, and no risk estimations comparing SCEs to baselines were conducted. A baseline represents an episode without a critical event. It is utilized to identify relevant factors that might have caused the critical event. The decision to not proceed with SCE and risk assessment (that need random and matched baselines in the analysis) came relatively late in the project. Therefore, a large part of the work in preparation of baseline selection and risk calculation methods had already been conducted at the time of the decision. This section is documenting part of this preparation process, even if it was not utilized in UDRIVE. Also note that parts of this section has been developed as working documents in UDRIVE and are in part written in present or future tense.

2.1 Detection methods

One of the core types of analysis performed in naturalistic driving studies (NDS) is event based analysis (EBA). The basic principal of event based analysis is to find surrogate events for crashes, since actual crashes occur rarely. The surrogate events are driving situation without an actual crash, but the situation still unfolds in a way that may be used as an indicator of crash risk. The first step of EBA is identifying shorter driving situations called safety critical events. SCEs are typically in order of 10 seconds during which the crash risk is judged to be higher. The second step is to analyse why these events occurred and whether they are safety-critical.

NDS has been mostly recognized for such analysis, with analyses of the 100-car naturalistic driving data (NDD) such as those documented in Klauer et al. (2014), and analysis of the US SHRP2 NDD in Victor et al. (2015) are examples. Some researchers, for example Kidd and McCartt (2015) and Knipling (2015), articulate critics to some forms of EBA. Bärgman (2016) further discussed the drawbacks and benefits of EBA, which is not in the scope of this deliverable.

EBA considers surrogate events, called safety critical events (SCE), for crashes, since actual crashes are rare. One of the challenges of NDS is identifying SCEs from a large data set in a cost-effective manner. In order to find potential safety critical events, static parameter thresholds have been commonly used (Najm & Stearns et.al. 2006; Guo & Hankey 2009; Simons-Morton et al., 2011). Nonetheless, the method is time-consuming as potential SCEs need to be visually (manually) inspected in order to be classified in terms of relevance for safety. A new approach probabilistic trigger, complementing traditional static trigger methods like EBA, has been developed aiming at lowering the time needed for manual validation of SCEs. These are the probabilistic trigger thresholds.

Both approaches, static and probabilistic trigger thresholds, have been applied to the UDRIVE data aiming at:

- Provide SCE-candidates that result in an unbiased selection of actual SCEs that are reasonably representative for actual crashes;
- Capture as many relevant SCEs as possible with minimum annotation resources, with minimum bias in selection.

In the following sections, it will be described how the approaches were implemented. In addition, the method used to compare and evaluate the methods will be described.

2.2 Static trigger thresholds

The approach in this section, referred to as static kinematic triggers, searches data points that exceed the threshold values within a pre-defined temporal window. This approach relies on the driver response to unexpected traffic situation and the extreme vehicle kinematics such as hard braking at a certain time to collision. That is, SCE-triggers describe some mechanism in kinematics of the ego-vehicle or interaction kinematics between the ego vehicle and other road-users or infrastructure components. Examples of the former are thresholds on longitudinal or lateral acceleration, possibly complemented with thresholds on derivatives of the same measures.

The choice of threshold values is a compromise between exhaustiveness and false events. Lower trigger values capture the maximum number of potential events, but the trade-off is a higher chance of false-positive events, non-conflict events, and less severe conflicts. Similarly, a higher trigger values result in a higher percentage of valid events but generates some omissions. The applied triggers in UDRIVE take into account the experience from previous projects and are combination of the triggers used in SHRP2 (Hankey et al., 2016) and euroFOT (Malta et al., 2012). Each SCE-trigger is associated with a "severity measure" to indicate how severe the candidate is. This is a simple one-value item, e.g. peak deceleration, minimum time to collision (TTC), see Table 2-1, for description. The SCE-triggers are implemented here as segments with 30s before and 10s after the thresholds are exceeded.

Table 2-1: SCE trigger definition based on SHRP2 and euroFOT definitions

Trigger type	Threshold	Segment name	Severity measure	Description
Low time-to-collision	TTC < 1.5 s	SEG_CRE_fromTTC	Min TTC	Time-to-collision is indicative of SCEs by proximity. That is, time-to-collision is a measure that, when small, indicates small safety margins. Drivers have certain safety margins they are comfortable with (comfort zones, e.g. Summala (2007)). If the time-to-collision is less than what drivers normally accept, it may be indicative to a SCE. Also note that using time-to-collision with a threshold is a SCE-trigger that does not require driver actions. That is, most kinematic SCE-triggers (accelerations, yaw rate, brake pedal movements etc.) require that the driver actually saw the hazard. Time-to-collision does not have this limitation.
Hard braking by the driver	Longitudinal acceleration < -3 | (brake pressure > 70 & longitudinal acceleration < 0)	SEG_CRE_HardBrake	Max longitudinal acceleration	The braking reaction is identified from the brake pedal activity or through high longitudinal acceleration as a proxy metric. It is assumed that an unusual hard braking reaction could indicate a SCE.
Hard steering reaction by the driver	Abs(steering wheel jerk) > 500 deg/s^2	SEG_CRE_SWJerk	Max steering wheel jerk	The steering activity can be directly measured at the steering wheel. An unusually hard steering action of the driver may indicate a swerving manoeuvre which may have been initiated to avoid a crash.
High longitudinal acceleration of the vehicle	Longitudinal acceleration < -5.5 m/s^2	SEG_CRE_fromLongAcc	Max longitudinal acceleration	An extreme longitudinal acceleration is likely indicative of a crash.
High	Longitudinal jerk	SEG_CRE_L	Max	Vehicle jerk is the derivative of

vehicle longitudinal jerk	< -1g/s	ongJerk	Longitudinal jerk	acceleration. When jerk is high it means that the rate of acceleration increases fast.
High brake pedal depression jerk	Brake pressure change rate > 80 bar/s	SEG_CRE_BrakePressureChangeRate	Max brake pressure change rate	Brake pedal jerk is the derivative of the brake pedal acceleration. Brake pressure is used as proxy for the brake pedal depression, since brake pedal activity is available as dichotomous measure (pressed\not pressed).
High lateral acceleration of the vehicle	Abs(lateral acceleration) > 0.75g	SEG_CRE_LatAcc	Max lateral acceleration	An unusually high lateral acceleration may indicate a swerving manoeuvre which may have been initiated to avoid a crash.
Quick changes in yaw rate	Changes in yaw rate of ±4 deg/s within 3 s	SEG_CRE_YawRate	Max absolute yaw rate	Situations where the yaw rate oscillates from neutral to a value outside ±4deg/s, to ±4deg/s, and back to neutral within 3 seconds.
ME triggers	NaN	SEG_PCW	Min TTC	Mobile Eye triggers refer to Pedestrian Detection and Collision Warning (PCW) system. In other words, the system alerts drivers of a forthcoming collision with a pedestrian. (Note: These warnings are not visible to the driver and are used as detected by the system and analysed in detail in D4.4, section 8.)

2.2.1 Analysis steps

First, all situations in the data where the given kinematic trigger is exceeded are identified. Then a stratified random sample of 1200 segments is drawn from all identified triggered situations. This sampled number of situations is reviewed by trained video analysts, so that only relevant SCEs remain.

For each batch of data (approximately every 2 months), the following steps are applied:

1. Prepare derived measures that are used in SCE triggers (e.g. TTC, THW, jerk, road type, etc.) (see Deliverable 4.1 for description of the processing of derived measures.)
2. Candidate events
    a. Automatically select candidate events (kinematic triggers criteria, see Table 2-1).
    b. Randomly select subset of candidate events for preliminary review (SCE-candidates are segments in data around a time (SCE-trigger-time) identified by applying SCE-triggers to time-series data).
    c. Review the subset of candidate events and classify them into two categories: relevant and not relevant SCE.

Annotate the relevant SCE according to the annotation codebook (see Deliverable 4.1).

2.2.2 Results

Total number of records used for SCE trigger detection (database status May 16th 2017) is 124011, distance driven is 1.4 million kilometres and duration is 31114 hours. The details per country are shown in Table 2-2. The number of the automatically generated SCE-triggers per country is shown in Table 2-3.

Table 2-2: Number of records, distance, and duration per country.

	Number of records	Duration (hours)	Distance (km)	Number of drivers
Germany	14464	3459	162986	26
France	44174	11391	517901	43
Netherlands	11643	3309	203159	31
Poland	15708	4060	179649	36
UK	38022	8935	387929	50

Table 2-3: SCE triggers per country.

	Germany	France	The Netherlands	Poland	UK	Total
SEG_CRE_fromTTC	46	152	32	113	74	417
SEG_CRE_fromLongAcc	438	1870	382	1093	3980	7763
SEG_CRE_HardBrake	10766	137553	24444	41924	206531	421218
SEG_CRE_SWJerk	5686	18743	3955	5796	20665	54845
SEG_CRE_YawRate	2243	25067	5030	3304	7388	43032
SEG_CRE_BrakePressureChangeRate	2764	11277	3688	4500	6019	28248
SEG_CRE_LatAcc	33352	539582	46573	80087	300102	999696
Total	55295	734244	84104	136817	544759	1555219

For each of the segments the following attributes are available (for more information on the attributes see D44), written in the format [name of attribute]-[description of attribute]:

- Record ID – unique record identification
- Driver ID – unique driver identification
- DAS Config – vehicle type
- Operation site – country
- Begin time – start time of the event (relative to the start of the record)
- End time – end time of the event (relative to the start of the record)
- Max/min – maximum/minimum of the severity measure in the segment
- Road user presence – proximity of other road user (yes or no)
- Road user lateral distance – lateral distance to the road user
- Road user longitudinal distance - longitudinal distance to the road user
- Absolute distance – distance to the road user

- VRU (vulnerable road user) pedestrian – presence of pedestrian
- VRU cyclist – presence of cyclist
- VRU PTW – presence of PTW
- Trip duration – duration of the trip in seconds
- Trip distance – distance of the trip in meters
- Locality type – urban or rural road (map way area type)
- Speed limit – Map speed limit
- Locality intersection – occurring in intersection (map intersection and map way type)
- Subject vehicle speed – subject vehicle speed at the trigger onset.

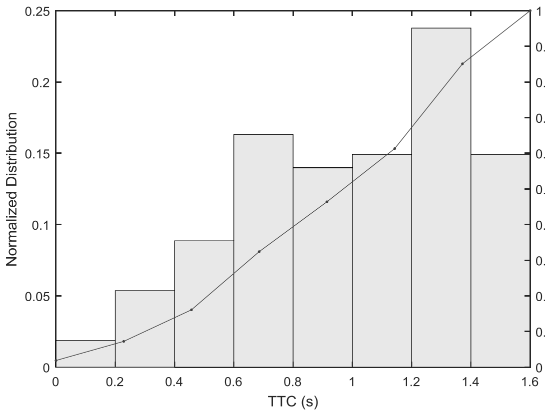


Figure 2-1: Normalized distribution (left axes) and cumulative distribution (right axes) of the minimum TTC for candidate events SEG_CRE_fromTTC

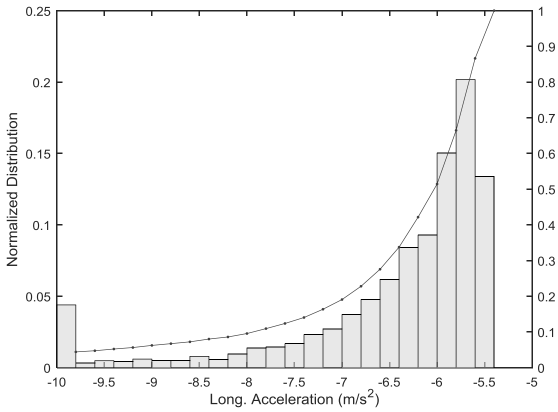


Figure 2-2: Normalized distribution (left axes) and cumulative distribution (right axes) of the minimum longitudinal acceleration for candidate events SEG_CRE_fromLongAcc

The cumulative distribution of minimum TTC (Figure 2-1) , shows that half of the events have TTC less than one second. The cumulative distribution of the minimum longitudinal acceleration (Figure 2-2) shows that half of the events have less than 6.3m/s^2 deceleration. These triggers are indicative for small safety margins however review of the video might show that the small TTC refers to low speeds and the hard braking refers to situations with speed bumps, which might not be critical. Therefore, first the presence of an SCE should be verified and then the detection rate of this method can be calculated.

The candidates were created by randomly selecting 200 candidates per type of SCE-trigger. Each candidate is extended with 20s before and 10s after around the threshold value is exceeded. The candidates are saved in segment “SEG_SCE_Candidates_B1”, where B1 refers to data of batch one, see Deliverable 41.1. To date there are 1200 generated for batch 1.

2.2.3 Summary

Even though the analysis steps 1, 2a, and 2b are implemented, the last steps 2c and 2d (review and annotation of SCEs) are not carried out in the project, since the baseline segments were prioritized for annotation and further analysis.

However, one of the advantages of the static triggers is having same triggers for the whole dataset. Another advantage is that the triggers can be associated with a scenario prior to the trigger scanning, for example, low TTC is associated with rear-end scenarios, while lateral acceleration rate may play a more important role in run-off-road scenarios. On the other hand, the static triggers do not take into account that the thresholds may vary by driver, vehicle, roadway, or environmental characteristics.

2.3 Probabilistic trigger thresholds

A key challenge of NDS is determining whether potential SCEs are safety-relevant requiring time-consuming and expensive manual video coding. In order to lower the time needed for manual coding, a new approach of identifying SCE is presented on the UDRIVE data. In this approach, SCE triggers are defined based on the likelihood of certain events countering the inherited high number of false-alarms of static thresholds (Simons-Morton et al., 2011). The approach provides a functional relationship between the threshold parameters (e.g. longitudinal acceleration and situational parameters such as speed).

Step 1: Estimation of the joint probability density distribution

This function is based on the joint probability density distribution (JPDD) of the involved trigger parameters. The JPDD is a multivariate probability distribution that defines the probability that a specific combination of values of its arguments falls in any particular range of values of those values. To give an example for the given context: the JPDD of speed and longitudinal acceleration (a bivariate probability distribution) specifies the probability of the situation that the ego vehicle drives at speed between 100-105 km/h and executes a deceleration between 2.0-2.5 m/s^2. If one knows this functional relationship one could estimate the probability of predefined events, as these are usually given in terms of static threshold values which specify a range of values deemed as potentially critical. Our primary motivation to use JPDD in this context however is to go the other way around: Setting the probability level of critical events and derive the corresponding ranges of values within the critical domains of the probability landscape. The critical domains are easily identified as one is usually looking for a certain type of event such as for instance hard braking manoeuvres. So if one wants to detect such events that occur with a specific probability one would naturally cumulatively sum up probability densities from the most negative values of acceleration towards the also more likely less negative regime of values, since hard braking corresponds to negative acceleration and the more negative the acceleration the harder the breaking. The first step is to estimate the JPDD from a representative sample. Because the linearly binned kernel density estimate approach (Wand, 1994) is computationally fast and allows for estimating two dimensional distributions, it was used for estimating the JPDD (Deng, 2011).

Step 2: Computation of the trigger function

The trigger function is computed as the percentile line of the estimated JPDD by computing the cumulative probability function. This can be calculated from the algorithmically estimated JPDD by sorting all pairs according to their corresponding probability density value in ascending order and then calculate the cumulative sum on this sequence. The line of a specific cumulated value is then determined as the set of points in the feature space that correspond to this probability value (the sum of probability density values). This curve represents a dynamically changing threshold of trigger parameters depending on a chosen set of situational parameters such as velocity. This approach allows for determining events, operationalized as points in the feature space occurring within a pre-defined range, with a specific probability. The following R code block demonstrates the use of this technique (Listing 2-1).

```
1    subdt <- alldt[, c("mSpeedCAN", "mLongitudinalAcceleration"),
                    with = FALSE]
2    subdt <- na.omit(subdt)
3    gs <- c(51L, 51L)
4    bw <- c(dpik(subdt[, mSpeedCAN]), dpik(subdt[, mLongitudinalAcceleration]))
5    dens <- bkde2D(x = subdt,
6    gridsize = gs,
7    bandwidth = bw,
8    range.x = list(c(0, 200), c(-7, 3)))
9
10   dx <- diff(dens[["x1"]][1:2])
11   dy <- diff(dens[["x2"]][1:2])
12   sz <- sort(dens[["fhat"]])
13   cs <- cumsum(sz) * dx * dy
14   level005 <- approx(cs, sz, xout = 0.05)$y
15   level0025 <- approx(cs, sz, xout = 0.02)$y
16   level001 <- approx(cs, sz, xout = 0.01)$y
17   level0005 <- approx(cs, sz, xout = 0.005)$y
18   level00025 <- approx(cs, sz, xout = 0.0025)$y
19   names(dens) <- c("x", "y", "z")
20   clevels <- c(level005,
21            level0025,
22            level001,
23            level0005,
24            level00025)
25
26   dt <- as.data.table(melt(dens$z))
27   dt[, mSpeedCAN := dens$x[Var1]]
28   dt[, mLongitudinalAcceleration := dens$y[Var2]]
```

Listing 2-1: Demonstration of how to calculate contour lines and the use of the linearity binned kernel density estimator package. The code is written in the R programming language. In lines 1 and 2, data retrieval is performed. The variable *gs* stores the number of grid points for each dimension (longitudinal acceleration and speed in this case; line 3). *dpik()* is an optimization algorithm called direct plug-in method for selecting the optimal value of the bandwidth parameter for kernel density estimation. The *bkde2D()*-function call performs the kernel density estimation. It returns the estimated density values and the corresponding locations at the grid points. In lines 10-13, the estimate of the density function is integrated. In lines 14-18, *approx()* performs a look-up of grid-points whose cumulated probability density corresponds closest to the given value of probability percentile stated by the *xout*, needed for visualizing the contour line.

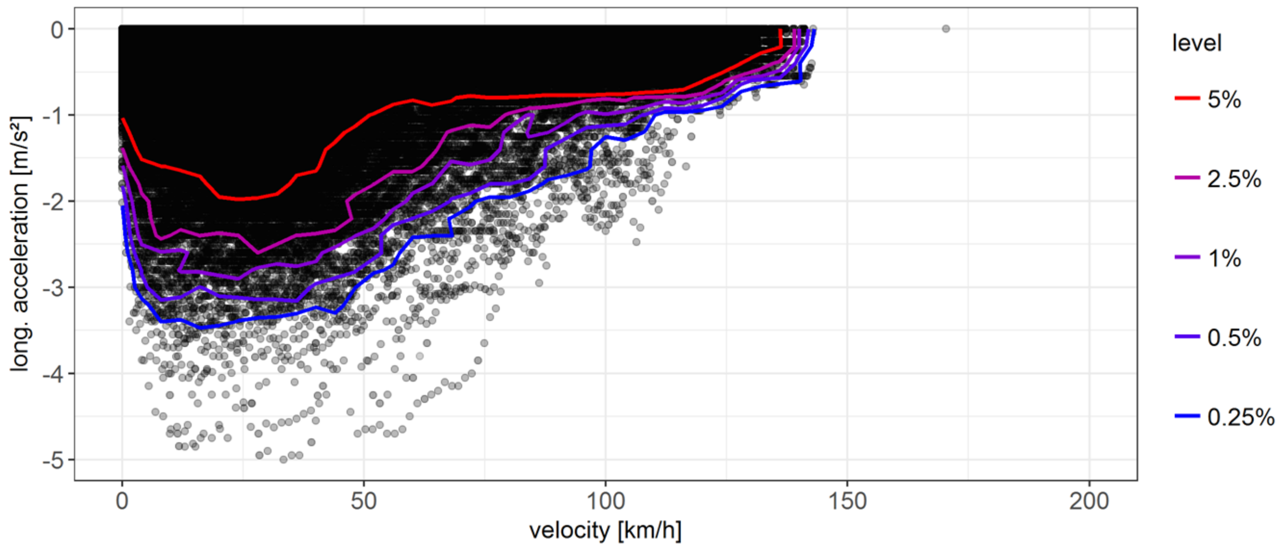


Figure 2-3: Scatter plot of longitudinal acceleration against velocity based on randomly drawn sample of trips. Coloured lines refer to different contour lines of equal cumulated probability levels from 0.25% up to 5%. These contour lines are used to approximate the trigger threshold functions.

Step 3: Determination of a polynomial fit

A polynomial fit can help determining an easy-to-implement representation of the trigger function. This approach has two advantages: Specific values of a static trigger threshold do not need to be chosen, but instead values are the result of a general parameter setting. Furthermore, it provides a dynamically changing threshold for different scenarios, such as a lower threshold on longitudinal deceleration values for driving on a highway (higher speed levels) compared to urban driving (lower speed levels). See the following figures for example combinations that have been implemented and calculated based on the given data of the UDRIVE project (Figure 2-4, Figure 2-5, Figure 2-6, Figure 2-7).

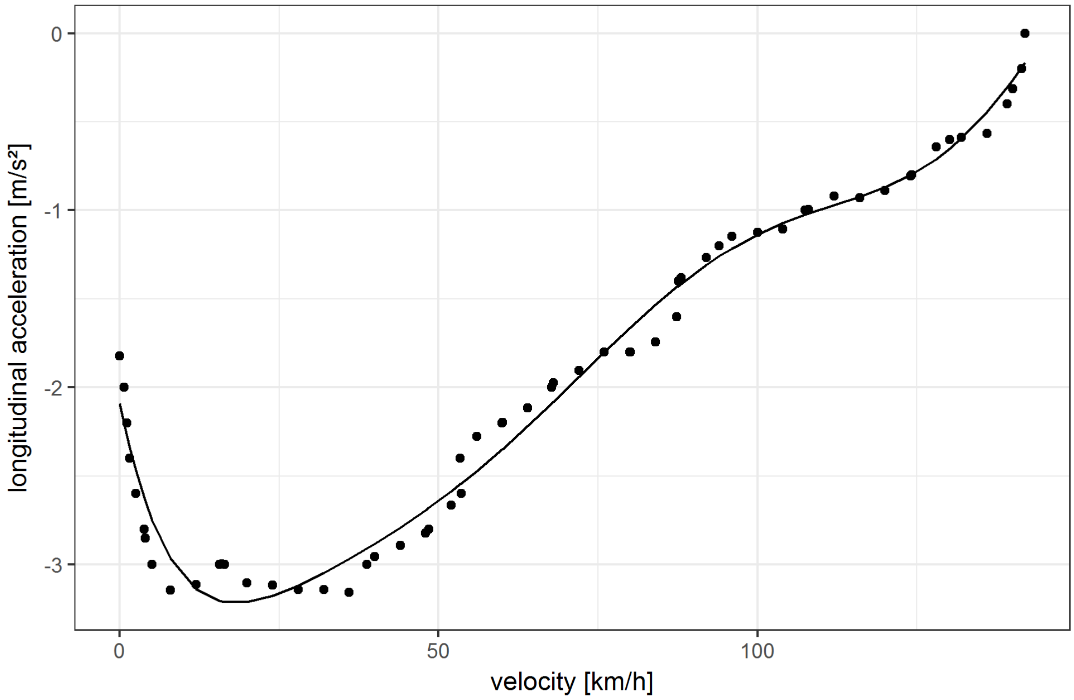


Figure 2-4: Plot of the 7th degree polynomial fitting of the points corresponding closest to the 0.5th percentile contour line of the bivariate joint probability distribution of longitudinal acceleration and velocity. The fitting was performed in terms of terms of minimization of mean square error. The number of coefficients was chosen based on a hierarchical regression, adding a polynomial degree after another until the difference in variance becomes not significant. The same procedure has been applied for each trigger type.

For implementation, the resulting values for the coefficients from the least-squared fitting process where used. Table 2-4 displays the coefficients representing the trigger threshold function for longitudinal acceleration and velocity.

Table 2-4: Coefficients resulting from the polynomial fitting of the contour line of the bivariate joint probability distribution of longitudinal acceleration and velocity.

Coefficient	Parameter
-2.09	intercept
-0.172	x
0.01	x^2
-0.00	x^3
4.20e-06	x^4
-3.68e-08	x^5
1.65e-10	x^6
-2.94e-13	x^7

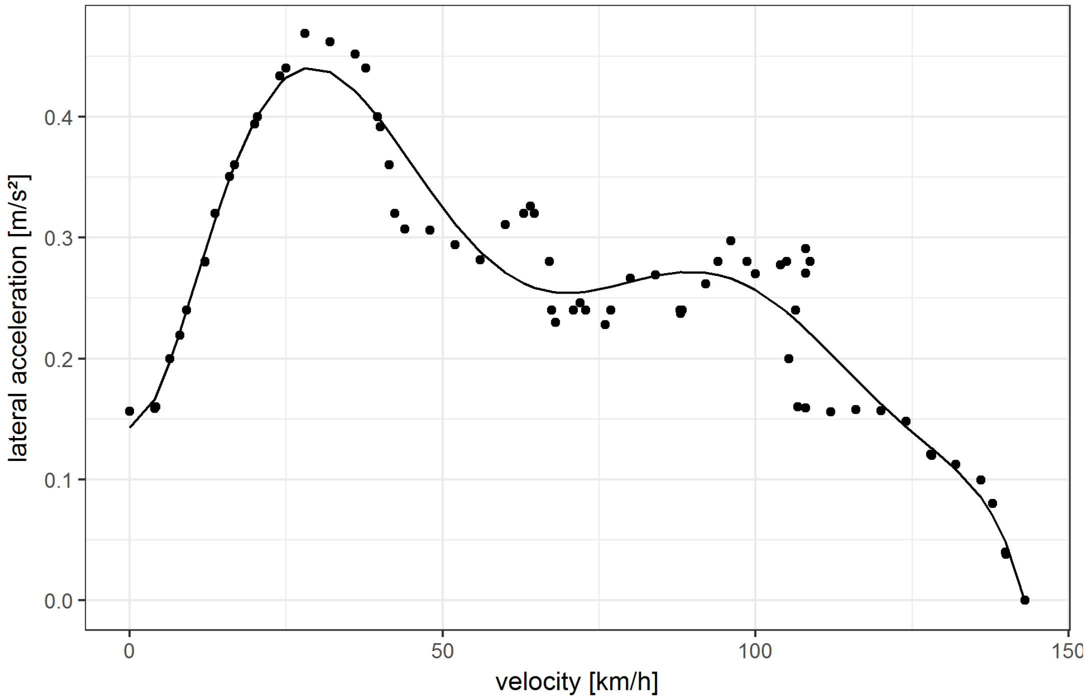


Figure 2-5: Plot of the 7th degree polynomial fitting of the points corresponding closest to the 0.5th percentile contour line of the bivariate joint probability distribution of lateral acceleration and velocity.

Table 2-5: Coefficients resulting from the polynomial fitting of the contour line of the bivariate joint probability distribution of lateral acceleration and velocity.

Coefficient	Parameter
0.14	intercept
-0.001	x
0.002	x^2
-9.91e-05	x^3
1.98e-06	x^4
-1.94e-08	x^5
9.38e-11	x^6
-1.78e-13	x^7

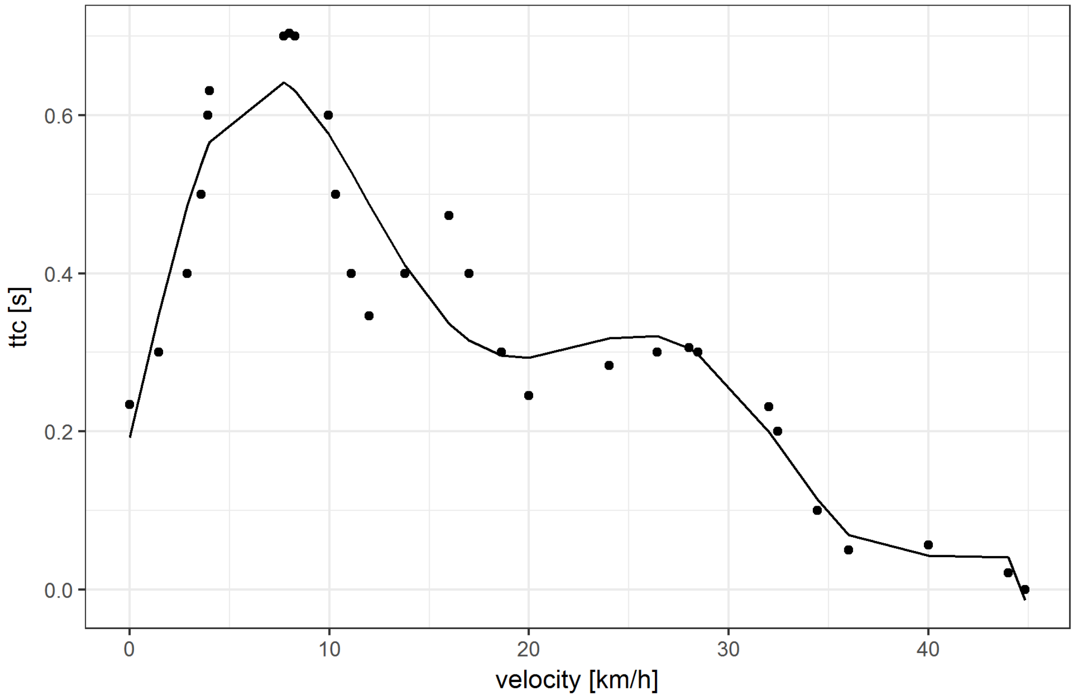


Figure 2-6 Plot of the 7th degree polynomial fitting of the points corresponding closest to the 0.5th percentile contour line of the bivariate joint probability distribution of time-to-collision and velocity.

Table 2-6 Coefficients resulting from the polynomial fitting of the contour line of the bivariate joint probability distribution of time-to-collision and velocity.

Coefficient	Parameter
0.19	intercept
0.11	x
0.005	x^2
-0.003	x^3
0.0002	x^4
-8.97e-06	x^5
1.54e-07	x^6
-1.01e-09	x^7

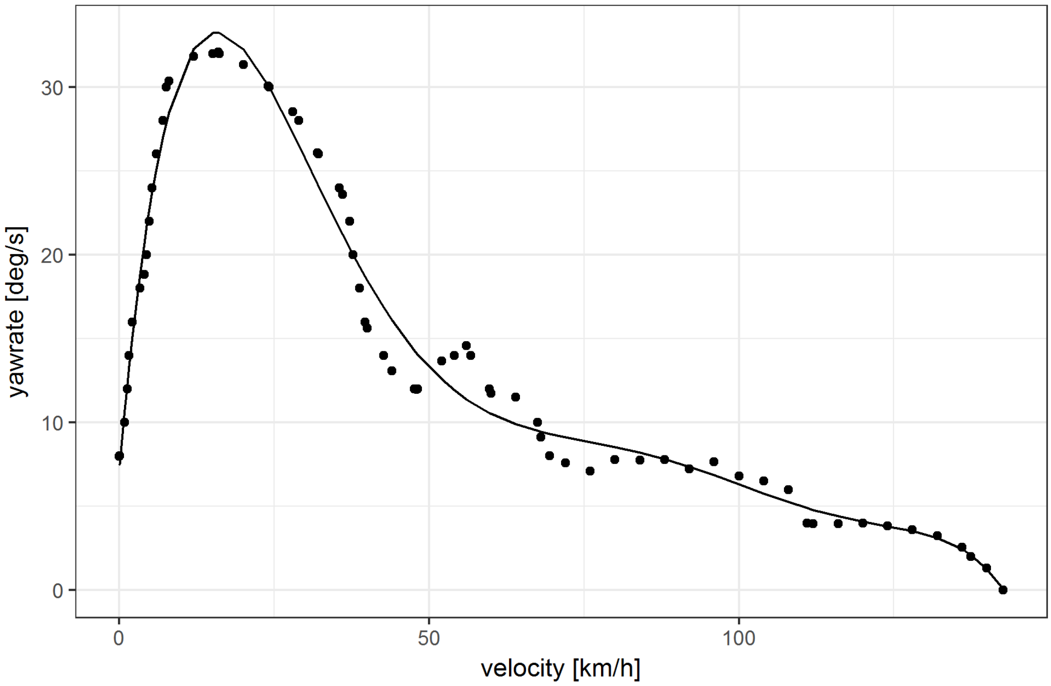


Figure 2-7 Plot of the 7th degree polynomial fitting of the points corresponding closest to the 0.5th percentile contour line of the bivariate joint probability distribution of yaw rate and velocity.

Table 2-7 Coefficients resulting from the polynomial fitting of the contour line of the bivariate joint probability distribution of yaw rate and velocity.

Coefficient	Parameter
7.46	intercept
4.12	x
-0.22	x^2
0.005	x^3
-5.41e-05	x^4
3.14e-07	x^5
-8.59e-10	x^6
7.53e-13	x^7

We expect that this framework closes the gap between discrete forms of SCE triggers and helps lowering the number of false-alarm detections and therefore relieves the burden of manual inspection.

2.4 Evaluation of selection methods

Static trigger thresholds have traditionally been used in order to find potential safety critical events in naturalistic driving data. As described above, this approach has its advantages and disadvantages. One of the major disadvantages is the high number of false alarms produced by static triggers. Each of them has to be visually inspection (videos) to verify the safety-critical event safety relevance. To improve the detection of potential SCE candidates another approach has been developed within the project UDRIVE – the detection of potential SCEs based on the probability of their occurrence. In addition, triggers have been defined dynamically, as criticality changes with speed.

This new approach needs to be evaluated. To do so, static and probabilistic trigger definition will be applied to the data. The number of identified events per trigger type and detection approach will be reported. Some of the events will be commonly triggered between the two methods. Those events will be visually inspected in the presence of SCEs verified. It is hypothesized that if an event is commonly triggered, the likelihood of an SCE is high. In addition to verifying these events, depending on the number of triggered events, either a subset (randomly selected) or all triggered events will be visually inspected and the presence of an SCE verified. Therefore, the detection rate of each method can be calculated. Additionally, it can be determined what type of SCEs can be found with each method and whether the approaches complement each other (i.e. finding different types of safety critical events).

Based on the analysis, it can be determined whether one approach leads to better results compared to other approach or whether the approaches complement each other. Therefore, the most efficient way of finding safety critical events can be determined.

Please note: the evaluation results of the approach are not part of the deliverable, but will be presented at the NDRS in June 2017 in The Hague.

The limitation of the new method lays primarily in potential selection bias. When SCE-triggers are discussed, the issue of individualization or regionalisation of triggers often comes up. Just because a specific driver or a driver in a specific region may have more aggressive driving style, should the trigger criteria be different between driver/regions?, or, more importantly (not in the scope of this deliverable), should what constitutes a valid safety critical (relevant) event be different between individuals/regions? The conclusion in such discussion is often than the triggers can be different, but what constitutes a safety relevant event should be as objective as possible, and not differ between regions. However, even if it is generally accepted to have regionalized triggers – which in an extension can be the probabilistic triggers used here – there is a question of generalizability when using such dynamic triggers. That is, when a probabilistic approach is used, where some combinations of triggers or trigger thresholds generate more critical events than other combinations (e.g., static triggers), analysis using for example odds ratios (ref) can result in biased results. That is, generalization may suffer and specific (but common) types of SCEs may be overrepresented in the analysis, potentially producing biases. However, even the traditional static trigger approach suffers from this issue if multiple (separate) triggers are used (the number of potential SCEs are then decided by thresholds and the annotation efforts per trigger) – although possibly not to the same degree as dynamic triggers. In both the use of static and dynamic triggers, care needs to be taken to issues of generalizability (Knipling, 2015; Bärgman, 2016).

3 SCE baseline selection

This section is based on a work document developed during the UDRIVE project and describes the process of estimating the number of SCEs in UDRIVE, as well as outlines the rationale between the selection criteria planned for random baseline (control) selection (baselines to be used in comparison with SCEs).

Selecting baselines/controls in NDS can be achieved in different ways. Nonetheless, any sampling process is aimed at enabling generalization to the population the samples are to represent. In terms of risk calculation, a matched and random baseline selection is proposed for UDRIVE and will be described. In addition to deciding on a baseline sampling strategy for the risk analysis, strategies for the analysis of everyday driving and secondary task have also been developed and are described in Section 5, respectively Deliverable 43.1 (section 4).

3.1 Estimated number of SCEs

A relatively low number of safety critical events was expected in UDRIVE. In part, because the amount of data available at the central data centre (CDC) turned out to be less than initially projected, and in part due to the rarity of SCEs overall. Nonetheless, in UDRIVE, the criteria of a near-crash would not be modified in order to find more safety critical events. As the level of severity diminishes, the safety-critical event becomes more similar to the baseline; therefore, the power of any analysis comparing baseline and SCE would have been reduced.

Based on what is expected from the UDRIVE database, descriptive statistics of SHRP2 and the Swedish part of the euroFOT project, an estimate of the number of SCEs in UDRIVE has been derived (Note: US drivers have a higher mileage. This may result in an even lower number of SCEs than in SHRP2).

Statistics of SHRP2 (at time of writing, May, 2017):

1. Descriptive statistics
   - 3247 drivers
   - 5 400 000 trips (with an additional 1 200 000 events with unconsented drivers)
   - 80 000 000 km driven
   - 3958 participant-years
   - 1465 crashes
   - 2710 near-crashes
   - "balanced sample baseline": 20000
   - Additional baseline: 12500
2. Estimates of number of crashes per km and participant per year
   - One (1) crash or near crash per 19 000 km
   - One (1) crash or near-crash per participant year (0.089 per participant month).
   - One (1) crash per 55 000km
   - One (1) near-crash per 29000km

Statistics of euroFOT (SCEs selection criteria yielded SCEs mainly from highway driving (in Sweden)

1. Descriptive statistics
   - ~2 700 000 km (average speed 58km/h)
   - 500 SCEs
     - 487 crash relevant events
     - 12 near-crashes
     - 1 crash

2. Estimates of number of crashes per km
   - One (1) SCE per 5400 km.
   - One near-crash per 225 000 km (much less than SHRP2, but then mainly highway driving).

Expected statistic of the UDRIVE data:

1. Descriptive statistics
   - 9 months of data from cars and trucks planned to be used for analysis.
   - An average trip with a car is 20min and with a truck 60min.
   - The total number of trips is estimated to be ~200 000
   - The estimated number of hours to be driven is 100 000 h (~50km/h average speed)
   - The estimated number of kilometres to be driven is 5 000 000 km.
2. Estimate of number of crashes per km (derived from SHRP2 statistics)
   1. Estimated number of UDRIVE crashes: 90
   2. Estimated number of UDRIVE near-crashes: 172

Estimated total number of SCEs in UDRIVE: 262

3.2 Estimated number of matched and random baselines

Depending on the effect size expected in the odds ratio calculation, the number of baseline needed can be estimated. Figure 3-1 below, is from a recent SHRP2 study (Victor et al., 2015) and outlines the relation between the number of baselines and safety critical event in terms of odds ration calculations. According to the figure, a 1:2 ratio of SCEs and baselines provides the best value for analysis using crude odds ratios. In fact, the number of SCEs is fixed while the number of baseline depends on budget, so this ratio ultimately depends on the budget for annotating baselines. Nevertheless, including more than two baselines in the analysis will likely not lead to a significant gain. Therefore, with an estimated 262 SCEs, 524 random and matched baseline epochs need to be selected, respectively.

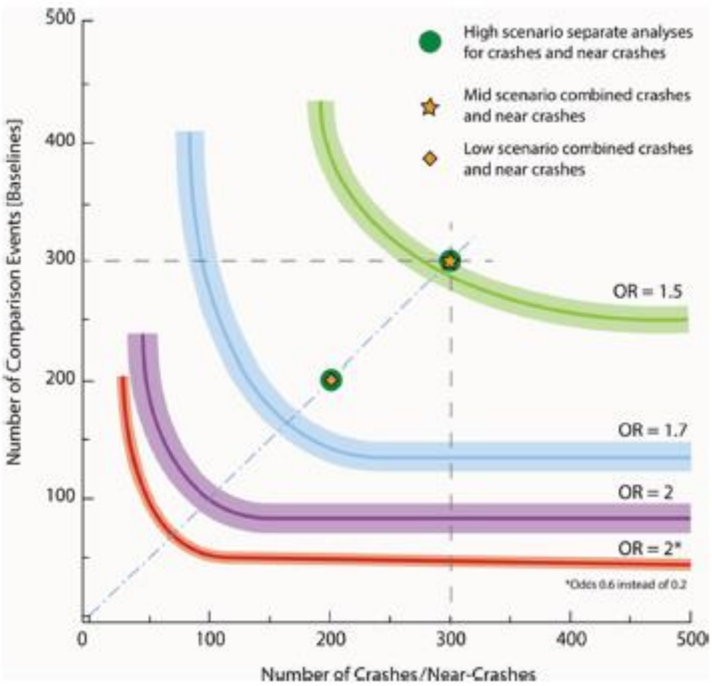


Figure 3-1: Estimates of needed controls/baselines for given odds ratios, when performing odds ratio calculation. Figure reprinted from Victor et al. (2015)

3.3 Random baseline selection procedure

Risk can be calculated and stated either in terms of hours or kilometres driven. Most analysis of naturalistic driving data (NDD) is likely to be sampled in terms of hours (US SHRP2). However, within the crash safety domain, the more common way to present safety numbers is per kilometre. Both strategies have their advantages and disadvantages and therefore reasons why sampling should be done either on kilometre or time.

Four reasons have been identified for sampling on kilometre:

1. It is traditionally used in crash safety, and in UDRIVE we want to be comparable (on the other hand, it limits comparability with previous naturalistic data analysis).
2. Higher speed is more dangerous in terms of injuries and fatalities, although not in number of crashes sampling per hour oversamples low speeds, where risk is lower.
3. Per definition, sampling on distance times means that when the host vehicle stands still, it is basically not part of the selection.
4. Drivers typically make decision in time, rather than distance, why several analyses investigating driver behaviour may be biased if distance is used as the sampling base.

However, the benefits of sampling on time include:

1. Results are more easily compared to.
2. It is, even with sampling on time, to weight the baselines with speed to simulate distance (kilometre) based selection.
3. Highway driving will be up-weighted compared to city driving.

Within the project, it has been decided to sample based on time. In addition to the mentioned advantages and disadvantages, segments are selected based on time, analysis will also be done based on time; therefore, sampling on time is in line with the overall approach.

3.4 Selection of random baselines

Within the project UDRIVE random baselines would have been selected in the following way: The time series are stacked across all trips in one vector, doing a cumulative sum across all. The total amount of hours to be driven in the entire project (e.g.100 000 hours) is estimated. Sampling using a Poisson approach, with the expected value of random baselines to be, for example, 524 across the 100 000 estimated hours would be considered. Instead of creating random numbers, exponential (independent) sampling with the number of hours and number of estimated random baselines across the entire data collection would have been used. For each batch, pick the baselines based on the exponentials. This will provide a truly random selection across time (not caring about driverID or other attributes). After some time, reconsidering the sampling (given budget) would be appropriate. The same procedure can be applied to the data collected until the next reconsideration, and new Poisson/exponential sampling for the remaining data administered. It is crucial to not aim for more totals than one can “afford” at the end.

This process can be performed several times. With this approach, standard methods of analysis can be used, not having to consider batch/stratified sampling.

3.5 Selection of matched baselines

Each SCE was planned to be matched to two matched baselines/controls. For example, 524 baselines would be matched to 262 SCEs (if the estimate turned out correctly). In a first step, SCEs need to be separated based on well-defined scenarios (e.g. according to Wassim et al., 2007). For each scenario a set of matching criteria need to be presented. For example, for rear-end scenarios (Wassim et al., 2007, scenarios 22-26), the match may be on the following:

- Driver ID
- Lead-vehicle present (e.g. based on TTC or THW)
- > 10 min before the SCE in the same trip (if possible). Alternatively, find the same driver, same (approximate) position a previous day.
- Time of day (light level)
- Host vehicle speed (e.g. $\pm$5km/h)
- Considerations to use host-vehicle braking as a matching criterion have also been discussed. This would then require the use of the smart camera MobilEye

Once matched, each matching criterion should be attached as attributes for traceability and to facilitate filtering. Doing so allows analysing the effects/sensitivity of the matching criterion on the results.

4 Overtaking other motorized vehicles on rural roads

4.1 Introduction: Driving on rural roads

Crash statistics show a 43% (i.e. a drop from 24.845 to 14.143) decrease on road fatalities between the years 2005 and 2014 in Europe (European Road Safety Observatory, 2016). This trend is in line with the overall decrease in road fatalities. Even though, road users feel safer on rural roads compared to motorways and urban areas (AXA, 2012; DEKRA, 2013), crashes on rural roads account for more than half of all fatal crashes. And even though the percentage of fatalities decreased, the fatality rate remained stable over the years. Figure 4-1 shows an overview of the percentage of fatalities on rural roads in Germany, France, Poland, the Netherlands and the UK between 2005 and 2014 (European Road Safety Observatory, 2016).

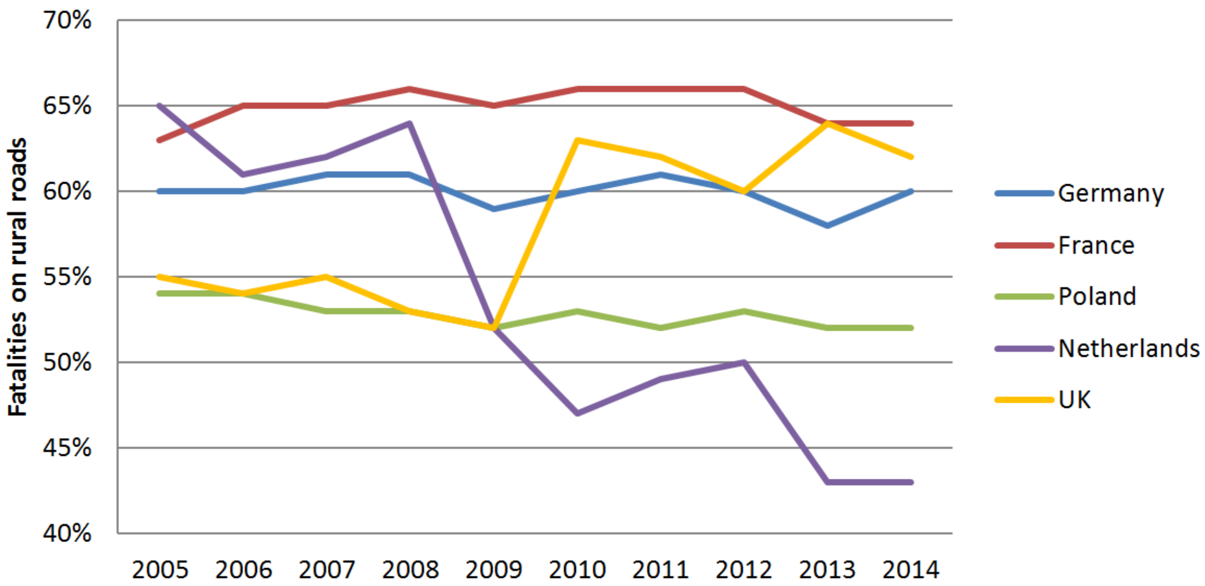


Figure 4-1: Overview of percentage of fatalities on rural roads in Germany, France, Poland, the Netherlands, and the UK; 2005-2014 (Source: European Road Safety Observatory, 2016)

While the percentage of fatalities remained stable for Germany, France, and Poland, the table also shows a 22% decrease of fatalities on rural roads in the Netherlands, while fatalities increased by 7% in the UK. Throughout Europe, 38% of all rural road fatalities fall into the age group of 18 to 49 years of age. In addition, about 78% of all fatally crashed are males. According to the European crash statistics, 57% of all fatalities on rural roads involve cars/taxis. Drivers and passengers make up 95% of the fatally crashed persons on rural roads. With regards to lighting condition, more than half of all fatalities occur during daylight, while 18% occur during darkness on roads without street lights. In 70% of the all fatal crashes on rural roads, the road surface was dry (European Road Safety Observatory, 2016).

Rural roads differ significantly from motorways and roads in urban areas. Rural roads are used by many different types of road users. Motorized traffic includes passenger vehicles, trucks, busses, motorcycle, and agricultural machinery; while, cyclists, pedestrians, and even horseback riders are also present on rural roads. Road users' intention to travel on rural roads ranges from commuting to recreational activities. No other road type serves a purpose to so many roads users. In addition, speed differences as prominent as they are on rural roads are not observed on motorways or in urban areas, especially not on combination with oncoming traffic and takeover manoeuvres. Besides speeding, obstacles at the roadside, blind bends, junctions, risky takeover manoeuvres are a main cause for fatal crashes on rural roads (DEKRA, 2013). For example, in Germany, in 2011, 751 of all 2.441 fatalities (31%) on rural roads occurred with oncoming traffic (Statistisches Bundesamt, 2012) indicating a collision followed by an unsuccessful takeover manoeuvre.

Not only rate drivers overtaking manoeuvres as dangerous (Harris, 1988), overtaking manoeuvres are rare occasions. They are hardly taught and learned in driving school. Therefore, during driving lessons, drivers cannot develop overtaking strategies, but need to acquire those after finishing driving school (Wilson & Best, 1982). In addition, those manoeuvres are executed rarely, so that processes are not as automatized as for other driving manoeuvres. Overtaking is a complex manoeuvre that requires drivers to perform two consecutive lane changes without endangering road users, including them. Before initiating a safe takeover manoeuvre, drivers have to assess the situation. This includes observing oncoming traffic and assessing whether the gap between oneself and the oncoming traffic is large enough (i.e. assessing the distance to oncoming traffic and the velocity of oneself, the lead vehicle, and the oncoming vehicle accurately). Drivers tend to underestimate the distance needed to safely overtake another road user (Clarke et al., 1998). For example, overtaking a truck traveling at a speed of 60km/h, will take about 10 seconds will require an overtaking distance of approximately 260 meters (DEKRA, 2013). In addition, a clear vision double the distance needed to overtake a road user is required to ensure that either no oncoming traffic is present or far enough away to complete the takeover manoeuvre safely. The further away oncoming traffic is, the more difficult it is to judge the speed of the oncoming traffic accurately as the change in visual size is very small (Farber et al., 1967; Hills, 1980). Moreover, traffic coming from behind also needs to be observed ensuring that no other vehicle already initiated an overtaking manoeuvre. When misjudging the presence or distance of oncoming traffic, drivers need to take corrective actions and either accelerate in order to complete the takeover manoeuvre as fast as possible or disrupt the manoeuvre and return to their lane. As rural roads are often narrower and curvier than other road types, it makes corrective actions as well as takeover manoeuvres more difficult to complete and to keep control over the vehicle, especially when exceeding the speed limit or having a high lateral acceleration. Narrow roads and high lateral acceleration might lead to losing control over the vehicle. Therefore, take-over manoeuvres on rural roads may be very risky and need to be analysed thoroughly and better understood.

In UDRIVE, naturalistic driving data of vehicles was collected throughout Europe. These data are used to analyse overtaking manoeuvres to better understand the mechanisms behind initiating and executing overtaking manoeuvres. In addition to analysing environmental factors, such as weather conditions, situational factors, such as overtaking when prohibited, and driving performance parameters such as lateral acceleration or time-to-collision at the time initiating a lane change, regional, age, and gender comparisons will also be made complementing the analysis.

4.2 Method

4.2.1 Participants

Based on the data query of March 31st, 2017, 2351 potential overtaking situations were extracted. Out of these potential events, 531 were visually inspected. Altogether 55 takeover manoeuvres were validated. Ten of these takeover manoeuvres were excluded from analysis as motorized traffic overtook VRUs (i.e. cyclists and a horse) in those situations. Out of the remaining 45 situations (13 OS-FR, 3 OS-DE, 17 OS-PL, and 12 OS-UK), 31 takeover manoeuvres were completed by male drivers, while the remaining 14 were executed by female drivers. The subsample of drivers was comprised of 16 drivers (11 male and 5 female).

4.2.2 Procedure identifying overtaking manoeuvres

The procedure to identify overtaking manoeuvres is based on its chosen definition – A takeover manoeuvre consists two consecutive lane changes in opposite travelling lane encompassing the passage of one or more vehicles travelling in the same direction on the neighbouring lane. In UDRIVE, the analysis of overtaking manoeuvres is limited to overtaking other motorized traffic on a two-lane rural road.

In order to find potential overtaking manoeuvres on rural roads, a detection algorithm needed to be developed. The following parameter where used to identify takeover manoeuvres:

1. Lateral acceleration of the ego vehicle,
2. Distance to left and right lane (as measured by the MobilEye smart camera)
3. Relative distance of vehicles within front view of the MobilEye smart camera.

In a first step, the algorithm identified single lane change. The position of the crossing points when the host vehicles lateral position crossed the middle of the lane was determined. The manoeuvre of the first lane change started at the last instance when the lateral acceleration value was below the 5th percentile of the zero-level lateral acceleration. Values up to the 95th percentile of a 15s time window ending at the cross point were considered as zero-level acceleration. Likewise, the lane change ended at the first instance when the lateral acceleration decreased below the 5th percentile of the zero-level lateral acceleration after the cross point. This again was calculated based on the 15s time window that starts at the cross point.

In a next step, with the help of the procedure, it was determined whether to consecutive lane changes took place within a 15 second time window travelling in the opposite lane. If this condition was true, the event was marked as possible overtaking candidate. If one or more vehicles were overtaken, the starting point of the segment was re-calculated. The MobilEye smart camera can detect lead vehicles and calculate the distance to those vehicles. The starting point of the overtaking manoeuvre was the point when the distance to the vehicle to be overtaken was 50 m.

Overtaking candidates identified by the algorithm were marked. The videos associated with the marked segment needed to be reviewed and the overtaking manoeuvre validated. Successfully identified overtaking manoeuvres were marked, further annotated and included in the analysis (except for situations when cyclists were overtaken).

4.2.3 Variables

The analysis involved both categorical and interval scaled variables. Most of the categorical variables were determined by manual video annotation, interval-scaled variables were derived from the CAN Bus and MobilEye smart camera data.

The overtaking manoeuvre was partitioned into the following 6 phases (see Figure 4-2: Approaching and following the lead vehicle (phase 0), initiating first lane change (phase 1), ending first lane change (phase 2), overtaking/driving in opposing lane (phase 3), initiating second lane change (phase 4), ending second lane change (phase 5).

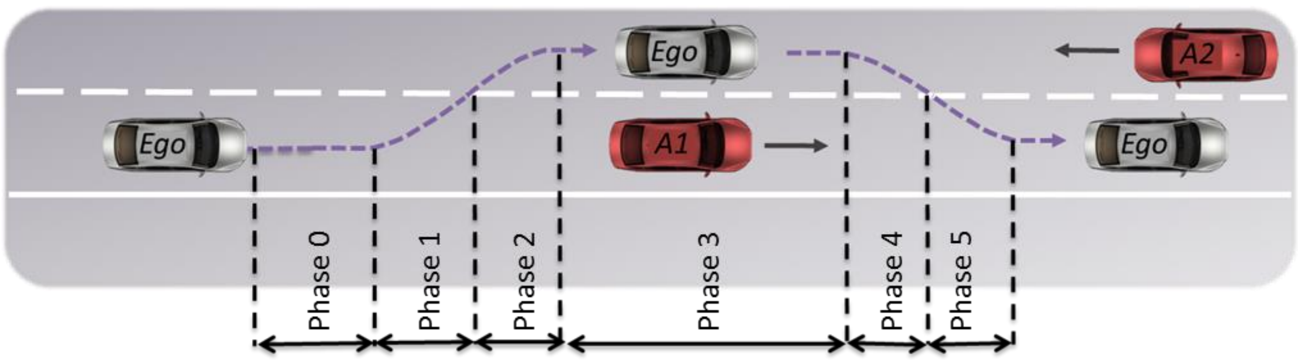


Figure 4-2: Schematic representation of an overtaking manoeuvre showing the division into several phases.

Interval scaled performance measures where determined for of those phases are presented in Table 4-1– duration, average and maximum lateral and longitudinal acceleration, time-to-collision with respect to the lead vehicle (TTC_lead) and time-to-collision with respect to the oncoming vehicle (TTC_oncoming), if

UDRIVE
European Naturalistic Driving Study

present. In addition to the variables listed below, the time since the last oncoming vehicle passed before the overtaking manoeuvre was calculated.

Table 4-1: Overview of driving performance parameters per overtaking phase

	Phase 0	Phase 1	Phase 2	Phase 3	Phase 4	Phase 5
Mean lat. acceleration		X	X		X	X
Max lat. acceleration		X	X		X	X
Duration	X	X	X	X	X	X
Mean long. acceleration	X	X	X	X	X	X
Max long. acceleration	X	X	X	X	X	X
Mean TTC_lead		X	X			
Min TTC_lead		x	x			
Mean TTC_oncoming					X	X
Min TTC_oncoming					X	X

Categorical variables were extracted both algorithmically and by video annotation. Algorithmically, information about the weather (raining or not, time of day), age and gender of the driver was extracted. The overtaking situations were further enriched with video annotations. The type of road user (e.g. car, bus, truck, bicycle) overtaken was annotated. The number of road users overtaken was also recorded. In addition, it was determined whether the overtaking manoeuvre was permitted (dotted lane markings vs. solid lane markings), whether passengers were present in the host vehicle, whether oncoming traffic and a bend was in sight at the beginning of the takeover manoeuvre, whether the takeover manoeuvre took place on an alley. Moreover, the type of takeover manoeuvre (i.e. normal, flying, normal and piggy backing, or flying and piggy backing) was determined and it was noticed whether the driver engaged in a secondary task while overtaking the lead vehicle(s) (Table 4-2).

Table 4-2: Definition of types of overtaking manoeuvres

Type of overtaking	Description
Normal	Following another vehicle, then overtaking it
Flying	Approaching and overtaking the other vehicle with vehicle speed relatively constant
Normal and Piggy backing	Normal with following another vehicle during overtaking
Flying and Piggy backing	Flying with following another vehicle during overtaking

4.2.4 Analysis

While this Deliverable was written, the huge amount of gathered data is still being analysed. The presented results reflect the data that was present at the time of analysis, 21st of April 2017 (about 60% of the data).

In order to analyse takeover manoeuvres on rural roads, data will be viewed and descriptive results will be shown. Those descriptive results will be presented by type of manoeuvre (normal, flying, normal-piggy backing, flying-piggy backing) and will provide information mainly on categorical variables describing the situation and the overtaking manoeuvre. In a second step chi-squared tests will be calculated for categorical variables. The chi-squared test is used to identify if two groups with categorical data are independent from each other.

Differences between *type of manoeuvre*, *gender*, and *country* will be investigated. Results will be reported on situational factors, factors describing the takeover manoeuvre and driving performance parameters. With regard to driving performance parameters, TTC values to oncoming traffic have been excluded from the analysis as those values were not recorded properly. Only four instances with oncoming traffic have been annotated while MobilEye detected oncoming traffic 28 times. In addition, duration of the takeover manoeuvre and the individual phases can also not be included in the analysis due to computational errors.

4.3 Results

4.3.1 Descriptive analysis of overtaking manoeuvres

Descriptive analysis of gender and country reveals that most overtaking manoeuvres were completed in Poland followed by France and the UK (see Table 4-3). For Germany only three overtaking manoeuvres were found, while none were found for the Netherlands. These results reflect the status of data as of March 1st, 2017. More data is now available, but not included in the analysis.

Looking at the total number of overtaking manoeuvres, more male than female driver overtook other motorized traffic on rural roads. The same trend is found when looking at the countries individually, more male than female drivers overtook vehicles on rural roads, except in England. Here most of the overtaking manoeuvres were driven by females (see Table 4-3).

Table 4-3: Overview of gender and country.

Type of manoeuvre	N		Country							
			FR		DE		PL		UK	
	M	F	M	F	M	F	M	F	M	F
Normal	14	11	7	1	3	0	4	3	0	7
Flying	13	2	3	0	0	0	8	0	2	2
Normal-piggy backing	1	0	1	0	0	0	0	0	0	0
Flying-piggy backing	3	1	1	0	0	0	2	0	0	1
Total	**31**	**14**	**12**	**1**	**3**	**0**	**14**	**3**	**2**	**10**

Table 4-4 provides information about situational factors. Results show that most of the takeover manoeuvres were normal takeover manoeuvres. One third of the manoeuvres were flying manoeuvres. In addition, a manoeuvre being a combination of normal and piggy backing was observed as well as a combination of flying and piggy backing in four instances. In this table, descriptive analysis also provides some more information on situational factor accompanying takeover manoeuvres. In about one third of the occasions, no passengers were present in the vehicle. Almost all, except for five overtaking manoeuvres, were completed when overtaking was permitted. In two instances, driver engaged in a secondary task while completing the manoeuvre. Those secondary tasks were either singing or interacting with passengers. In addition, most overtaking manoeuvres were initiated and completed with a clear view. In five instances was a bend coming up. More overtaking manoeuvres took place on rural roads without trees along the road.

Table 4-4: Overview of frequencies of type of manoeuvre in total and by situational factors.

Type of manoeuvre	N	Passenger present		Overtaking prohibited		Secondary task		Bend		Alley	
		Yes	No	Yes	No	Yes	No	Yes	No	Yes	No
Normal	25	11	14	3	22	1	24	3	33	10	15
Flying	15	5	10	0	15	1	14	2	13	4	11
Normal-piggy backing	1	0	1	1	1	0	1	0	1	0	1
Flying-piggy backing	4	1	3	1	3	0	4	0	4	2	2
Total	**45**	**17**	**28**	**5**	**41**	**2**	**43**	**5**	**51**	**16**	**29**

The descriptive analysis in Table 4-5: Overview of total number of overtaken vehicles and number of oncoming traffic provides more information of the actual takeover situation. On average, drivers overtook only one vehicle at the time and in about 90% of the occasions no oncoming traffic was observed.

Table 4-5: Overview of total number of overtaken vehicles and number of oncoming traffic.

Type of manoeuvre	N	Total number of overtaken vehicles		Oncoming traffic	
		M	Sum	Yes	No
Normal	25	1.1	27	3	22
Flying	15	1.1	17	0	15
Normal-piggy backing	1	2	2	0	1
Flying-piggy backing	4	1	4	1	3
Total/Average	45	1.1	50	4	41

Most of the time (45%), car-like vehicles were overtaken (Figure 4-3). Second and third are non-motorized traffic participants such as bicycles (18%, not included in the analysis) and agriculture machinery such as tractors (13%). Trucks were overtaken in 9% of the cases and powered two wheelers even more rarely (2%).

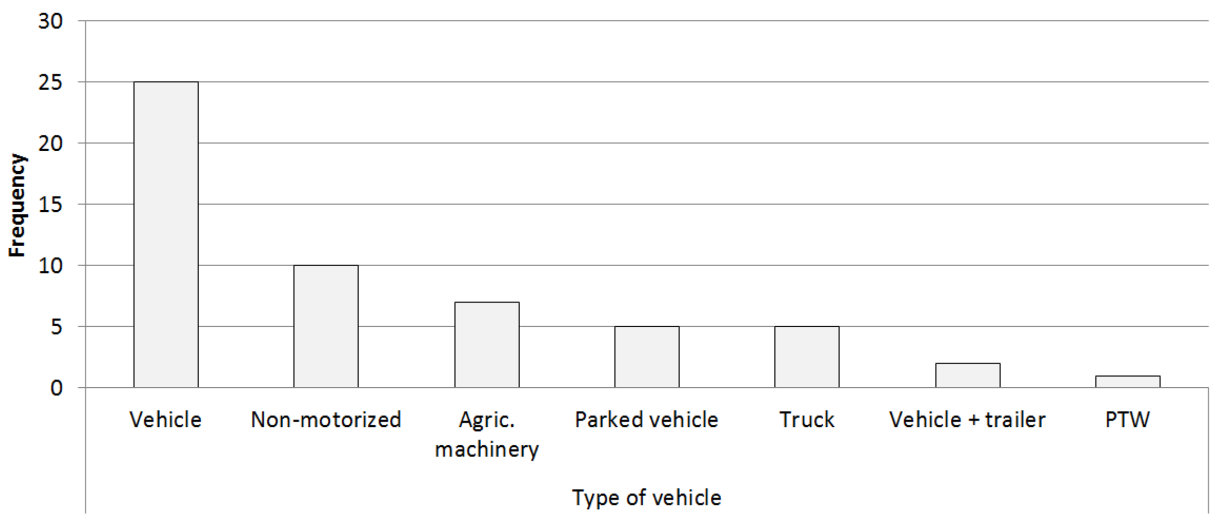


Figure 4-3: Overview of overtaken vehicle types

4.3.2 Analysis of takeover manoeuvre by type of manoeuvre

Chi-squared (Pearson, 1900) analyses were calculated for the situational factors *passenger present, overtaking prohibited secondary task, bend, and alley.* No differences between types of manoeuvre were found for those variables. In addition, further chi-squared tests were calculated examining the difference between *oncoming traffic,* and *type of manoeuvre.* No differences were revealed. A one-way ANOVA was used to test for differences between *total number of overtaken vehicles,* and *type of manoeuvre*. Here again, no significant differences were observed (for detailed statistics, please see Annex A.1.1)

In addition to the comparisons mentioned above, comparisons of driving performance parameters were also made. Mean values of mean lateral acceleration and maximum lateral acceleration were derived for phases 1, 2, 4, 5 (see Section 4.2.3) as phases 0 and 3 should have a lateral acceleration of zero. No differences in mean and maximum lateral acceleration could be observed for any of the phases. Mean and maximum longitudinal acceleration were also derived from the data. Values of longitudinal acceleration were calculated for all six phases of the takeover manoeuvre. A significant difference was observed in mean longitudinal acceleration of phase 1 $F(3,40) = 2.82$, $p = .049$. As shown in Figure 4-4, mean longitudinal acceleration is lowest in flying and highest in normal takeover manoeuvres. No other mean or maximum values differ between type of manoeuvre (detailed statistics can be found in Annex A.1.1).

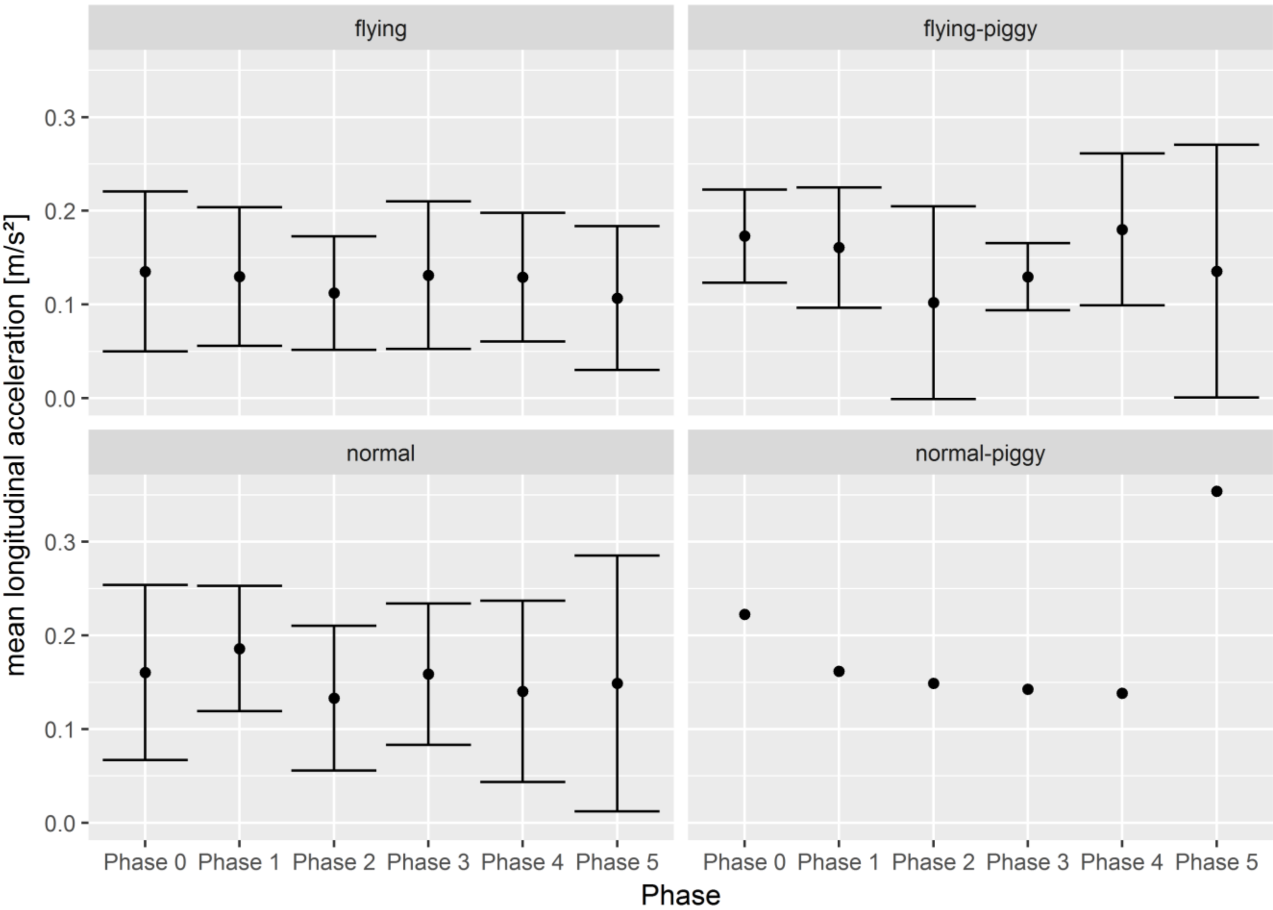


Figure 4-4: Mean and standard deviation of mean longitudinal acceleration during the phases of the overtaking manoeuvre. Only a case of normal-piggy backing was observed, hence no standard deviation was plotted.

4.3.3 Analysis of takeover manoeuvre by gender

Differences between genders were also investigated. Chi-squared analyses did not reveal significant differences between male and female drivers with regard to *passenger present, overtaking prohibited, secondary task, bend, and alley.*

Examining differences between genders with regard to variables used to describe the actual takeover situation shows significant differences in the total number of overtaken vehicles, $F(1,43) = 4.75$, $p = .038$. On average, male drivers overtook fewer vehicles (M= 1.03) than female drivers (M= 1.29)

Analysing driving performance parameters, differences between gender and mean lateral acceleration led to significant differences in Phases 1, $F(1,42) = 28.3$, $p > .001$, phase 2, $F(1,42) = 16.7$, $p > .001$, phase 4, $F(1,42) = 7.4$, $p = .009$, and phase 5, $F(1,42) = 8.5$, $p = .006$. On average, female drivers showed a slight negative acceleration, while male drivers showed a positive acceleration. This is true across all four phases (Figure 4-5).

In addition, maximum lateral acceleration differed in phase 1, $F(1,42) = 12.1$, $p = .001$, and phase 2, $F(1,42) = 8.6$, $p = .005$. In phase 1, positive maximum lateral acceleration was higher for male drivers than for female drivers, while in phase 2, a positive maximum lateral acceleration was observed for male drivers. Female drivers, on the other hand, showed a negative maximum lateral acceleration.

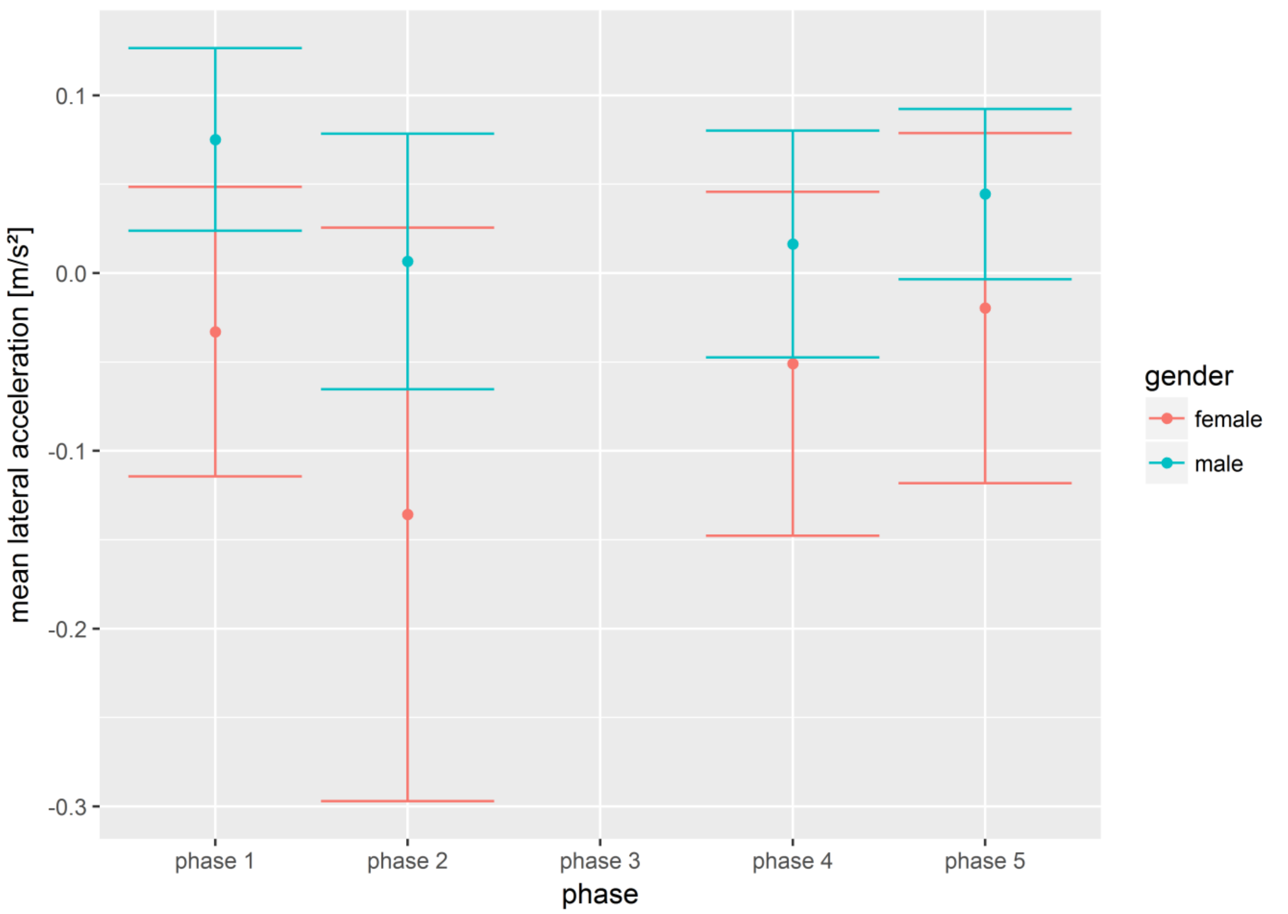


Figure 4-5: Mean and standard deviation of mean lateral acceleration during the phases of the overtaking manoeuvre.

With regard to mean longitudinal acceleration, a trend was observed for phase 1, $F(1,42) = 3.7$, $p = .06$. For male drivers, mean longitudinal acceleration was marginally higher than for female drivers. With regard to mean longitudinal acceleration, no other significant differences across the overtaking phases were observed. Significant differences between genders were observed with regard to maximum longitudinal acceleration.

Male and female drivers differed significantly in phases 3, $F(1,42) = 4.2$, $p = .046$ and phase 5, $F(1,42) = 10.4$, $p = .002$. Female drivers showed a higher maximum lateral acceleration in both phases (Figure 4-6).

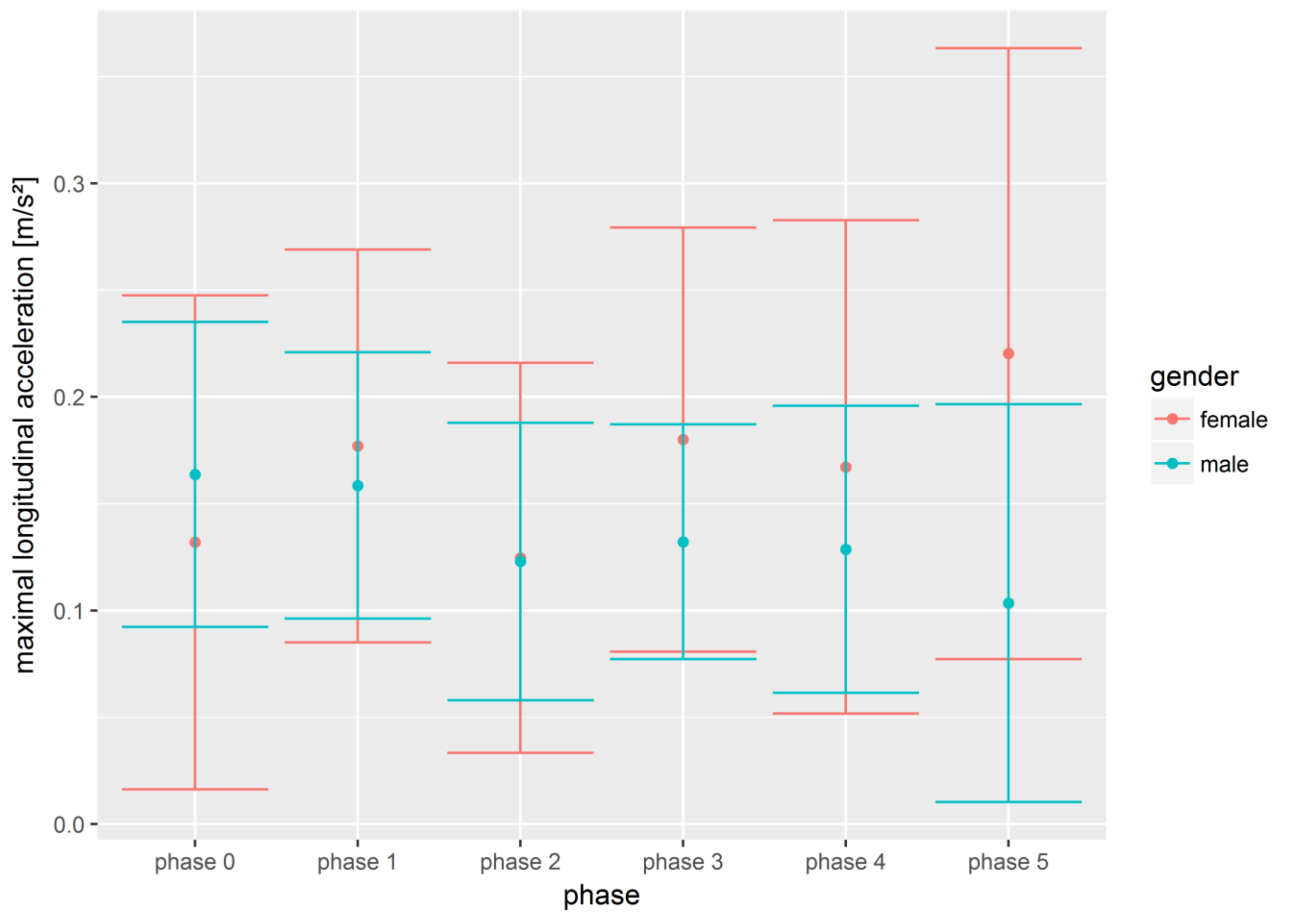


Figure 4-6: Mean and standard deviation of maximal longitudinal acceleration during the phases of the overtaking manoeuvre.

4.3.4 Analysis of takeover manoeuvre by country

Takeover manoeuvres were also analysed by country (France, Poland, Germany, and United Kingdom). With regard to situational factors, no differences between countries were observed. With regard to variables describing the actual overtaking manoeuvres, no significant differences were revealed.

With regard to driving performance parameters several differences between countries were found. Significant differences between countries were observed for mean lateral acceleration. They differed in phase 1, $F(3,40) = 19.9$, $p > .001$, phase 2, $F(3,40) = 5.1$, $p = .005$, phase 4, $F(3,40) = 3$, $p = .04$, and phase 5, $F(3,40) = 4.9$, $p = .005$. Differences are displayed in Figure 4-7 and means and SDs are reported in the Annex A.1.2. Differences in maximum lateral acceleration were observed in phase 1, $F(3,40) = 6$, $p = .002$, and phase 2, $F(3,40) = 3.1$, $p = .037$. In phase 1, the average maximum lateral acceleration was lowest in the UK (M = .05, SD = .13), while Germany (M = .19, SD = .1), Poland (M = .17, SD = .06), and France (M = .18, SD = .06) showed similar maximum lateral accelerations. In the following phase, the UK showed a negative maximum lateral acceleration while the other countries had positive lateral accelerations.

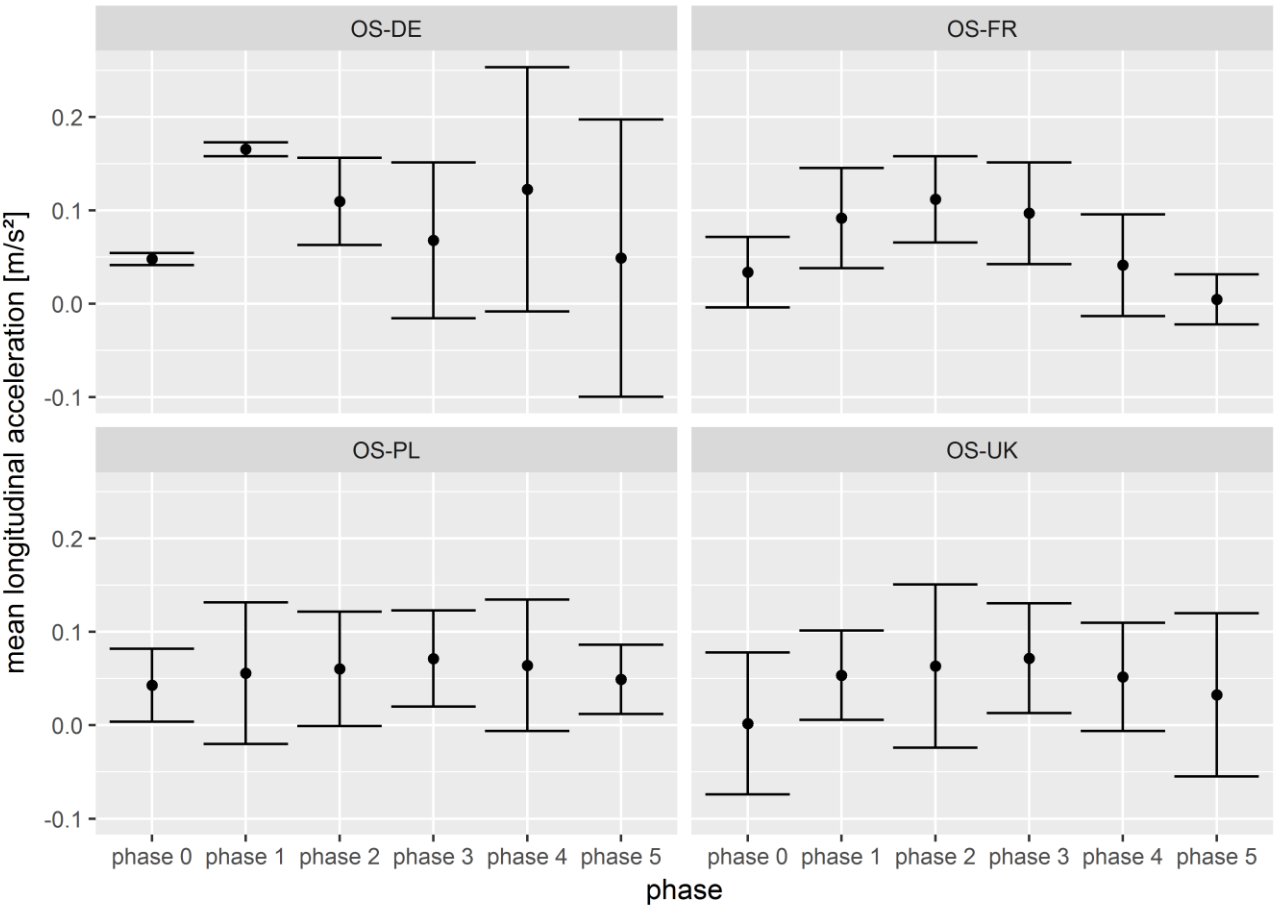


Figure 4-7: Mean and standard deviation of mean longitudinal acceleration during the phases of the overtaking manoeuvre.

With regard to mean and maximum longitudinal acceleration, differences were revealed for mean longitudinal acceleration in phase 1, $F(3,40) = 3.5$, $p = .02$, a trend for maximum longitudinal acceleration in phase 1, $F(3,40) = 2.5$, $p = .07$, and significant difference of maximum longitudinal acceleration in phase 5, $F(3,40) = 3.5$, $p = .02$. On average, France showed the highest longitudinal acceleration in phase 1, while Germany had the lowest (Figure 4-8). With regard to maximum longitudinal acceleration in phase 1, Germany showed the highest longitudinal acceleration and Poland the lowest. In phase 5, though, UK had the highest maximum while France had the lowest (please see Annex A.1.2 for means and standard deviations).

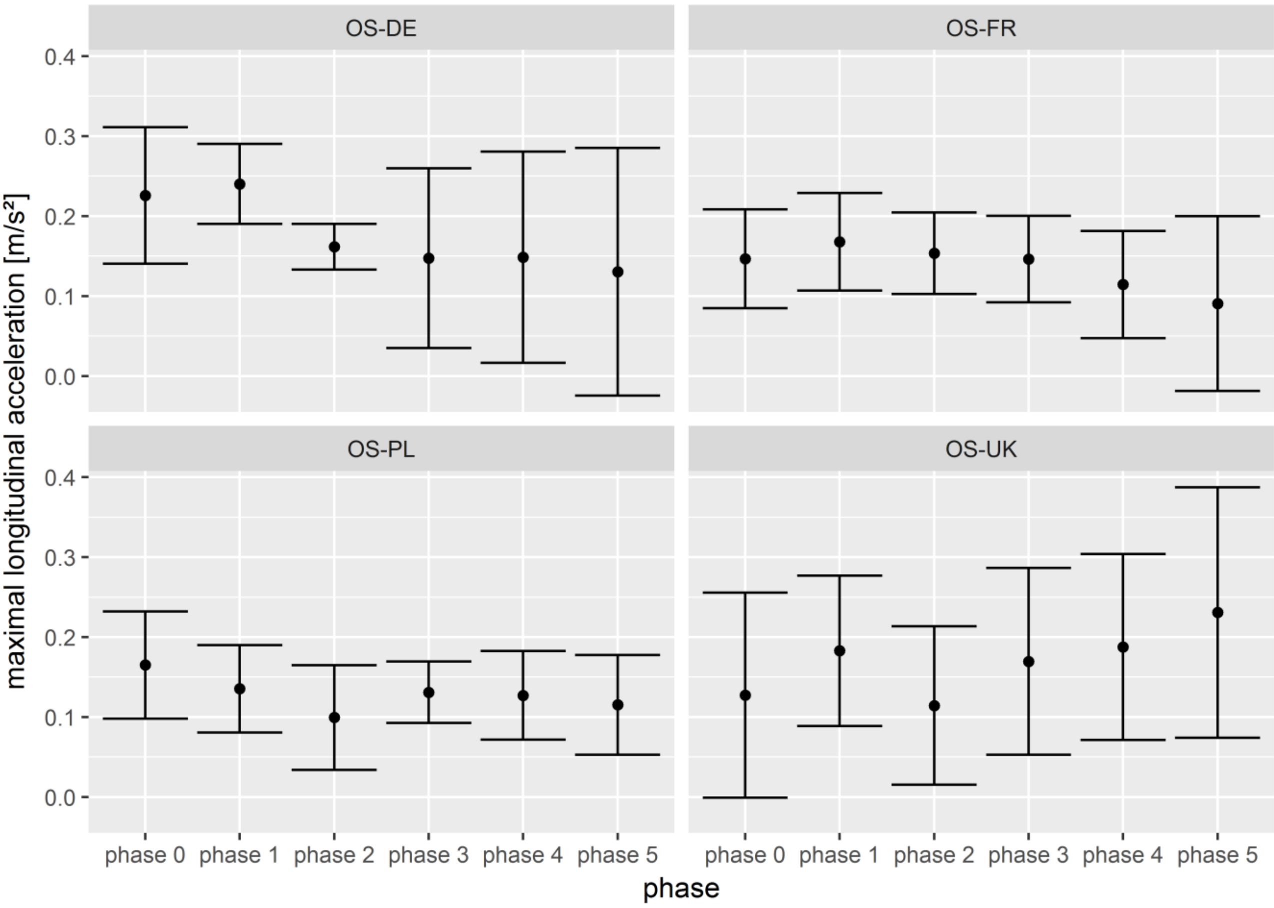


Figure 4-8: Mean and standard deviation of mean longitudinal acceleration during the phases of the overtaking manoeuvre.

4.4 Discussion

Overtaking manoeuvres were described focussing on the effects in type of manoeuvre, gender, and country. The analysis of type of manoeuvre did not reveal any evidence for an effect on situational factors. This indicates that drivers did not make the decision on how to overtake another vehicle dependent on the presence of passengers, overtaking regulations, road curvature, vegetation, or secondary tasks. On the other hand, it was observed that drivers were conscious of the surrounding circumstances. For example, bends are not suited for overtaking, since the vision range is restricted and the acceleration needed for overtaking would risk sliding off the road while turning. Less than 10% overtaking manoeuvers involved driving through a bend. Also, drivers generally respected the overtaking regulation. Other factors may influence the decision of overtaking in a normal, flying, or other fashion. In contrast to normal overtaking, flying overtaking does not include following the lead vehicle. If drivers travel very fast and difference in speed between the lead and ego vehicle is large, it is unlikely drivers would accept slowing down for another vehicle. Thus, the decision to perform a flying overtaking manoeuvre might be influenced by the speed difference between the host and lead vehicle. Investigated speed differences of the host and overtaken lead vehicle in dependence of the type of manoeuver may answer the question. This analysis is planned for future publications.

In almost every other case, the road user overtaken was a car. Trucks, on the other hand, were only overtaken once in every 11th overtaking manoeuvre. This is surprising, since trucks usually travel slower hindering other road users. It appears that fewer trucks were encountered on rural roads than cars. In order to draw more exclusive conclusions, the typical presence of different vehicle types on rural roads needs to be considered. A selection bias is also possible. Occasions in which trucks are overtaken by cars might be in the remaining dataset.

It appears that drivers avoid dangerous situations while overtaking. The observed drivers followed traffic regulations. Only 12% of overtaking manoeuvers were done in prohibited sections. Most of the time only one vehicle was overtaken. This minimizes the time spend in the opposite lane and decreases the risk of having a frontal collision. Also more than 90% of the overtaking manoeuvres did not involve encountering oncoming traffic. This indicates that drivers plan their overtaking manoeuvre carefully. The rare occasions of drivers deviated from the safe behaviour are worth investigating in more detail as factors such as impatience or inaccurate estimation of the length of the takeover manoeuvre might contribute to initiating unsafe takeover manoeuvres.

The lateral acceleration was distributed equally over the four relevant phases of the overtaking manoeuvre. No peak was observed during lane change initiation or completion. This indicates that lane changes are usually done smoothly without peaks at the beginning or the end of the manoeuvre.

Differences in longitudinal acceleration were only observed in phase 1. It was highest during flying overtaking and lowest during normal ones. In normal overtaking, drivers have to accelerate to bridge the speed difference to the lead vehicle. Flying overtaking involves keeping the travel speed. Drivers usually can benefit from their initial velocity and do not have to accelerate much to complete the manoeuvre.

Only a few differences between genders were found. A small difference in the mean longitudinal acceleration during phase 1 may mean that male drivers are a bit more enthusiastic in starting the manoeuvre. More profound differences are found in phases 3 and 5 in the maximal longitudinal acceleration. Female drivers showed negative, while male drivers a positive lateral acceleration. A positive lateral acceleration is more pronounced at the start of an overtaking manoeuvre and a negative lateral acceleration is associated with the end. So, this indicates that males were enthusiastic in leaving their lane while females were sporty in adjusting back to the lane they came from. One could also say males are keen on starting the overtaking manoeuvre while females are keen on ending it. Males initiate the initial turn of the steering wheel fast, while females overcompensate a bit in the counter motion with the steering wheel.

Between countries, differences were observed in the lateral and longitudinal acceleration. Lateral acceleration gives information about the lane change, longitudinal acceleration about closing in to another vehicle. The UK's lower maximum lateral acceleration in phase 1 compared to Germany, Poland and France may indicate British drivers initiate their lane change smoother than drivers of the other countries. In phase 2, British drivers stopped the lane change harsher, while the other drivers initiated it more abruptly.

Average longitudinal acceleration indicates that French drivers accelerated faster compared to German drivers who had the lowest values in the beginning of the lane change. However, the fastest lane changes were initiated in Germany and the slowest in Poland. French drivers were always fast in the beginning of their lane changes while German drivers started slower, but the peaks are done by Germany again. This could be due to the fact that German drivers are used to a wider speed range (i.e., great variance of speed on motorways). The absence of speed limits in specific parts leads to great differences in travel speed in different driving lanes. Great speed ranges may require very quick lane changes, occasionally. For example, a driver may be travelling with 130 km/h on the middle lane and needs to change into the left lane to overtake a slower vehicle. In the left lane, travel speeds above 160 km/h are not unusual in Germany. Thus, the driver may have to overtake quickly to free the left lane for an approaching vehicle travelling significantly faster. This behaviour could be learned and manifest on rural roads as well. UK drivers were quick in accelerating back into their lane compared to French drivers. A possible reason for the effects could also be different horse power setups of the vehicles between countries. This factor could not be included in this analysis but might be considered for future analyses. Differences may also result from the rather small and unevenly distributed sample size. It may also be the case that other factors (not included in the analysis) contribute to these differences.

UDRIVE strived to use global definitions and taxonomy. But in some cases adaptations had to be made to account for different analysis needs. The analysis of VRUs and risky driving were adapted to their type of road user and thus different approaches were taken. For example, the approach for analysing cars differs from the one taken in D44.1 Interactions with Vulnerable Road Users. An overtaking manoeuvre for bicycles includes only 4 phases. This is due to the fact, that cyclists can be overtaken while staying in the lane. Thus the lane changes are rarely necessary. Overtaking of another vehicle takes longer than overtaking a cyclist. Therefore, the video coding of an upcoming bend was adapted to the purpose. For cyclists the information “is straight road” (Yes/No) for a given overtaking manoeuvre was coded, while for cars it was coded “Yes” if a bend was in sight of the vehicle. Another result from the longer overtaking manoeuvre for cars overtaking other vehicles is that piggy-backing, the following of another vehicle during overtaking, is much more common when no cyclist is being overtaken. Thus, this additional type of overtaking was used in this analysis. Cyclists are not protected by a huge metal cage and thus are far more concerned about safety margins during the overtaking manoeuvre, than car drivers. This makes the lateral distance a major factor in overtaking assessment for VRUs. However, this is not the case when a vehicle overtakes another one. Here, the comfort of the manoeuvre is influenced more by lateral and longitudinal acceleration.

5 Self-confrontation

5.1 Introduction: Self-confrontation

The study of the driving behaviours over long periods of time and in natural conditions has been widely recommended within the last 10 years, to complete experiments carried out on driving simulator or tracks (Stutts & al., 2001 ; Young & Regan, 2007; Kircher, 2007 ; Ranney, 2008), or even restricted time-span monographs of crashes (e.g. Crashworthiness Data System). While revealing the actual motivations of the drivers in their risk-taking is of obvious concern, naturalistic driving studies also aim at highlighting the strategies that drivers display to organize their life on board in a way that is subjectively compatible with the driving task, whether their compensatory behaviours are effective or not. As such, the UDRIVE project constitutes an interesting basis for the study of situations of secondary tasks/distraction, as well as of the related or not situations of risk-taking/pre-crashes. To our knowledge, the self-confrontation method to produce and explore such data has been rarely (if at all) used before.

5.2 Method

In the UDRIVE project, drivers were interviewed on the day when their car was being unequipped. Each private interview was conducted with different steps:

- A verbal exchange on preselected video sequences recorded between January and February 2017 : 2 to 3 events for each participant
- Participant’s free recall of critical sequences, even out of UDRIVE recording period. Those event were raised by the participant because of their high emotional impact

At the end of the interview, more far-reaching questions were asked guided by the participant’s answers to a questionnaire completed before the interview.

5.2.1 Participants

16 of the 30 participants of the UDRIVE French panel were individually interviewed during the de-installation of their vehicle, in April 2017. With one exception (participant’s refusal), each interview was fully recorded for subsequent analysis. The sample was balanced in terms of genders (8 male, 8 female). The distribution of age ranges from 23 to 70 years old with a mean value of 47.1 years. Due to recruitment bias, in this naturalistic driving studiy, older drivers are over-represented.

5.2.2 Selection of situations

The critical methodological node was the choice of the clips to be used for self-confrontation. Self-confrontations were designed to extract material from the driver’s experience (subjective data) in relation with 2 research questions in UDRIVE:

- Near crash situations extended to safety critical event and/or risky situations. The selection was based on hard braking events reviewed and selected by the psychologist to address highly emotional situations
- Life on board (secondary tasks), more specifically the use of a mobile phone. These sequences were also tracked manually in the data and the selection was done by the psychologist

Our strategy consisted in combining objective and subjective cues: the statistical analysis of behavioural patterns (pointing out hard braking events), the distribution of trips (choice of a usual route and exceptional drive), finally the drivers’ responses to a questionnaire asking them to pinpoint remarkable situations encountered during the experimentation.

5.2.3 Experimental design and procedure

The interview began with presenting at least 2 sequences selected from the most recent available videos (dating back to January or February 2017) to the driver one presented: a double task situation (typically a phone call), the other a hard braking event. The objective was to reconstruct as precisely as possible the episode as seen by the driver (Theureau, 2003), using the classical techniques of the self-confrontation method. This first phase was meant to stimulate the drivers' evocation of situations experienced as critical, whether these situations took place during or before the UDRIVE project. The bet was that priming the interviewed with videos permitted to place and maintain drivers in a situated storytelling state, avoiding the pitfall of the make-believe, social desirability bias or even general attitudes which may differ from actual practices.

In this self-confrontation method, the psychologist introduces at present tense the global context ("Here, Thursday, 5:00, you just quit your job"), launches the video a few seconds before the critical event– shares the driver's field of interest (e.g. looking at what he/she is looking at) – then suspends it to let the memory take over, use empty content ("and here...") and echoes phrases (resumption of the last terms of the interviewed; Barbier, 2010).

The interview ended with a questionnaire covering all the above issues (ex: "I rarely/sometimes/often/always write text messages while driving").

For this study, self-confrontation method is in its boundary conditions of application in two ways:

- Due to the large size of video material, the video-sequences selection involved analyst's subjective selection

The selected events took place two months before the interview thereby infringing the 'recency' principle. This time-gap (between the video and the past experience) may have diluted the evocative power of the chosen sequences, preventing participants from diving into their actual past experiences. On another hand, the sequences may have prompted a self-analysis, namely a reflexive stand triggered by the viewing of oneself that could have interfered with participant's free-recall part of the interview.

5.2.4 Analysis

On the one hand, this studied aimed at providing ecological data regarding secondary tasks and uncomfortable driving situations. The goal was to get detailed chronicles (monographs) exposing the situation dynamics: how and when secondary tasks or risky sequences occur, what was the sequence of events having led to them and what was the driver's role in this sequence (active or reactive). The study was also meant to assess the efficiency of the self-confrontation method as used in the framework of Naturalistic Driving Studies.

5.3 Results

5.3.1 Uncomfortable situations

From an epistemic point of view, the interviews are informative about at least two situations of driving judged as uncomfortable:

- on specific occasions (e.g., being scared of having misjudged the distance when passing another car)
- or recurrently (e.g., trepidation at the entrance in a tunnel).

The 60 verbatim collected on uncomfortable situations point out in five main items:

1. Adverse driving conditions (night driving, rain)

   Such situations were mainly raised by drivers who think they have low vision (wearing corrective glasses, seniors) or are prone to drowsiness

2. Interaction with a vulnerable user

Pedestrian, bike-/motorcycle-rider were judged by the driver to be self-involved and/or unpredictable, in particular in winter (by night, with poor visibility and with reduced breaking conditions). This situation was frequently raised (pedestrians as a main cause of stress was mentioned by 3 out of the 16 participants) and is often related to negative experiences (witness of or actor in an accident or in a quasi-accident);

3. Dangerous (rather than vulnerable) other driver

Mostly a truck or other vehicle whose behaviour seems to be unpredictable (sudden and unexpected manoeuvre, no turn signal, unexpected breaking)

4. Road infrastructure

Structural configuration of the infrastructure judged as ill-conceived:

- roundabouts and/or intersections with poor visibility (due to trees or to hairpin turns)
- road bends, particularly those that occur together with highway on/off ramps (paradox between speeding along and poor visibility
- Exits on highway, requiring drivers to cross lanes
- unusual road configurations, typically slippery road on the left side of the road

5. Limits of driver vehicle or driver himself

Poor visibility, reverse gear, improper acceleration and passing
Or self-awareness of drowsiness, vigilance drop, realization of having had a mind blank (followed by "I am already there!") after long driving or due to performing a secondary task (mind wandering, being absorbed by a radio show, etc.), particularly so during regular journeys.

Several strategies to avoid uncomfortable situations are explicitly raised in entries (1), (4) and (5) above: the driver doesn't drive at night anymore or shuns away from doing so if possible, decides to take another (exit) lane if it is judged to not be exceedingly time-consuming (e.g. a dangerous exit), avoids parking in reverse gear and prefers parking lots unfrequently occupied.

An interesting and specific case of an uncomfortable situation is the use of GPS driving in the city (the GPS paradox): in such contexts the use of GPS is at the same time helpful – as it provides the needed information – and a handicap – as it diverts the driver's attention from the road to the interior of the cockpit precisely when he or she should pay attention to the (unrecognized) route..

5.3.2 Crashes

Besides these situations of awkwardness, three crashes that occurred during the experiment were described in more detail.

- frontal collision with a tram

This crash was a due to an "emotional black-out" caused by the announcement of a medical diagnostic by the driver's passenger. The driver described a sudden withdrawal from the awareness of the collision event.

- collision from behind at an intersection with poor visibility

The driver (female) wanted to move forward slightly for better visibility while the vehicle behind really started up and crashed into the rear of the driver's car. Noteworthy, this driver, a petite, has already had two similar crashes: she now uses at this road point the handbrake.

- roundabout collision

The car was hit by an entering vehicle that first stopped to give way – hence demonstrating that third party's vigilance – but failed subsequently to yield to actually give way. Remarkably, this driver had already a similar collision event due to such over-/under-vigilance in a situation of giving priority. The detailed report completes these three cases with 16 other accidents that occurred before the UDRIVE period as described by the participants.

5.3.3 Risk-taking

In a way that is transverse to the situations of discomfort described above, and beyond these configurations (situations not experienced as uncomfortable), the study raises the main risk-taking factors (objective or subjective):

- the widely shared struggle against boredom, which leads the driver to engage in another, potentially absorbing activity
- distraction, whether or not it follows from boredom
- the pleasure of driving fast, of passing other vehicles, marginal in our panel (1 driver out of 16)
- the discrepancy between objective and subjective risk in the sense of an under-awareness of the risk (over-confidence in the braking capacities of the vehicle, obscuring the possibility of obstacles)
- 'de-realization' – the driver has had many positive reinforcement experiences and no longer consider an incident (specific example: driving with a cat free to move in the vehicle; more general case : a familiar route) and/or easily engages in a compromise ("I know it is not prudent but ...")
- pseudo-control, which consists in symbolically countering the feeling of risk-taking by a derisory measure (e.g., "I slow down a little")
- the magical thought that the crash, experienced as a punishment, occurs only if one crosses moral limits (e.g., "I agree to burn two orange lights but not three").

Furthermore, the role of expectations is ambivalent in a risky behaviour, more than in risk taking: the near crash situations are alternatively related to

- a lack of anticipation : driver caught between two trucks after a phone call with emotional content
- or, on the contrary, to an over-anticipation, particularly risky because it leaded to the dismissal of a possible scenario. In a roundabout a driver over-interpreted the behaviour of a third party: identified the entry of the other vehicle, inferred the last possible exit to avoid a full roundabout, but full round eventually performed by the other vehicle.

5.3.4 Distraction

Finally, the study initiates a typology of distraction factors, partly complementary to the above situations of risk taking.

Ambivalent role of the passenger

Passenger is considered as a source of distraction when he/she adds a load to the driver's attention (e.g., when the driver asks for help to monitor the central console).On the contrary, it is experienced as a protection from distraction when secondary task can be delegated (co-piloting, answering the phone).

A drifting attention and its subsequent capture by emotional event or cognitive related situation is very often raised, whether mediated by a third party (conversations in the cockpit or by telephone), carried by a medium (radio broadcasts or even CD-books) or even by free mind wandering (lost in thoughts).

Contemplation of the environment

Drivers also frequently mention the contemplation of the environment, with mentioned attractors such as appealing landscapes (e.g., trees in bloom), changing (works), nostalgic (a farm where one learned to mount a Pony as a child) or unknown landscapes.

Secondary task

The distraction can finally result from a compulsion to check mobile devices (informational addiction) or from affordance to respond to their solicitations (lack of inhibition of response).

It is interesting to note, for these various scenarios, that the consciousness withdrawal from the driving situation can alternately be brutal or very progressive, which presumably calls, as the case may be, for specific assistance and strategy.

5.4 Discussion

Global assessment of the methodology is positive. The memory revival failed with only one of the 16 participants. For the remaining subjects, it was evaluated as acceptable to excellent by the analyst: the recollection was more or less complete, covered more or less extended periods of time. It appears, the video sequences have facilitated driver's free-recall (participants who most talked on the video sequences were also those who were more prone to recollect other driving situations absent from the sequences). Hence, the self-confrontation technique appeared to have merit particularly for risky situations for the naturalistic driving studies.

Driver experience of uncomfortable situations was addressed in an innovative way using the self-confrontation technique with UDRIVE subjects. As a result, drivers attribute their uncomfortable feeling mainly to (misleading) infrastructure factors (e.g., roundabout without visibility). However, this type of limit has the advantage of giving rise to clear recommendations: improve road infrastructure when possible.

When focused on 'secondary tasks', self-confrontation appeared less promising: apart from two cases of highly anxiety-provoking telephone calls (notification of a sanction, announcement of a sick child), video images of secondary tasks were less evocative, with participants merging unmarked recollections presumably actually experienced but not necessarily related to the viewed video sequences. Moreover, such video sequences appear to trigger a self-analytic posture (adoption of a reflexive stand) more easily, presumably because the situation pertains more to the life on board than to the external environment: participants' gaze frequently moved from the front scene view to the image of their own face and body. Most of the epistemic results in that study replicate data already highlighted in previous studies (e.g., factors of driving distraction).

The contrasting success of the self-confrontation method is not surprising. Indeed, risky situations and dual-task situations differ by their emotional valence: the former are by definition highly emotional while the latter, are barely significant for the drivers. Yet, cognitive theories highlight the strong facilitating effect of emotion on memorization and recall (Broadhurst, 1957; Kensinger, 2004; LaBar & Cabeza, 2006). Given its limitations, the use of self-confrontation method should be primarily focused on striking events when mobilized in longitudinal naturalistic driving studies.

Finally, we believe together with many others that our most promising results are those related to drivers' distributions of attentional resources, particularly related to the on-board task management along with different length and environment of the trip (Horrey & Lesch, 2009; Tivesten & Dozza, 2015): departure, city context, tiredness sequence, turbulent traffic zone, etc. In our view, this will allow for assessing the most accurate danger value of any secondary task activity, being a potential vigilance enhancement and a source of pleasure for the life on board.

6 Everyday driving

6.1 Introduction

One of the key benefits of a naturalistic driving study is that it provides the opportunity to gain a detailed insight into the everyday driving behaviour on European roads that does not necessarily result in safety critical events. There is much to be learnt from an in-depth study of the driver, the vehicle and their environment on a day-to-day basis, as drivers go about their normal driving tasks and routines, with unobtrusive observation and an absence of experimental control. This section contains an analysis of everyday driving behaviours, specifically risky driving behaviours that do not result in SCEs. This is an important addition to the knowledge base, and it is an area that can be overlooked by attention to the causes of on-road incidents only.

This analysis uses data collected from both cars and trucks and considers a wide range of risky behaviours including speeding, close following, high lateral g, and harsh acceleration and braking. Everyday usage of ADAS (specifically cruise control systems and speed limiter systems) is investigated as well as seatbelt usage. These behaviours are presented in a largely descriptive analysis, firstly to get an overall perspective of the prevalence of these behaviours in the UDRIVE driver sample, and then to further understand the occurrence of each of these behaviours in relation to combinations of driver and environmental factors. The car and truck data analyses are presented separately with no comparison of the two vehicle types.

The major research questions to be answered in the descriptive analysis of everyday driving are:

- Who engages in risky driving behaviour?
- What driver characteristics influence the occurrence of risky driving behaviours?
- What environmental characteristics influence the occurrence of risky driving behaviours?
- Where and when do drivers use ADAS?
- What driver and environmental characteristics influence seatbelt usage?
- Do risky driving behaviours co-occur in an individual driver?

The dataset analysed for each research question was extracted from the UDRIVE database and validated by the partner responsible for each research question. The data used shows some variation depending on the specific research question, and therefore is described in more detail in each of the following sections of this chapter. The commonalities of the analysis approach are now described before moving on to each research question in detail.

6.1.1 Definition of risky behaviours

The risky behaviours selected for investigation are those with a strong link to vehicle crash risk, such as speeding and close following. Similarly, driver inattention has been shown to be an important contributory factor in crashes. ADAS systems are primarily designed to enhance driver safety, however, their use – or misuse – has the potential to degrade safety. Hence the use of two ADAS (cruise control and speed limiter) is investigated here. Seatbelt usage is mandatory throughout the UDRIVE countries, and yet rates of non-compliance and reasons for this are not fully understood.

Risky behaviour thresholds were identified through a review of the literature and discussion amongst project partners such that behaviours were captured that reflected unsafe driver behaviours without the need for these behaviours to develop into a safety critical event. The following risky behaviours were identified and analysed in the remainder of this chapter.

- **Speeding** – events were identified during which a driver was exceeding the speed limit for a defined period of time. A minimum time period was set so as to avoid the inclusion of momentary over-speeding events which could reflect speed control errors rather than over-speeding. Different levels of speeding were defined dependent on the percentage by which the speed limit was exceeded (11-

15%, 16-20% etc.). Within each speeding event, mean speed and maximum speed were also considered. Duration and frequency of speeding were also used as metrics of risky speeding behaviour. These variables used speed data recorded directly from the vehicle CAN and speed limit and road type information from the map database.

- **Close following** – events were identified during which a driver was following the lead vehicle with a short time headway for a defined period of time. Various thresholds were piloted to ensure that risky rather than normal driver behaviour was being captured. Variables considered in this analysis included mean and minimum time headway, duration of close following, and mean and minimum time-to-collision.
- **Harsh acceleration and deceleration** – events were identified during which a driver exceeded a specific threshold of acceleration or deceleration. These thresholds were piloted to ensure that the analysis was capturing aggressive or harsh driver behaviours rather than everyday safe behaviours. Variables considered for this analysis were maximum acceleration and number of exceedances of the acceleration or deceleration threshold.
- **Harsh steering** – events were identified during which a driver exceeded a specific threshold of lateral g or steering wheel angle. These thresholds were piloted to ensure that the analysis was capturing true harsh steering events. Intersections were identified and removed from the analysis so as to focus on events that occurred during driving on curves. Variables considered for this analysis were maximum lateral g and number of exceedances of the lateral g threshold
- **Overtaking –** the dataset described in Section 4.2.2 was re-used in this everyday driving behaviours analysis to identify links between driver personality and overtaking behaviours. The variable of interest was whether or not a driver performed an overtaking manoeuvre, and hence this behaviour was described as a binary variable for the purposes of this analysis.
- **Seatbelt use –** seatbelt switch information was directly recorded by the vehicle and was used to determine whether the seatbelt was used and for what proportion of the trip it was fastened.
- **ADAS use** – interactions with ADAS systems were directly recorded by the vehicle. Variables of interest included frequency, duration, and number of interactions.

6.1.2 Driver factors

Driver factors collected in the UDRIVE project included driver age, gender and driving experience. For analysis purposes, drivers have been further subcategorised into groups based on their age (e.g. young, mid and older drivers).

6.1.3 Driver personality

A wide range of subjective measures were administered to drivers at the outset of the study to capture aspects of their personality:

1. Driver attitudes questionnaire (20 items)
2. Driver behaviour questionnaire (19 items)
3. Driver style questionnaire (15 items)
4. Traffic locus of control (17 items)
5. Arnett Inventory of Sensation Seeking (20 items)
6. Driving history: accidents (4 items)
7. Driving history: violations (7 items)
8. Vehicle equipment
9. Nomadic device use (5 items)
10. ADAS use (5 items)
11. Experience (1 item)

Drivers were categorised based on the scores given on each subscale of the aforementioned question and these categories were submitted into the analysis. These personality metrics were also used to produce an overall composite personality measure. K-means clustering was attempted to define clusters of drivers. However, the heterogeneity within the sample and cross-country variation made this clustering approach challenge. A summary personality measure was defined by combining a selected range of the personality metrics above.

6.1.4 Environmental factors

The vast UDRIVE dataset ensures that there is an array of environmental factors that can be considered in this analysis. For the most part, we are interested in the impact of variables such as road type, speed limit, and road curvature on the performance of risky driving behaviours. The use of these variables is explained in each analysis below, including whether or not they were directly recorded in-vehicle or coded retrospectively by the analysts.

6.1.5 Trip characteristics

For some analyses trip length has been considered as an additional variable, with a specific interest in whether drivers behave different during short and long trips.

6.1.6 Traffic culture

UDRIVE operation site, or *country*, is used as a proxy variable for traffic culture.

6.2 Research question 2.1: To what extent are driver and environmental factors associated with risky behaviours?

6.2.1 Introduction

In driving school, drivers learn a complex set of rules aiming at minimising the risk of being involved in a crash. Among others, key rules of conduct are driving within legal speed limits, keeping a safe following distance, driving carefully and anticipating the situation, not engaging in secondary tasks such as texting or dialling. Crash statistics and other research show that not following these rules of conducts may lead to crashes as speeding, close following, and distraction have been identified as main causes of crashes. Being caught overstepping rules is unlikely, crashes are rare events as well; therefore, drivers may not receive direct feedback on their action making it more cumbersome to follow the rules. Little is known about the prevalence of these risky behaviours. Questions, such as whether drivers speed on rare occasion or deliberately and whether drivers always follow on close distance, are of research interest as the answers provide information on the prevalence and risk of disregarding safety precautions. In (semi-) controlled experimental settings, participants will not engage in such behaviour, naturalistic driving studies (NDS), on the other hand, provide a setting that lets drivers feel at ease and behave naturally. The NDS administered within the EU project UDRIVE offers the opportunity to investigate the prevalence and risk of disregarding safety precautions.

Risky driving behaviours such as speeding and close following distance are extracted from the data, their prevalence determined, situational factors added. In addition, driver factors will be included in the analysis. Results will provide more insight on for example, how age, gender, and annual mileage affect risky behaviour and what type of situational factors contribute to engaging in risky behaviour. In addition, because data were collected throughout Europe, regional differences can also be analysed. Understanding the prevalence of risky behaviour will allow for proposing countermeasures that will improve road traffic safety.

6.2.2 Method

Data used for the analysis of speeding and close following in everyday driving is based on the data query of April 21st, 2017.

Altogether, 11297 events of speeding were detected. 5614 were committed by males and 5683 by females. Drivers were between the ages of 20 and 77 years of age with an average age of 41 years (SD = 10.9). Altogether 75 male drivers between the ages of 21 and 77 (M = 44.4, SD = 12.3) committed the speeding. Out of these 75 drivers, 20 came out France, 9 the Netherlands, 11 Germany, 16 Poland, and 19 the UK. Female drivers accounted for the remaining 64 drivers (20 French, 9 Dutch, 6 German, 5 Polish, and 24 British) and were between the ages of 20 and 66 with an average of age 37.7 years (SD = 8.2).

Close following

A total of 9824 events of close following were detected. 4226 were committed by male drivers and 5366 by female drivers. Drivers were between the ages of 18 and 80 years of age with an average age of 42.72 years (SD = 11.25). Here, 84 male drivers between the ages of 18 and 80 (M = 46.02, SD = 14.54) were following closely. Out of these 84 drivers, 19 came from France, 13 the Netherlands, 13 Germany, 18 Poland, and 21 the UK. Female drivers accounted for the remaining 71 drivers (21 French, 13 Dutch, 7 German, 4 Polish, and 26 British) and were between the ages of 22 and 66 with an average of age 42.07 years (SD = 11.1).

Definitions of risky behaviour

Speeding and close following behaviour were used as surrogates of risky behaviour and are defined as followed:

1. Speeding is defined as travelling at least 11% over the speed limit for more than 10 seconds. For the purpose of the analysis, speeding was grouped into three categories:
   - Light speeding = speed exceeding 11%-15% of the speed limit
   - Severe speeding = speed exceeding 16-20% of the speed limit
   - Extreme speeding = speed exceeding 21% and more of the speed limit.

   Please note: Events with speed exceedance between 1 and 10% have not been included in the analyses as those speeding events are not necessarily sanctioned by law. In addition, extreme speeding events were not detected.
2. Close following is defined as travelling behind another vehicle with a time-headway smaller than 1.5 seconds for at least 1.5 seconds.

Performance parameters used to assess risky behaviour are:

1. Speeding
   - Maximum speed during a speeding segment
   - Mean speed during a speed segment
   - Duration of speeding
   - Degree of speeding
   - Frequency of speeding
2. Close following
   - TTC (minimum, mean)
   - THW (minimum, mean)
   - Duration of close following

6.2.3 Analysis

Speeding

In a first step data will be presented descriptively providing a general overview of speeding events in UDRIVE. In addition, one- way ANOVAs are calculated in order to assess the effect of age group (18-24, 25-49, and 50-99), gender, country, weather condition, and time of day on maximum speed, mean speed, and duration of speeding.

In addition, categorical variables are tested using chi-squared tests. Therefore, the effects of age group, gender, country, weather condition, and time of day on the frequency of speeding within defined speed limit categories (0-30, 31-50, 51-70, 71-90, 91-110, and 111-130) are tested. The same tests are administered testing the effect on degree of speeding (11-15% vs. 16-20%).

Close following

As in speeding, a descriptive overview of close following events distributed by age group, country and gender is given. Statistical analysis is performed. To build the sample for the statistical analysis, all found close following events were aggregated per driver ID. The resulting sample was then investigated for effects on age, country, and gender. For age and country homogenity of variance was tested with Bartlett tests and depending on the result either a Kruskal-Wallis rank sum test or an ANOVA for independent groups was performed. For gender the normality was tested with Shapiro-Wilk tests and depending on the result either a Wilcoxon-Rang Sum Test or t-test for independent was performed. Each factor (age, gender, and country) was tested on all relevant performance indicators (minTTC, meanTTC, meanTHW, minTHW and duration) for significant differences.

6.2.4 Results

Descriptive analysis of speeding

Table 6-1 summarizes the speeding events with respect to frequencies weighed by duration of driving under a specific speed limit. Altogether, more speeding events were detected for exceeding the speed limit between 11 and 15%. More than half of all speeding events happened in sections with posted speed limits of up to 50 km/h. When the posted speed limit was between 51 and 90 km/h, 37% of the speeding events were committed, while about ten percent took place when the speed was between 91 and 130 km/h. Within all speed limit categories except 110-130 km/h, more speeding events exceed the speed limit between 11 and 15%.

Table 6-1: Number of speeding events per posted speed limit split by severity of the speeding event.

		Speed Limit [km/h]					
		0-30	**31-50**	**51-70**	**71-90**	**91-110**	**111-130**
Degree of speeding	Light	1918,9	857,1	1155,8	935,1	525,4	57,0
	Severe	1431,5	672,9	739,1	560,7	272,6	64,9

Table 6-2: Number of speeding events per posted speed limit by country and severity of the speeding event. The number is weighed by duration driven in each country.

		Germany	France	The Netherlands	Poland	UK
Degree of speeding	Light	565,2	1150,4	797,4	1084,7	1455,2
	Severe	435,4	1052,7	293,9	893,8	796,8

Table 6-3: Number of speeding events per posted speed limit by time of day and severity of the speeding event. The number is weighed by duration driven at each time of day.

		Morning	Pre-noon	Noon	Afternoon	Evening	Night	Late-night
Degree of speeding	Light	977,4	663,4	355,9	1145,9	310,5	163,3	1678,4
	Severe	689,6	487,4	265,1	676,9	248,2	120,3	1254,6

The Table 6-2 shows that about 26% of all speed violations took place in France and the UK, 23% in Poland, and 13% in the Netherlands and Germany. Investigating the time of day speed violations were committed, data shows that about 32% occurred at late night, 18% in the morning, 13% pre-noon, 7% around noon, 20% in the afternoon, and about 3% at night (Table 6-3).

The frequency of speeding events declined with increased speed limit as displayed in in Figure 6-1 with the exception of speed limits between 31-50 km/h. While the light speeding events (11-15%) were more frequent for almost all speed limit categories, both light and severe speeding categories were about the same frequency when the posted speed limit was 110-130 kmh.

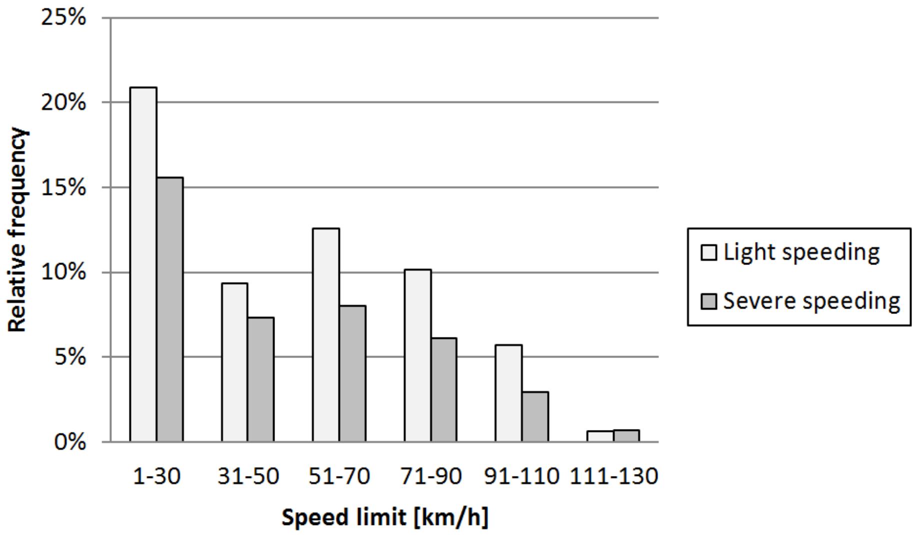


Figure 6-1: Relative frequency of speeding events by posted speed limit (from map data). Frequencies are weighed by exposure of time driven under a specific speed limit.

Effects of age on speeding

The analysis of age on speeding revealed significant effects of mean speed $F(2, 11294) = 410.8$, $p < .001$, maximum speed $F(2, 111294) = 426.5$, $p < .001$, and duration $F(2, 11294)= 8.9$, $p < .001$. On average, age group 1 (18-24 years) had the highest mean and maximum speed, followed by age group 2 (25-49 years). Drivers falling into age group 3 (50-99 years) had the lowest mean and maximum speed. The distribution of average duration of a speeding event follows the same trend. Drivers of age category 1 sped the longest, while drivers of age category 3 sped the shortest (for means and standard deviations, Annex A.2.1). Chi-squared analysis revealed that the number of speeding events per age groups differed $\chi^2(10) = 87.8$, $p < .001$ (see Figure 6-2).

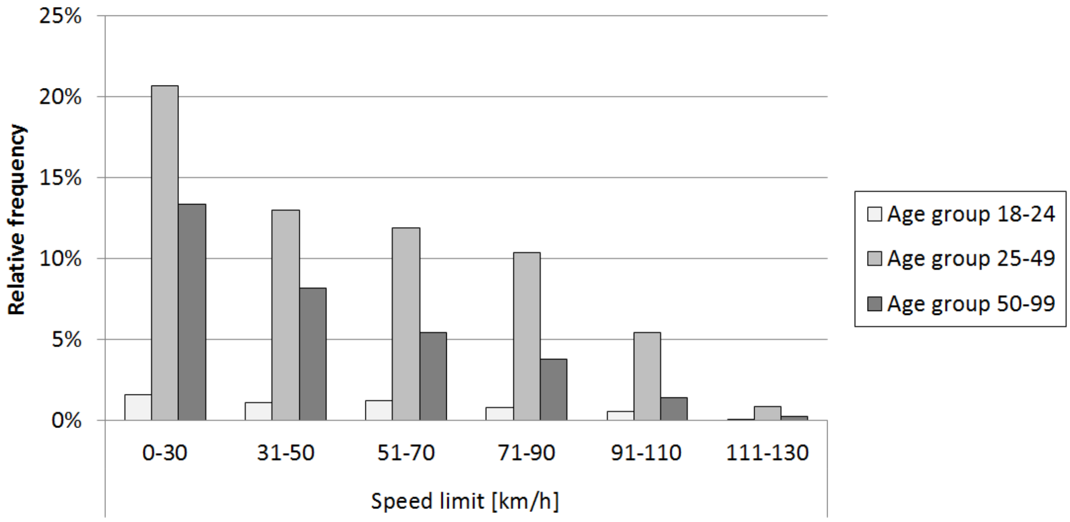


Figure 6-2: Distribution of frequencies in terms of speed limit category by age category.

In Table 6-4 below, relative frequencies were calculated for speed events within speed limit categories for each age group. Overall, the frequency of speeding events reduced when the speed limit increased. The distribution between age groups 1 and 2 appears to be similar, while age group 3 shows a higher distribution of speed events in the low speed limit category and has fewer events in higher speed limit categories compared to age groups 1 and 2.

Table 6-4: Relative frequencies of speeding events within age groups.

		Age group		
		18-24	25-49	50-99
Speed limit category	0-30	29.8%	33%	41%
	31-50	20.7%	21%	25%
	51-70	22.5%	19%	16.8%
	71-90	15.2%	16.6%	11.7%
	91-110	10.2%	8.7%	4.3%
	111-130	1.4%	1.4%	0.7%

Additionally, chi-squared analysis were administered in order to analyse the effects of age group on level of speed exceedance (11-15% vs 16-20%). A significant effect was revealed, $\chi^2(2) = 9.1$, $p = .010$. The youngest age group committed equally as many speed violations in both speeding categories while the other age groups had fewer excessive speeding (16-20%) events (Figure 6-3).

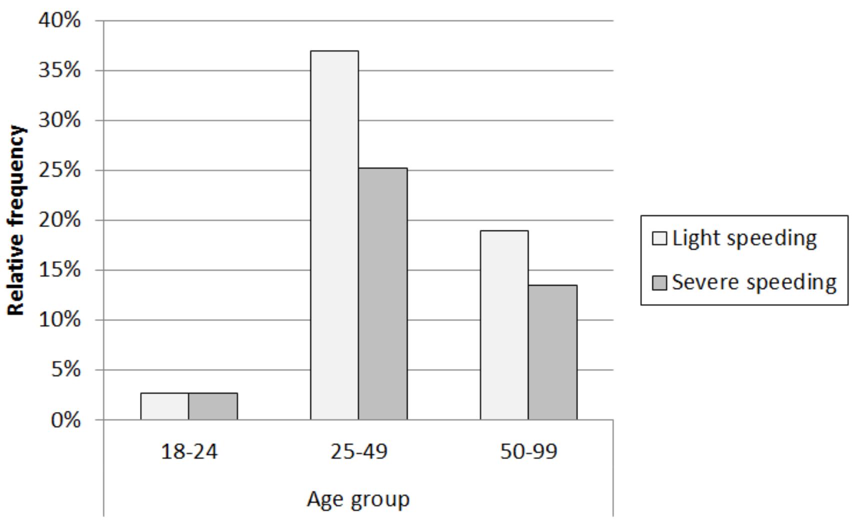


Figure 6-3: Frequencies of speeding events by age group, weighed by duration driven for each age group.

Effects of gender on speeding

One-way ANOVAs were calculated in order to determine the effects of gender on mean and maximum speed as well as duration of speeding. Analyses show a significant effect of gender on mean $F(1, 11295) = 101.1$, $p < .001$ and maximum speed $F(1, 11295) = 108.3$, $p < .001$. While male drivers showed a mean speed of 97.1 km/h (SD = 27.9), female drives travelled at an average speed of 102.1 km/h (SD = 26.3) while speeding. The same trend was observed for maximum speed values. Females' (M = 112.1; SD = 23.6) maximum value was higher than males' (M = 107.1, SD = 27.37). Nonetheless, no differences in the duration of speeding were found. On average, a speeding event took 21 seconds.

Chi-squared analysis revealed significant differences of gender on speed limit category $\chi^2(5) = 29.6$, $p < .001$. More female than male driver sped in speed limit zones 0-30 km/h, 31-50 km/h, and 51-70 km/h, and 111-130 km/h. Male drivers on the other hand, had more speeding events in the speed limit categories 71-90 km/h and 91-110 km/h (see Figure 6-4 and Annex A.2.2).

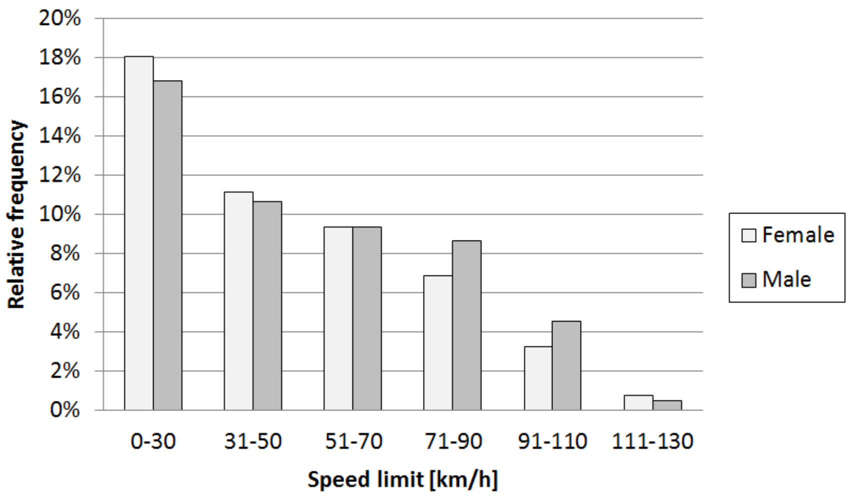


Figure 6-4: Distribution of frequencies in terms of speed limit category by gender. Frequencies are weighed by duration driven per gender.

Gender differences were observed in speeding behaviour. While the distributions between the genders were similar, small differences were found between groups. Chi-squared analysis of severity of speeding and gender revealed a significant difference $\chi^2(1) = 10.2$, $p < .001$. Males committed more severe speed violations

(16-20%) than females. Females on the other hand, exceeded the speed limit by 11-15% more often than male drivers (Figure 6-5).

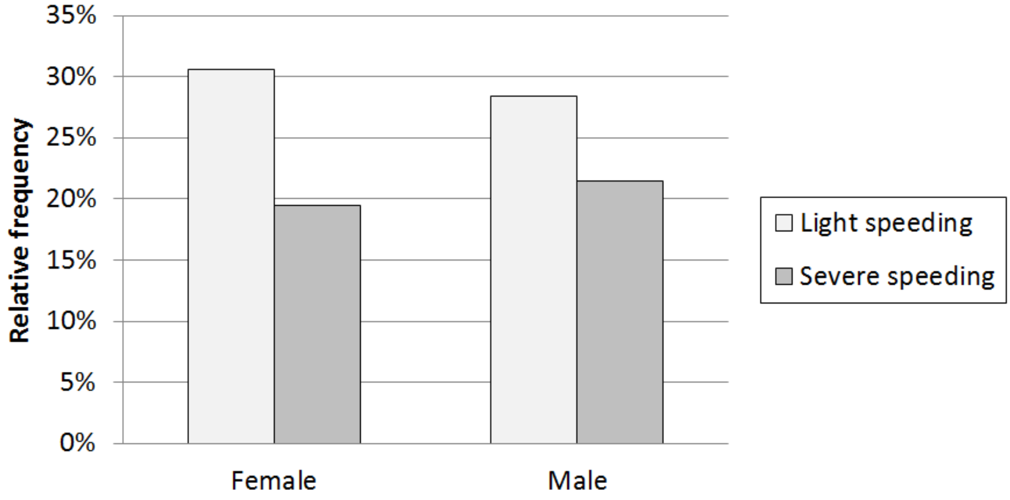


Figure 6-5: Frequencies of severity of speeding by gender weighed by duration driven per gender.

Effects of country on speeding

One-way ANOVAs revealed a significant effect of country on mean speed (F4, 11292) = 187.6, p < .001), maximum speed $F(4, 11292) = 193.6$, $p < .001$, and duration $F(4, 11292) = 60$, $p < .001$. Mean speed was lowest in Germany followed by France and the Netherlands. The UK showed the highest average speed while speeding. Maximum speed follows the same trend as mean speed (for means and standard deviations, please see Annex A.2.3). While speeding events took about 19 seconds in the France, Germany, and the UK, they took 21 seconds in Poland and 32 seconds in the Netherlands (Figure 6-6).

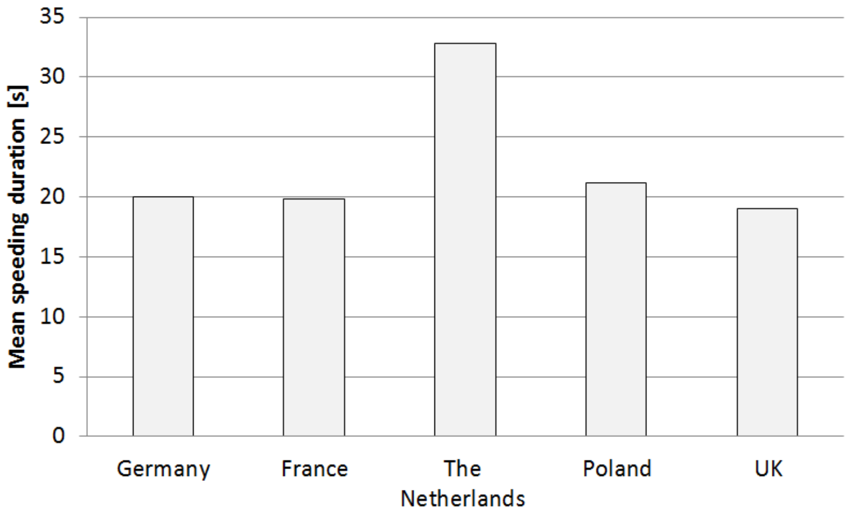


Figure 6-6: Mean speeding duration by country.

Chi-squared analysis also revealed a significant difference in frequencies in terms of speed limit category and country, $\chi^2(20) = 254.5$, $p < .001$. Table 6-5 below summarizes the relative frequencies per country. In comparison, in Germany, speeding was more frequent in the speed limit categories 0-30 km/h and 31-50 km/h compared to France, the Netherlands, Poland, and the UK, while speeding in the Netherlands was higher between 71-90 km/h. The UK showed the highest relative frequency in the speed limit category 91-111 km/h.

While speeding events in all countries declined with increasing speed limit, most speeding events occurred in France, followed by the UK and Poland. The least events occurred in The Netherlands and Germany (Figure 6-7).

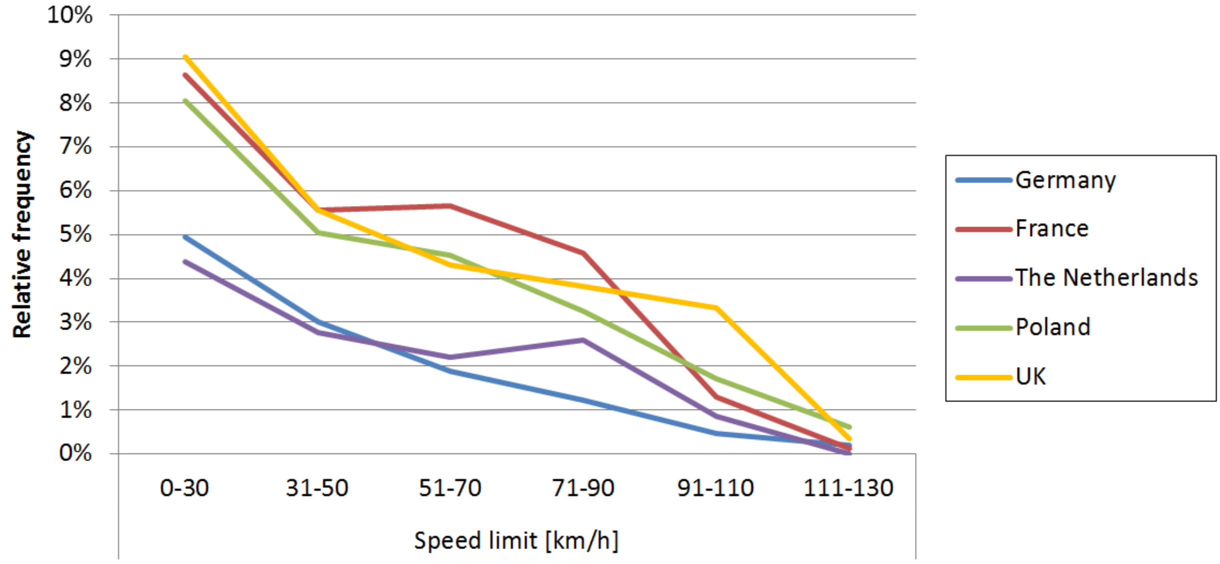


Figure 6-7: Frequency of speeding events per country by posted speed limit. Frequencies are weighed by duration of time driven in each country.

Table 6-5: Summary of relative frequencies (in percent) within country per speed limit category.

		Country				
		France	The Netherlands	Germany	Poland	UK
Speed limit category	0-30	33.4	34.2	42.2	34.6	34.3
	31-50	21.5	21.5	25.6	21.7	21
	51-70	21.8	17.1	16	19.5	13.6
	71-90	17.6	20.3	10.5	14	14.5
	91-110	5	6.7	3.9	7.4	12.6
	111-130	0.5	< 0.1	1.6	1.9	1.3

Chi-squared analysis revealed significant differences in the degree of speed exceedance per country, $\chi^2(4) = 177.4$, $p < .001$. In all countries, the number of speed exceedance was higher for 11-15% compared to 16-20%. Table 6-6 also shows that the proportion of speed exceedance in the category 11-15% was highest in the Netherlands, while excessive speeding was more prominent in France.

Table 6-6: Summary of frequencies (in percent) per country by degree of speeding.

		Country				
		France	The Netherlands	Germany	Poland	UK
Degree of speeding	Light	52.2	73.1	56.5	54.8	64.6
	Severe	47.8	29.9	43.5	45.2	35.4

Effects of time of day on speeding

ANOVA analysis also revealed significant effects of time of day on mean speed, $F(6, 11290) = 13.9$, $p < .001$, maximum speed $F(6, 11290) = 13.7$, $p < .001$, and duration, $F(6, 11290) = 6.4$, $p < .001$. Mean values of mean speed show that speed was above 100 km/h during late night, morning and pre-noon hours, while below 100 km/h during 12:00 pm and 0:00 am. Therefore, maximum speed was also higher between 0:00 am and 12:00 pm. Speeding events lasted about 19 seconds around noon and during night hours, while lasting around 20 seconds between 0:00 am and 12:00 pm. In the afternoon and evening, average speeding was done for about 23 seconds.

No significant effect was observed for time of day and speed limit category ($\chi^2(30) = 35.9$, p = .211). The relative frequencies in Table 6-7 show that the frequency speeding events was highest between 0-30 km/h across all day.

Table 6-7: Summary of frequencies (in percent) with day of time per speed limit category

		Time of day						
		late night	morning	Pre-noon	noon	afternoon	evening	night
Speed limit category	0-30	34.2	35.4	34	38.3	34.2	39	34.9
	31-50	21.9	21.9	21.5	23.5	21.6	23.1	21.8
	51-70	17.7	19.9	19.1	18	19.2	14.5	18.7
	71-90	15.6	14.4	16	14.5	16.3	15.2	15.6
	91-110	8.9	7.3	7.9	5.1	7.8	7.6	7.8
	111-130	1.7	1.1	1.5	0.6	1	0.7	1.2

Analysing the effects of time of day on degree of speeding revealed a significant effect, $\chi^2(6) = 18.9$, $p = .004$. Table 6-8 shows that the relative frequencies and distributions are similar across the day, except for the afternoon hours. More speed violations between 11-15% were committed than between 16-20%.

Table 6-8: Summary of degree of speeding by time of day (frequencies in percent).

		Time of day						
		late night	morning	Pre-noon	noon	afternoon	evening	night
Degree of speeding	Light	57.2	58.6	57.6	57.3	62.9	55.6	57.6
	Severe	42.8	41.4	42.4	42.7	37.1	44.4	42.4

In respect to time of day, speeding events were mostly observed during late night hours (0:00-6:59; ~32%). In addition, speeding was also frequently observed in the afternoon (13:00-18:00), but also in the morning (06:00-10:00; Figure 6-8). About 40% of all speeding events were observed during these two time frames.

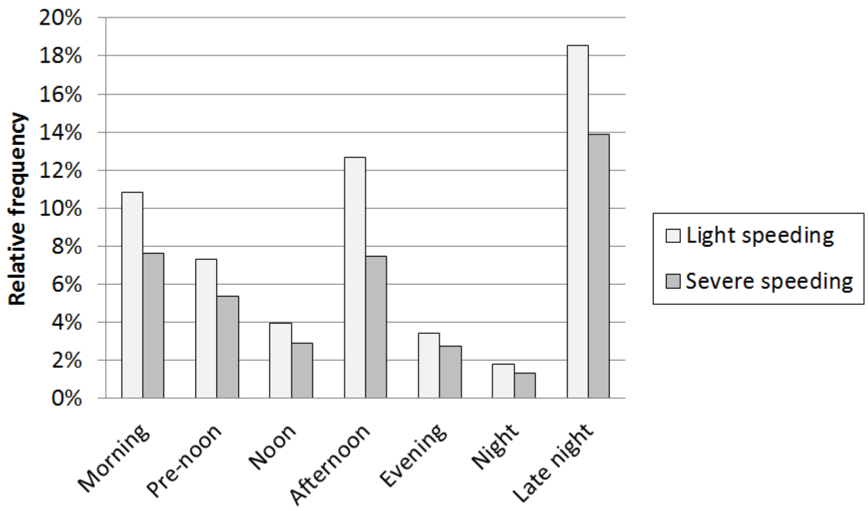


Figure 6-8: Relative frequency of speeding events by time of day. The frequencies are weighed by exposure of time driven during a specific time of day.

Effects of weather conditions on speeding

The presence of rain did not affect mean speed, maximum speed, and duration of the speeding event significantly, but more events were detected when it was not raining (n =10164) compared to when it was raining (n = 1133). In addition, chi-squared analyses of the effects of weather condition of the frequency per speed limit category as well as degree of speed exceedance did not result in significant differences.

Descriptive analysis of close following

Table 6-9 summarizes the frequencies of close following events. Altogether 120 close following events were found. More than 57% of close following events were observed in drivers over 50 years of age. Still more than 38% of drivers aged between 25 and 49 drove too close. The most close following events were observed in the UK (39%) and France (33%). In regard to gender, close following events were distributed equally.

Table 6-9: Descriptive overview of close following

		France		Germany		Poland		UK		The Netherlands	
	Gender	F	M	F	M	F	M	F	M	F	M
Age Group	18-24	1	1	0	0	0	0	1	3	0	0
	25-49	9	6	2	1	2	7	11	4	3	1
	50-99	11	11	2	4	1	6	14	14	3	2

Looking at close following events in regard to the posted speed limit, in can be seen that most close following events occurred within a speed limit of 30-50 km/h (N = 3223), accounting for almost 75% of the events. Only 10% of close following events were observed for the speed limit 50-70 km/h.

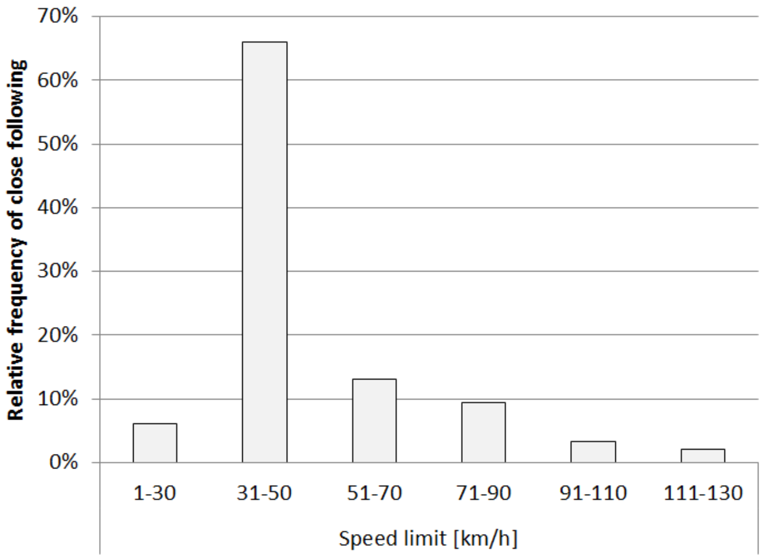


Figure 6-9: Frequency of close following events per posted speed limit weighed by duration driven under each speed limit.

No significant differences were found for age , weather and gender, but for duration, $t(117.88) = -2.21$, $p = 0.029$, $r = 0.2$. Male drivers followed closely longer (M = 2.34 s) then female drivers (M = 2.25 s).

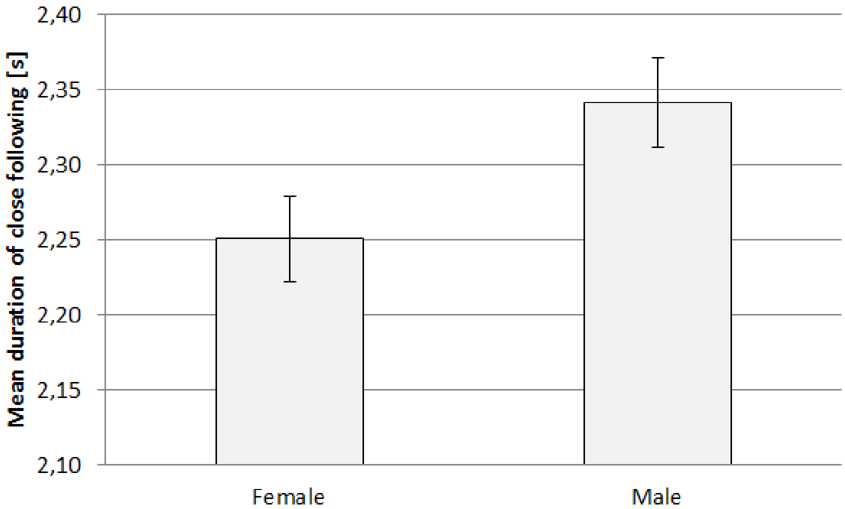


Figure 6-10: Mean duration for close following episodes by gender.

Regarding country, a Kruskal-Wallis Rank Sum test revealed a tendency for an effect of the minimum time headway (MinTHW), $\chi^2(4) = 8.2361$, $p = 0.083$. The minimal time headway was smaller in France (M = 1.13 s) and the Netherlands (M = 1.13 s) compared to Germany (M = 1.18 s), Poland (M = 1.19 s), and the UK (M = 1.16 s; Figure 6-11). Significant effects were found for the country.

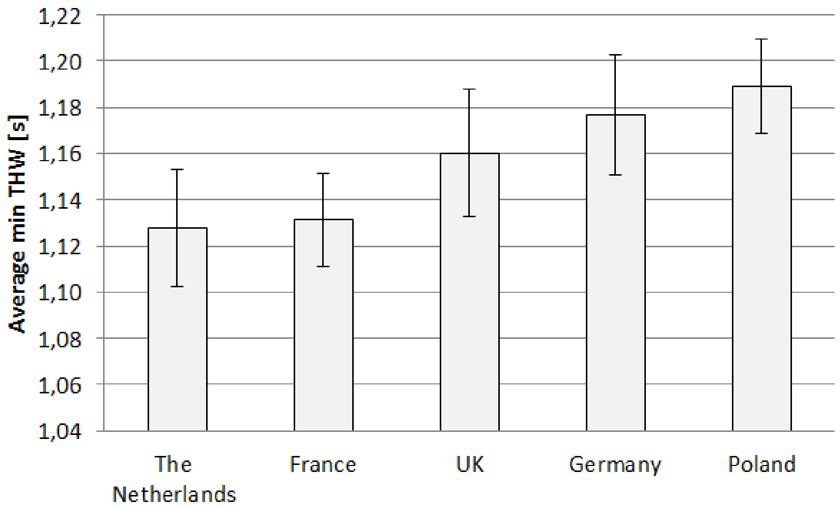


Figure 6-11: Average of the minimal time headway per operation site.

6.2.5 Discussion

The analysis of speeding behaviour showed that drivers were more tempted to drive faster than allowed when the speed limit was low. An almost linear decline towards higher speed limits indicates travelling speed preferences. It appears, drivers are to bend the regulations in favour of getting closer to this target. While doing this, drivers appear to be aware of the increased crash risk and/or penalties that go along with speeding. Light speeding events were 15-30% more frequent than severe ones. The only exception was the speed limit of 110-130 km/h. Frequencies were almost the same. Especially younger drivers drive faster and also speed longer compared to the other age groups. This finding is in line with existing literature about driving behaviour of young drivers who have less driving experience and drive more reckless than other drivers (Teese & Bradley, 2008). Even though speeding was observed within the female and male population, differences were found in their speeding behaviour. Females sped more cautiously meaning more light speed violations were seen at lower speed limits. Males, on the other hand, drove slower than females in general, but their violations were more severe. In general, males tend to have a higher probability of risk taking compared to females (Byrnes, Miller & Schafer, 1999). So, it is no surprise to find this relationship in NDS data as well.

Differences in speeding were observed between countries. While all countries had a higher ratio of light speed violations, France had the smallest offset closely followed by Poland. In both countries severe speeding was almost as frequent as light speeding. This indicates that drivers in other European countries are more cautious, compared to the French and Polish drivers. For some reason speeding events lasted about 10 s longer in the Netherlands than in any other country. It is unclear for now, whether this is due to differences in the road infrastructure or maybe the lived traffic culture. Further research is needed to identify a likely explanation for this effect.

The most popular times of day for speeding were late night (32%), the afternoon (20%), and the morning (18%). This observation may be explained by the fact that traffic density at night is low and the chance of being caught speeding is low. A clear road ahead may provoke speeding. Other prominent times of day were morning and afternoon: rush hours. It appears drivers ignore speed limits when they drive to and from work. Time pressure may contribute to speed limit violations. This hypothesis could be tested by comparing the speeding events during the week with weekends. Since the UDRIVE dataset will be available after the project, this can be addressed in follow-up analyses. Overall, speeding results indicate that drivers are interpreting the regulations depending on their current assessment of the situation and their current needs. Further research is needed to verify this theory.

Contrary to the speeding events, close following was observed more frequently in drivers over 50 years old with the highest numbers in the UK and France. This appears to contradict other findings that youngsters are the risk takers (Teese & Bradley, 2008). However, it is not unheard of, since the effect was observed before. Rajalin, Hassel and Summala (1997) found that drivers between 35-54 years had the most close following events (~53%). However, it is unclear why close following is more common in drivers in this age group. The effect appears to wane again with increasing age. In the study of Rajalin et al. (1997), drivers over 55 years accounted for the least amount of close following violations (9%). It is unlikely that general risk taking is responsible, since the age group did not stand out in the speeding analysis. Also, effects on eyesight or cognitive abilities due to age manifest themselves later in life. A closer inspection of the circumstances is needed to understand this relationship. Most close following events were observed within posted speed limits of 31-50 km/h. Speed limits between 31 and 50 km/h are common on main roads in urban areas. Those main roads are also characterized by high traffic volume and their proneness to congestions. Commuting bumper to bumper may explain the high frequency of close following events. Low speed zones are often minor roads in urban areas with far less traffic volume. This may explain the low number of close following events. In addition, as seen in Figure 6-9, the higher the posted speed limit is the lower the observed frequency of close following events. When being outside of urban areas, it appears that drivers adjust their following distances to travelling speed (i.e., the higher the speed, the greater the following distance).

6.3 Research question 2.2: To what extent are driver personality factors associated with risky behaviour?

6.3.1 Introduction

This section looks at the impact of personality on risky driving behaviours. This considers both individual personality components (as measured by a range of questionnaires completed by drivers at the outset of the study) and also a summary measure of driver personality – a negative driver personality traits score – calculated from the suite of personality metrics applied. The potential utility of this score as a means of predicting risky driving behaviours is considered.

Drivers were presented with a suite of questionnaires at the beginning of the project, including:

- Driver Attitude Questionnaire (DAQ) (20 items assessing attitudes towards speeding and close following behaviours);
- Driver Behaviour Questionnaire (DBQ) (19 items assessing the prevalence of errors and violations in the driver's everyday behaviours);
- Driver Skills Questionnaire (DSQ) (15 items assessing how drivers behaviour in a series of described driving scenarios including speeding behaviour, travelling with passengers, engaging with distractions, journey planning etc.)
- Traffic Locus of Control (TLOC) Questionnaire (17 items assessing views towards the factors that cause road accidents);
- Arnett Inventory of Sensation Seeking (AISS) (20 items assessing the risk-taking and sensation-seeking nature of a driver's personality).

The five questionnaires above have been subjected to factor analysis in prior work, which has been assumed valid for the purposes of this subsequent analysis (Warner et al., 2010; Department for Transport, 2005; Özkan & Lajunen, 2005; Lajunen et al., 2004; Parker et al., 1996; Arnett, 1994; French et al., 1993; West et al., 1993). The subscales derived from factor analysis have been used as independent variables to represent attributes of driver personality.

For each subscale of a questionnaire (e.g. Driver Attitude Questionnaire, speeding subscale) and for the overall summary scale (e.g. DAQ, overall), drivers were ranked by their score before being split into two groups. The break point between the two groups was selected such that all drivers fell into one group or the other, with no overlap in scores between the two groups. In some cases, this results in slight disparities in the number of members of these two groups. This process was done so as to dichotomise a continuous variable for subsequent exploratory data analysis (Table 6-10). This was selected in preference to a median split as it resulted in more balanced groups in terms of the number of members. These groups were as follows:

Table 6-10: Splitting driver sample along personality dimensions

Questionnaire sub-scale	Group membership				Group description
	Group 1: Mean score /# members		Group 2: Mean score /# members		
DAQ speeding	27.8	46	38.2	46	High score = more negative attitude towards speeding
DAQ close following	34.4	49	43.1	43	High score = more negative attitude towards close following
DAQ overall	62.6	45	79.9	47	High score = more negative attitude towards speeding and close following behaviours
DBQ errors	1.4	49	2.0	39	High score = more reported driving errors
DBQ aggressive violations	1.5	42	2.4	36	High score = more reported aggressive driving violations
DBQ ordinary violations	1.2	38	2.2	53	High score = more reported ordinary driving violations
DBQ all violations	1.5	38	2.3	40	High score = more reported driving violations
DSQ speed	6.1	43	11.2	49	High score = more reported speeding behaviour
DSQ calmness	11.3	39	15.5	52	High score = more reported calm driving behaviour
DSQ social resistance	5.7	52	9.6	38	High score = more reported resistance to others' advice
DSQ focus	10.2	52	14.8	38	High score = more reported cautious driving and resistance to distraction
DSQ planning	6.1	48	10.4	45	High score = more reported planning ahead before and during driving
DSQ deviance	2.0	35	4.0	58	High score = more reported rule-breaking and deviant behaviours
TLOC self	2.2	42	3.5	50	High score = rate their own driving as contributing to the cause of road accidents
TLOC other	3.6	47	4.5	45	High score = rate the driving of others as contributing to the cause of road accidents
TLOC vehicle and environment	2.9	47	4.0	44	High score = rate vehicle and environmental factors as contributing to the cause of road accidents
TLOC fate	1.9	46	3.3	45	High score = rate coincidence or fate as contributing to the cause of road accidents
AISS novelty	22.4	48	29.3	42	High score = drivers seek out novel experiences

AISS intensity	18.4	50	25.9	42	High score = drivers seek out high intensity experiences
AISS overall	41.8	42	52.6	47	High score = drivers seek out highly novel or high intensity experiences

Composite personality measure

The composite personality score includes six questionnaires given to drivers at the outset of the study, measuring driver attitudes, driver behaviour, driving skill, traffic locus of control, sensation-seeking and driving history. The raw data for each of these questionnaires involved an ordinal scale with 4-6 intervals. These were converted to a three-point scale in each case as detailed in Table 6-11. The cut-off points for the revised scale were determined based on inspection of the raw data to assess which response categories were logically grouped together or infrequently used. The objective was to produce a new scale from 0-2, which could be considered to have three distinct categories (e.g. negative, neutral, positive). All scales were recoded such that a high score reflected the most negative driver attitude or behaviour, such as high acceptance of close-following behaviour, high reported speeding, high reported mobile phone use.

Table 6-11: Overview of applied questionnaires and scale revisions.

Questionnaire	Original scale	Revised scale
Driver Attitudes Questionnaire	1 - I strongly disagree 2 - I disagree 3 - I neither agree or disagree 4 - I agree 5 - I strongly agree	0 – Disagree 1 – Neutral 2 – Agree
Driver Behaviour Questionnaire	1 - Never 2 - Hardly ever 3 - Occasionally 4 - Quite often 5 - Frequently 6 - Nearly all the time	0 – Never 1 – Rarely 2 – Regularly
Driver Skills Questionnaire	1 – Never or very infrequently 2 - Infrequently 3 – Quite infrequently 4 – Quite frequently 5 - Frequently 6 – Very frequently or always	0 – Never 1 – Rarely 2 – Regularly
Traffic Locus of Control Scale	1 – Not at all possible 2 – Not fairly possible 3 - Possible 4 – Fairly possible 5 – Highly possible	0 – Not a factor in accident causation 1 – Weak factor in accident causation 2 – Strong factor in accident causation
Arnett Inventory of Sensation Seeking	1 – Does not describe me at all 2 – Does not describe me very well 3 – Describes me somewhat 4 – Describes me very well	0 – Does not describe me 1 – Describes me to some extent 2 – Describes me well
Driving History Question – Driver at Fault Accidents	0 1 2, 3, 4, or 5	0 – No accidents 1 – 1 accident 2 – More than 1 accident
Driving History Question – Reported bad	1 - Never 2 - Hardly ever	0 – Never 1 – Rarely

behaviours	3 - Occasionally 4 - Quite often 5 - Frequently 6 - Very often	2 – Regularly

This recoding procedure resulted in 17 personality subscales scored on a 0-2 scale. Traffic locus of control and sensation-seeking are not inherently positive or negative characteristics so these 5 subscales were removed from further consideration. The remaining twelve subscale scores relating to driver behaviour, attitudes and skills were summed to give a personality score out of 24. This score was used as a continuous variable in subsequent analyses. Ninety-three drivers with trips annotated for secondary task engagement submitted sufficient personality data to calculate a summary score. The distribution of scores amongst the sample is positively skewed with a mean personality score of 7.6 (range: 2.7-14.4; IQR: 6.2-8.8). High scores reflect high reporting of negative behaviours, attitudes and skills. The dataset shows few instances of drivers giving themselves high ratings, suggesting a reluctance to criticise themselves during self-report.

6.3.2 Rural over-taking and personality

4551 rural overtaking manoeuvres were identified in the database. Of the 197 unique drivers involved in the project, 143 performed at least one rural overtaking manoeuvre during the data collection phase, with 137 of these drivers providing a complete subjective dataset. Drivers were categorised based on whether they performed an overtaking manoeuvre (0 = No, 1 = Yes). These groups were subjected to independent samples t-tests to determine whether there was any difference in the dependent variable (overall personality score or personality metric subscale) for two between-subjects groups (over-takers vs. non-over-takers). Prior exploratory data analysis showed the dependent variable to be normally distributed (Kolmogorov-Smirnov test) hence parametric analysis methods were applied.

There was no significant difference in overall personality score between the over-taking and the non-overtaking groups. When considering the individual questionnaire subscales, there were significant differences between the overtaking and non-overtaking groups on the subscales listed in

Table 6-12: Personality subscale scores comparison between overtaking and non-overtaking drivers

Questionnaire subscale	No overtake group score	Overtake group score	p-value	Scale direction
DAQ Speeding	30.84	33.02	0.048	High score = more negative attitude towards speeding behaviours
DAQ Close Following	36.00	38.65	0.010	High score = more negative attitude towards speeding behaviours
DBQ Errors	1.50	1.63	0.043	High score = more reported driving errors
DSQ Speeding	8.19	9.03	0.023	High score = more reported speeding behaviours
DSQ Focus	11.29	12.10	0.038	High score = higher reported focus on driving task
TLOC Other	3.76	4.00	0.049	High score = high reported contribution of other road users to accidents
AISS Intensity	21.82	23.12	0.015	High score = higher reported sensation seeking behaviours

Drivers who performed at least one rural over-taking manoeuvre reported more speeding behaviour on the DSQ than non-overtaking drivers. This finding suggests that drivers who have a propensity to speed are more likely to engage in over-taking manoeuvres. However, surprisingly, drivers who performed a rural over-taking manoeuvre reported significantly higher negative attitudes towards speeding than those individuals who did not engage in an over-taking manoeuvre. Of course, it is not necessarily the case that an over-taking manoeuvre will involve exceedance of the speed limit, but these two findings do appear to contradict each other to some extent. It was also observed that drivers who performed an over-taking manoeuvre had significantly higher negative attitudes towards close-following behaviours. It could be posited that these drivers are choosing to over-take a slow moving vehicle rather than following it at a short headway.

Drivers who performed at least one rural over-taking manoeuvre reported significantly more errors on the driver behaviour questionnaire (DBQ). Two of the error-related items on the DBQ refer specifically to the misjudgement of the speed or turning manoeuvre of surrounding vehicles whilst over-taking, so it is concerning that these individuals seem to be more likely to engage in over-taking manoeuvres.

Drivers who performed at least one rural over-taking manoeuvre also scored higher on the 'Intensity' subscale of the sensation-seeking questions. This suggests that individuals with a high risk-seeking propensity are also those who are most likely to overtake on rural roads. Drivers who overtook also scored more highly on the Focus subscale of the DSQ, indicating that a greater reported ability to focus on the driving task and ignore distractions than their non-overtaking counterparts. It could be that a higher focus on the driving task indicates higher driving enjoyment, or lower willingness to perform competing tasks. Rather than being more dangerous, it could be suggested that these more focused individuals are those who may perform more safely if and when they do choose to overtake.

Drivers who performed at least one rural over-taking manoeuvre scored more highly on the traffic locus of control 'Other' subscale than their non-overtaking counterparts. It could be that a disassociation of the link between the driver's own actions and accident causation can lead to a greater confidence in their ability to safely perform an overtaking manoeuvre on a rural road.

Overall, it is interest to observe that particular negative personality traits seem to be linked to a willingness to engage in an overtaking manoeuvre. It would seem that questionnaire measures of risk-taking, speeding and driving errors can indicate propensity to overtake on rural roads to some extent.

A cross-tabulation exploratory analysis was used to identify behavioural trends for each driver personality group. Drivers in each personality category were subsequently categorised simply as over-takers or non-over-takers, depending on whether they performed an overtaking manoeuvre during any drive (Table 6-13).

Table 6-13: Percentage of drivers who performed a rural over-taking manoeuvre by personality sub-group

Questionnaire subscale	Group 1 overtakers (%)	Group 2 overtakers (%)	Difference (Group 2 % Use – Group 1 % Use)
DAQ speeding	68	78	10
DAQ close following	64	83	19
DBQ errors	69	74	5
DBQ aggressive violations	66	79	13
DBQ ordinary violations	70	71	1
DSQ speed	69	77	8
DSQ calmness	75	69	-6
DSQ social resistance	72	73	1
DSQ focus	69	75	6
DSQ planning	74	71	-3
DSQ deviance	69	75	6
TLOC self	67	78	11
TLOC other	67	76	9
TLOC vehicle and environment	69	77	8
TLOC fate	73	72	-1
AISS novelty	74	70	4
AISS intensity	79	65	13

Table 6-13 suggests that driver personality can affect whether or not drivers choose to over-take on rural roads:

- High negative attitudes towards both speeding and close following are linked to a higher likelihood of overtaking on rural roads.
- Drivers who report a higher number of aggressive violations are more likely to overtake on rural roads. Aggressive violations involve expressing annoyance or dissatisfaction towards other road users. If over-taking is often an act of frustration or annoyance at the speed choice of the lead vehicle, then this trend makes sense. The same pattern is not observed for the ordinary violations subgroups.

- Drivers who report more speeding behaviour on the DSQ show higher propensity to over-take on rural roads compared to those drivers who report less speeding behaviours.

Overall, then the suite of driver personality questionnaires appear to have some value for predicting driver overtaking behaviour on rural roads.

6.3.3 Over-speeding by 16-20kmh and personality

Excessive speed is a significant contributory fact in crashes; hence it was investigated as an example of risky driver behaviour. An over-speeding event was classified as any instance speed limit exceedance, where drivers were driving in the range of 16-20kmh over the speed limit for greater than 10s. Each instance where these conditions were true was recorded as a new over-speeding event. 12,111 over-speeding events were extracted from the database for analysis. Within this sample, 169 out of 197 drivers were represented. A number of summary statistics were calculated for the 85.7% of the driver sample who performed at least one over-speeding event of this type (Table 6-14).

Table 6-14: Summary statistics for participants who over-speed by 16-20kmh

Statistic (per participant)	**Minimum**	**Maximum**
# over-speeding events	1	652
Mean duration of over-speeding event	10.8s	75.8s
Time proportion over-speeding	0.25%	5.69%
# over-speeding events/hr	0.45	6.86

The participant group were categorised based on whether they were an over-speeder (1) or not (0). The personality scores of these two categories were compared for both personality questionnaire subscales and the composite personality measure described above. This analysis showed that those drivers who performed an over-speeding event gave significantly higher ratings on the DSQ speed subscale (three items measuring self-reported speeding generally, on motorways and in built-up areas), $t(33.664) = -2.354$, $p=.025$. This shows that drivers who report speeding behaviours are more likely to drive at excessive speed on the roads. No further significant effects were observed in this analysis.

The threshold for being categorised as an 'over-speeder' was adjusted to reflect that a single over-speeding instance may not be indicative of a driver's personality or habitual behaviour. Two further thresholds were applied whereby over-speeders were categorised as committing at least 20 over-speeding events or at least 50 over-speeding events.

In the former analysis, those drivers who committed more than 20 over-speeding events had a significantly higher negative personality traits score (Mean = 7.96, SD = 2.18) than those drivers who had fewer over-speeding events (Mean = 7.23, SD = 2.12). This suggests that drivers who are more prone to exceed the speed limit are more likely to exhibit other negative driving personality traits. The over-speeding group also gave significantly greater scores on the DSQ Speed subscale (Mean = 9.60, SD = 3.15) than the group that was less prone to over-speeding (Mean = 7.91, SD = 3.00) indicating that drivers are aware of their over-speeding behaviour and thus suggesting that it they may be choosing to exceed the speed limit by such a large amount. Drivers with more over-speeding events also gave higher ratings on the Fate subscale of the Traffic Locus of Control questionnaire (Mean = 2.68, SD = 0.91) than the less frequent over-speeders (Mean = 2.31, SD = 0.82). This may suggest that the more frequent over-speeders perceive less of a connection between their own actions and the cause of road accidents, and thus may be more willing to exceed the speed limit significantly.

When considering the participant group categorised based on whether the number of over-speeding events was more or less than 50, a similar pattern of results is observed, although the difference is more marked in many cases. The difference in negative personality traits score remained, with more frequent over-speeders obtaining a significantly higher score (Mean = 8.38, SD = 2.20) than the less frequent over-speeders (Mean = 7.29, SD = 2.10). The frequent over-speeder group also reported more aggressive violations (Mean = 1.85, SD = 0.76) and ordinary violations (Mean = 2.07, SD = 0.64) on the DBQ than the less frequent over-speeder group (Mean = 1.64, SD = 0.61; and Mean = 1.84, SD = 0.52 respectively). This is an interesting finding as the speed-related violations are confined to the ordinary violations subscale, suggesting that a driver who excessively speeds more often is also more likely to engage in aggressive violations (such as showing anger or annoyance to other road users). This ties in with the significant different in negative personality score, showing that drivers are likely to commit multiple bad behaviours if they commit one. The drivers with more over-speeding offences also reported high levels of speeding (Mean = 10.23, SD = 3.22) and deviant behaviours (Mean = 3.74, SD = 1.65) via the DSQ compared to the less frequent over-speeders (Mean = 8.20, SD = 2.98; and Mean = 3.08, SD = 1.23 respectively). In keeping with this pattern, the frequent over-speeders group also reported more bad behaviours in a questionnaire reporting on past at-fault accident occurrence, speeding, red light violations, drink driving, mobile phone use and non-use of seat belts.

Overall, the analysis of over-speeding events shows that there appears to be a link between driver personality and likelihood of exceeding the speed limit by 16-20kmh. Furthermore, the frequency of over-speeding becomes greater in those drivers that report more negative driving behaviour traits, specifically speeding behaviours, but also other types of violations and risky behaviours.

A cross-tabulation exploratory analysis was used to identify behavioural trends for each driver personality group. Drivers were split into two groups along each personality subscale. The proportion of drivers in each group who performed greater than 20 over-speed events was calculated (Table 6-15).

Table 6-15: Percentage of drivers who performed greater than 20 over-speeding manoeuvres by personality sub-group

Questionnaire subscale	Group 1 over-speeders (%)	Group 2 over-speeders (%)	Difference (Group 2 % Use – Group 1 % Use)
Composite personality score	48	59	11
DAQ speeding	57	51	-6
DAQ close following	54	53	-1
DBQ errors	55	52	-3
DBQ aggressive violations	55	53	-2
DBQ ordinary violations	51	57	6
DSQ speed	43	62	19
DSQ calmness	52	54	2
DSQ social resistance	54	51	-3
DSQ focus	51	54	3
DSQ planning	56	51	-5
DSQ deviance	47	58	11
TLOC self	52	54	2
TLOC other	54	52	-2
TLOC vehicle and environment	55	50	-5
TLOC fate	45	50	5
AISS novelty	57	48	-9

AISS intensity	52	54	2

Table 6-15 suggests that driver personality is linked to likelihood of over-speeding by 16-20kmh:

- The group of drivers with higher negative personality scores showed a 11% higher incidence of more frequent over-speeding.
- Drivers who report more speeding behaviour on the DSQ show higher propensity to over-speed by 16-20kmh more frequently compared to those drivers who report less speeding behaviours.
- Drivers who report more deviant behaviours on the DSQ show higher propensity to over-speed by 16-20kmh more frequently compared to those drivers who report less deviant behaviours.

Self-reported speeding behaviour appears to be a good indicator of the likelihood that a driver will frequently overshoot the speed limit of the road by 16-20kmh.

6.3.4 Discussion

Driver personality types were investigated for two types of risky driving behaviours: exceeding the speed limit – specifically speed limit exceedances of 16% or greater – and overtaking on rural roads. For rural road overtaking, this is considered to be a rare event, so drivers were categorised based on whether or not they had performed an overtaking manoeuvre at any point in the recorded dataset. Drivers who performed an overtaking manoeuvre exhibited certain differences in personality traits when compared with those who did not overtake. Overtaking drivers reported significantly higher levels of speeding behaviours, indicating that their over-speeding tendencies are not simply confined to instances where they are overtaking a slow moving leader. The overtaking drivers also showed high sensation-seeking tendencies, which is an expected finding – drivers who score highly on a scale that measures risk-taking are also significantly more likely to perform a manoeuvre that has the potential to bring them into the path of oncoming traffic. The overtaking drivers also exhibited significantly stronger negative attitudes towards speeding and close following (as indexed by high DAQ scores). This may suggest that they perceive others' unsafe behaviour more negatively than their own. It would be useful to gain some measure of an individual's driving (over)confidence to more fully understand the discrepancy between attitudes towards speeding and the driver's own tendency to speed. It could be that this effect is partly explained by the significantly higher TLOC 'Other' scores seen in the overtaking group. Again, the propensity to overtake may be underpinned by the driver not perceiving a strong a link between their own behaviours and the likelihood of accident occurrence, compared to the non-overtaking group. It is a potential cause for concern that the drivers who performed an overtaking manoeuvre reported more speed misjudgement and turning manoeuvre errors than the non-overtaking group. The combination of high sensation-seeking and high error rate are likely to raise the likelihood of an overtaking-related accident amongst this subset of the driving sample. These drivers could represent a target group for remedial actions such as speed and vehicle control education. Encouragingly, the overtaking group drivers had higher ratings for focus (and ignoring distractions), suggesting that these individuals may be more engaged with the primary driving task. It is possible that focused attention may mitigate some of the aforementioned negative behavioural traits that could inflate accident risk.

Exceeding the speed limit is a common event in the UDRIVE database, hence risky behaviour was defined as greater than 20 exceedances of the speed limit by more than 16% for longer than 10 seconds. This threshold was applied so as to capture habitual rather than one-off risky behaviour. Drivers were categorised into categories based on whether they fulfilled this criterion or not. Drivers who were categorised as 'speeders' were shown to self-report significantly more speeding behaviours. The fact that drivers are aware of their speeding behaviour to some extent, demonstrates a lack of awareness of the potential negative consequences of excessive speed and highlights that an educational or enforcement countermeasure may be required to trigger a behavioural improvement. Interestingly, speeders also showed significantly higher scores on other measures of negative personality traits, including the overall negative personality score. This suggests that drivers who perform one risky behaviour may also be those who are more likely to commit additional risky behaviours. This is supported by the higher reporting of general deviant behaviours and additional aggressive violations by the speeder group relative to the non-speeder group. An analysis of the overlap between the driver samples for each risky behaviour analysis would allow conclusions to be drawn about whether drivers are 'risky' in general, or whether individual risky behaviours can occur in isolation. Of some concern is also the fact that speeders appear to be more likely to attribute the cause of traffic accidents to bad luck or fate, meaning that those individuals who are engaging in more risky driving are also those who are less able to connect the actions with negative safety-related outcomes such as accidents.

Overall, it appears that subjective measures of personality may have some utility in predicting real-world overtaking and speeding behaviours. Specifically, metrics that assess speeding behaviour and sensation-seeking behaviour may be useful in identifying those drivers who are more willing to perform risky overtaking manoeuvres, whilst self-reported speeding behaviour and other deviant behaviours more generally may give an indication of an individual's propensity to regularly brake the speed limit. Both overtaking and speeding may be linked to some extent to an individual's with high sensation-seeking tendencies and traffic locus of control. However, it should be noted that these findings require further investigation across a larger dataset to more conclusively link personality traits to the occurrence of risky behaviours.

6.4 Research question 2.3: To what extent are driver assistance systems used?

6.4.1 Introduction

In this research question, we want to describe to what extent are driver assistance systems used and in particular, where do drivers use ADAS (e.g. urban, rural, motorway etc.)?

To better understand ADAS use, we will also study what drivers really know about their equipment from the questionnaire they all had to answer at the beginning of the recordings.

The Cruise control/ speed limiter is the only ADAS available in the UDRIVE cars. The questionnaire also included question about other existing ADAS such as Lane Departure warning (LDW) and automatic high beam/ low beam ADAS even if none of the studied cars was equipped. It will produce information about driver knowledge on such systems.

6.4.2 Method

The analysis was conducted over the data set available on March 15th 2017, consequently, the countries and number of drivers are unbalanced. The first step consisted in excluding the trips of drivers that are not equipped with CC/SL. 83% of the cars in the dataset was equipped with CC/SL. The total number of trip is 45 344.

For each trip, CC or SL use is coded as 1 if the ADAS was used at least once during the trip.

Road type classification was based on map data (max speed) as well as speed limit in each country. The variable was validated by each country. The Road types coding is presented in **Error! Reference source not found.**.

Table 6-16 Road type coding

Road type	Coding
Undefined	0
Urban road	1
Rural	2
Country motorway	3

Urban motorway	4
Slip road	5

2.2% of the segments composing the trips in our data set have undefined road type and have been excluded from analysis.

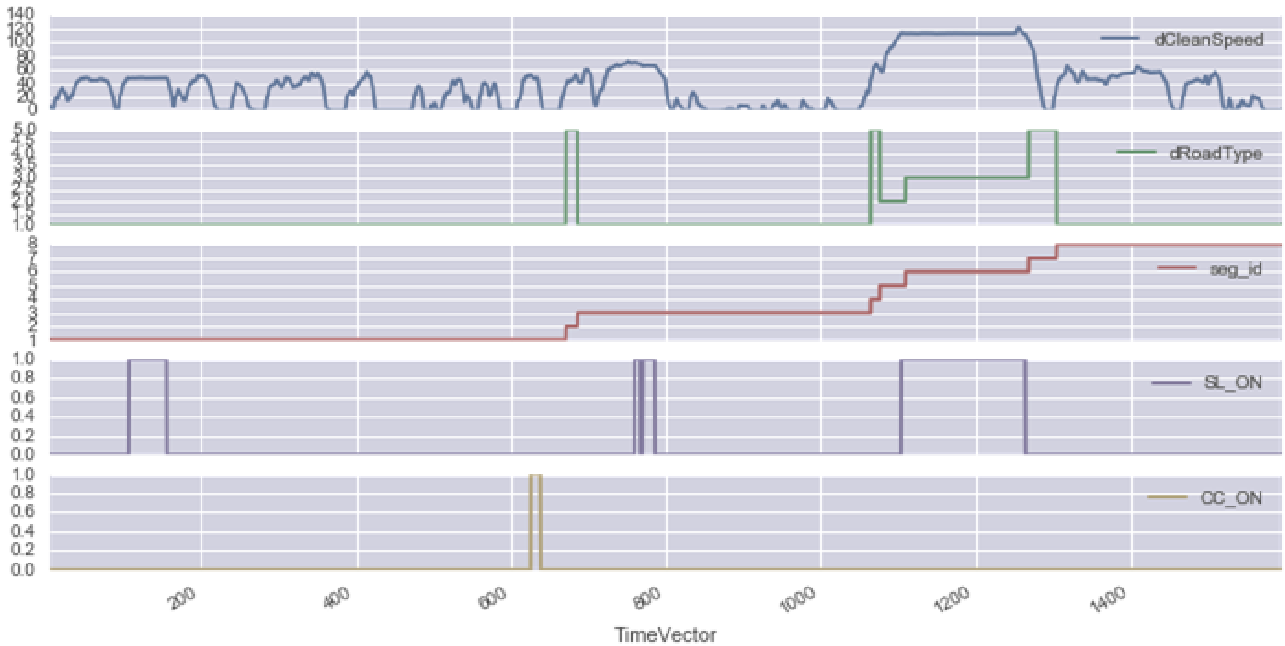


Figure 6-12: Sample view of data used for the analysis. X axis is time and each window shows one data: speed (CAN data), road type (calculated), segment id (map data), SL_ON, CC_on

Figure 6-12 shows a sample of the data used for our analysis. The segment is defined by the map data. For each segment, the following features are calculated: road type, duration, distance, CC activation: 0 if CC was never ON during the segment, else 1, SL activation: 0 if SL was never ON during the segment, else 1. Figure 6-13 shows the distribution of segment road types in our data set. Urban motorway represents a small part of the data and the greater part of the trips includes more urban roads (number of segments, time and distance driven on that type of road).

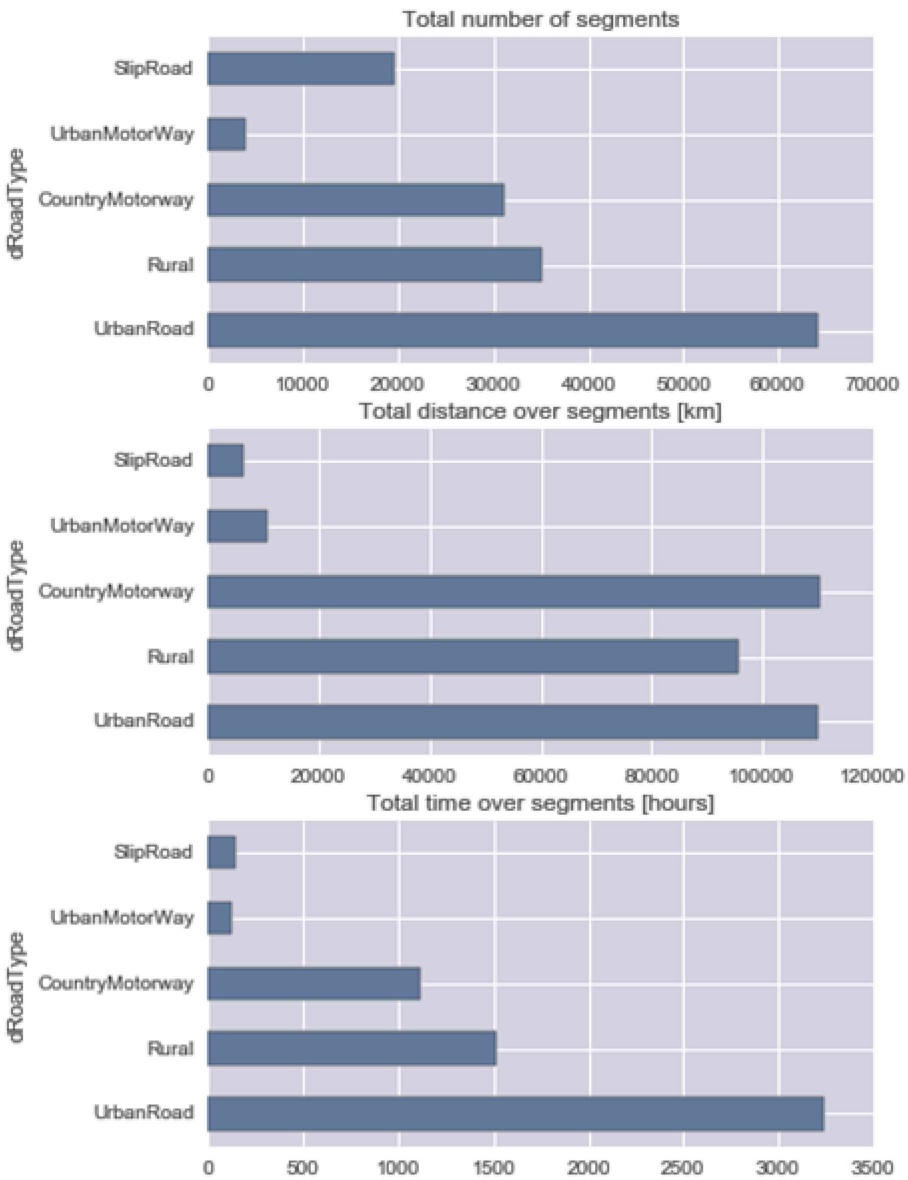


Figure 6-13: Distributions (in number of segment, distance, time) of segments road types

The distribution of the data set in terms of trip per driver is presented in Figure 6-14. Each bar plot represents the number of trip for a driver. Drivers were ranked by number of trip.

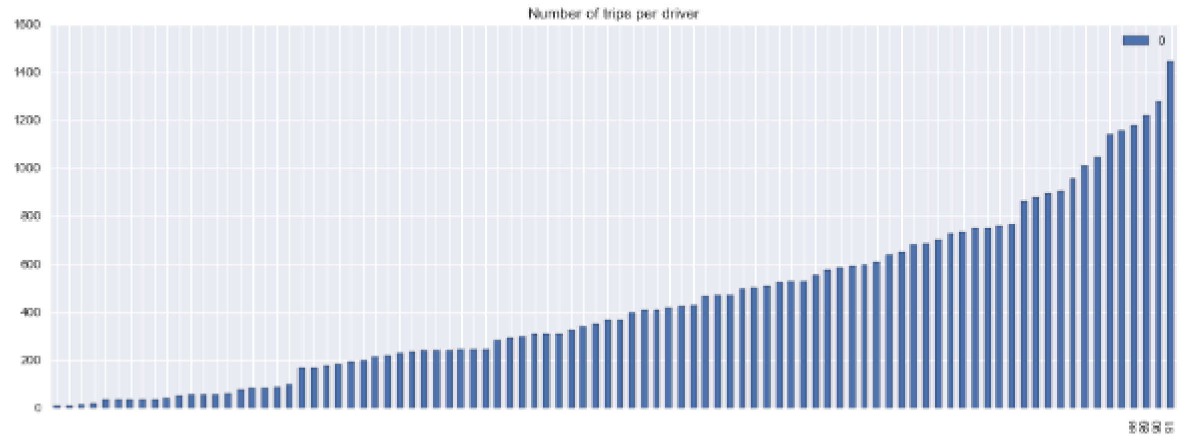


Figure 6-14: Distribution of number of trips for each driver

6.4.3 Results

<u>CC/SL usage frequency</u>

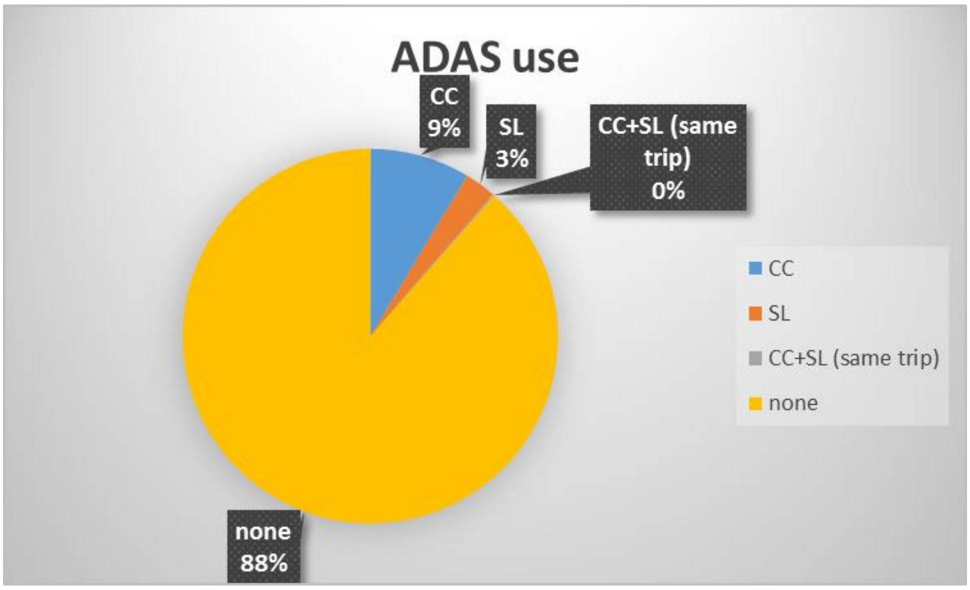


Figure 6-15: Distribution of ADAS used for trips (equipped drivers)

Figure 6-15 shows that 88% of the trips of equipped drivers were done without any activation of CC or SL. Only a small number of trips included CC AND SL activation (0.14%). Finally, SL is used only in 3% of trips.

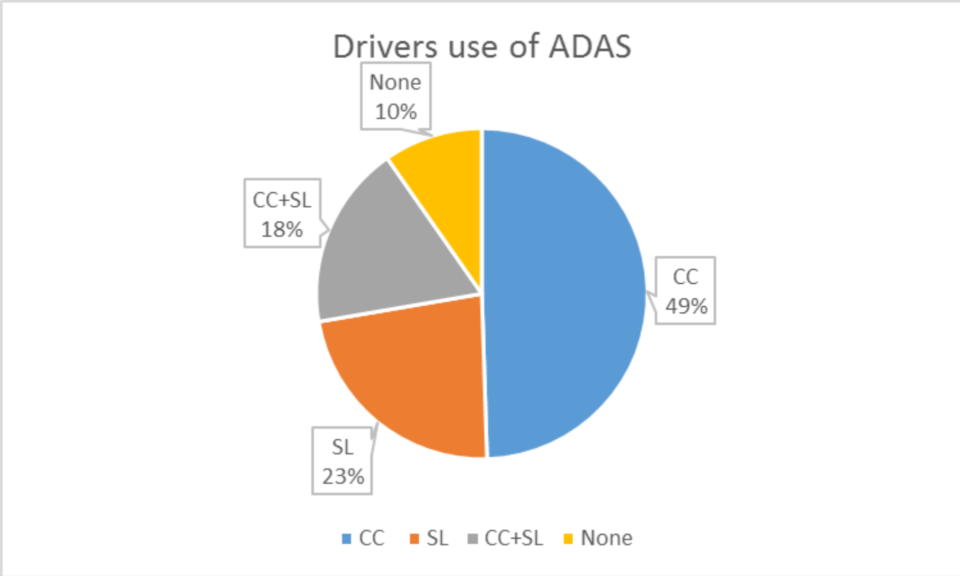


Figure 6-16: Repartition of ADAS use for drivers

49.5% of the drivers used CC whereas only 22.8% used SL (Figure 6-16). Only 18% of drivers use both CC and SL in their driving. This shows that drivers tend to specialize in the use of one specific ADAS and that cruise control is more used than speed limiter.

<u>Trip characteristics of CC/SL usage</u>

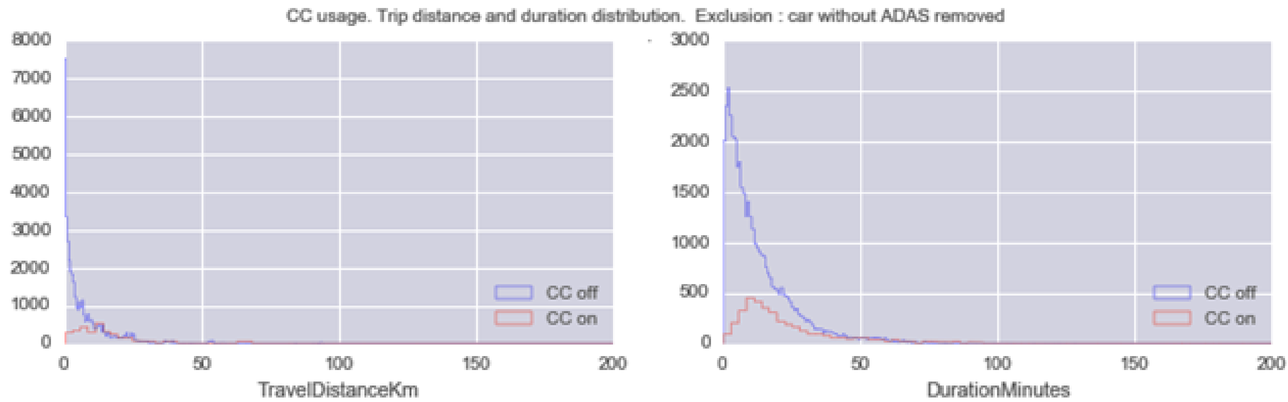


Figure 6-17: Distribution of trip distance and duration when CC ON

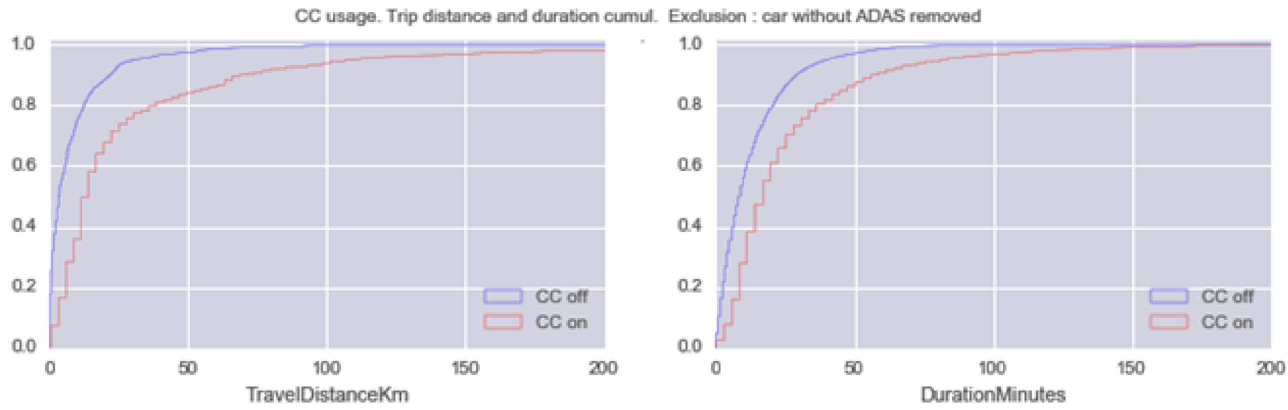


Figure 6-18: Cumulative plot of total distance and duration of trips with CC on

Figure 6-17 and Figure 6-18 describe the type of trip (distance and total time) where CC is used. Both total distance and total time of trips including CC used are higher than when CC was not used. For example, 80% of the trips with CC activation are approximately 40km long whereas 80% of the trips without any activation are approximately 10 km long. The same can be seen for total time of the trip leading to the conclusion that drivers use CC on longer trips.

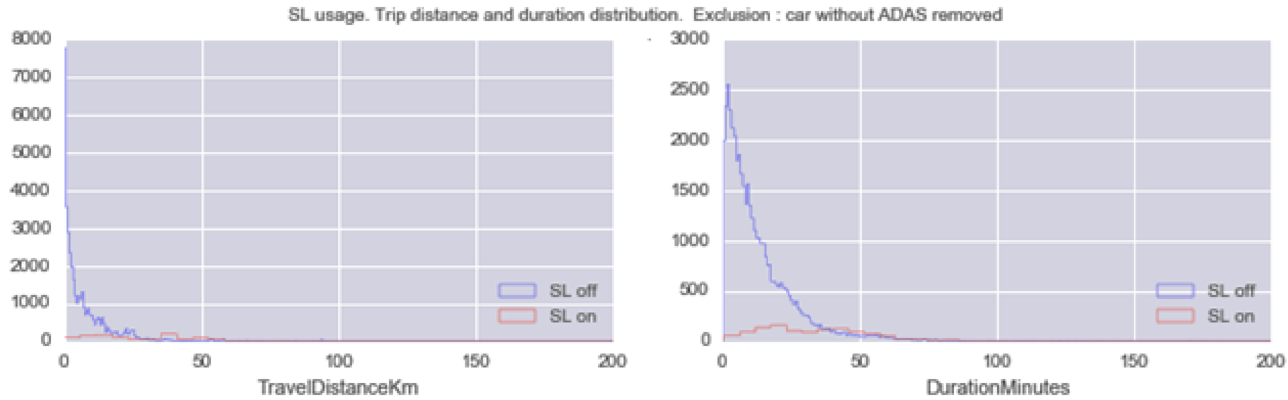


Figure 6-19: Distribution of trip distance and duration when SL on

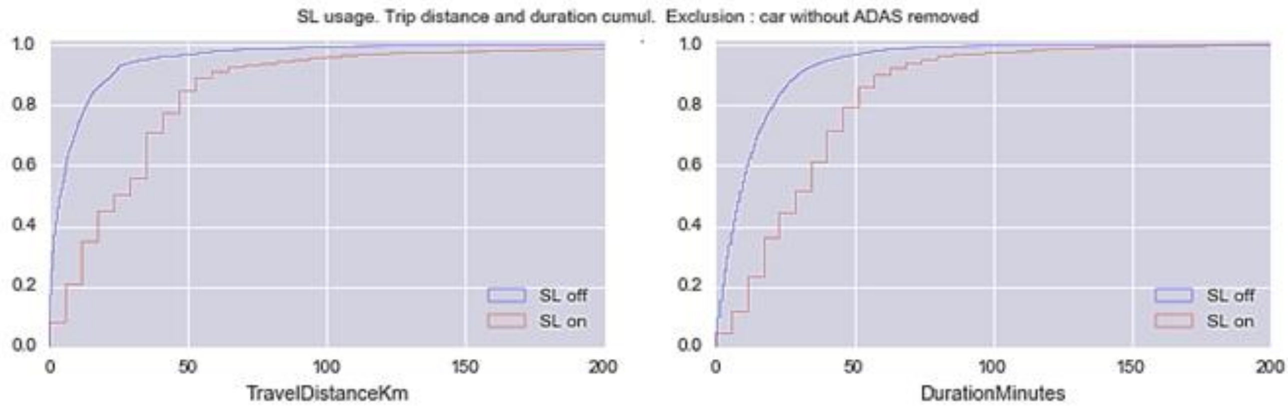


Figure 6-20: Cumulative plot of total distance and duration of trips with SL on

Figure 6-19 and Figure 6-20 give the same information for SL. The same kind of result raises which means there is no real difference on total distance neither trip duration between CC and SL use. Box plots are another way to represent this kind of information.

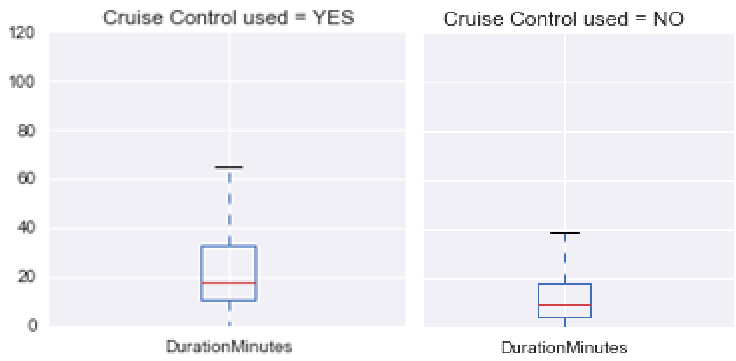


Figure 6-21: Boxplot of trip duration with/without CC use

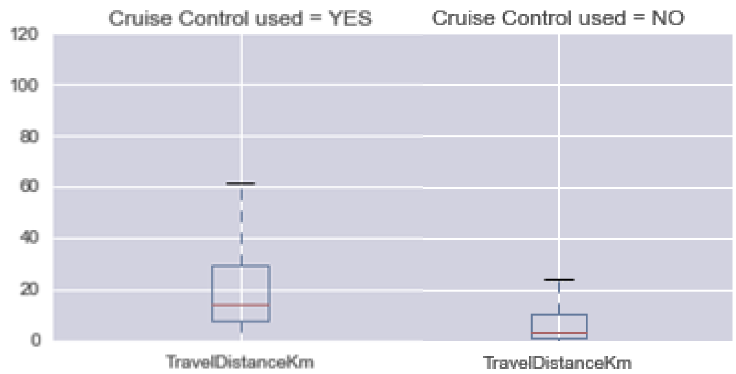


Figure 6-22: Boxplot of total distance of trips with/ without CC use

CC is used on longer trips in terms of duration (Figure 6-21) or total distance (Figure 6-22).

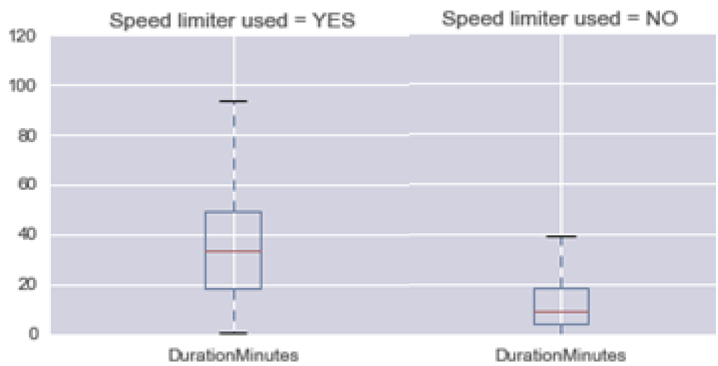


Figure 6-23: Boxplot of trip duration with/without SL use

Figure 6-24: Boxplot of total distance of trips with/without SL use

SL is also used on longer trips (Figure 6-23 and Figure 6-24). The median duration of trips including SL use (around 30 min) is even higher than the median duration of trips including CC use (around 20 min). The same is true in terms of total distance

<u>Road type of CC/SL use</u>

Figure 6-25 shows the road type of the segment with and without CC/SL. The first striking information is that segments road type distribution is very similar whether we speak about CC activation or SL activation. The distribution of segment road type with and without CC looks different in particular for urban roads: the largest number of segments is urban roads (Figure 6-13) whereas the number of urban road segment where CC is used is not so big. This indicates that drivers tend to use CC less frequently on that type of road. The proportion of urban motorway is much smaller than country motorway or rural roads (nearly equivalent in terms of number of segment) whereas activation of CC is nearly equivalent on those 3 types of roads. This could be explained by the fact that drivers use CC more often on an urban motorway than on country motorway/rural road. This is more or less different to the general idea that the CC is more often used on country motorways

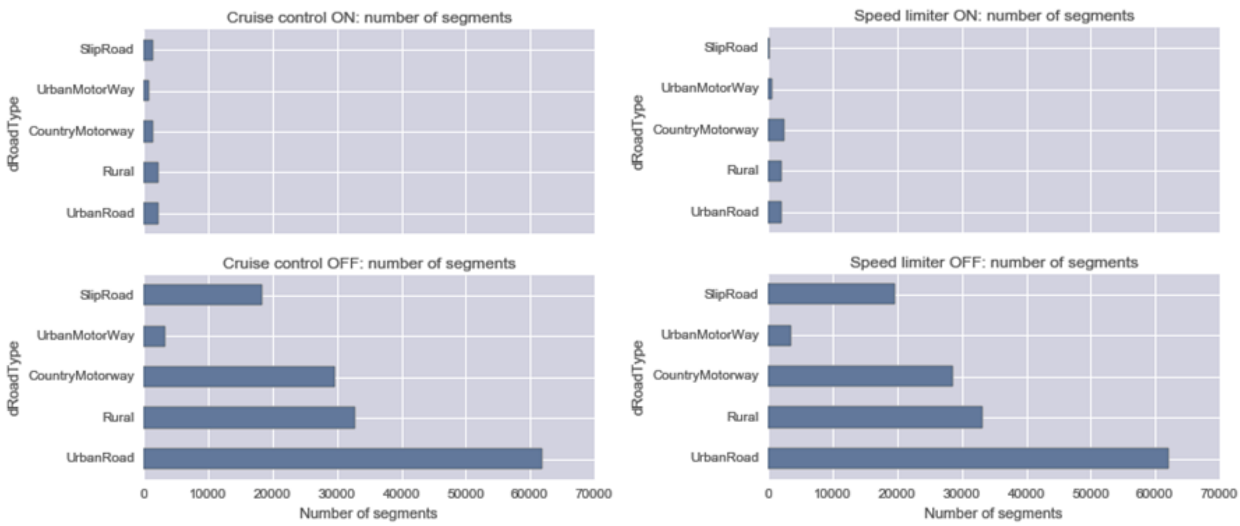


Figure 6-25 CC and SL use by segment type of road

Figure 6-26 shows the frequency of CC/SL use by road type. It should be read as follows*: “19 % of Urban Motorway segments included CC use”*.

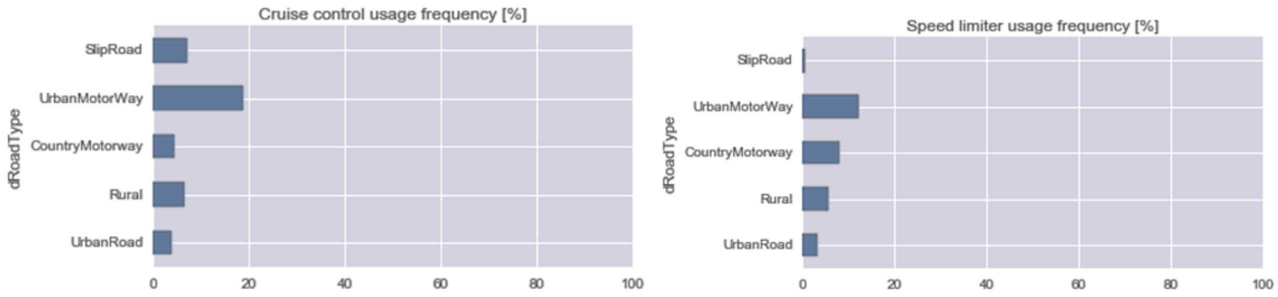


Figure 6-26 Frequency of CC/SL use by road type

The numbers are summed up in Table 6-17. The main difference between CC and SL is for Country motorway with a higher activation rate for SL

Road type	CC usage [%]	SL usage [%]
Slip road	7.1	0.6
Urban motorway	18.7	12.3
Country motorway	4.5	8
Rural	6.4	5.6
Urban road	3.7	3.2

Table 6-17: Frequency of CC/SL use by road type

The same information can be represented in terms of percentage in Figure 6-27.

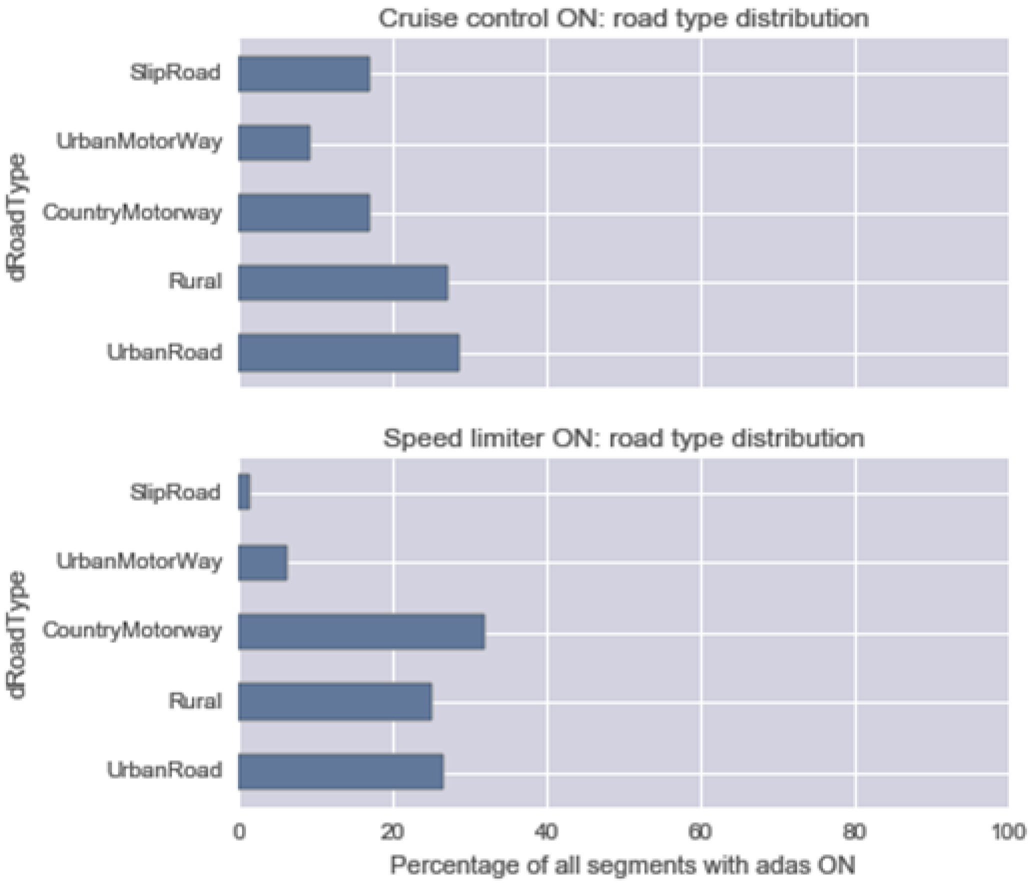


Figure 6-27: Comparison between segment road type of CC and SL activation

CC activation is more frequent than SL activation on rural roads. The use of CC on urban road is quite surprising and should be investigated further. In particular, some misclassification of urban roads could lead to over estimating the rate of CC use. Anyway, the distributions of road type for CC and SL still are comparable: there is no real difference in the use of the two ADAS.

Duration of CC/SL use

To study duration of CC/ SL use, the database available on May, 19th was used. Distance driven and time of each regulation period was processed for each activation window. Example of CC processing in presented in Figure 6-28.

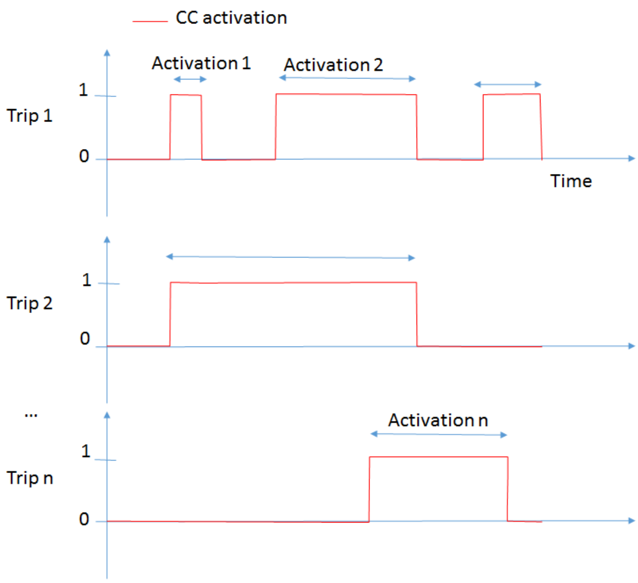


Figure 6-28 Processing of activation periods for CC study

The same processing was applied to SL activation and the results are presented in Table 6-18 .21507 CC activation windows were studied and 3289 SL windows.

	CC		SL	
	Distance (km)	Duration (min)	Distance (km)	Duration (min)
mean	2,8	1,6	11,1	11,1
std	6,5	3,19	16,4	14,6
min	0,005	0,008	0	0,01
25%	0,3	0,3	2,3	2,23
50%	0,9	0,7	6,2	5,8
75%	2,6	1,6	13	13,7
max	170,7	77,9	344,9	179,9

Table 6-18 Duration of CC/ SL activation over UDRIVE database

The mean distance driven for each CC activation (2.8km) is shorter than for SL (11.1km). Equally for distance and for time of each period of activation, SL is used longer than CC.

Driver awareness on ADAS

Drivers' responses to the question: "is your car equipped with CC/SL"? are summed up in the confusion matrix (Table 6-19).

"is your car equipped with CC/SL"	Answer= no	Answer= yes
Car not equipped with CC/SL	32	1
Car equipped with CC/SL	21	67

Table 6-19: Confusion matrix (number of answers) on CC/SL awareness. Correct answers in green, incorrect answers in orange.

This matrix shows that 17% of the drivers are not aware that their car is equipped with CC/SL. No specific driver characteristic was found for this group (neither country, nor type of car). It has to be noted that some drivers did not answer this question. This could be explained by misunderstanding the question maybe because they actually don't even know the ADAS name. Car manufacturer (especially Renault) could maybe better communicate to drivers about ADAS.

Another question was included in the questionnaire about other ADAS equipment. It concerned park assist AHL and LDW. Figure 6-29 shows the answer.

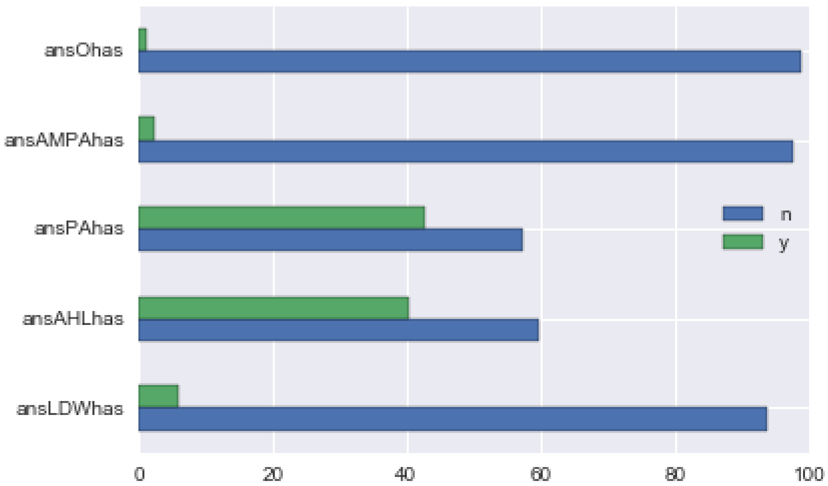


Figure 6-29: Other ADAS knowledge

None of the cars were equipped with AHL whereas 40% of drivers think they have this ADAS. An explanation could be that drivers confuse this ADAS with auto light function.

6.4.4 Summary

In the studied dataset, 88% of the trips were driven without any ADAS activation. Only 2 % of the trips included speed limiter use. Cruise control and speed limiter were used in comparable conditions of road type and trip distance. 17% of the drivers don't know that their cars are equipped with CC/SL and more generally, 40% of the drivers confuse automatic high beam low beam function with automatic lighting function.

6.5 Research question 2.4: To what extent are seatbelts used?

6.5.1 Introduction

This study is focused on understanding the use of driver seat belt. Of course, the first question is how often drivers are using their seat belts and the differences between countries. But the analysis below provides a more detailed study on driver characteristics and environmental factors (e.g. urban/rural/motorway, time of day, etc.) or trip characteristics that influence seatbelt usage.

The analysis is divided into two parts:

- Driver study provides a description of the drivers with seat belt locked during the whole trip and reading between the lines, those driving unbelted part of their trips. The characteristics common to these groups of drivers are described

Trip study describes what sort of trip drivers choose to drive without seat belt from the beginning until the end

6.5.2 Method

Data mining algorithms

These types of algorithms are useful to study large amount of data. Data must be prepared in a matrix (a table). The matrix rows represent the perimeter: the observable to be studied. Perimeter can be the driver or records. Each row must be identified by a distinct value (key): the observable identification ('Driver ID' for example). One column of the matrix must include the target variable: the variable to be explained ('belt locked' for example). The other matrix columns represent the features: the explanation variables (example: age, gender, etc.). The number of samples is too low in our case to study the behaviour of drivers without seat belt because this event is infrequent

Tree

The aim of the current study is to identify recommendations to prevent drivers from driving without seat belt. They have to be easy to interpret by human factor experts. If variables that could explain the wrong behaviour (feature) are not too numerous, descriptive statistics are relevant. On the contrary, if the number of features increases, it becomes more and more difficult to understand all the features contributions to explain the behaviour.

Random Forrest (RF) (Breiman, 2001) or Gradient Boosting algorithms (GB) (Friedman, 1999) are the state of the art to predict a behaviour from a set of features taking into account all the interaction between features. Unfortunately, even if these algorithms are useful to identify the most contributing features, they produce "black-box" models impossible to interpret by human factor experts.

In between those two data processing methods, the decision trees can handle an unlimited number of feature and their interactions. They are less accurate than RF or GB algorithm for prediction but they produce a decision tree compound by a set of rules easy to understand by researchers. That is why this type of algorithm was selected.

The decision trees were proposed by Breiman et al. in 1984. They aim to explain a target variable (behaviour in our case) from a set of explaining variables also called features. These features can be either quantitative or qualitative.

This algorithm searches the best rule to split a sample in two branches (sub-samples) with the highest distance between each other. A rule is composed of a feature and a threshold. The criterion used to measure the distance between the two resulting branches is the Mean Squared Error if the target variable is quantitative (regressor usage). In the other case, if the target variable is quantitative (classifier usage), the criteria could be either the Gini or the entropy index.

Once the first rule is found, the original sample (root) is then split in two branches. Iteratively, the algorithm is applied again on both branches to build a tree. The iterations stop when a branch contains only one value of the target variable. It may also stop earlier under conditions such as a minimum sample size per branch or a maximum depth in the tree. The aim of the decision tree is to learn by this way a model based on a set of rules.

The aim of the current study is not to predict the driver behaviour. On the contrary, the research question is "what feature is linked to the wrong behaviour (driving without seat belt)"? In particular, the set of rules leading to the wrong behaviour is relevant to identify recommendations easy to interpret by stakeholders.

In the tree development, the rule used to split the tree root is the most important one (it is the best to explain the behaviour). On the second step, rules used at first branch levels are the best remaining one to explain the behaviour, and iteratively

Methodology for driver characteristics study

To study characteristics of subjects driving with seat belt locked during the whole trip, we compare trips with seat belt locked 100% to those not locked 100% of the trip. The perimeter is defined by driverID. The target is the percentage of records with belt locked (100% of the trip). The available features are: gender, age, driving style indicator (Guyonvarch et al. 2017), country. Table 6-20 shows an overview of the data.

Table 6-20: Matrix used for driver characteristic study

Perimeter key **DRIVERID**	**Target** **Belt_full_Locked**	**Feature 1** **Age**	**Feature 2** **Gender**	**…**	**Feature N** **Country**
Driver-1	98%	33	M	…	OS-DE
Driver-2	96%	76	F	…	OS_FR
…	…	…	…	…	…
Driver-N	89%	45	M	…	OS-UK

A decision tree was applied to the matrix to explain the target variable with the set of features [Feature1 – Feature N]. The result of this algorithm provides the more significant variable (among driver characteristics) that are linked to trip with seat belt locked from the beginning to the end

Methodology for trip characteristics study

To study which trips are driven without seat belt locked at all (0% of the trip), we will study 0% seat belt locked trips with the others. For that study, the record characteristics, the perimeter is defined by 54867 distinct recordID. For each trip, the target is a binary variable taking following values:

- 0 = Less than 100 meters driven seat belt unlocked
- 1 = More than 100 meters driven seat belt unlocked

Table 6-21 shows the list of feature used to study what type of trip is driven without seat belt

Table 6-21: Overview of features describing the record

Variable	**Description**
Day of the week	from 1 (Monday) up to 7 (Sunday)
Weekend	1 if Saturday or Sunday, else 0
Hour	Hour of the record start
Distance	Total trip distance in meter
Time	Total trip time in minute
Usage_Distance_n	This describes how “usual” (in terms of distance) the trip is. It is the percentage of distance used more than n time by the driver. N=2,3,5,10,20
Usage_Time_n	This describes how “usual” (in terms of frequency) the trip is. It is the percentage of time driven in segment used more than n time by the

	driver, N=2,3,5,10,20
dDayNight_Day_Dist	distance driven by daylight
dDayNight_Night_Dist	distance driven by night
dDayNight_Day_Time	time driven by daylight
dDayNight_Night_Time	time driven by night
dRainState_heavyRain_Dist	distance driven under heavy rain
dRainState_Rain_Dist	distance driven under rain
dRainState_No_Rain_Dist	distance driven without rain
dRainState_heavyRain_Time	time driven under heavy rain
dRainState_Rain_Tim	time driven under rain
dRainState_No_Rain_Time	time driven without rain
dRoadType_Country_MotorWay_Dist	distance driven on country motorway
dRoadType_Rural_Dist	distance driven on rural road
dRoadType_Rural_Time	time driven on rural road
dRoadType_Slip_Road_Dist	distance driven on slip road
dRoadType_Slip_Road _Time	time driven on slip road
dRoadType_Undefined_Dist	distance driven on undefined road
dRoadType_ Undefined _Time	time driven on undefined road
dRoadType_Urban_Motorway_Dist	distance driven on urban motorway
dRoadType_ Urban_Motorway _Time	time driven on urban motorway
dRoadType_ Urban_Road _Dist	distance driven on urban road
dRoadType_Urban_Road_Time	time driven on urban road

Features linked to the frequency of the trip were obtained using the Mapdata and the segments ID. Features linked to the light condition/weather were calculated using the can data (light sensor, wipers). Road types were extracted using map data on speed limit as well as CAN data (speed) and local regulation on max speed. This automatic classification was validated by each country. All these features are exclusively describing the trip. All features linked to the driver (age, gender, habits, etc.) are excluded since this study focuses on the influence of the environment to the belt usage.

For this study, we excluded trips with total distance <100m because they represent only very short manoeuvres which couldn't be used to understand when drivers do not choose to lock seat belt.

6.5.3 Sample description

The analysis provided was conducted over the data set available on March 15th 2017, consequently, the countries and number of drivers is unbalanced. The data is described below.

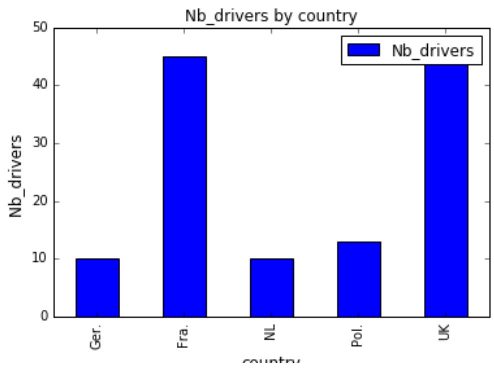
Table 6-22 provides the corresponding Operation Site codes used in Figure 6-30 to Figure 6-38.

Table 6-22: Country and operation site coding

Country	OS Code

Germany	OS-DE
France	OS-FR
The Netherlands	OS-NL2
Poland	OS-PL
United Kingdom	OS- UK

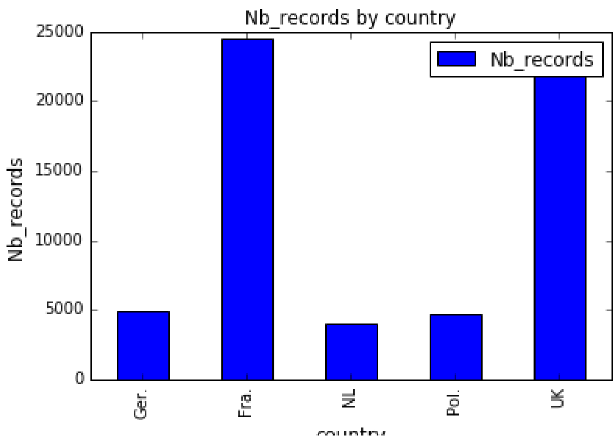
The number of driver is unbalanced between the countries in this data set. Figure 6-30 shows the distribution of number of drivers over countries. France and UK have the higher number of drivers with respectively 45 and 48 drivers. German, Dutch and Polish driver are fewer (10, 10 and 13 drivers).



Country	Nb Drivers
Germany	10
France	45
NL	10
Poland	13
UK	48

Figure 6-30: Number of drivers by country

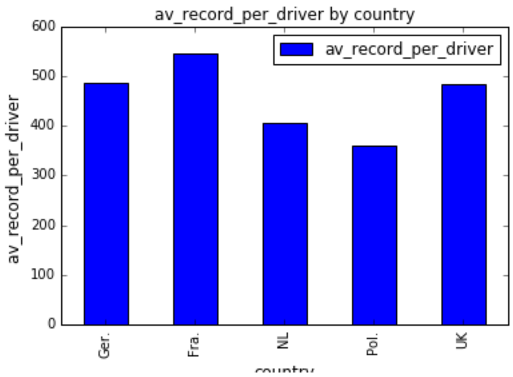
The total number of records of the data set is 61417 [1]. The relative number of records by country (Figure 6-31) is similar to the Figure 6-30. The average number of trip by driver (Figure 6-32) is equivalent from one country to the other, even if French drivers drove 550 records in average and Poland one only 350.



Country	Nb Records
Germany	4879
France	24504
NL	4063
Poland	4697
UK	23274

Figure 6-31: Number of record per country

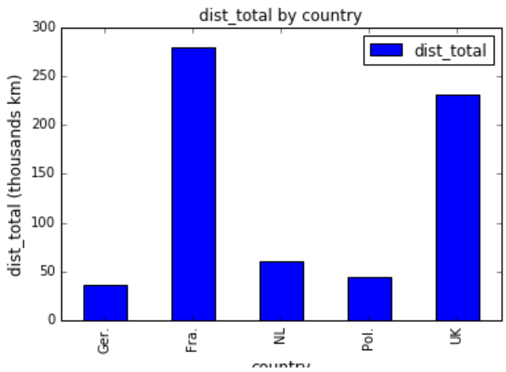
[1] Warning, part of these records are not taken into account in further investigation according to filters applied to the sample.



Country	Record/Driver
Germany	487
France	544
NL	406
Poland	361
UK	484

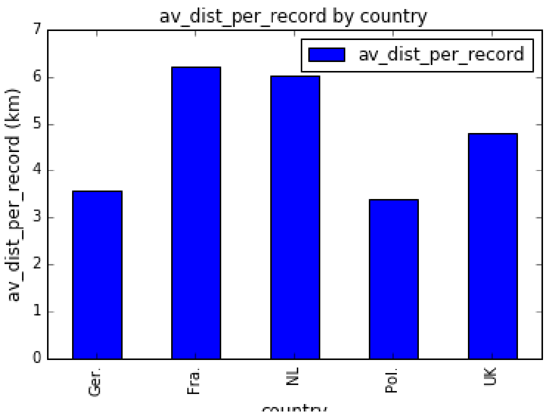
Figure 6-32: Average number of record by driver per country

The total distance for all trips in the data set is 649 thousand kilometres. Highest country contributors are France (279 000 km) and UK (230 000 km) (Figure 6-33). For the average distance by trip, France and the Netherlands are close to 6 km for each trip, Germany and Poland close to 3.5 km and UK is intermediate at 4.4 km per trip (Figure 6-34).



Country	Distance (km)
Germany	35 641
France	279 066
NL	60 344
Poland	43 916
UK	230 652

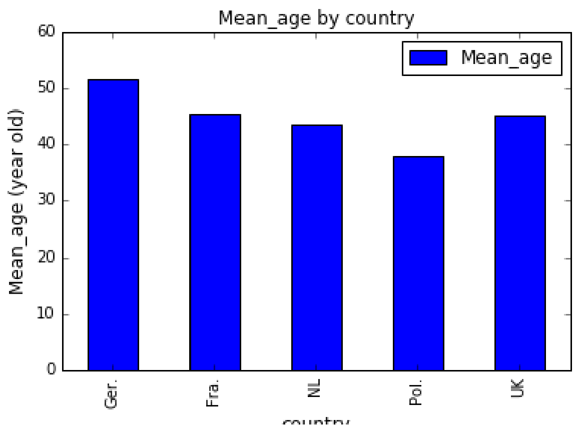
Figure 6-33: Distance per country



Country	Av distance (m)
Germany	3564
France	6201
NL	6034
Poland	3378
UK	4805

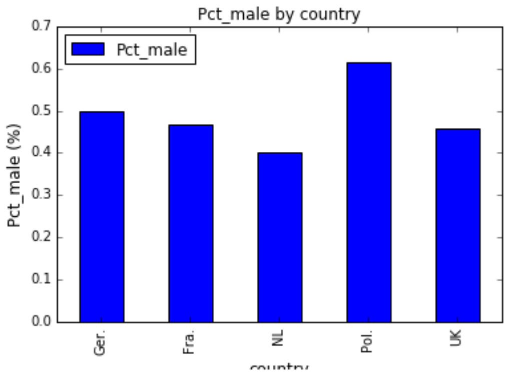
Figure 6-34: Average record distance per country

The mean age of drivers per country is balanced between 37 years old (Poland) and 51 (Germany). The genders are balanced for most of the countries excepted Poland with 60% of male drivers (Figure 6-36).



Country	Mean age
Germany	51.6
France	45.3
NL	43.5
Poland	37.8
UK	45.1

Figure 6-35: Mean driver age per country



Country	% Male
Germany	50.0 %
France	46.6 %
NL	40.0 %
Poland	61.5 %
UK	45.8 %

Figure 6-36: Percentage of male drivers per country

The data set includes 126 drivers. The following variables were selected for each driver: age, gender, country, belt usage and record characteristics (trips, time, mileage, etc.).

The number of records per driver increases regularly from 13 records up to 1508 (Figure 6-37) when the total distance increases from 30 km up to 20896 km (Figure 6-38).

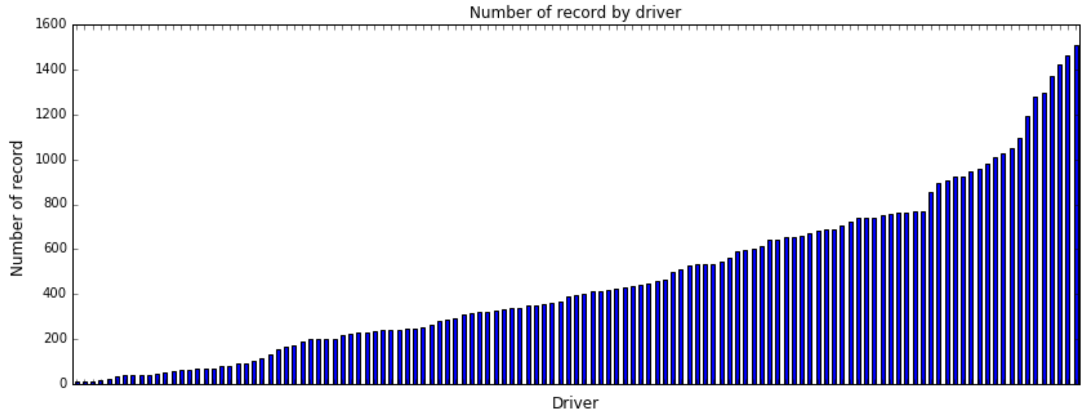


Figure 6-37: Number of records travelled by drivers

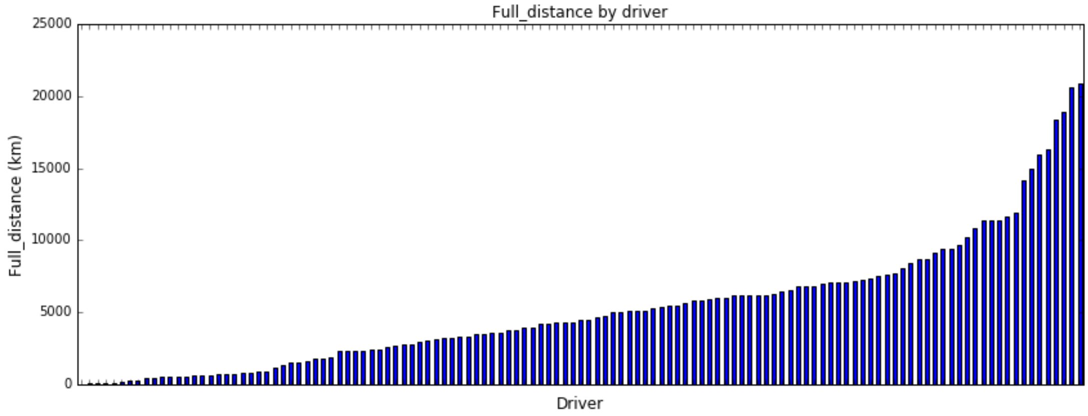


Figure 6-38: Total distance driven by drivers

The mean distance by record per driver is spread between 2.3 km and 40 km by record (Figure 6-39).

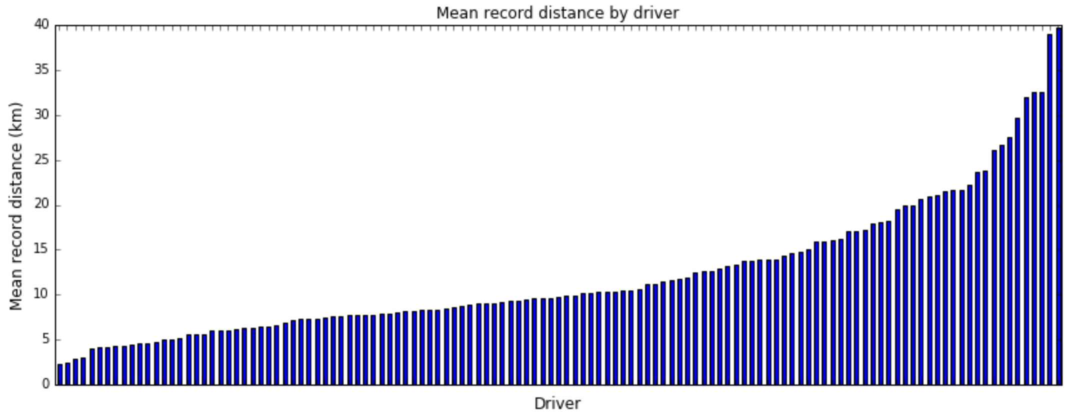


Figure 6-39: Mean records distance by drivers

The mean time by record per driver is spread between 4 mn and 33 mn by record (Figure 6-40).

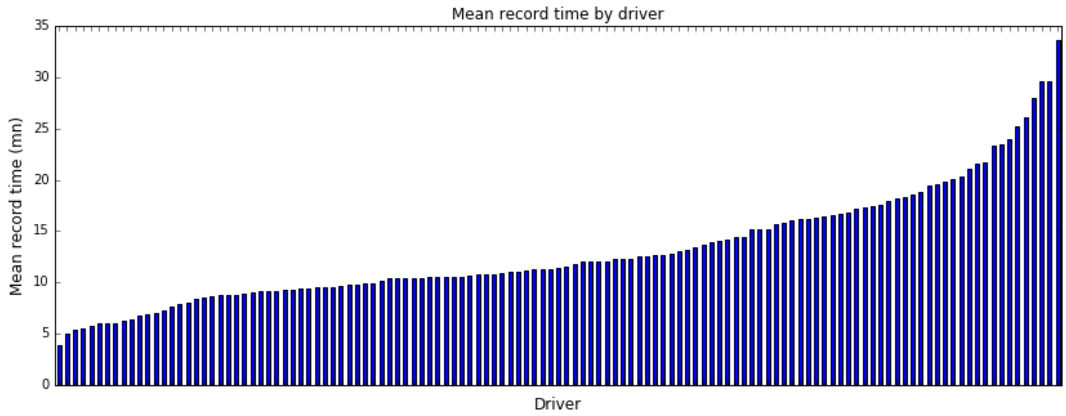


Figure 6-40: Mean record time per driver

On Figure 6-41, for each driver, the rate of trips including some driving without seat belt is presented. This distribution can be compared to Reagan et al. (2013).The UDRIVE distribution is more flat than the one presented by Reagan whereas number of drivers is comparable. Generally speaking, the seat belt rate is higher for UDRIVE participant than for 100-cars participants. This difference could be explained by several factors:

- 100 cars included a lot of young drivers which is really different with UDRIVE subjects (mean age = 37 years)
- All UDRIVE cars (like the majority of actual person cars in Europe) are equipped with a seat belt reminder. For Renault cars, SBR includes 2 sound levels to encourage driver to wear the seat belt.
- The way 100-car measured seat belt wearing was biased by the fact that only the first video frame was coded and recorded as seat belt status for the whole trip. In Udrive data, seat belt status is a continuous variable. For example, the trip where driver starts engine then locks seat belt and after that, begin to move will be classified as an unbelted trip in the 100 car study and as a full belted trip in UDRIVE.
- Finally, a change in behaviour may have occur since 100-car data was recorded (2006)

Figure 6-41 displays the percentage of records with belt locked during the whole trip by country

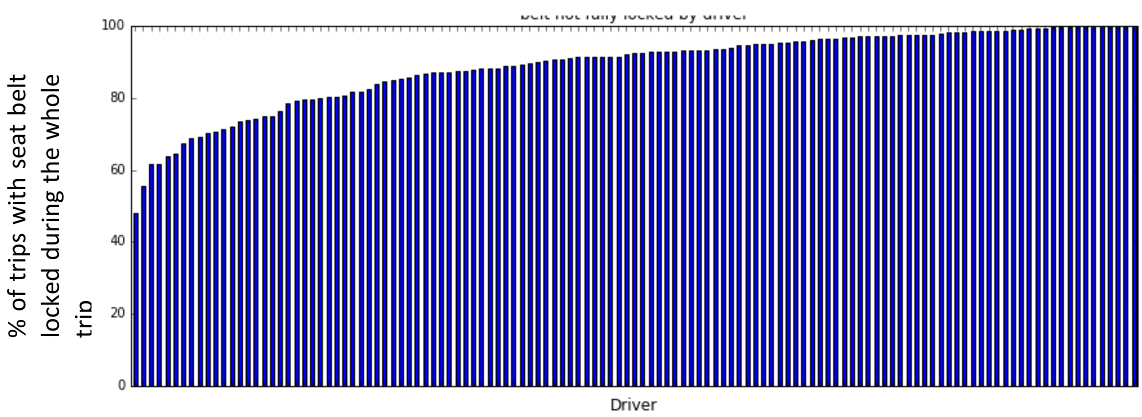
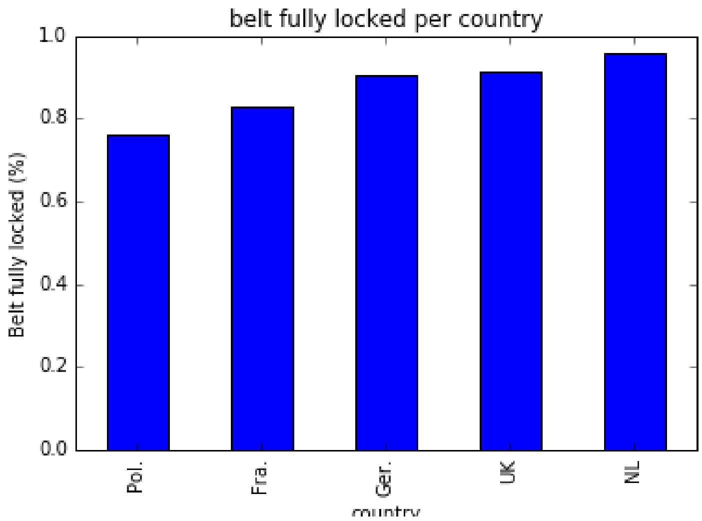


Figure 6-41: Percentage of records with belt unlocked part of the time



Country	Percentage of belt fully locked
Poland	76.1%
France	82.9%
Germany	90.4%
UK	91.3%
NL	95.6%

Figure 6-42: Percentage of records with belt locked all the time

Globally speaking, 87.3% of the trips are driven with driver seat belt on from start to end. This percentage is different from one country to the other with the lower rate for Poland (76%) and the higher for the Netherlands (95.6%). This measurement represents a real naturalistic behaviour and is quite different from the numbers actually used in road safety study. For example, in France, ONISR (2014) published a report on driver behaviour announcing a rate between 95% and 98% depending on the road type. In the UDRIVE data, only 82.9% of the trips are done with seat belt on for the whole trip. Moreover, the sample of drivers studied in the UDRIVE database is composed of volunteers to participate in a road safety study; it does not include young drivers suggesting this rate is even lower on average.

6.5.4 Results

One of the drivers had a very specific behaviour concerning seat belt used. The following analysis aimed to give general information on driver/trip characteristics with a wrong seatbelt use. Since the OS of this driver included a small number of drivers, this specific driver was removed for the following analysis to prevent from a bias in the analysis.

<u>Driver characteristics study</u>

The tree algorithm provided the results presented in Figure 6-43

The tree displays the main rules automatically extracted by Decision Tree algorithm.

1. The most significant rule is "French drivers have different behaviour compared to other countries":
83% of records with belt locked during all trips of French drivers versus 90% for other countries
2. For people living in non-French countries, the Gender is the most discriminant variable:
86% of trips driven seatbelt on for the whole trip for men versus 94% for women
3. For men living in non-French countries, the DSI is then the most discriminant variable:
82% drivers having dynamic driving style (DSI>0.14) versus 92% for drivers with smooth driving style

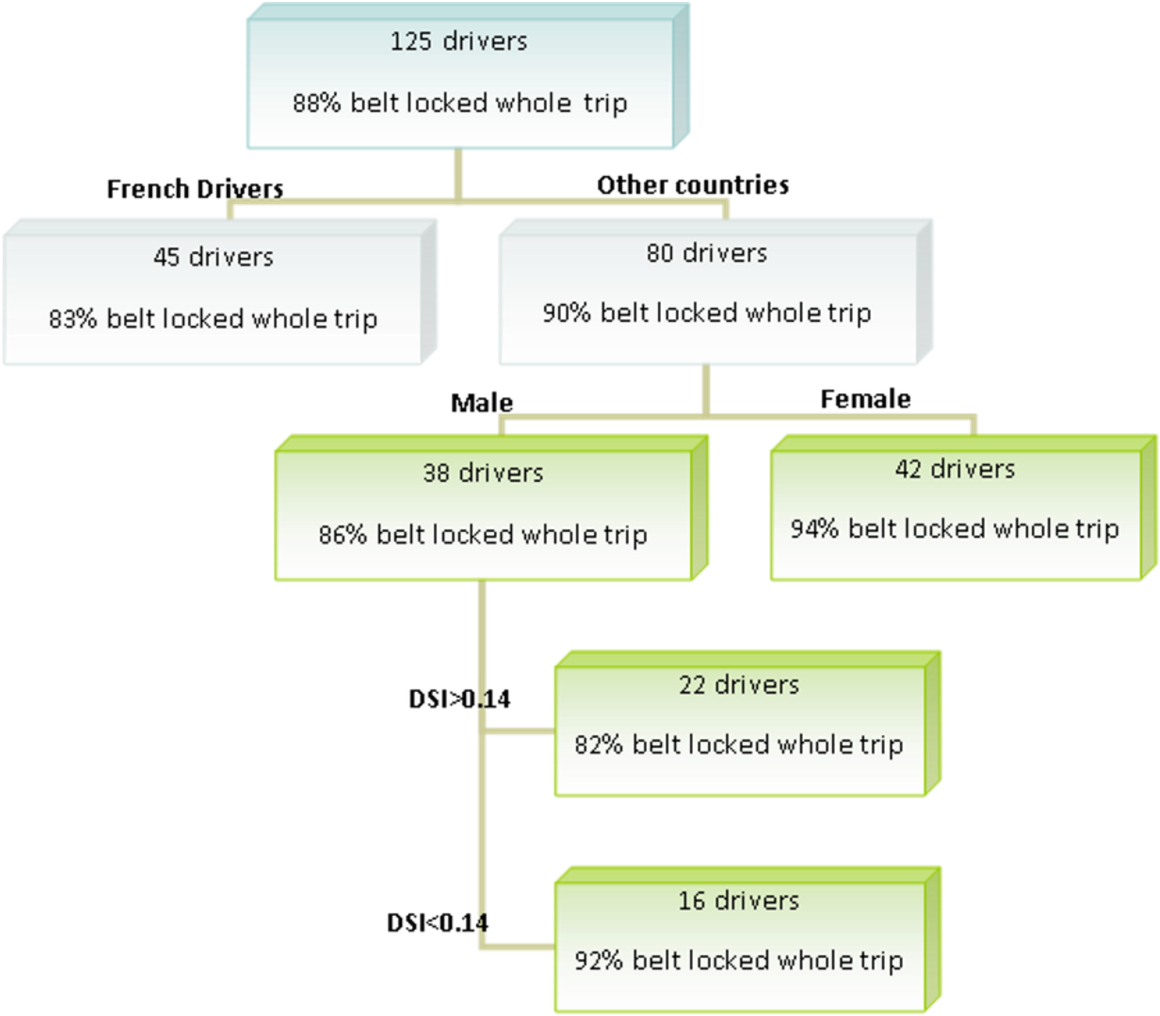
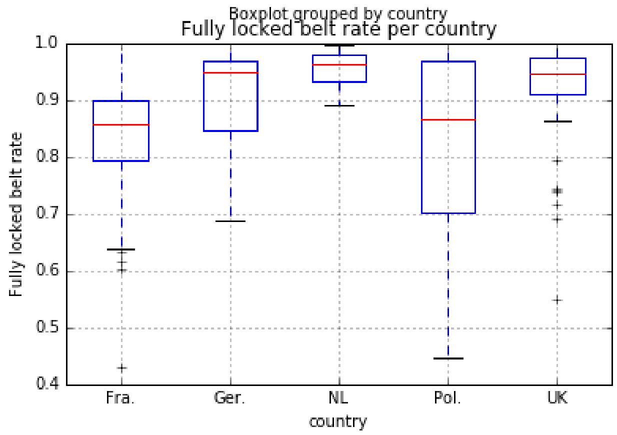


Figure 6-43: Tree for driver characteristics linked to seat belt use

This tree was parametrized to produce only four leaves (max_leaf_nodes parameter in sklearn.tree.DecisionTreeRegressor). The number of leaves was limited to make it easy to interpret and explain. It is the minimum one to explain the hierarchy between features: country, gender then DSI. Last but not least, there are obvious differences of belt locking rate between resulting leaves. It means that the features are relevant to explain belt locking behaviour differences.

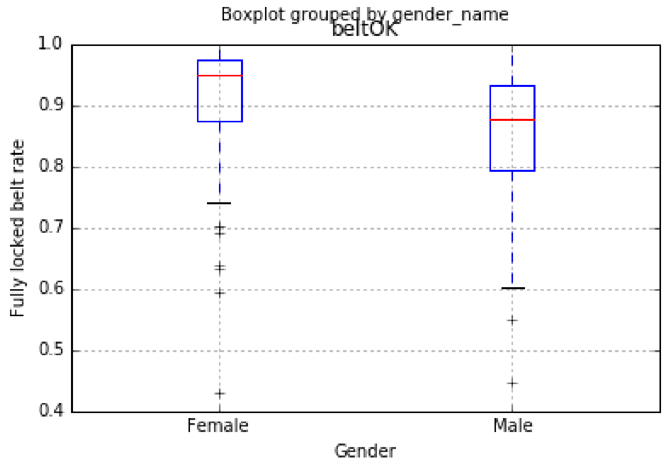
Once these segments and rules are identified, target distribution is analysed by features to confirm these preliminary conclusions.



Country	Fully belt locked median
France	85.8%
Germany	95.0%
NL	96.3%
Poland	86.7%
UK	94.7%

Figure 6-44: Rate of trips with seat belt locked whole trip

Figure 6-44 confirms the difference of trip fully locked belt rate between countries and particularly the low belt locking rate of French drivers. Polish drivers have a low median value close to the French one. Poland has not been detected by the decision tree because France average value is a little bit lower. The first split is then applied to France only. Once this split is done, 'Male vs Female' rule is more discriminant than 'Poland vs other country'. In the other hand, German, Nederland and UK people have a common high rate of belt locking.



Gender	Fully belt locked median
Male	87.8%
Female	94.9%

Figure 6-45: Seat belt locked during whole trip for other countries by gender

Figure 6-45 confirms the variation of fully locked belt rate per gender. Men have lower belt locking rate than Women.

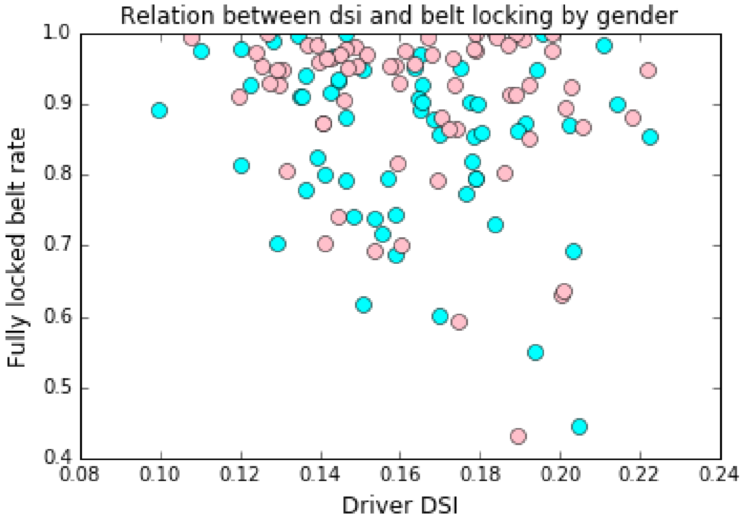


Figure 6-46: Belt locked rate as a function of driving style indicator

Figure 6-46 is a scatter plot because Driving Style Indicator is a quantitative variable. Each point represents one distinct driver: pink for women, blue for men. The lack of point in the bottom left of this scatter means that driver having a 'smooth' driving style (low DSI value on the left) are more likely to used their seatbelt. On the contrary, 'dynamic' drivers (high DSI value on the right) may unlock their belt more frequently even if it is not systematic.

Trip study

We now focus on trips without seatbelt on all along.

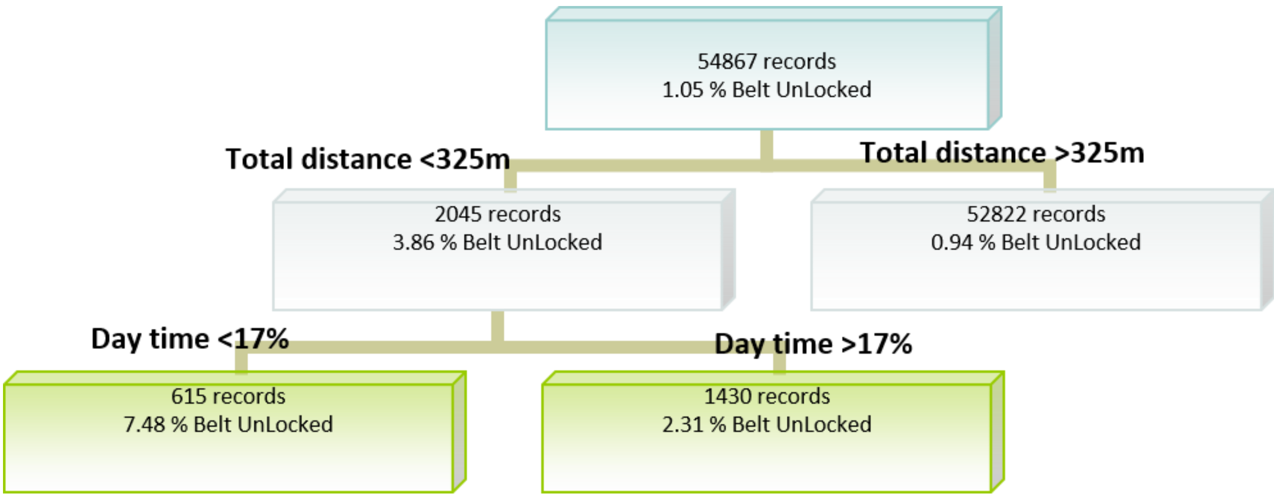


Figure 6-47: Decision tree for trip characteristics influencing seat belt use

The decision tree applied to the 54867 trips collected displays the following rules (Figure 6-47).

1. The most significant feature is "record distance lower than 325 meter": 3.86 % of trips are detected with belt unlocked for short record distance (lower than 325 meter). Note than only 2045 records are so pointed
2. For lowest distance records, the day time is the most discriminant variable: 7.48% for trips mainly

driven by night (daylight lower than 17%) versus 2.31% for the other. This sample contains only 615 records : this represent only 1% of the initial sample

This tree was parametrized to produce only 3 leaves (*max_leaf_nodes* parameter in sklearn.tree.DecisionTreeRegressor). As for previous tree, the number of leaves have been voluntary be limited to make it easier to understand and explain

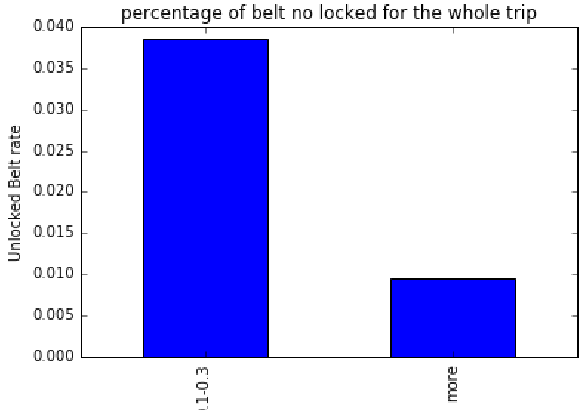


Figure 6-48: Comparison of rate of full trip unlocked between trip with total distance <325m and 325m

Figure 6-48 illustrates the first rule: belt are not locked 3 time more if the record distance is below 325m (3.86 % vs 0.94%).

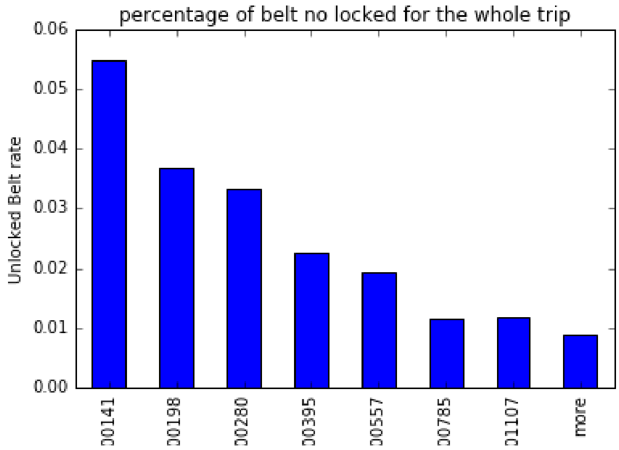


Figure 6-49: Unlocked belt rate for specific distances (x axis in meters)

Figure 6-49 shows more in detail the relationship between the belt-lock rate and the distance travelled during the record: there is a continuous decrease of the belt-lock rate from 100m up to 1km. Over 1km, this rate does no longer decrease and stays at the same value.

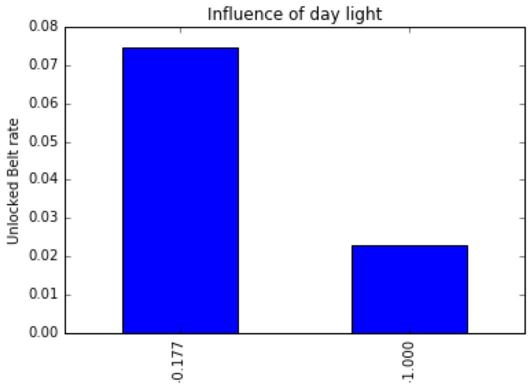


Figure 6-50: Belt unlocked rate of trips with less than 17% of day light (left) compared with trips with more than 17% of day light (right)

For the 2045 records with a total distance below 325m , Figure 6-50 shows the influence of the day light : belt are not locked 3 time more if the day light distance is below 17% (7.48% vs 2.31%).

6.5.5 Summary

Trees were used to track the most significant factors to explain wrong seatbelt behaviour (trip including driving without seat belt). The same method was used to study trip characteristic for full trips without belt. Driver country is the most significant parameter explaining the behaviour of driving for some distance without wearing the seatbelt. The second contributing factor is gender. More male drivers do not wear a seatbelt. For whole trips without driver seatbelt use, the most significant variable is trip distance: this behaviour is more frequent for very short trips (<325m) and especially at night.

6.6 Research question 2.5: What is the impact of road context, weather condition, time of day, leading behaviour and ADAS use on hard braking occurrence?

6.6.1 Introduction

In order to analyse potential events that involve create risky situations, we are interested in hard braking. Indeed, this type of braking can indicate when a situation becomes critical and how the driver reacts to control the situation. Hard braking can also be a hazard for drivers. When we examined many of these braking situations, we realised that many of them were due to abrupt driving and could be considered normal driving in certain types of driving situations. In addition, braking does not happen everywhere. It may be interesting to consider the number of brakings per kilometre travelled in order to identify their occurrence. When considering the number of brakings per kilometre, we will not identify the explanatory factors that may explain their occurrence. We therefore decided to have a reference in terms of exposure to create non-hard braking events that will serve as a baseline. The objective of this study is to investigate which factors are more often present in hard braking than in non-hard braking. It will allow us to better explain which factors may explain the presence of hard braking.

In this analysis, we will present:

- The pre-processing that was done on the database to extract events and the creation of attributes to characterise them.
- The treatments that were made to select valid events by removing events with inconsistent values that are often due to data collection problems.
- The descriptive analysis of the data to identify the disparities between data collection in different countries.

The aim of the analysis is to identify factors affecting the occurrence of hard braking versus non-hard braking events.

6.6.2 Method

Participants

To prepare the analyses, we have already compared the ages and genders of the participants whose data were used in this analysis. This was done on the data present in the database on April 18, 2017, the day of extraction. It is different from that of the participants recruited due to the fact that some data were missing in the database on the day of the extraction.

Table 6-23 : Distribution of participants by gender

Gender	Germany	France	The Netherlands	Poland	UK	Total
Female	6	22	8	10	26	72
Male	11	20	6	17	25	79
Total	17	42	14	27	51	151
Proportion of women	35%	52%	57%	37%	51%	48%
Proportion of men	65%	48%	43%	63%	49%	52%

Table 6-23 shows that German and Polish drivers have a strong difference between the proportion of men and women while the UK panel is balanced. The other panels are close to 50% +/- 7%. A large proportion of the participants (Table 6-24) are in the age groups 30-40 and 40-60 years. There are few young or older people.

Table 6-24 : Distribution of participants according to their age

Age	Germany	France	The Netherlands	Poland	UK	Total
<30 years	5	5	4	3	9	26
30-40 years	2	12	4	14	13	45
40-60 years	7	19	5	9	21	61
> 60 years	3	6	1	1	8	19
	17	42	14	27	51	151
<30 years	27%	10%	33%	0%	20%	17%
30-40 years	9%	25%	33%	53%	20%	30%
40-60 years	36%	45%	33%	41%	40%	40%
> 60 years	27%	20%	0%	6%	20%	13%

Creation of events and their characteristics

To create the events, we calculated three time signals in SALSA. For this, we used the signal "longiAcc" which measured the longitudinal acceleration in the algorithm ALY_IFSTTAR_BrakeSignalV2 (Figure 6-51). We have averaged this acceleration over 0.5 seconds to have less noise in the signal. Then we looked at segments with a value below the threshold for a minimum duration. The first signal was used to detect hard braking. The first signal had accelerations less than $-5 m/s^2$ (i.e., deceleration greater than $5 m/s^2$) for at least 0.3 seconds. The second signal was used to detect hard braking which had acceleration of less than $-3 m/s^2$ (i.e. deceleration greater than $3 m/s^2$) for at least 0.5 seconds. The third signal was used to detect hard braking which had acceleration less than $-1 m/s^2$ (i.e. a deceleration greater than $1 m/s^2$) for at least 2 seconds.

```
function [ TS_BrakeSignal, BrakeSignal ] = ALY_IFSTTAR_BrakeSignalV2( mlongitudinalAcceleration,mSpeedCAN , thresholdAcceleration,
thresholdMinTime)

% Recovery of signals and their timebase
TS_BrakeSignal = mlongitudinalAcceleration.Time;
LongiAcc=mlongitudinalAcceleration.Value;
Speed = mSpeedCAN.Value;

sizetrip= length(LongiAcc);
% remove bad values
LongiAcc(LongiAcc>30) = NaN;
LongiAcc(LongiAcc<-30) =NaN;
%do the mean between 5 points to remove step effect
LongiAcc(3:sizetrip-2) = (LongiAcc(1:sizetrip-4)+LongiAcc(2:sizetrip-3)+   LongiAcc(3:sizetrip-2)+    LongiAcc(4:sizetrip-1)
+LongiAcc(5:sizetrip))/5;

% Thresholding of the signal according to the chosen threshold
LongiACCDiscret = 2*ones(length(LongiAcc),1);
LongiACCDiscret(LongiAcc>thresholdAcceleration ) = 0;
LongiACCDiscret(LongiAcc<=thresholdAcceleration ) = 1;
LongiACCDiscret(Speed<=0.1) = 0;

% Segment identification and removal of too short
Time = TS_BrakeSignal;
signal = LongiACCDiscret;
MinDuration=thresholdMinTime;
changes = find(diff(signal)~=0 );
Occurences = [];
for i = 1: length(changes)
   if changes(i) < length(signal)% if the last change is on the last sample we do not process
      if ((signal(changes(i)) == 0) && (signal(changes(i)+1) ==1)) % front up
         if (i==length(changes)) % jusqu'a la fin du fichier
            if  Time(length(Time)) - Time(changes(i)) >= MinDuration
               Occurences = [Occurences; [(changes(i)) (length(Time))]]; % Segment storage
            end;
         elseif ((signal(changes(i+1)) == 1) && (signal(changes(i+1)+1) ==0)) % follow by front down
            if Time(changes(i+1)-1) - Time(changes(i)) >= MinDuration
               Occurences = [Occurences; [(changes(i)) (changes(i+1)-1) ]]; % Segment storage
            end;
         elseif ((signal(changes(i+1)) == 1) && (signal(changes(i+1)+1) ==2)) % follow by undefined front
            if  Time(changes(i+1)-1)- Time(changes(i)) >= MinDuration
               Occurences = [Occurences; [(changes(i)) (changes(i+1)-1) ]]; % Segment storage
            end;
         end;
      elseif ((signal(changes(i)) == 0) && (signal(changes(i)+1) ==2)) %front undefined nothing done
      elseif ((signal(changes(i)) == 1) && (signal(changes(i)+1) ==0)) %front down  processed before
      elseif ((signal(changes(i)) == 1) && (signal(changes(i)+1) ==2)) %front undefined to be process after
      elseif ((signal(changes(i)) == 2) && (signal(changes(i)+1) ==1)) %front undefined
         if (i==length(changes)) % jusqu'a la fin du fichier
            if  Time(length(Time)) - Time(changes(i)) >= MinDuration
               Occurences = [Occurences; [(changes(i)) (length(Time))]]; % Segment storage
            end;
         elseif ((signal(changes(i+1)) == 1) && (signal(changes(i+1)+1) ==0)) % follow by front down
            if Time(changes(i+1)-1) - Time(changes(i)) >= MinDuration
               Occurences = [Occurences; [(changes(i)) (changes(i+1)-1) ]]; % Segment storage
            end;
         elseif ((signal(changes(i+1)) == 1) && (signal(changes(i+1)+1) ==2)) % follow by undefined front
            if Time(changes(i+1)-1) - Time(changes(i)) >= MinDuration
               Occurences = [Occurences; [(changes(i)) (changes(i+1)-1) ]]; % Segment storage
            end;
         end;
      end;
   end;
end;

BrakeSignal=zeros(sizetrip,1);
for i=1:size(Occurences,1)
    BrakeSignal(Occurences(i,1):Occurences(i,2)) = 1;
end;

end
```

Figure 6-51 : Algorithm to create brake signals

Once these signals have been created, segments have been extracted from the database by the SALSA software. Then for each of these signals, initial attributes were created using the algorithm, ALY_BrakeEvent_AttributeCalculation Analysis (Figure 6-51).

```
function [EventDuration, minBrake, speedLimit10sBefore,speedLimit10sAfter, dRoadType, speed5sBefore, speed5sAfter,
    meanBrake, minTTC, meanTTC, speedVariation, TIVStartBrakeThreshold, TIVEndBrakeThreshold,TTCStartBrakeThreshold,
TTCEndBrakeThreshold, LeadVehDistStartBrakeThreshold, LeadVehDistEndBrakeThreshold,ABSActive,ASRActive,AYCActive, roudaboutPresence,
intersectionPresence, speedlimitReduction, DayNight, RainState,TrafficCongestion, Flashing_Indicator, CCSLState5s, CCSLState30s,
maxRadiusClass10s,minRadiusClass10s, brake3, brake5,diffMaxMeanGyroym21, nextStop, durationNextStop] = ...
    ALY_BrakeEvent_AttributCalculationAnalyse(TSmlongitudinalAccelerationSeg, TSmSpeedCANRec, TSdLVRangeRec, TSdRoadTypeRec,
    TSdLVRelSpeedRec, TSSpeedLimitRec, TSMAP_WAY_TYPERec, TSVeh_ABS_RegulationSeg, TSVeh_ASR_RegulationSeg,
TSVeh_AYC_RegulationSeg, TSVeh_PhidgetSpatial_gyro_yRec,TSVeh_Flashing_Indicator_Status, TSdDayNightSeg, TSdRainStateSeg,
TSdTrafficCongestionRec,TSdCCSLStateRec, TSdRadiusClassRec, TSMAP_IntersectionRec, TSBrake3Seg, TSBrake5Seg )
```

Figure 6-52 : Matlab Function heading to create initial attributes

Other attributes, derived attributes, were also created with SPSS software (Figure 6-52) by deriving these initial attributes. For more clarity, we present the variables in terms of what they represent and not how they are calculated.

```
RECODE speedLimit10sBefore (Lowest thru 0.5=0)  (1 thru 34=1)  (34 thru 62=2) (62 thru 98=3) (98 thru Highest=4) INTO CatSpeedLimit.
RECODE diffMaxMeanGyroym21  (Lowest thru -0.9 =0)  (-0.9  thru 3=1)  (3 thru 8=2)   (8 thru Highest=3) INTO CatdiffMaxMeanGyroym21.
RECODE TIVStartBrakeThreshold  (Lowest thru 0.01 = 0)  (0.01  thru 1 =1)  (1 thru  2 =2)  (2 thru  4  = 4)  (4 thru  10  = 10)
 (10  thru Highest=0) INTO CatTIVStartBrakeThreshold.
RECODE TIVStartBrakeThreshold  (Lowest thru 0.01 = 0)  (0.01  thru  4  = 1)  (4 thru Highest=0) INTO CatAlone.
RECODE Age (0=0) (MISSING=0) (17 thru 30=1) (30 thru 40=2) (40 thru 60=3) (60 thru Highest=4) INTO
COMPUTE InfraType2=((roudaboutPresence=1) *1) +
(( roudaboutPresence=0) *  (IntersectionPresence=1) * (speed5sAfter <1)  *1.5) +
(( roudaboutPresence=0) *  (IntersectionPresence=1) * (speed5sAfter>=1)  *2) +
(( roudaboutPresence=0)  *  (IntersectionPresence=0) * (speedlimitReduction=1)  *4) +
(( roudaboutPresence=0)  *  (IntersectionPresence=0) * (speedlimitReduction=0)  * (Curve10s=1)*6) +
(( roudaboutPresence=0) * (IntersectionPresence=0)  * (speedlimitReduction=0)  * (Curve10s=0) *8).
```

Figure 6-53 : SPSS code to create derived variables

To characterise the driver, we have:

- IdDriver (Initial numeric attribute): Unique identifier for each driver, number
- Gender (Nominal derived attribute): Driver type, M male, F Female.
- Age (Numeric derived attribute): Age of driver in year.
- CatAge: age category of the driver (0: Undefined, 1: 17/30, 2: 30/40, 3: 40/60, 4: 60+).

To locate the event temporally, we have:

- BeginDate: The date and time of the start of the record containing this event.
- BeginTime: The start time of the event in the second record.
- EventDuration: The duration of the event in seconds.
- DayNight: Type of time slot during event (0: Undefined; 1: Dawn; 2: Day; 3: Twilight; 4; Night)

To characterise the intensity of the event, we have:

- BrakeType: event type (0: undefined, 1: 1m/s^{2}for at least 2 seconds, 3: 3m/s^{2}for at least 0.5 seconds, 5: 5m/s^{2}for at least 0.3 second) .
- MinBrake, meanBrake: The minimum and average intensity of the event, calculated on the time interval between the start and the end of the event.
- SpeedVariation: the difference in speed between the speed at the end of the event and the speed at the beginning of the event ?

UDRIVE

European Naturalistic Driving Study

To characterise interactions with other users during the event, we have:

- MinTTC, meanTTC: Minimum and average time at collision between the participant's vehicle and the previous vehicle, calculated on the time interval between the start and the end of the event.
- TTCStartBrake, TTCEndBrake: The value of the collision time between the participant's vehicle and the lead vehicle, recorded at the beginning and at the end of the event.
- TIVStartBrake, TIVEndBrake: The value of the inter-vehicular time between the participant's vehicle and the lead vehicle, recorded at the beginning and at the end of the event.

To characterize the dynamics of the vehicle, we have:

- ABSActive, ASRActive, AYCActive: Activation of vehicle safety systems during braking (-1: undefined, 0: No activation, 1: Activation)
- Speed5sBefore, Speed5sAfter: Vehicle speeds 5s before start and brake and 5s after braking ends.
- CCSLState5s: indication of the status of the speed management systems 5 seconds before the start of braking (0: off; 1: speed controller; 2: speed limiter; 7: undefined)
- FlashingIndicator: Indicates the use of turn signal lights (0: none, 1: right flashing light, 2: left turn signal light, 3: both flashing lights.) In the time interval of 10 seconds before the start of braking and its end.

To charactersze the road situation of the event, we have:

- RoudAboutPresence, (Initial Nominal Attributes): Indicator to indicate whether a roundabout or intersection was present (-1: undefined, 0: No presence, 1: Presence) a certain distance after the end of braking (20 metres if the speed limit was less than 55 km/h if not 50 metres)
- Curve10s: Indicator to determine whether a curve was present (-1: undefined, 0: No presence, 1: Presence) during the 10 s after the braking was completed.
- SpeedLimitReduction: Indicator to indicate whether the limit speed was reduced (-1: undefined, 0: No reduction, 1: reduction) in the time interval of 10 seconds before the start of braking and 10 seconds after the end of braking.
- RainState: Indicator to indicate the presence of rain (0: No rain, 1: Rain, 2: Significant rain).
- CatdiffMaxMeanGyroym21 (Nominal derived attribute): Pavement status category (-1: Undefined, 0: Plane, 1: Fairly flat, 2: Small deformation such as small holes or kerb crossing, speed humps);
- DiffMaxMeanGyroym21: Indicator to measure pavement condition;

This indicator was created by processing the Y-accelerometer signal. To do this, a Hilbert transform was used to calculate the envelope area of this signal (Yger A., 1999). Then, we averaged this signal on 21 points corresponding to 0.66 seconds. We calculated an indicator that is the difference between the maximum value of the averaged envelope and its average on a 30 second window. We can see in the Figure 6-54 the Hilbert transform averaged over 21 points of a segment during which the vehicle passed over a speed hump.

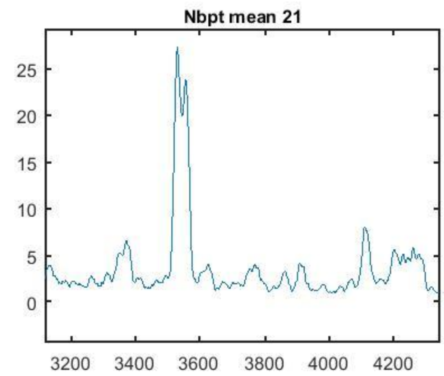


Figure 6-54 : Example of indicator when passing over a speed hump

Descriptive analysis of data

The first check was to see whether the panels of participants in all countries had similar braking behaviours.

Table 6-25: Number of braking types per operation site

BrakeType * OperationSite Crosstabulation

			OperationSite					
			OS-DE	OS-FR	OS-NL2	OS-PL	OS-UK	Total
BrakeType	1ms2During2s	Count	31838_{a}	215919_{b}	28520_{c}	82575_{d}	183529_{d}	542381
		% within OperationSite	94.5%	88.2%	87.6%	89.5%	89.7%	89.2%
	3ms2During0.5s	Count	1782_{a}	28214_{b}	3901_{b}	9411_{c}	20663_{c}	63971
		% within OperationSite	5.3%	11.5%	12.0%	10.2%	10.1%	10.5%
	5ms2During0.3s	Count	61_{a}	717_{b}	135_{c}	$322_{b, c}$	424_{a}	1659
		% within OperationSite	0.2%	0.3%	0.4%	0.3%	0.2%	0.3%
Total		Count	33681	244850	32556	92308	204616	608011
		% within OperationSite	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Each subscript letter denotes a subset of OperationSite categories whose column proportions do not differ significantly from each other at the ,05 level.

subset a for braking 3 m/s^2

subset b for braking 3 m/s^2

subset c for braking 3 m/s^2

The German participants behave much more calmly than the other panels. Indeed, they have 94.5% of slow braking and 5.5% of hard braking while the Polish and UK drivers have 89.5% and 89.7% slow braking and 10.5% and 10.3% of hard braking. The most abrupt behaviour is observed for the French and Dutch with 11.8% and 12.4% of hard braking for 88.2% and 87.6% of slow braking. These similar behaviours are found in the proportional comparison test (Test -z), which indicates that there are no significant differences in our data for French and Dutch (subset b for 3 m/s^2 braking) and UK and Polish (subset c for braking 3 m/s^2) (Table 6-25).

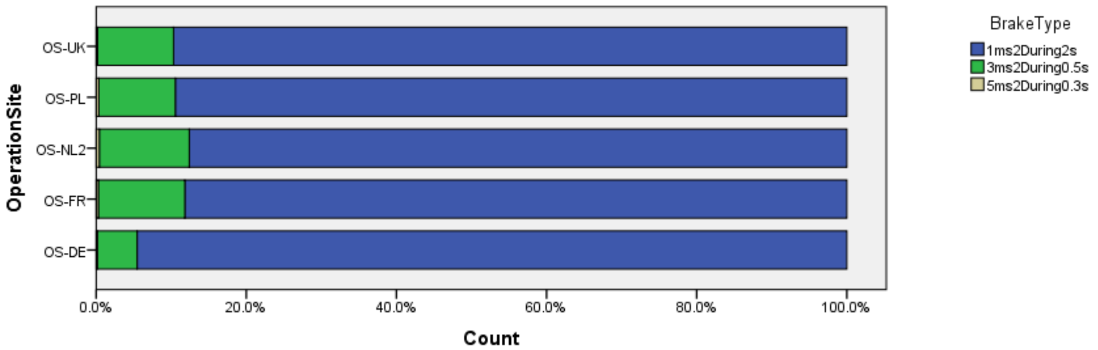


Figure 6-55: Distribution of braking types by OS

Figure 6-55 above shows these differences for braking 3m/s^2 but not for the 5 m/s^2 which are far too few in number. We have therefore enlarged the values between 0 and 3% in Figure 6-56 below.

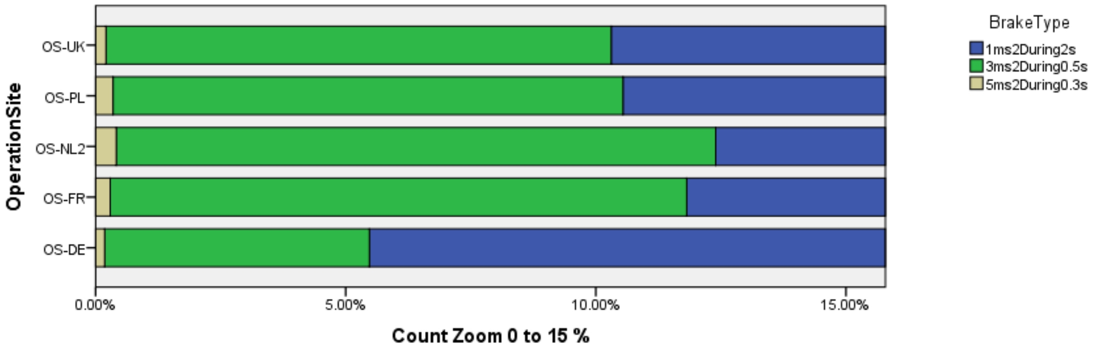


Figure 6-56: Zoom on the distribution of braking types by OS

In Figure 6-55, we can observe that the Dutch have proportionally more braking at 5 m/s^2 than the others (subset c for braking 5 m/s^2), UK and Germans least (subset a for braking 5 m/s^2) whereas the Poles and the French are in the middle (subsets b and c for braking 5 m/s^2).

The second check was to see if the compositions of panels of participants in all countries were homogeneous (Table 6-23 and Table 6-24 in section 0). The gender distribution shows that the German and Polish panels have few women. The age distribution shows that it would be difficult to study age by separating groups by OS. The UK and French panels are the only ones to have enough older drivers. Indeed, it is impossible to give conclusions on the data with less than 5 persons in a group.

Table 6-26: Distribution of participants according to the proportion of hard braking by OS

% Hard braking	OS_DE	OS_FR	OS_NL	OS_PL	OS_UK	Total
<10%	15	26	8	11	36	96
10-20%	2	10	4	11	11	38
> 20%	0	6	2	5	4	17
	17	42	14	27	51	151
	88%	62%	57%	41%	71%	
	12%	24%	29%	41%	22%	
	0%	14%	14%	19%	8%	

The histograms of distribution of hard braking for each panel (see annex B.1) are summarised in Table 6-26. This table indicates how many participants have less than 10% of hard braking, between 10% and 20% and more than 20% of hard braking. Both Germans and the UK groups have a proportion of hard braking much less than the other groups. The Polish group have a little more hard braking while the Dutch and the French groups have a distribution similar to that of all participants.

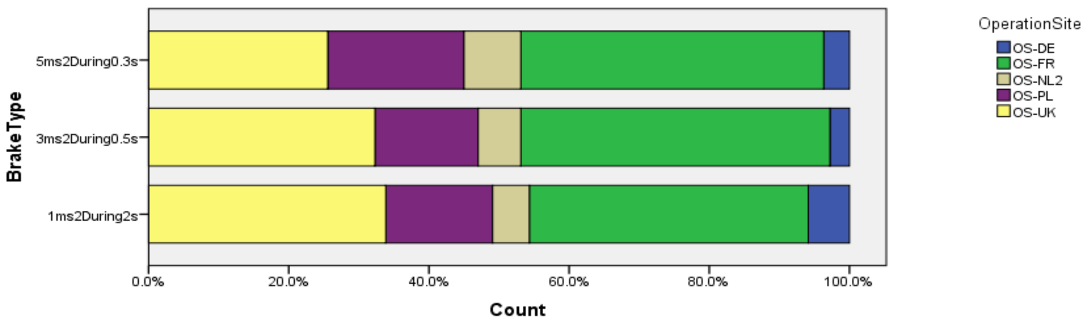


Figure 6-57 : Breakdown of braking by operational sites.

We can observe in Figure 6-57 that almost half of the braking results come from French data, one-third from the UK data and the rest from the other three operational sites.

In conclusion, we will therefore have to check in the following analyses:

- If the factors identified for all the dataset are significantly different for each data collection site.
- It will be difficult to carry out analyses by site on the age factor, in particular for the classes of young and old people, except for the French and UK sites.

Analysis of the influence of speed limits

To create the speed limit categories, we looked at the speed limits in each country and we recoded the speed limit recorded 10 seconds before the braking into four categories:

- From 1 to 34 km/h, category 1 which corresponds mostly to low speed limits in the city.
- From 34 km/h to 62 km/h, category 2 which corresponds mostly to zones of normal speed limit in the city.
- From 62 km/h to 98 km/h, category 3 which corresponds mostly to zones of speed limits in rural areas.
- More than 98 km/h, category 4 which mostly correspond to zones of speed limit on motorways or expressways.

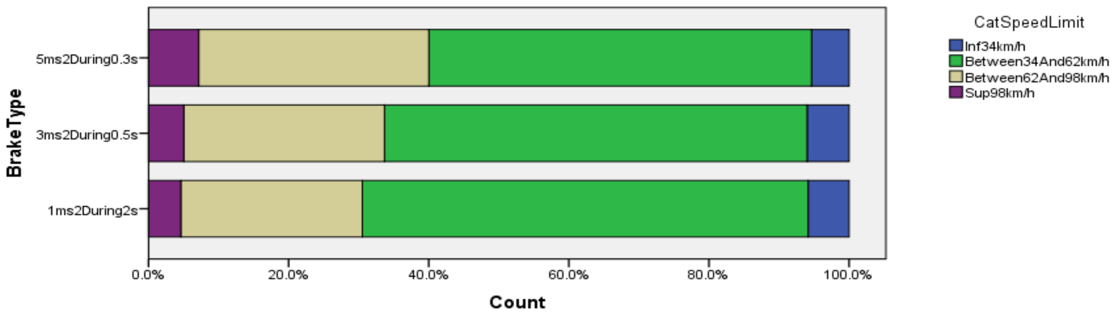


Figure 6-58: Breakdown of braking according to the speed limit.

Table 6-27: Distribution of the hard braking as a function of the speed limits

BrakeType * CatSpeedLimit Crosstabulation

			CatSpeedLimit				
			Inf34km/h	Between34 And62km/h	Between62 And98km/h	Sup98k m/h	Total
BrakeType	1ms2During2s	Count	29430a	320508b	130519c	23164c	503621
		% within CatSpeedLimit	88.95%	89.71%	88.13%	88.18%	89.18%
	3ms2During0.5s	Count	3573a	35941b	17069c	2996a, c	59579
		% within CatSpeedLimit	10.80%	10.06%	11.53%	11.40%	10.55%
	5ms2During0.3s	Count	82a	838a	504b	110b	1534
		% within CatSpeedLimit	0.25%	0.23%	0.34%	0.42%	0.27%
Total		Count	33085	357287	148092	26270	564734
		% within CatSpeedLimit	100.0%	100.0%	100.0%	100.0%	100.0%

Each subscript letter denotes a subset of CatSpeedLimit categories whose column proportions do not differ significantly from each other at the ,05 level.

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	333.203[1]	6	.000
Likelihood Ratio	326.101	6	.000
Linear-by-Linear Association	194.080	1	.000
N of Valid Cases	564734		

1. 0 cells (,0%) have expected count less than 5. The minimum expected count is 71,36.

The distributions (Figure 6-57) are significantly different according to the different speed limit categories (Table 6-27). The 5m/s^{2}hard braking has two significantly different groups:

- one group (subset a for braking 5 m/s^{2}with a proportion less than 0.3%) for speed limits below 62 km/h corresponding to the city, most of the time
- a second group (subset b for 5 m/s^{2}braking with a proportion greater than 0.3%) for speeds exceeding 62 km/h corresponding mostly to extra-urban areas.

The other brakings have three significantly different groups:

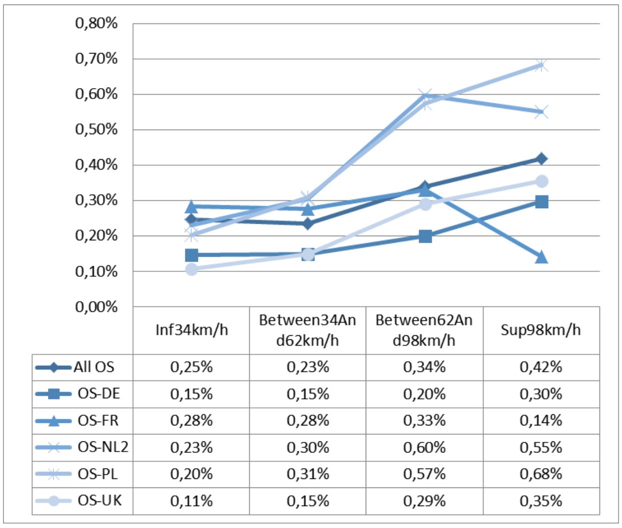
- one group (subset a with a proportion of 10.8% for braking 3 m/s^2) for speed speeds below 34 km/h, which mostly correspond to zones of low speed limits in the city,
- an intermediate group (subset b with a proportion of 10.1% for braking 3 m/s^2) for speed limits between 34 km/h and 62 km/h which corresponds mostly to zones of normal speed in the city,
- another group (subset c with a proportion greater than 11% for braking 3 m/s^2) for speeds exceeding 62 km/h which corresponds mostly to extra-urban areas.

Since the different operational sites are not homogeneous from the point of view of the distribution of braking, we have carried out these analyses by country (see Annex B.2).

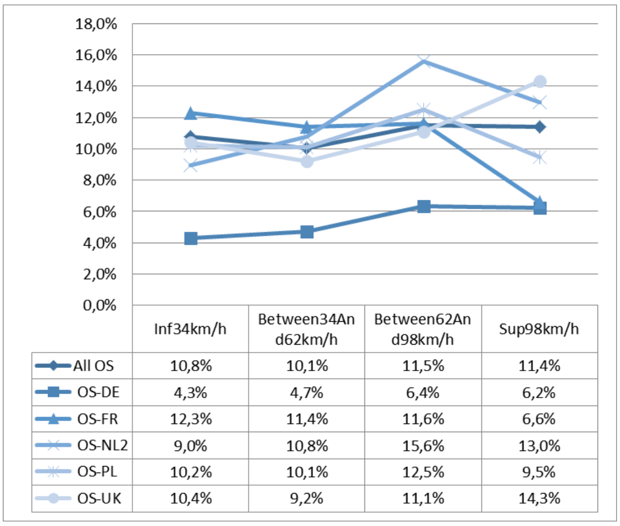
Table 6-28 : Percentage of braking according to the speed limit by OS.

OS	Below 34km/h	Between 34 and 62km/h	Between 62 and 98km/h	Above 98km/h
All OS	5,9%	63,3%	26,2%	4,7%
OS-DE	11,4%	63,4%	11,7%	13,5%
OS-FR	9,2%	61,4%	28,4%	0,9%
OS-NL2	10,5%	50,1%	24,4%	15,1%
OS-PL	5,8%	76,6%	12,2%	5,4%
OS-UK	0,5%	61,6%	32,2%	5,7%

Table 6-28 describes the distribution of braking (all braking events) by speed limits for each OS. The UK participants drove twenty times less and the Polish half as much as the others in zones with low speed limits. The German and Dutch participants drove three times more than the others on highways. The French participants drove five times less than the others on highways. In France, speed limits on motorways approaching the cities are limited to 90 km/h.



	Inf34km/h	Between34And62km/h	Between62And98km/h	Sup98km/h
All OS	0,25%	0,23%	0,34%	0,42%
OS-DE	0,15%	0,15%	0,20%	0,30%
OS-FR	0,28%	0,28%	0,33%	0,14%
OS-NL2	0,23%	0,30%	0,60%	0,55%
OS-PL	0,20%	0,31%	0,57%	0,68%
OS-UK	0,11%	0,15%	0,29%	0,35%



	Inf34km/h	Between34And62km/h	Between62And98km/h	Sup98km/h
All OS	10,8%	10,1%	11,5%	11,4%
OS-DE	4,3%	4,7%	6,4%	6,2%
OS-FR	12,3%	11,4%	11,6%	6,6%
OS-NL2	9,0%	10,8%	15,6%	13,0%
OS-PL	10,2%	10,1%	12,5%	9,5%
OS-UK	10,4%	9,2%	11,1%	14,3%

Figure 6-59 : Breakdown of brakings of 5 m/s^2 and 3 m/s^2 by speed limits.

We can see in Figure 6-59 that two different groups for 5m/s^2 braking are provided by the Polish and Dutch data versus the rest. For brakings of 3 m/s^2, the difference between categories 1 (inf34) and 2 (34-62) is provided by the French and UK data, while the different behaviour of these two panels in category 4 is neutralised. Indeed, the strong increase for the UK and the sharp decrease for the French compensate for each other. The French seem to have a more driving on highways.

<u>Analysis of the influence of the time of day</u>

To create the categories of periods of the day, we used the existing time series in SALSA which defined four categories: sunrise, daytime, sunset and night.

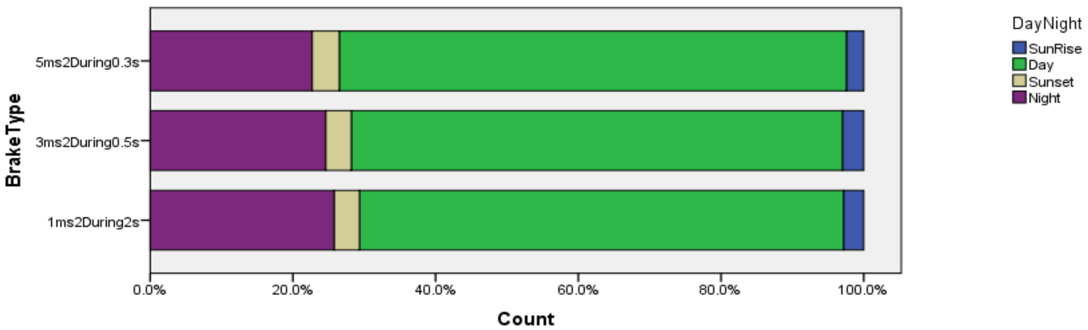


Figure 6-60 : Breakdown of braking by time of day.

Table 6-29: Distribution of hard braking by time of day.

BrakeType * DayNight Crosstabulation

			DayNight				
			SunRise	Day	Sunset	Night	Total
BrakeType	1ms2During2s	Count	14258a	340514a	17915a, b	129357b	502044
		% within DayNight	88.7%	89.1%	89.1%	89.7%	89.2%
	3ms2During0.5s	Count	1773a	40761a	2126a, b	14558b	59218
		% within DayNight	11.0%	10.7%	10.6%	10.1%	10.5%
	5ms2During0.3s	Count	37a, b	1087b	58a, b	347a	1529
		% within DayNight	0.2%	0.3%	0.3%	0.2%	0.3%
Total		Count	16068	382362	20099	144262	562791
		% within DayNight	100.0%	100.0%	100.0%	100.0%	100.0%

Each subscript letter denotes a subset of DayNight categories whose column proportions do not differ significantly from each other at the ,05 level.

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	49.894[1]	6	.000
Likelihood Ratio	50.377	6	.000
Linear-by-Linear Association	45.363	1	.000
N of Valid Cases	562791		

1. 0 cells (,0%) have expected count less than 5. The minimum expected count is 43,65.

The distributions (Figure 6-60) are significantly different according to the different categories of periods of the day (Table 6-29). Whatever the type of braking, the Day and Night categories are significantly different. On the other hand, the category Sunset does not differ from either of the two groups for all braking. The Sunset category does not differ from the Day category but differs from the Night category except for braking at 5m/s^2.

Since the different operational sites are not homogeneous from the point of view of the distribution of braking, we have carried out these analyses by country (see Annex B.3).

Table 6-30 : Percentage of braking according to time of day by OS.

OS	Sunrise	Day	Sunset	Night
All OS	3%	68%	4%	26%
OS-DE	2%	78%	2%	18%
OS-FR	4%	62%	4%	30%
OS-NL2	2%	76%	4%	19%
OS-PL	2%	78%	3%	17%
OS-UK	3%	67%	3%	27%

Table 6-30 describes the distribution of braking (all brakings) by period of day for each OS. We can see that two-thirds of the data were collected from daytime periods and a quarter of the data were collected at night. The French data have more data at night because the collection began in 2016 and the participants drove during two winters.

Table 6-31 : Breakdown of braking by 5 m/s^2 according to the period of the day by OS

OS	SunRise	Day	Sunset	Night
All OS	37a, b	1087b	58a, b	347a
OS-DE	1a	40a	0a	10a
OS-FR	21a	432a	34a	171a
OS-NL2	1a	92a	6a	17a
OS-PL	5a	233a	8a	52a
OS-UK	9a	290a	10a	97a

Table 6-31 shows that the effect on the braking of 5 m/s^2 is not confirmed in any of the sites. We cannot therefore assert that the occurrence of braking of 5 m/s^2 is affected by the time of day.

Table 6-32: Breakdown of braking by 3 m/s^2 according to the period of the day by OS.

OS	SunRise	Day	Sunset	Night
All OS	1773a	40761a	2126a, b	14558b
OS-DE	31a, b	1211b	41a, b	214a
OS-FR	921a	16336a	1007a	7585a
OS-NL2	50a	2623a	189b	618a
OS-PL	121a, b	7169b	236a	1178a
OS-UK	650a	13422b	653b, c	4963c

By contrast, when looking at brakings of 3 m/s^2, Table 6-32 shows that the effect day/night is significant for the German, Polish and UK sites. The French data do not show any difference. Finally, Dutch data do not show a Day/Night difference but a difference of the category Sunset compared to all other categories.

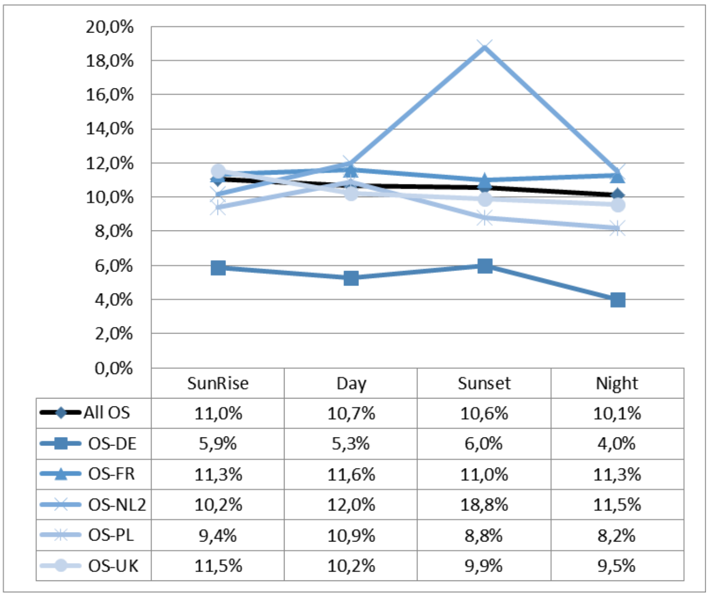


Figure 6-61: Distribution of the braking percentages 3m/s^2 by categories of periods of the day.

Figure 6-61 shows that the difference between the types day and night is provided by the German data (5.3% day, 4.0% night), Poland (10.9% day, 8.2% night) and UK (10.2% day, 9.7% night). We also did not show any effect for the French data. An interesting effect on the sunset category was found on the Dutch data (18.8% sunset, 12% day, 11.5% night). Given the small amount of Dutch data compared to the other sites, we should check that this effect is not due to another factor such as traffic jams.

Analysis of the influence of rain

To create the categories of weather, we used the existing time series in SALSA which defined three categories: No rain, Rain and Heavy Rain. These data was created using the wiper signal. It is not possible to identify a distinction between snow and rain with this signal, and it also cannot evaluate whether the road surface was slippery due to rain or ice when rain has stopped.

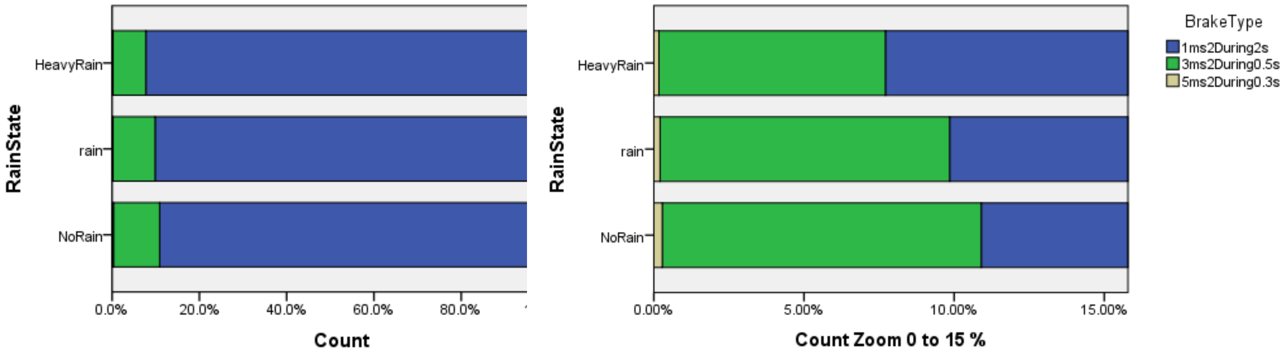


Figure 6-62 : Breakdown of braking by rain categories.

Table 6-33 : Distribution of braking by rain categories

BrakeType * RainState Crosstabulation

			RainState			
			NoRain	rain	HeavyRain	Total
BrakeType	1ms2During2s	Count	466073a	36459b	1089c	503621
		% within RainState	89.1%	90.1%	92.3%	89.2%
	3ms2During0.5s	Count	55589a	3901b	89c	59579
		% within RainState	10.6%	9.6%	7.5%	10.5%
	5ms2During0.3s	Count	1449a	83b	2a, b	1534
		% within RainState	0.3%	0.2%	0.2%	0.3%
Total		Count	523111	40443	1180	564734
		% within RainState	100.0%	100.0%	100.0%	100.0%

Each subscript letter denotes a subset of RainState categories whose column proportions do not differ significantly from each other at the ,05 level.

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	57.899[1]	4	.000
Likelihood Ratio	60.605	4	.000
Linear-by-Linear Association	56.222	1	.000
N of Valid Cases	564734		

1. 1 cells (11,1%) have expected count less than 5. The minimum expected count is 3,21.

The distributions (Figure 6-62) are significantly different according to the rain categories (Table 6-33). Regardless of the type of braking, the No Rain and Rain categories are significantly different. The heavy Rain category does not differentiate between the two groups for the braking of 5m/s^{2}because the number of brakings in this category is relatively limited.

Since the various operational sites are not homogeneous from the point of view of the distribution of braking, we have made these analyses by country (see annex B.4).

Table 6-34 : Percentage of braking according to rain categories by OS

OS	No rain	Rain	Heavy rain
All OS	92.6%	7.2%	0.2%
OS-DE	92.1%	7.6%	0.3%
OS-FR	96.3%	3.5%	0.1%
OS-NL2	100.0%	0.0%	0.0%
OS-PL	91.6%	8.0%	0.4%
OS-UK	87.8%	12.0%	0.2%

Table 6-34 describes the distribution of braking (all braking events) by rain categories for each OS. Situations with heavy rainfall are not sufficiently numerous to take into account in the analyses. In addition, Dutch data do not have rain data. The French data collected in the south of France have less rain than the others and those from the UK have more.

Table 6-35 : Breakdown of braking by 5 m/s^2 as a function of rain by OS

OS	NoRain	Rain	HeavyRain
All OS	1449a	83b	2a, b
OS-DE	46a	6a	0a
OS-FR	641a	18a	0a
OS-NL2	117	0	0
OS-PL	281a	18a	1a
OS-UK	364a	41a	1a

Table 6-35 shows that the effect for braking of 5 m/s^2 is not confirmed in any of the sites. We cannot therefore assert that the occurrence of 5m/s^2 braking is impacted by rain.

Table 6-36 : Breakdown of braking by 3 m/s^2 as a function of rain by OS

OS	NoRain	Rain	HeavyRain
All OS	55589a	3901b	89c
OS-DE	1398a	100a	7a
OS-FR	25131a	832b	19a, b
OS-NL2	3492		
OS-PL	8137a	592b	22a, b
OS-UK	17431a	2377a	41a

Table 6-36 shows that the effect Rain / NoRain is significant for the Polish and French sites. The German and UK data show no difference. The effect of heavy rain did not show up in any of the OS for braking at 3 m/s^2. It is possible that the number of brakings with heavy rain is not enough to detect an effect.

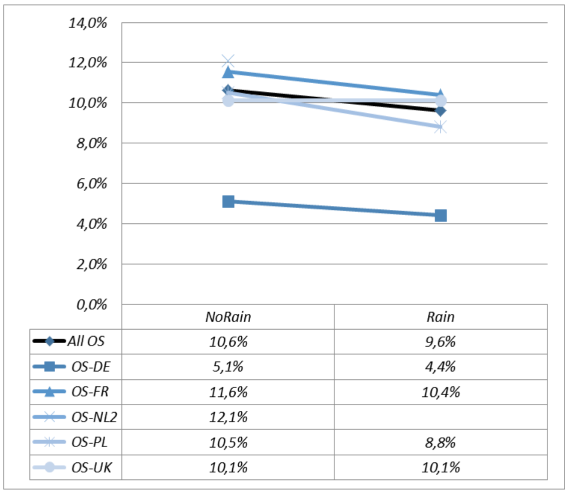


Figure 6-63 : Distribution of braking percentages 3m/s^2 by rain categories.

Given the fact that the effect of the two groups is different for braking of 3 m/s^2, the difference between Rain and No Rain categories (Figure 6-63) is provided by the Polish data (10.5% NoRain, 8.8% Rain) and French (11.6 % NoRain, 10.4% Rain).

Analysis of the influence of the use of assistance system

To create categories of assistance system usage, we used the existing time series in SALSA which defined three categories: System off, Cruise control on, Speed limiter on.

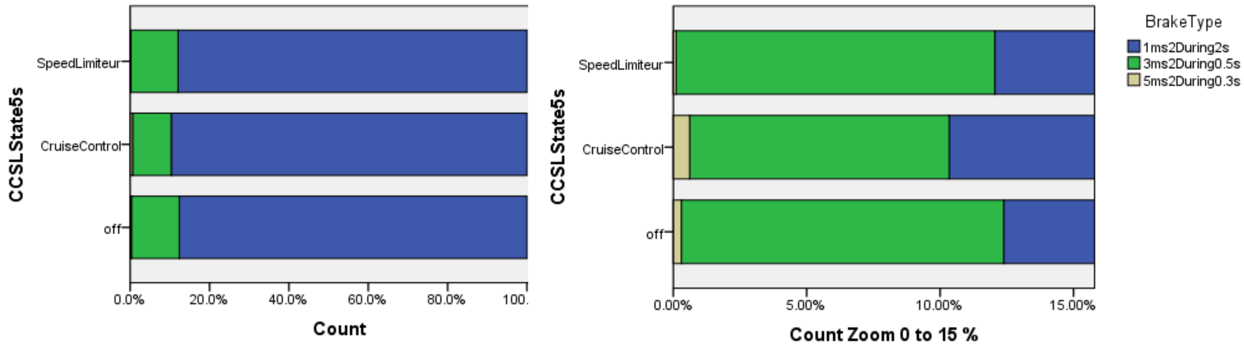


Figure 6-64 : Breakdown of braking according to the use of the assistance system.

Table 6-37: Distribution of hard braking according to the use of assistance system.

BrakeType * CCSLState5s Crosstabulation

			CCSLState5s			
			off	CruiseControl	SpeedLimiteur	Total
BrakeType	1ms2During2s	Count	334083a	3077b	11611a	348771
		% within CCSLState5s	87.6%	89.7%	87.9%	87.6%
	3ms2During0.5s	Count	46078a	334b	1577a	47989
		% within CCSLState5s	12.1%	9.7%	11.9%	12.1%
	5ms2During0.3s	Count	1143a	21b	14c	1178
		% within CCSLState5s	0.3%	0.6%	0.1%	0.3%
Total		Count	381304	3432	13202	397938
		% within CCSLState5s	100.0%	100.0%	100.0%	100.0%

Each subscript letter denotes a subset of CCSLState5s categories whose column proportions do not differ significantly from each other at the ,05 level.

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	45.358[1]	4	.000
Likelihood Ratio	49.017	4	.000
Linear-by-Linear Association	5.663	1	.017
N of Valid Cases	397938		

1. 0 cells (,0%) have expected count less than 5. The minimum expected count is 10,16.

The distributions (Figure 6-64) are significantly different (Table 6-37) according to the different categories of assistance system use only between the Off category and the speed controller category. Regardless of the type of braking, the Off and Cruise Control categories are significantly different. The speed limiter category does not differ from the group Off for the braking of 3m/s^{2}but it differs for the braking of 5m/s^{2}from the two other categories.

Since the various operational sites are not homogeneous from regarding the distribution of braking, we made these analyses by country (see annex B.5)

Table 6-38 : Percentage of braking according to the use of driving assistance systems by OS

OS	Off	Cruise control	Speed limiter
All OS	95.8%	0.9%	3.3%
OS-DE	97.0%	3.0%	
OS-FR	93.8%	0.2%	6.0%
OS-NL2	94.6%	4.5%	1.0%
OS-PL	99.0%	1.0%	
OS-UK	99.1%	0.7%	0.2%

This distribution (Table 6-38) does not represent the use of driving systems in general. It is only the distribution of brakings according to system use. The French use the cruise control less but they drove very little on roads with speeds above 98 km/h. The Dutch and Germans, who drove more on roads with speeds above 98 km/h, use the cruise control more.

Table 6-39: Breakdown of braking by 5 m/s^{2}depending on the use of ADAS system by OS.

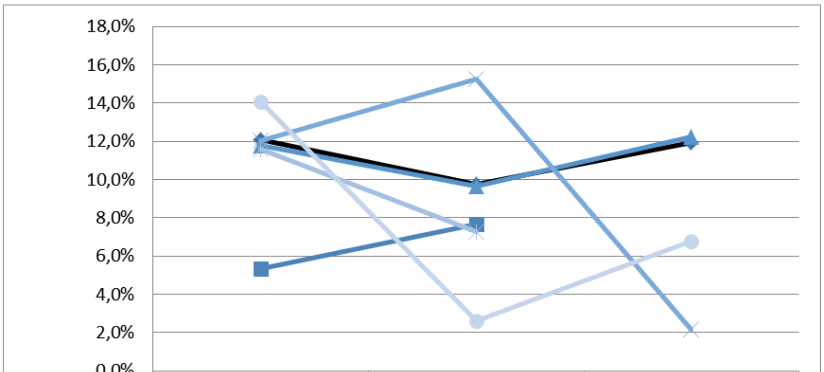
OS	off	CruiseControl	SpeedLimiteur
All OS	1143a	21b	14c
OS-DE	27a	1a	0
OS-FR	620a	2a, b	13b
OS-NL2	104a	13b	0a, b
OS-PL	144a	3a	0
OS-UK	248a	2a	1a

Table 6-39 shows that the effect of ADAS system use on brakings of 5 m/s^2 is confirmed only for the French site for the speed limiter and the Dutch site for the cruise control.

Table 6-40: Breakdown of braking by 3 m/s^2 depending on the use of the ADAS system by OS.

OS	off	CruiseControl	SpeedLimiteur
All OS	46078a	334b	1577a
OS-DE	850a	38b	0
OS-FR	23694a	49a	1560a
OS-NL2	3289a	197b	6c
OS-PL	5006a	32b	0
OS-UK	13239a	18b	11c

However, Table 6-40 shows that the braking effect of 3 m/s^2 for condition Off, Cruise Control, and Speed limiter is not significant for the French site but it is significant for all other sites for cruise control. It is significant for the UK and Dutch sites for the speed limiter but it should be noted that the numbers of brakings are very low. More data is needed to confirm this result.



	off	CruiseControl	SpeedLimiteur
All OS	12,1%	9,7%	11,9%
OS-DE	5,4%	7,6%	
OS-FR	11,8%	9,6%	12,2%
OS-NL2	12,0%	15,3%	2,2%
OS-PL	11,6%	7,3%	
OS-UK	14,1%	2,6%	6,7%

Figure 6-65 : Distribution of the braking percentages 3m/s^2 according to the use of the OS support system.

In Figure 6-65 we examine whether ADAS system use affected braking of 3 m/s^2. For all OS apart from the German and the Netherlands sites, the use of cruise control decreased the percentage of hard braking. The speed limiter was only used in French OS and had no influence on occurrence of 3 m/s^2 braking.

<u>Analysis of the influence of age</u>

To create the categories of use of support system, we used the categorization presented above which defined four categories:

- Younger than 30 years
- Between 31 and 40 years
- Between 41 and 60 years
- Older than 60 years

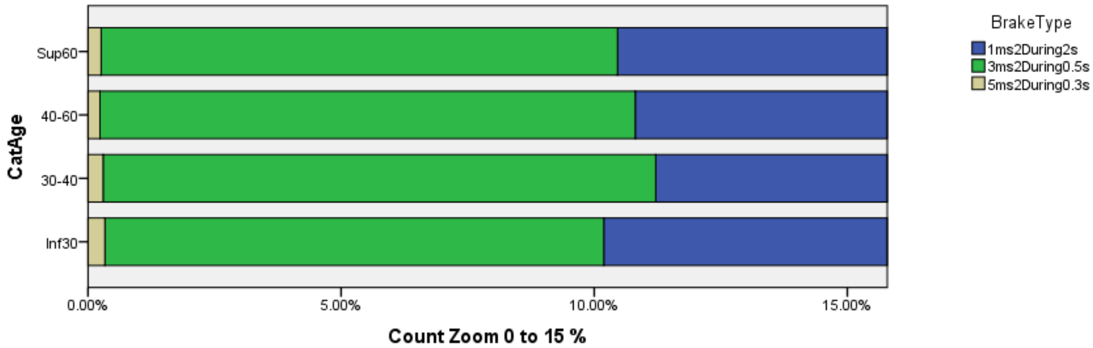


Figure 6-66 : Breakdown of braking according to the age.

Table 6-41 : Distribution of hard braking by age

BrakeType * CatAge Crosstabulation

			CatAge				
			Inf30	30-40	40-60	Sup60	Total
BrakeType	1ms2During2s	Count	85383a	159799b	236597c	60585a, c	542364
		% within CatAge	89.8%	88.8%	89.2%	89.5%	89.2%
	3ms2During0.5s	Count	9368a	19648b	28048c	6902a	63966
		% within CatAge	9.9%	10.9%	10.6%	10.2%	10.5%
	5ms2During0.3s	Count	318a	535a, b	630c	176b, c	1659
		% within CatAge	0.3%	0.3%	0.2%	0.3%	0.3%
Total		Count	95069	179982	265275	67663	607989
		% within CatAge	100.0%	100.0%	100.0%	100.0%	100.0%

Each subscript letter denotes a subset of CatAge categories whose column proportions do not differ significantly from each other at the ,05 level.

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	112.325[1]	6	.000
Likelihood Ratio	112.496	6	.000
Linear-by-Linear Association	.000	1	.999
N of Valid Cases	607989		

1. 0 cells (,0%) have expected count less than 5. The minimum expected count is 184,63.

The distributions (Figure 6-66) are significantly different (Table 6-41) for all braking only between the first three categories. As we anticipated by studying the numbers of participants, there are not enough seniors to draw conclusions about this population.

Since the various operational sites are not homogeneous regarding the distribution of braking, we made these analyses by country (see annex B.6).

Table 6-42 : Percentage of all braking according to age by OS

OS	Inf30	30-40	40-60	Sup60
All OS	16%	30%	44%	11%
OS-DE	23%	11%	47%	19%
OS-FR	18%	24%	48%	10%
OS-NL2	28%	9%	52%	11%
OS-PL	6%	57%	31%	5%
OS-UK	14%	30%	42%	14%

Table 6-42 describes the distribution of braking (all braking events) by age categories for each OS. The most represented age groups are 30-40 and 40-60 years old. It is therefore normal that they have more braking. The distribution of braking for all data are similar to the gender participant distributions (Table 6-24).

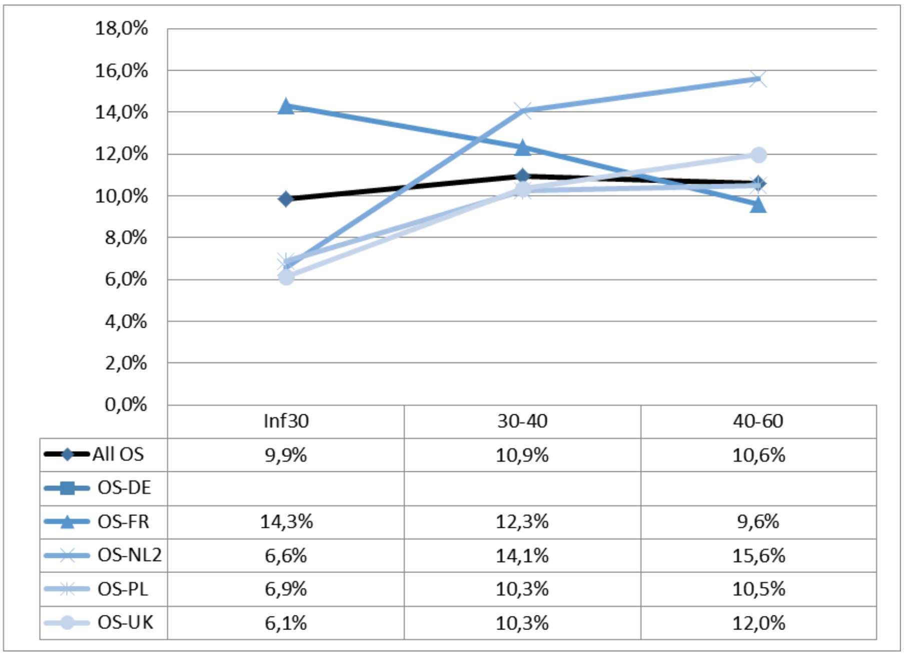


Figure 6-67 : Distribution of braking percentages 3m/s^2 as a function of age by OS

Given the fact that the effect of the two groups is different for braking of 3 m/s^2, the difference between the categories (Figure 6-67) between the lower category to 30 years and 40-60 years is provided by the Dutch data (6.6% < 30 years, 15.6% 40-60 years) and UK data (6.1% <30 years,12.0% 40-60 years) whereas French data are the opposite way round (14.3% <30 years, 9.6% 40-60 years). Young French drivers seem to be have more abrupt driving.

Analysis of the influence of gender

We used the categorisation into two categories: Male and Female

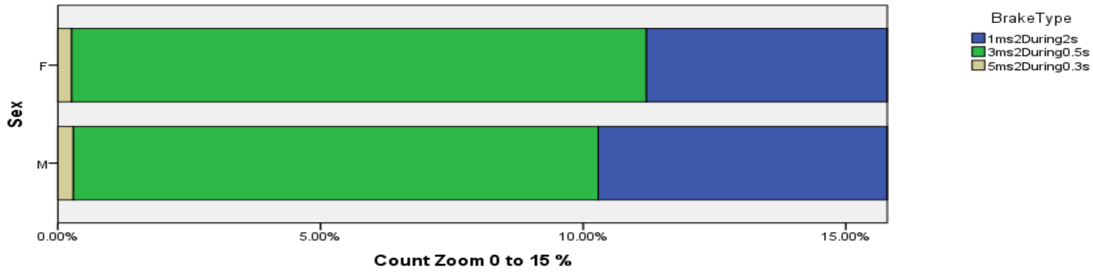


Figure 6-68 : Breakdown of braking by gender.

Table 6-43 : Breakdown of braking by gender

BrakeType * Sex Crosstabulation

			Sex M	Sex F	Total
BrakeType	1ms2During2s	Count	242316a	300048b	542364
		% within Sex	89.7%	88.8%	89.2%
	3ms2During0.5s	Count	26992a	36974b	63966
		% within Sex	10.0%	10.9%	10.5%
	5ms2During0.3s	Count	787a	872b	1659
		% within Sex	0.3%	0.3%	0.3%
Total		Count	270095	337894	607989
		% within Sex	100.0%	100.0%	100.0%

Each subscript letter denotes a subset of Sex categories whose column proportions do not differ significantly from each other at the ,05 level.

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	148.695[1]	2	.000
Likelihood Ratio	149.143	2	.000
Linear-by-Linear Association	112.502	1	.000
N of Valid Cases	607989		

1. 0 cells (,0%) have expected count less than 5. The minimum expected count is 737,00.

The distributions (Figure 6-68) are significantly different (Table 6-43) for all braking according to the different gender categories. The women seem to do more hard braking that men.

Since the various operational sites are not homogeneous regarding the distribution of braking, we made these analyses by country (see annex B.7 Breakdown of braking by genre and operational site

Table 6-44 : Percentage of all braking according to gender by OS

OS	M	F
All OS	44%	56%
OS-DE	65%	35%
OS-FR	44%	56%
OS-NL2	23%	77%
OS-PL	72%	28%
OS-UK	32%	68%

Table 6-44 describes the distribution of braking (all brakings) by gender categories for each OS. The distributions of braking for German and French data are similar to their gender participant distributions (Table 6-23). The UK women seem to do less driving than UK men. The Dutch women seem to do more driving than Dutch men.

Table 6-45 : Breakdown of braking by 5 m/s^2 as a function of gender by OS

OS	M	F
All OS	787a	872b
OS-DE	48a	13b
OS-FR	374a	343b
OS-NL2	33a	102a
OS-PL	218a	104a
OS-UK	114a	310b

The Table 6-45 below shows that the gender effect on the braking of 5 m/s^2 is confirmed only for the French, UK and German sites between the under 30 and 40-60 year groups.

Table 6-46 : Breakdown of braking by 3 m/s^2 as a function of gender by OS

All OS	26992a	36974b
OS-DE	1348a	434b
OS-FR	12310a	15904b
OS-NL2	888a	3013a
OS-PL	6729a	2677a
OS-UK	5717a	14946b

However, Table 6-46 shows that the gender effect on the braking of 3 m/s^2 is not significant for the Polish site but it is significant for all other sites.

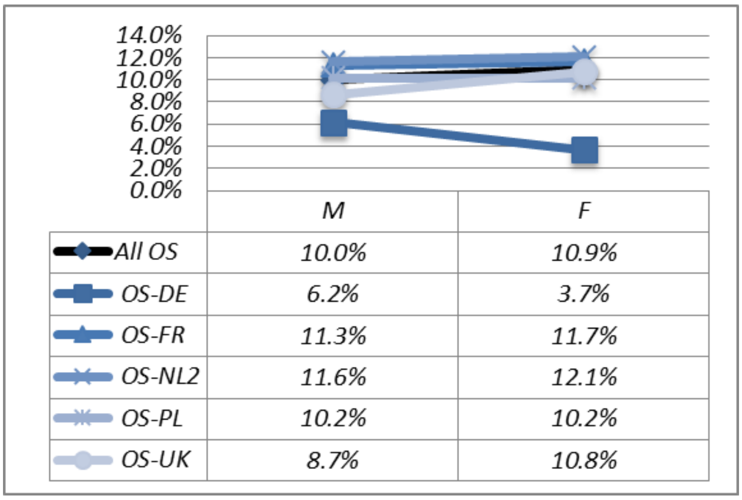


Figure 6-69 : Distribution of braking percentages 3m/s^2 as a function of gender by OS

All sites except the German and Polish sites show that women have more 3m/s^2 brakes, especially the UK women (Figure 6-69). The UK women drive less than men and brake more abruptly. We can have two explanations: less driving experience leads to harder braking or perhaps the equipped vehicles are not the usual ones for the women. The Polish data show no differences between the men and women. The German data show that women have less hard braking than men.

Analysis of the influence of the type of infrastructure

To create the infrastructure categories, we used cartographic information and our algorithm which detect deceleration to identify six infrastructure categories:

- Roundabout: the braking is associated with this category if the cartographic data indicate that a roundabout is present within less than 20 metres if the speed limit is under 55 km/h or within 50 metres (if speed limit is above 55 km/h) after the end of braking.
- Intersection with a speed of zero after 6 seconds: the braking is associated with this category if the cartographic data indicate that a roundabout is not present but an intersection is present within less than 20 metres if the speed limit is under 55 km/h or within 50 metres if the speed limit is above 55 km/) after the end of braking end and if the speed is less than 0.1 km/h 6 seconds after the end of braking. Most of the time, these situations correspond to a stop at a traffic light.
- Intersection: the braking is associated with this category if the cartographic data indicate that a roundabout is not present but an intersection is present within less than 20 metres if the speed limit is under 55 km/h or within 50 metres if the speed limit is above 55 km/h after the end of braking and if the speed is above 0.1 km/h 6 seconds after the end of braking end. This situation occurs when the car is driving across an intersection which is not a roundabout.
- Speed reduction: the braking is associated with this category if the cartographic data indicate that roundabout and intersection are not present but the speed limit 10 seconds before the braking is higher than the speed limit 10 seconds just before the brake. This situation occurs when the car is driving on a road without intersection but with a speed limit reduction.
- Curve: the braking is associated with this category if the cartographic data indicate that a roundabout, intersection or speed reduction are not present but a curve is present within less than 20 metres if the speed limit is under 55 km/h or within 50 metres if the speed limit is above 55 km/h after the end of braking. This situation occurs when the car is driving on a road with a curve and without intersection and speed reduction.
- NoCurve: all other situations. This situation occurs when the car is driving on a road without curve, roundabout, intersection or speed reduction.

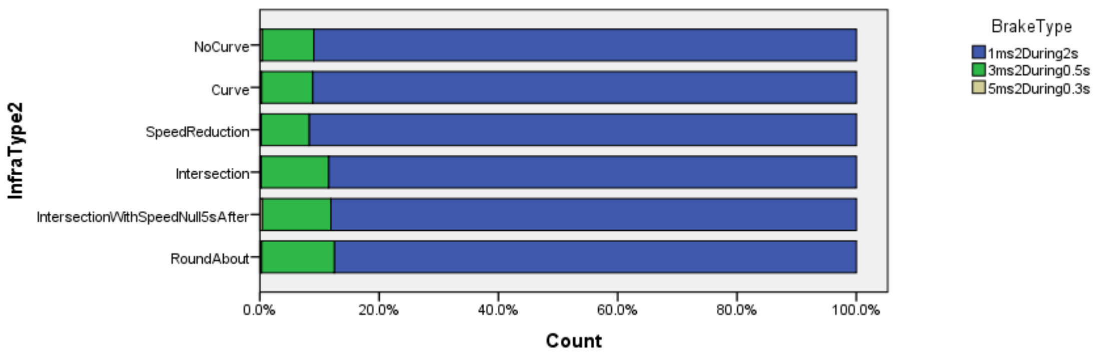


Figure 6-70 : Breakdown of braking by infrastructure type.

The situations which involve crossing other roads have more hard brakings, especially roundabouts (Figure 6-70). The speed reduction situation seems to be more anticipated than the others.

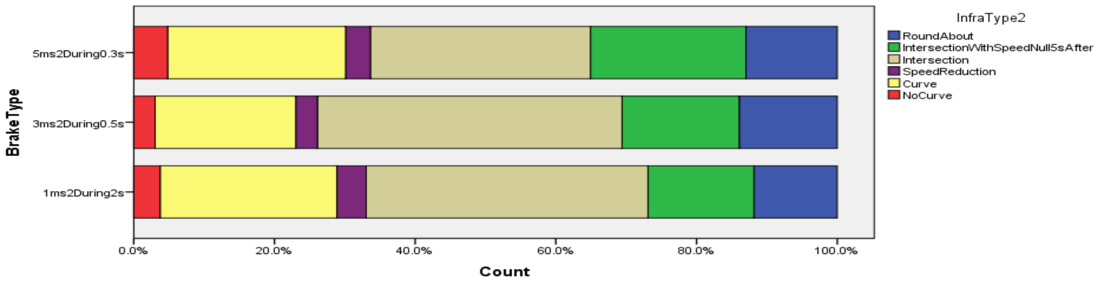


Figure 6-71 : Breakdown of the infrastructure type according to braking.

The braking of 5 m/s^2 occurs more on intersections with a zero speed and the braking of 3 m/s^2 more on roundabouts (Figure 6-71).

Table 6-47 : Distribution of braking by infrastructure type

BrakeType * InfraType2 Crosstabulation

			InfraType2						
			Round About	Intersection WithSpeed Null5sAfter	Intersection	SpeedRed uction	Curve	NoCurve	Total
BrakeType	1ms2During2s	Count	59628a	75907b	201651c	20975d	126573e	18886e	503620
		% within InfraType2	87.5%	88.1%	88.5%	91.8%	91.1%	91.0%	89.2%
	3ms2During0.5s	Count	8302a	9929b	25794b	1832c	11917c	1804c	59578
		% within InfraType2	12.2%	11.5%	11.3%	8.0%	8.6%	8.7%	10.5%
	5ms2During0.3s	Count	199a	339b	480c	54a, c	388a	74a, b	1534
		% within InfraType2	0.3%	0.4%	0.2%	0.2%	0.3%	0.4%	0.3%
Total		Count	68129	86175	227925	22861	138878	20764	564732
		% within InfraType2	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Each subscript letter denotes a subset of InfraType2 categories whose column proportions do not differ significantly from each other at the ,05 level.

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	1310.697[1]	10	.000
Likelihood Ratio	1345.356	10	.000
Linear-by-Linear Association	968.809	1	.000
N of Valid Cases	564732		

1. 0 cells (,0%) have expected count less than 5. The minimum expected count is 56,40.

In Table 6-47, we have three groups for the $3m/s^2$ brakings:

1. The first one corresponds to roundabout with more than 12% (subset a).
2. The second one corresponds to intersection with around 11.4% (subset b).
3. The third one corresponds to others situations with between 8% and 8.7% (subset c).

We have 3 groups for the $5m/s^2$ brakes:

1. The first one corresponds to roundabout at 0.3% (subset a).
2. The second one corresponds to intersection with zero speed at 0.4% (subset b)
3. The third one corresponds to intersection without zero speed at 0.2% (subset c)

The other situations are not significantly different from roundabout.

Since the various operational sites are not homogeneous regarding the distribution of braking, we made these analyses by country (see annex B.8).

Table 6-48 : Percentage of braking according to the infrastructure categories by OS

OS	Rounda bout	Intersection with speed zero 5s after	Intersection	Speed reduction	Curve	No curve
All OS	12%	15%	40%	4%	25%	4%
OS-DE	3%	22%	40%	7%	25%	4%
OS-FR	15%	14%	40%	5%	23%	3%
OS-NL2	11%	15%	38%	4%	26%	7%
OS-PL	3%	14%	39%	4%	32%	8%
OS-UK	14%	16%	42%	3%	23%	2%

Regarding Table 6-48, roundabouts in France, England and Netherlands are more often encountered by participants. German participants have more speed limit reductions and intersections with zero speed than other OS. The Polish participants have more curves than the other OS.

Table 6-49 : Breakdown of braking by 5 m/s^2 as a function of infrastructure by OS

OS	Roundabout	Intersection with speed zero 5s after	Intersection	Speed reduction	Curve	No curve
All OS	199a	339b	480c	54a, c	388a	74a, b
OS-DE	3a	13a	20a	5a	8a	3a
OS-FR	94a	141b	218a	23a	165a, b	18a, b
OS-NL2	18a	18a	30a	4a	39a	8a
OS-PL	8a, b	76b	89a	13a, b	86a	28a, b
OS-UK	76a	91a	123b	9a, b	90a, b	17a

The effects of infrastructure (Table 6-49) on all the data are not confirmed by OS for the 5 m/s^2 brakings. The German and Dutch data do not show significant differences between the categories. The French data differentiate intersections with zero speed from roundabouts, intersections and speed reduction but not from curve and no-curve. The Polish data differentiate intersections with zero speed from roundabouts, intersections and curve but not from speed reduction and no-curve. The UK data differentiate intersections from roundabouts, intersection with zero speed and no-curve but not from speed reductions and curves.

Table 6-50 : Breakdown of brakings of 3 m/s^2 as a function of infrastructure by OS

OS	Roundabout	Intersection with speed zero 5s after	Intersection	Speed reduction	Curve	No curve
All OS	8302a	9929b	25794b	1832c	11917c	1804c
OS-DE	53a, b	363b	536a	90a, b	407a, b	56a, b
OS-FR	4139a	4152a	11677a	950b	4524b	540b
OS-NL2	440a, b	727b	1139c	132a, c	849c	205c
OS-PL	365a	1297b, c	3638c	298b, c, d	2493d	660b, d
OS-UK	3305a	3390a, b	8804b	362c	3644c	343c

For the 3 m/s^2 brakings (Table 6-50), in the French and UK OS the differences are significant only between the infrastructure with or without a crossing road. In the German Polish and Dutch OS, nothing seems to be significantly different.

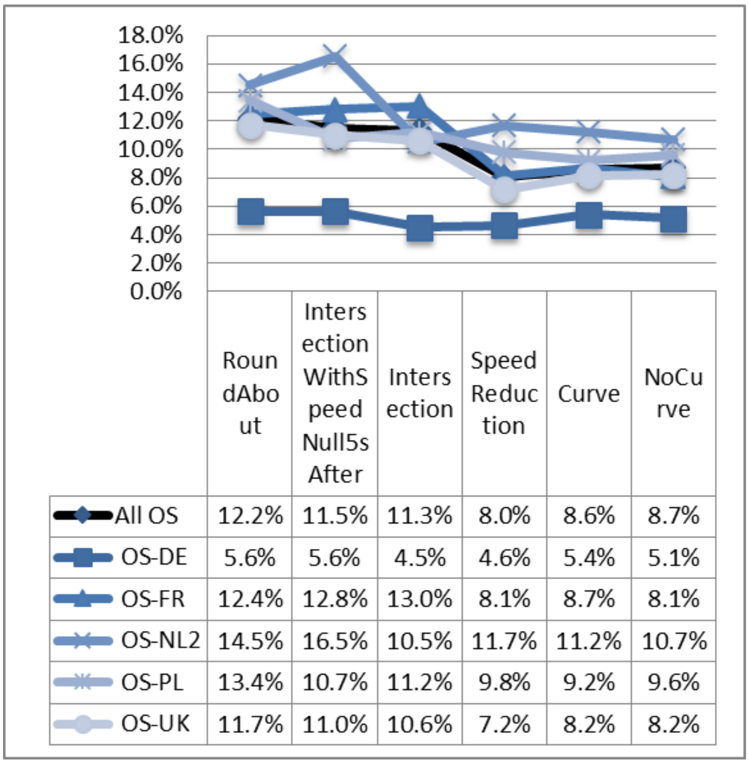


Figure 6-72 : Distribution of braking percentages 3m/s^2 as a function of infrastructure by OS

The roundabout situations have more hard brakings for most of the OS (Figure 6-72). The intersections with zero speed have also a lot of hard braking especially for the French and Dutch OS. The last three situations have few hard brakings except for German data.

Analysis of time headway time

To evaluate the urgency of the situation, we can use the THW (Time Headway) which corresponds to the time needed for the participant car to arrive at the position where the leading vehicle is. To calculate this time, the distance between the participant car and the leading vehicle is divided by the participant car speed.

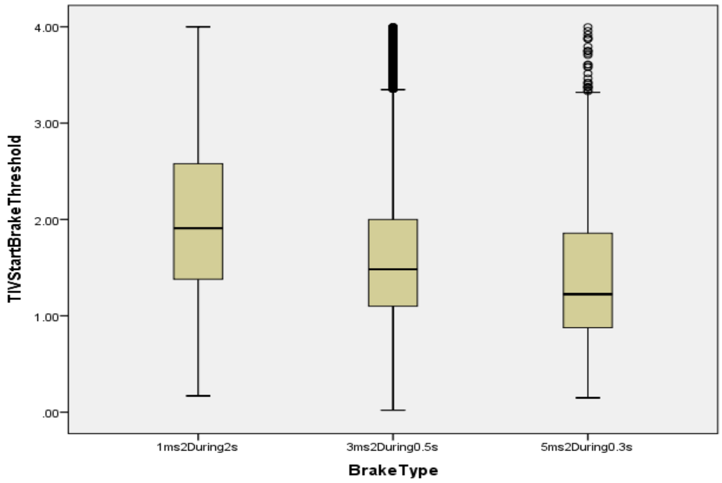


Figure 6-73 : Time headway time by type of braking

These times (Figure 6-73) are significantly different (Table 6-51) between the three types of braking. The post hoc tests show that all brakings are significantly different from the others. For the 1 m/s² brakings, the mean of TIV is 2.01 seconds; for the 3 m/s² brakings the mean of TIV is 1.63 seconds; and for the 5 m/s² brakings, the mean of TIV is 1.47 seconds.

Table 6-51 : ANOVA on TIV and Post Hoc Tests by Brake type

ANOVA

TIVStartBrakeThreshold

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	2650.210	2	1325.105	2026.946	.000
Within Groups	135822.680	207761	.654		
Total	138472.890	207763			

Multiple Comparisons

Dependent Variable: TIVStartBrakeThreshold

Bonferroni

(I) BrakeType	(J) BrakeType	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
1ms2During2s	3ms2During0.5s	.39028*	.00628	.000	.3753	.4053
	5ms2During0.3s	.54621*	.03707	.000	.4575	.6349
3ms2During0.5s	1ms2During2s	-.39028*	.00628	.000	-.4053	-.3753
	5ms2During0.3s	.15592*	.03750	.000	.0661	.2457
5ms2During0.3s	1ms2During2s	-.54621*	.03707	.000	-.6349	-.4575
	3ms2During0.5s	-.15592*	.03750	.000	-.2457	-.0661

*. The mean difference is significant at the 0.05 level.

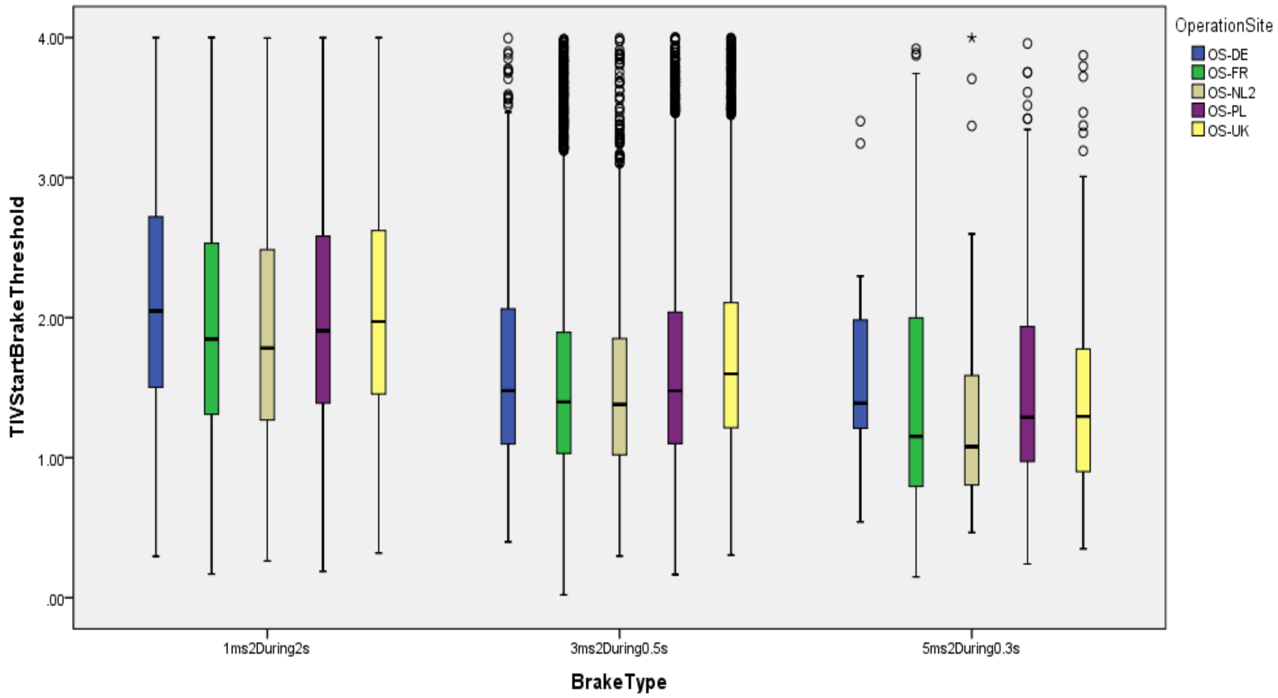


Figure 6-74 :Time headway by type of braking by OS

Table 6-52 : ANOVA on TIV by Brake type by OS

ANOVA

TIVStartBrakeThreshold

OperationSite		Sum of Squares	df	Mean Square	F	Sig.
OS-DE	Between Groups	111.042	2	55.521	85.317	.000
	Within Groups	6497.886	9985	.651		
	Total	6608.928	9987			
OS-FR	Between Groups	1035.168	2	517.584	770.807	.000
	Within Groups	51482.631	76670	.671		
	Total	52517.800	76672			
OS-NL2	Between Groups	234.632	2	117.316	177.827	.000
	Within Groups	7447.588	11289	.660		
	Total	7682.220	11291			
OS-PL	Between Groups	510.530	2	255.265	388.700	.000
	Within Groups	26424.262	40237	.657		
	Total	26934.792	40239			
OS-UK	Between Groups	677.719	2	338.859	546.262	.000
	Within Groups	43154.720	69568	.620		
	Total	43832.438	69570			

For all OS (Figure 6-74), the ANOVA tests (Table 6-52) show that the TIV for the three types of braking are significantly different, but the post hoc tests (Table 6-53) show that Braking 1 is significantly different from Braking 3 and Braking 5. Braking 3 is not significantly different from Braking 5.

Table 6-53 : Post hoc tests by brake type and by OS

	OS-DE	OS-FR	OS-NL2	OS-PL	OS-UK
1ms2During2s /3ms2During0.5s	0,000	0,000	0,000	0,000	0,000
1ms2During2s /5ms2During0.3s	0,012	0,000	0,000	0,000	0,000
3ms2During0.5s /5ms2During0.3s	1,000	0,346	0,147	0,629	0,002
	Significance > 0.05		0.001 < Significance < 0.05		Significance < 0.001

Table 6-53 summarises the post hoc tests by OS presented in annex B.9 Post Hoc tests for leading time analyse by infrastructure type and operational site

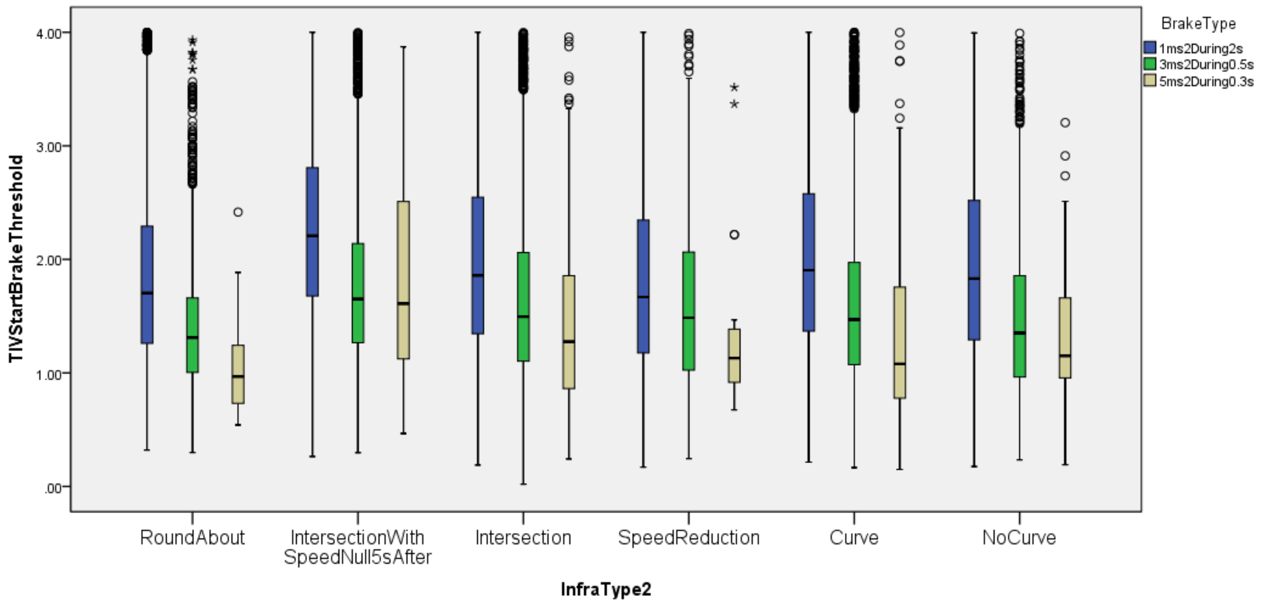


Figure 6-75 : Description of inter vehicle time at the beginning of braking by Infrastructure and OS

Table 6-54 : ANOVA on infrastructure by braking type

ANOVA

TIVStartBrakeThreshold

InfraType2		Sum of Squares	df	Mean Square	F	Sig.
RoundAbout	Between Groups	334.827	2	167.413	310.142	.000
	Within Groups	10830.463	20064	.540		
	Total	11165.290	20066			
IntersectionWithSpeedNull5sAfter	Between Groups	685.858	2	342.929	604.610	.000
	Within Groups	20088.757	35418	.567		
	Total	20774.616	35420			
Intersection	Between Groups	663.663	2	331.832	497.050	.000
	Within Groups	46314.281	69374	.668		
	Total	46977.945	69376			
SpeedReduction	Between Groups	26.371	2	13.185	19.079	.000
	Within Groups	5666.961	8200	.691		
	Total	5693.332	8202			
Curve	Between Groups	801.780	2	400.890	605.817	.000
	Within Groups	40605.317	61362	.662		
	Total	41407.097	61364			
NoCurve	Between Groups	232.092	2	116.046	168.574	.000
	Within Groups	9174.980	13328	.688		
	Total	9407.072	13330			

This effect is the same for all types of infrastructure (Figure 6-75) and the ANOVA tests show that in all OS, the TIV for the three braking are significantly different (Table 6-54). But the post hoc tests show that the differences between 3 m/s^2 and 5 m/s^2 are not significant at 0.05 for Intersection with zero speed and for No-curve (Table 6-55). The differences between 1 m/s^2 and 5 m/s^2 are not significant for Speed reduction.

Table 6-55 : Post hoc tests on infrastructure by brake type

	RoundAbout	Intersection WithSpeed Null5sAfter	Inter section	Speed Reduction	Curve	NoCurve
1ms2During2s /3ms2During0.5s	0,000	0,000	0,000	0,000	0,000	0,000
1ms2During2s /5ms2During0.3s	0,000	0,000	0,000	0,053	0,000	0,000
3ms2During0.5s /5ms2During0.3s	0,008	1,000	0,039	0,605	0,003	0,680

Table 6-55 summarises the annex B.9 Post Hoc tests for leading time analyse by infrastructure type and operational site

6.6.3 Discussion

In conclusion, we can say that all the factors studied have a significant effect on the percentage of hard braking compared to the set of brakes. On the other hand, these factors do not all have a significant effect on this percentage of hard braking on the data collected in each country. The German data show a very significant effect for the gender and the categories of speed limit, a significant effect for age, time of day and type of infrastructure but show no effect of rainfall and ADAS use. The French data show a significant effect for time of day and rainfall and a very significant effect on all other factors. The Dutch data show a very significant effect on all factors except the gender. The rain effect was no tested due to lack of data. The Polish data show no effect on the gender, a very significant effect for time of ADAS use and a very significant effect on all other factors. The UK data show no effect of the rain and a very significant effect on all other

UDRIVE European Naturalistic Driving Study factors. Theses effect are summarised in Table 6-56, which provides the asymptotic significance (2-sided) calculated by the Pearson method.

Table 6-56: Asymptotic significance (2-sided) calculated by the Pearson method for each factor by OS

	All OS	DE	FR	NL2	PL	UK
Age	0,000	0,048	0,000	0,000	0,000	0,000
Gender	0,000	0,000	0,000	0,556	0,316	0,000
Speed limit	0,000	0,000	0,000	0,000	0,000	0,000
DayNight	0,000	0,006	0,032	0,000	0,000	0,000
RainState	0,000	0,312	0,001		0,000	0,568
ADAS	0,000	0,083	0,000	0,000	0,009	0,000
Infrastructure	0,000	0,031	0,000	0,000	0,000	0,000
	Significance > 0.05		0.001 < Significance < 0.05		Significance < 0.001	

The ANOVA tests on the times to reach the position of the previous vehicle (TIV) showed that they were significantly different between the brakings 1 and 3 for each operational site (Table 6-53) and for each type of infrastructure in which braking was performed (Table 6-54).

7 Conclusions

The goal of this deliverable was to identify safe and unsafe behaviours in Everyday Driving. The results are grouped by research question area:

1. Development and implementation of triggers for safety critical events (SCE) and a method for baseline selection
2. Results of everyday driving, overall and for different driver groups
3. Results of driving behaviour in specific situations (i.e. vehicle overtaking on rural roads) with respect to safe and unsafe behaviours
4. Study results using the self-confrontation technique

The development of SCE triggers is crucial to calculate risks for driving related factors. Especially risk calculations depend on video annotation since most of the secondary task engagements cannot be detected automatically. Based on the advice of the UDRIVE Advisory Board, only very limited analyses of SCEs were made, and no risk estimations comparing SCEs to baselines were conducted. Instead, the focus was shifted to laying the groundwork for performing this analysis after the project. The definition of SCE-triggers is an important requirement for the upcoming SCE analysis. It lays the foundation by developing two approaches for SCE trigger identification. Two methods were developed: The static trigger thresholds and probabilistic trigger thresholds. Both promise advantages in trigger detection. These methods will be used in follow up studies with the data to calculate the risks for secondary tasks and situational factors. It may turn out that one of the methods is better suited to reliably find SCE triggers or even that a combination of both works best on NDS data. In any case the upcoming analysis after the project will be facilitated by this groundwork.

For some analysis, it was necessary to draw random baselines from the population of data. It was decided to sample based on time driven instead of distance. There is no perfect selection for this. Both measures have advantages. Distance is the more common way to present safety estimates and it avoids oversampling of data were the vehicle drives very slow or stands still. Sampling by time avoids oversampling of road types where drivers usually travel fast. Since NDS data is traditionally usually sampled by time, it is easier to compare the UDRIVE data to previous results by going with this approach. Additionally the drawback of oversampled slow travelling episodes can be negated by weighing the baseline by travel speed.

Within the everyday driving analysis, the prevalence was calculated of two aspects of risky driving behaviour: speeding and close following. The data showed that light speeding is more frequent than severe speeding, and that the night and rush hours are most popular for speeding. Close following was most frequent under a very specific speed limit. The identified accompanying situational factors give insight about when drivers decide to take risks. Closer inspection of these events might reveal possible reasons for this behaviour. A first step has been done by performing follow-up interviews with the self-confrontation technique. Furthermore, there is evidence to suggest that certain types of driving personality are more likely to engage in frequent risky behaviours such as speeding. Future studies can build on the data and the groundwork laid in UDRIVE to understand driver’s motivations for risk taking even better. The strength of NDS data is to develop hypothesis about relations between observed behaviour and possible causing factors. Once revealed, each relation can be verified in experimental studies. In this way different approaches in traffic safety research complement each other to gain a deep understanding about causes for crashes. This again is the basis to develop measures to reduce or even avoid crashes altogether.

Safe and unsafe behaviours were also investigated in regard to overtaking on rural roads. The results of this deliverable showed that driver generally take care to avoid dangerous situations in overtaking and helps to identify the few occasions where they do not. There is evidence to suggest that drivers who frequently exceed the speed limit and have a sensation-seeking personality may be more likely to perform a risky overtaking manoeuvre. This is a valuable contribution to the research about a road type that seems to be a blind spot in traffic research until now. It offers a high validity since the results are based on field data with almost no intrusions in the driving situation. Factors that seem to influence driver's overtaking behaviour could be revealed. The next step is to identify the remaining situations in which drivers take risks while overtaking. This deeper analysis will help to understand why drivers sometimes choose to engage in these manoeuvres and in deriving measures to support them in avoiding these. The driving behaviour analysis in this document focused on overtaking on rural roads, but the dataset allows investigation of many more research questions such as high-beam usage during the night and day-light usage during the day, usage of turn signals and driver's reactions to regulating signs, to only name a few. Since the UDRIVE dataset will be available even after the project, future studies can reveal even more insight than what could be presented here.

The self-confrontation technique presented in this report revealed a unique insight about how drivers perceive challenging driving situations. The broad range of experienced situations due to the NDS setup allowed learning about driver's thoughts about the behaviour of other road users, infrastructure, bad driving conditions and her or his own limits, but also about possible measures to mitigate uncomfortable situations. The interview character allowed to complementing the top down approach by the researchers who usually focus on the broad picture. It helped to identify situational configurations that are of special interest for investigation. In general the method worked very well and almost all drivers could recollect the events that they were asked about.

The analysis of the ADAS usage revealed that most trips do not include the use of the two systems available in UDRIVE vehicles, CC or SL. Also there was no trip, were both systems were used at the same time. Questionnaire data indicated that many drivers did not know that their cars are equipped with these systems. A better communication of the availability of these systems from the manufacturer or car dealer could help improving the usage. In regard to seat belts about 90% of drivers buckled up for the whole trip. This also means that about 10% of drivers are not using their seat belt correctly. This is alarming since crashes even at low speeds can lead to severe injuries without the use of a seatbelt. Clearly measures are needed to reduce these numbers. Public awareness campaigns or electronic regulation within the vehicle are possible ways to address this issue. The data also revealed that the counter measures are mostly needed in France and Poland, since the numbers of not using a seatbelt are highest here.

The analysis of hard braking events revealed that all the factors studied have a significant effect on the percentage of hard braking compared to the set of brakes. On the other hand, these factors do not all have a significant effect on this percentage of hard braking on the data collected in each country. German data show a very significant effect for the gender and the categories of speed limit, a significant effect for age, time of day and type of infrastructure but show no effect of rainfall and ADAS use. French data show a significant effect for time of day and rainfall and a very significant effect on all other factors. Dutch data show a very significant effect on all factors except the gender. The rain effect was no tested due to lack of data. Polish data show no effect on the gender, a very significant effect for time of ADAS use and a very significant effect on all other factors. English data show no effect of the rain and a very significant effect on all other factors.

Overall, the results presented in this deliverable give insight into natural driving behaviour on a level which is unprecedented in Europe. Even though many research questions were addressed, there is still much to be found in the data. There are still many research questions that could not be addressed within this project. Also the data used for analysis had to be frozen to make it into the deliverable in time. At time of freezing data was still pre-processed and made available. This means that 100% of the data will only be available after the project, inviting to replicate the performed analysis with the whole dataset and investigating new research questions with it. Most of the data will be made available after the project. Details about how to access the data will be published at www.udrive.eu.

References

AXA Konzern AG (2012). AXA Verkehrssicherheitsreport. Eine Studie zum Verhalten der Deutschen im Straßenverkehr. Köln.

Barbier, C. (2010). Autoconfrontation et analyse des activités humaines. In G. Le Meur & M. Hatano (coord.) Approches pour l'analyse des activités, Paris : L'Harmattan.

Bärgman, J. (2016). *Methods for analysis of naturalistic driving data in driver behavior research: From crash-causation analysis using expert assessment to quantitative assessment of the effect of driver behavior on safety using counterfactual simulation.* (Doctoral thesis), Chalmers University of Technology, Göteborg, Sweden. (ISBN: 978-91-7597-501-6).

Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). *Classification and regression trees*. CRC press.

Breiman, L. (2001). Random forests. *Machine learning*, *45*(1), 5-32.

Broadhurst, P.L. (1957). Emotionality and the Yerkes-Dodson law. *Journal of Experimental Psychology, Vol. 54, 5*, 345- 352.

Byrnes, J. P., Miller, D. C., & Schafer, W. D. (1999). Gender differences in risk taking: A meta-analysis. *Psychological Bulletin, 125*(3), 367–383.

Clarke, D. D., Ward, P. J., & Jones, J. (1998). Overtaking road-accidents: Differences in manoeuvre as a function of driver age. *Accident Analysis & Prevention*, *30*(4), 455-467.

Deng, H., & Wickham, H. (2011). Density estimation in R. *Electronic publication*.

DEKRA Automobil GmbH (2013) Road Safety Report Rural Roads: Strategies for preventing accidents on European roads.

European Road Safety Observatory. (2016). *Roads outside urban areas* (Traffic Safety Basic Facts 2016). Retrieved from https://ec.europa.eu/transport/road_safety/sites/roadsafety/files/pdf/statistics/dacota/bfs2016_rou_areas.pdf

Farber, E., & Silver, C. A. (1967). *Conceptualization of overtaking and passing on two-lane rural roads*. Clearinghouse for Federal Scientific and Technical Information (US).

Friedman, J. H. (1999). Greedy function approximation: a gradient boosting machine. *Annals of statistics*, 1189-1232. Retrieved from: http://statweb.stanford.edu/~jhf/ftp/trebst.pdf

Gini, C. (1914). L'Ammontare e la Composizione della Ricchezza delle Nazioni,Bocca, Torino.

Guo, F. and J. Hankey (2009). Modeling 100-Car Safety Events: A Case-Based Approach for Analyzing Naturalistic Driving Data, The National Surface Transportation Safety Center for Excellence.

Guyonvarch et al. (2017). *Driving Style Indicator using UDRIVE NDS data*. AAAM 2017, Las Vegas

Harris, D. (1988). The subjective probability of involvement in an overtaking accident in the vicinity of a junction. *Road Users and Traffic Safety*, 27-59.

Hills, B. L. (1980). Vision, visibility, and perception in driving. *Perception*, *9*(2), 183-216.

Horrey, W. J., & Lesch, M. F. (2009). Driver-initiated distractions: Examining strategic adaptation for in-vehicle task initiation. Accident Analysis & Prevention, 41(1), 115-122.

Najm, W. G., M. D. Stearns, et al. (2006). Evaluation of an Automotive Rear-End Collision Avoidance System, Research and Innovative Technology Administration, Volpe National Transportation Systems Center, National Highway Traffic Safety Administration: 425p

Hankey, J., Perez, M., McClafferty, J., Description of the SHRP 2 Naturalistic Database and the Crash, Near-Crash, and Baseline Data Sets (2016), VTTI, 2016-04 http://hdl.handle.net/10919/70850

Malta L, et. al. (2012) euroFOT Deliverable 6.4 Final results: Impacts on traffic safety.

Kensinger, E. (2004). Remembering Emotonal Experiences: The Contribution of Valence and Arousal. *Reviews in the Neurosciences, 15*, 241-251.

Kidd, D. G., & McCartt, A. T. (2015). The relevance of crash type and severity when estimating crash risk using the SHRP2 naturalistic driving data*. Proceedings of the 4th International Conference on Driver Distraction and Inattention.* Sydney, Australia.

Kircher, K. (2007). Driver distraction, a review of litterature. *VTI rapport 594A*

Klauer, S. G., Guo, F., Simons-Morton, B. G., Ouimet, M. C., Lee, S. E., & Dingus, T. A. (2014). Distracted driving and risk of road crashes among novice and experienced drivers. *New England Journal of Medicine. 370*(1), 54-59. doi:10.1056/NEJMsa1204142.

Knipling, R. R. (2015). Naturalistic driving events: No harm, no foul, no validity *Proceedings of the Eighth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design* Snowbird, Salt Lake City, UT. (pp. 197-203).

LaBar, K.S. & Cabeza, R. (2006). Cognitive neuroscience of emotional memory. *Nature Reviews Neuroscience, 7*, 54-64

ONISR (2014), internal report "*L'évolution du comportement des conducteurs: 2012 et années précédentes* (hors vitesse) note de synthèse"

Pearson, K. (1900). On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Philosophical Magazine Series 5, 50*(302), 157–175. doi:10.1080/14786440009463897

Rajalin, S., Hassel, S.-O., & Summala, H. (1997). Close-Following Drivers on two-lane highways. *Accident Analysis & Prevention, 29*(6), 723–729.

Ranney, T. A. (2008). Driver distraction: A review of the current state-of-knowledge. *National Highway Traffic Safety Administration Repor*t. No. HS-810 787.

Simons-Morton, B. G., Ouimet, M. C., Zhang, Z., Klauer, S. E., Lee, S. E., Wang, J., ... & Dingus, T. A. (2011). The effect of passengers and risk-taking friends on risky driving and crashes/near crashes among novice teenagers. *Journal of Adolescent Health*, *49*(6), 587-593.

Statistisches Bundesamt (2012). Verkehrsunfälle 2011. Wiesbaden.

Stutts, J. C., Reinfurt, D. W., Staplin, L., & Rodgman, E. A. (2001). The role of driver distraction in traffic crashes. AAA Foundaton for Traffic Safety

Teese, R., & Bradley, G. (2008). Predicting Recklessness in Emerging Adults: A Test of a Psychosocial Model. *The Journal of Social Psychology, 148*(1), 105–126.

Theureau, J. (2003). Course-of-action analysis and course-of-action centered design. In Hollnagel (ed.), Handbook of cognitive task design, Lawrence Erlbaump Ass.

Tivesten, E. & Dozza, M. (2015) Driving context influences drivers' decision to engage in visual–manual phone tasks: Evidence from a naturalistic driving study, Journal of Safety Research Volume 53, Pages 87–96

Victor, T., Dozza, M., Bärgman, J., Boda, C.-N., Engström, J., Flannagan, C., . . . Markkula, G. (2015). *Analysis of naturalistic driving study data: SAFER glances, driver inattention, and crash risk*. (SHRP 2 SO8A). Transportation Research Board of the National Academy of Sciences. Washington, D.C. ISBN: 978-0-309-27423-4. Retrieved from http://onlinepubs.trb.org/onlinepubs/shrp2/SHRP2_S2-S08A-RW-1.pdf.

Wassim G. Najm, John D. Smith, and Mikio Yanagisawa (2007). *Pre-Crash Scenario Typology for Crash Avoidance Research.* DOT HS 810 767.

Wilson, T., & Best, W. (1982). Driving strategies in overtaking. *Accident Analysis & Prevention*, *14*(3), 179-185.

Young, K. & Regan, M. (2007). Driver distraction: A review of litterature. In: I.J. Faulks, M.Regan, M. Stevenson, J. Brown, A. Porter & J.D. Irwin (Eds.). *Distracted driving*. Sydney, NSW: Australian College of Road Safety, 379-405.

Yger, A. (1999). Théorie et analyse du signal: Cours et initiation pratique via MATLAB et SCILAB. Paris: Ellipses-Marketing.

List of abbreviations

ADAS - Advanced driving assistance systems

AHL - Adaptive head light

ANOVA - Analysis of variance

CAN - Controller area network

CC - Cruise control

CDC - Central data centre

DAS - Data acquisition system

EBA - Event based analysis

FP7 - Framework Program 7

FR - France

GE - Germany

GPS - Global positioning system

JPDD - Joint probability density distribution

LDC - Local data centre

LDW - Lane departure warning

NDS - Naturalistic driving study

NL - The Netherlands

OMT - Online monitoring tool

OS - Operation site

PL - Poland

PTW - Power two wheelers

SCE - Safety critical event

SL - Speed limiter

SP - Spain

THW - Time headway

TIV - Time headway

TTC - Time-to-collision

UK - United Kingdom

VRU - Vulnerable road user

List of Figures

List of Tables

Appendix A Additional data for the overtaking analysis

A.1 Overtaking manoeuvres on rural roads

A.1.1 Additional tables of the analysis of takeover manoeuvre by type of manoeuvre

Table A-1 Crosstabs manoeuvre by gender

Count

		Gender		Total
		male	female	
Manoeuvre	normal	14	11	25
	flying	13	2	15
	normal-piggy	1	0	1
	flying-piggy	3	1	4
Total		31	14	45

Table A-2 Crosstabs manoeuvre by bend

Count

		Bend		Total
		no	yes	
Manoeuvre	normal	22	3	25
	flying	13	2	15
	normal-piggy	1	0	1
	flying-piggy	4	0	4
Total		40	5	45

Table A-3 Crosstabs manoeuvre by alley

Count

		Alley		Total
		no	yes	
Manoeuvre	normal	15	10	25
	flying	11	4	15
	normal-piggy	1	0	1
	flying-piggy	2	2	4
Total		29	16	45

Table A-4 Crosstabs manoeuvre by secondary task

Count

		Secondary task		Total
		no	yes	
Manoeuvre	normal	24	1	25
	flying	14	1	15
	normal-piggy	1	0	1
	flying-piggy	4	0	4
Total		43	2	45

Table A-5 Crosstabs manoeuvre by overtaking regulation

Count

		Overtaking regulation		Total
		1	2	
Manoeuvre	normal	22	3	25
	flying	14	1	15
	normal-piggy	1	0	1
	flying-piggy	4	0	4
Total		41	4	45

Table A-6 Crosstabs manoeuvre by passenger present

Count

		Passenger present		Total
		1	2	
Manoeuvre	normal	14	11	25
	flying	10	5	15
	normal-piggy	1	0	1
	flying-piggy	3	1	4
Total		28	17	45

Table A-7 Crosstabs manoeuvre by oncoming traffic

Count

		Oncoming traffic		Total
		1	2	
Manoeuvre	normal	22	3	25

	flying	15	0	15
	normal-piggy	1	0	1
	flying-piggy	3	1	4
Total		41	4	45

Table A-8 Mean, N, SD, and sum of total number overtaken

Manoeuvre	Mean	N	Std. Deviation	Sum
normal	1.08	25	.277	27
flying	1.13	15	.516	17
normal-piggy	2.00	1	.	2
flying-piggy	1.00	4	.000	4
Total	1.11	45	.383	50

Table A-9 Mean, N, SD, and sum of mean lateral acceleration

Manoeuvre		LatAcc_mean_1	LatAcc_mean_2	LatAcc_mean_3	LatAcc_mean_4	LatAcc_mean_5
normal	Mean	.033213438	-.073252646		-.0228436646	.016979238
	N	24	24	25	24	24
	SD	.0793049128	.1288025451		.08383076449	.0681711522
	Sum	.7971225	-1.7580635		-.54824795	.4075017
flying	Mean	.057022973	.019344067		.0271658460	.037970667
	N	15	15	15	15	15
	SD	.0768395887	.1070125564		.06822248536	.0773220714
	Sum	.8553446	.2901610		.40748769	.5695600
normal-piggy	Mean	.092502000	-.038194000		-.0170410000	.042059000
	N	1	1	1	1	1
	SD	.	.		.	.
	Sum	.0925020	-.0381940		-.01704100	.0420590
flying-piggy	Mean	.038840000	-.014173150		.0000212500	.025217250
	N	4	4	4	4	4
	SD	.1011631212	.0990474181		.09195430772	.0965339248
	Sum	.1553600	-.0566926		.00008500	.1008690
Total	Mean	.043189298	-.035517934		-.0035844605	.025454311
	N	44	44	45	44	44
	SD	.0786191926	.1231824082		.08004077912	.0719974070
	Sum	1.9003291	-1.5627891		-.15771626	1.1199897

Table A-10 Mean, N, SD, and sum of maximum lateral acceleration

Manoeuvre		LatAcc_max_1	LatAcc_max_2	LatAcc_max_3	LatAcc_max_4	LatAcc_max_5
normal	Mean	.1470483	-.0138383		.0946396	.1286404
	N	24	24	25	24	24
	SD	.10227387	.16112800		.14443108	.11263116
	Sum	3.52916	-.33212		2.27135	3.08737
flying	Mean	.1531660	.0933747		.1134840	.1404453
	N	15	15	15	15	15
	SD	.10006976	.15327797		.09339647	.11037601
	Sum	2.29749	1.40062		1.70226	2.10668
normal-piggy	Mean	.1928700	.0081800		.0603000	.3914800
	N	1	1	1	1	1
	SD	.	.		.	.
	Sum	.19287	.00818		.06030	.39148
flying-piggy	Mean	.1127325	.0496525		.1336050	.2368150
	N	4	4	4	4	4
	SD	.08218083	.05522804		.05236547	.21676533
	Sum	.45093	.19861		.53442	.94726
Total	Mean	.1470557	.0289839		.1038257	.1484725
	N	44	44	45	44	44
	SD	.09745679	.15578336		.11998071	.12795083
	Sum	6.47045	1.27529		4.56833	6.53279

Table A-11 Mean, N, SD, and sum of mean longitudinal acceleration

Manoeuvre		LongAcc_ mean_0	LongAcc_ mean_1	LongAcc_ mean_2	LongAcc_ mean_3	LongAcc_ mean_4	LongAcc_ mean_5
normal	Mean	.028475504	9.66160E-002	.0925497804	.0849252592	.0567120167	.036717450
	N	24	24	24	24	24	24
	SD	.0548864850	6.52112E-002	.06935888402	.05897224889	.07818046595	.0598306158
	Sum	.6834121	2.3188E+000	2.22119473	2.03820622	1.36108840	.8812188
flying	Mean	.022585520	3.79565E-002	.0671853733	.0731463067	.0586452860	.025768500
	N	15	15	15	15	15	15
	SD	.0350004230	5.80727E-002	.05925808373	.05485910061	.06030427122	.0759976689
	Sum	.3387828	5.6935E-001	1.00778060	1.09719460	.87967929	.3865275
normal-piggy	Mean	.013526000	1.00350E-001	.1216900000	.1128500000	.0300260000	.029115000

	N	1	1	1	1	1	1
	SD	.	.	.	.	.	.
	Sum	.0135260	1.0035E-001	.12169000	.11285000	.03002600	.0291150
flying-piggy	Mean	.074541650	5.93967E-002	.0381087500	.0545940200	.0734910000	.026362800
	N	4	4	4	4	4	4
	SD	.0734102775	5.94593E-002	.07508929952	.04238825895	.04573070024	.0277261513
	Sum	.2981666	2.3759E-001	.15243500	.21837608	.29396400	.1054512
Total	Mean	.030315625	7.33211E-002	.0796159166	.0787869750	.0582899475	.031870739
	N	44	44	44	44	44	44
	SD	.0509648680	6.62834E-002	.06677204816	.05548817502	.06812290756	.0622727862
	Sum	1.3338875	3.2261E+000	3.50310033	3.46662690	2.56475769	1.4023125

Table A-12 Mean, N, SD, and sum of maximum longitudinal acceleration

Manoeuvre		LongAcc_max_0	LongAcc_max_1	LongAcc_max_2	LongAcc_max_3	LongAcc_max_4	LongAcc_max_5
normal	Mean	.1603696	.1860363	.1330008	.1587013	.1402142	.1487888
	N	24	24	24	24	24	24
	SD	.09348885	.06682726	.07717728	.07547849	.09674158	.13667505
	Sum	3.84887	4.46487	3.19202	3.80883	3.36514	3.57093
flying	Mean	.1352293	.1297753	.1121493	.1311867	.1292073	.1068520
	N	15	15	15	15	15	15
	SD	.08527816	.07402992	.06053022	.07875426	.06875101	.07680131
	Sum	2.02844	1.94663	1.68224	1.96780	1.93811	1.60278
normal-piggy	Mean	.2225300	.1616200	.1490500	.1427000	.1381800	.3540000
	N	1	1	1	1	1	1
	SD	.	.	.	.	.	.
	Sum	.22253	.16162	.14905	.14270	.13818	.35400
flying-piggy	Mean	.1729125	.1607375	.1020500	.1295475	.1800825	.1355000
	N	4	4	4	4	4	4
	SD	.04975746	.06419092	.10289105	.03581427	.08104261	.13489389
	Sum	.69165	.64295	.40820	.51819	.72033	.54200
Total	Mean	.1543520	.1640016	.1234434	.1463073	.1400400	.1379480
	N	44	44	44	44	44	44
	SD	.08662197	.07170274	.07259990	.07312987	.08481320	.12112230
	Sum	6.79149	7.21607	5.43151	6.43752	6.16176	6.06971

A.1.2 Additional tables of the analysis of takeover manoeuvre by country

Table A-13 Summary of N, mean, SD by country

		N	Mean	Std. Deviation	Std. Error
LatAcc_mean_1	OS-FR	13	.068606115	.0474434915	.0131584570
	OS-DE	3	.130735000	.0648607553	.0374473745
	OS-PL	17	.073594765	.0387303443	.0093934883
	OS-UK	11	-.057715127	.0687392894	.0207256756
	Total	44	.043189298	.0786191926	.0118522892
LatAcc_mean_2	OS-FR	13	-.014074577	.0462893817	.0128383646
	OS-DE	3	-.009490667	.1013164558	.0584950830
	OS-PL	17	.013807082	.0808715492	.0196142317
	OS-UK	11	-.144188000	.1795329834	.0541312312
	Total	44	-.035517934	.1231824082	.0185704468
LatAcc_mean_4	OS-FR	13	.0057156231	.04112775078	.01140678571
	OS-DE	3	-.0303553333	.00114916593	.00066347126
	OS-PL	17	.0282290206	.07448897446	.01806622998
	OS-UK	11	-.0564406100	.10732697013	.03236029907
	Total	44	-.0035844605	.08004077912	.01206660147
LatAcc_mean_5	OS-FR	13	.030045692	.0291876336	.0080951930
	OS-DE	3	.034741333	.0888120867	.0512756821
	OS-PL	17	.059273382	.0451917827	.0109606173
	OS-UK	11	-.034770527	.1021506035	.0307995658
	Total	44	.025454311	.0719974070	.0108540175
LatAcc_max_1	OS-FR	13	.1817808	.06184201	.01715189
	OS-DE	3	.1918967	.09671478	.05583830
	OS-PL	17	.1725641	.05580336	.01353430
	OS-UK	11	.0543655	.12911459	.03892951
	Total	44	.1470557	.09745679	.01469216
LatAcc_max_2	OS-FR	13	.0517869	.11596784	.03216369
	OS-DE	3	.0946067	.18373849	.10608147
	OS-PL	17	.0738888	.09964515	.02416750
	OS-UK	11	-.0852609	.21406679	.06454357
	Total	44	.0289839	.15578336	.02348523
LatAcc_max_4	OS-FR	13	.0827823	.07700536	.02135744
	OS-DE	3	.0654700	.03593117	.02074487
	OS-PL	17	.1224565	.12067768	.02926864
	OS-UK	11	.1103627	.17196229	.05184858
	Total	44	.1038257	.11998071	.01808777
LatAcc_max_5	OS-FR	13	.1176946	.09914795	.02749869
	OS-DE	3	.1324867	.06608045	.03815157

	OS-PL	17	.1474665	.11659730	.02827900
	OS-UK	11	.1907609	.18082728	.05452148
	Total	44	.1484725	.12795083	.01928931
LongAcc_mean_0	OS-FR	13	.033848754	.0377891280	.0104808184
	OS-DE	3	.047897333	.0065795201	.0037986877
	OS-PL	17	.042875453	.0389886989	.0094561485
	OS-UK	11	.001934455	.0758369976	.0228657151
	Total	44	.030315625	.0509648680	.0076832429
LongAcc_mean_1	OS-FR	13	9.169105E-002	5.3624803E-002	1.4872844E-002
	OS-DE	3	1.654167E-001	7.4904361E-003	4.3246053E-003
	OS-PL	17	5.578864E-002	7.5809807E-002	1.8386579E-002
	OS-UK	11	5.359003E-002	4.7965741E-002	1.4462215E-002
	Total	44	7.332116E-002	6.6283490E-002	9.9926121E-003
LongAcc_mean_2	OS-FR	13	.1117813846	.04604934574	.01277179056
	OS-DE	3	.1096630000	.04660870677	.02690954940
	OS-PL	17	.0603146665	.06132661810	.01487388965
	OS-UK	11	.0632367273	.08735361860	.02633810700
	Total	44	.0796159166	.06677204816	.01006626501
LongAcc_mean_3	OS-FR	13	.0968799169	.05433708081	.01507039469
	OS-DE	3	.0678956333	.08355405981	.04823995892
	OS-PL	17	.0714064518	.05143101974	.01247385452
	OS-UK	11	.0717810364	.05882322891	.01773587084
	Total	44	.0787869750	.05548817502	.00836515713
LongAcc_mean_4	OS-FR	13	.0413198615	.05439291535	.01508588041
	OS-DE	3	.1226253333	.13071606935	.07546895782
	OS-PL	17	.0641649818	.07031615161	.01705417178
	OS-UK	11	.0517198909	.05791603571	.01746234180
	Total	44	.0582899475	.06812290756	.01026991473
LongAcc_mean_5	OS-FR	13	.004791354	.0268405343	.0074442248
	OS-DE	3	.048996267	.1484289942	.0856955197
	OS-PL	17	.049142571	.0371428830	.0090084723
	OS-UK	11	.032510218	.0874530485	.0263680862
	Total	44	.031870739	.0622727862	.0093879757
LongAcc_max_0	OS-FR	13	.1465885	.06182936	.01714838
	OS-DE	3	.2257500	.08518864	.04918368
	OS-PL	17	.1651165	.06695393	.01623871
	OS-UK	11	.1274191	.12825959	.03867172
	Total	44	.1543520	.08662197	.01305875
LongAcc_max_1	OS-FR	13	.1677531	.06099690	.01691750
	OS-DE	3	.2402367	.05004085	.02889110

	OS-PL	17	.1354259	.05457331	.01323597
	OS-UK	11	.1829391	.09398014	.02833608
	Total	44	.1640016	.07170274	.01080960
LongAcc_max_2	OS-FR	13	.1535831	.05086856	.01410840
	OS-DE	3	.1617467	.02855828	.01648813
	OS-PL	17	.0994435	.06554765	.01589764
	OS-UK	11	.1144682	.09916001	.02989787
	Total	44	.1234434	.07259990	.01094485
LongAcc_max_3	OS-FR	13	.1462408	.05411979	.01501013
	OS-DE	3	.1472967	.11233532	.06485683
	OS-PL	17	.1310612	.03833732	.00929817
	OS-UK	11	.1696782	.11666123	.03517468
	Total	44	.1463073	.07312987	.01102474
LongAcc_max_4	OS-FR	13	.1145400	.06699760	.01858179
	OS-DE	3	.1485200	.13202805	.07622643
	OS-PL	17	.1271912	.05535737	.01342613
	OS-UK	11	.1877209	.11621933	.03504145
	Total	44	.1400400	.08481320	.01278607
LongAcc_max_5	OS-FR	13	.0907831	.10929336	.03031252
	OS-DE	3	.1305333	.15480404	.08937616
	OS-PL	17	.1152629	.06228870	.01510723
	OS-UK	11	.2307691	.15648756	.04718278
	Total	44	.1379480	.12112230	.01825987

A.2 Speeding and close following behaviour

A.2.1 Effects of age on speeding

Table A-14 Summary of N, mean, SD of speed and duration by age

		N	Mean	Std. Deviation	Std. Error
Max speed	18-24	275	115.092131	24.8555149	1.4988439
	25-49	8596	113.128602	24.2455202	.2615070
	50-99	2426	96.638707	26.4815896	.5376488
	Total	11297	109.635240	25.6730630	.2415439
Mean speed	18-24	275	105.85141734918	24.915471851037	1.5024594837191
	25-49	8596	103.14092629708	24.944842246199	.26904972863456
	50-99	2426	86.581300252743	26.958945274085	.54734039194474
	Total	11297	99.650772473568	26.295091961065	.24739625982983
Duration	18-24	275	23.2545	46.72812	2.81781
	25-49	8596	21.7965	29.95859	.32313
	50-99	2426	19.1791	15.14947	.30758
	Total	11297	21.2699	28.04265	.26384

A.2.2 Effect of gender on speeding

Table A-15 Summary of N, mean, SD of speed and duration by gender

		N	Mean	Std. Deviation	Std. Error
Max speed	male	5614	107.118097	27.3740762	.3653451
	female	5683	112.121821	23.6143960	.3132478
	Total	11297	109.635240	25.6730630	.2415439
Mean speed	male	5614	97.158924578404	27.914007013224	.37255121073429
	female	5683	102.11236566087	24.344688182811	.32293520184759
	Total	11297	99.650772473568	26.295091961065	.24739625982983
Duration	male	5614	21.5186	28.53895	.38089
	female	5683	21.0243	27.54392	.36537
	Total	11297	21.2699	28.04265	.26384

A.2.3 Effects of country on speeding

Table A-16 Summary of N, mean, SD of speed and duration by country

		N	Mean	Std. Deviation	Std. Error
Max speed	OS-FR	3338	107.625491	20.7240771	.3587005
	OS-NL	1203	107.124026	18.3922623	.5302764
	OS-DE	1110	92.936626	28.7895521	.8641187
	OS-PL	2311	113.429147	27.2558824	.5669703
	OS-UK	3335	115.481499	27.3416249	.4734525
	Total	11297	109.635240	25.6730630	.2415439
Mean speed	OS-FR	3338	97.553742504897	21.367401687141	.36983545920349
	OS-NL	1203	97.127640626681	19.432651027033	.56027241172815
	OS-DE	1110	82.861953088947	29.359256081576	.88121840185563
	OS-PL	2311	103.49004496609	27.917530419141	.58073369517319
	OS-UK	3335	105.58727755120	27.871039645733	.48261992735803
	Total	11297	99.650772473568	26.295091961065	.24739625982983
Duration	OS-FR	3338	19.8337	27.43315	.47482
	OS-NL	1203	32.7842	49.31903	1.42194
	OS-DE	1110	19.9792	13.39826	.40215
	OS-PL	2311	21.2079	30.81282	.64096
	OS-UK	3335	19.0266	16.16063	.27984
	Total	11297	21.2699	28.04265	.26384

Appendix B Additional data for the analysis of hard braking

B.1 Breakdown of braking by participants and by operational site

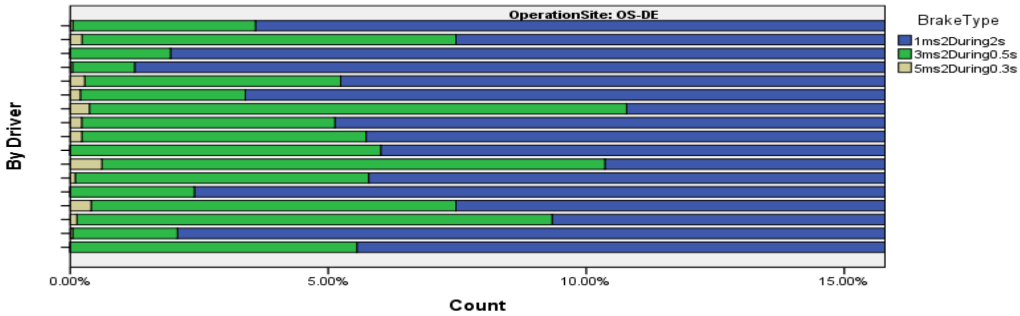


Figure B-1 : Breakdown of braking by participants for German operational site

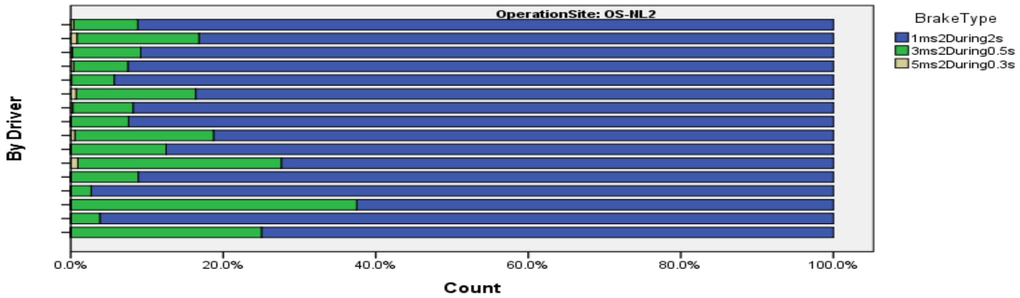


Figure B-2 : Breakdown of braking by participants for Dutch operational site

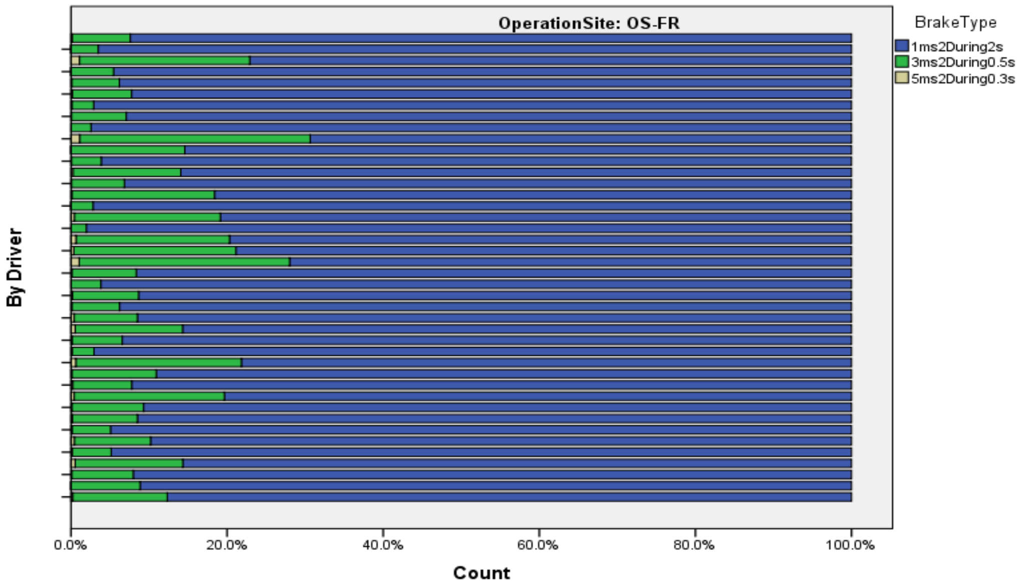


Figure B-3 : Breakdown of braking by participants for French operational site

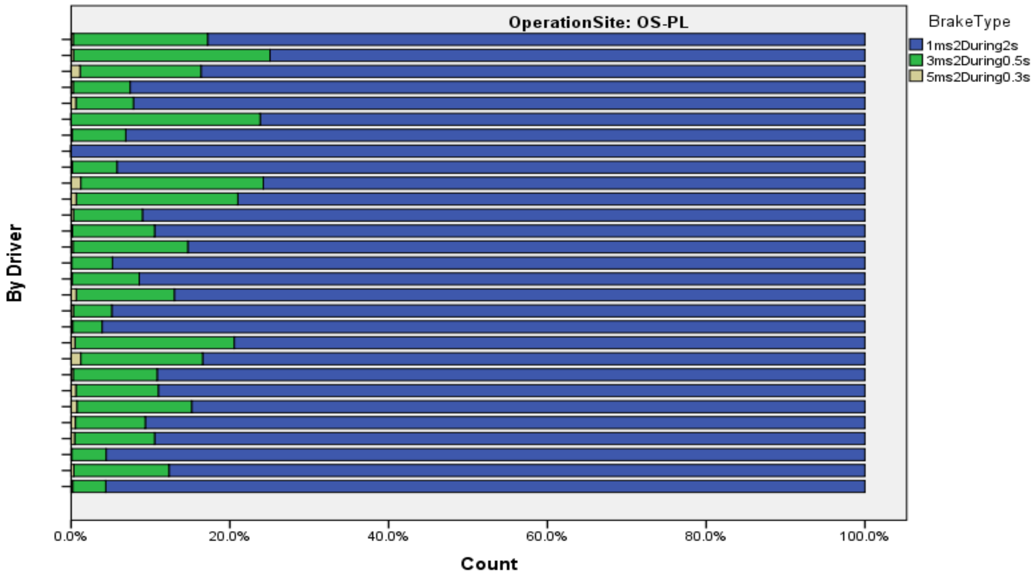


Figure B-4 : Breakdown of braking by participants for Polish operational site

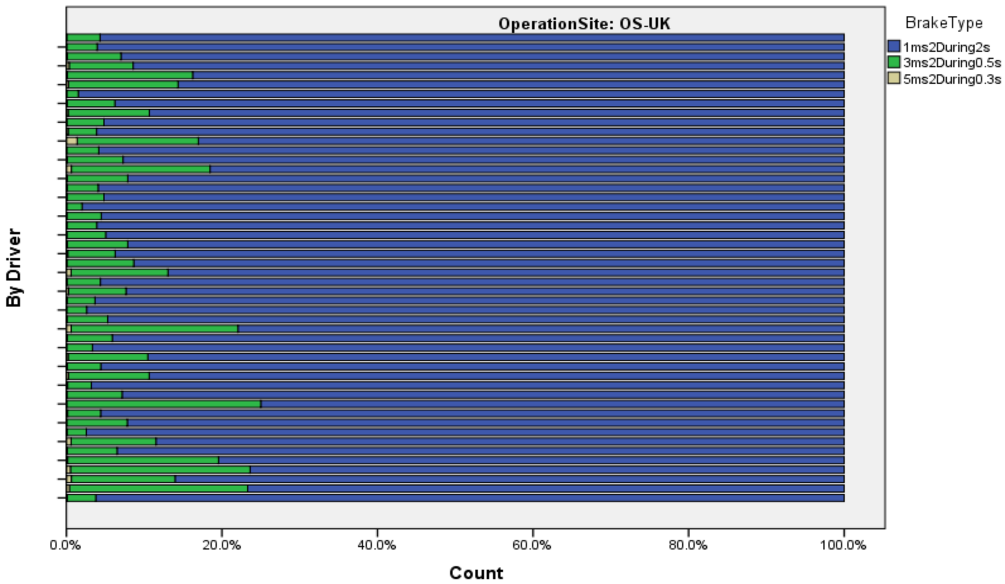


Figure B-5 : Breakdown of braking by participants for English operational site

B.2 Breakdown of braking by speed limit category and operational site

Table B-0-17 : Z-test on speed limit category by operational site

BrakeType * CatSpeedLimit Crosstabulation

OperationSite				CatSpeedLimit Inf34km/h	Between34 And62km/h	Between62And98km/h	Sup98km/h	Total
OS-DE	BrakeType	1ms2During2s	Count	3231a	17933a	3262b	3756b	28182
			% within CatSpeedLimit	95.6%	95.1%	93.4%	93.5%	94.8%
		3ms2During0.5s	Count	145a	888a	222b	250b	1505
			% within CatSpeedLimit	4.3%	4.7%	6.4%	6.2%	5.1%
		5ms2During0.3s	Count	5a	28a	7a	12a	52
			% within CatSpeedLimit	0.1%	0.1%	0.2%	0.3%	0.2%
	Total		Count	3381	18849	3491	4018	29739
			% within CatSpeedLimit	100.0%	100.0%	100.0%	100.0%	100.0%
OS-FR	BrakeType	1ms2During2s	Count	18212a	122418b	56496a, b	1955c	199081
			% within CatSpeedLimit	87.4%	88.3%	88.0%	93.3%	88.2%
		3ms2During0.5s	Count	2557a	15819b	7468a, b	138c	25982
			% within CatSpeedLimit	12.3%	11.4%	11.6%	6.6%	11.5%
		5ms2During0.3s	Count	59a	385a	212a	3a	659
			% within CatSpeedLimit	0.3%	0.3%	0.3%	0.1%	0.3%
	Total		Count	20828	138622	64176	2096	225722
			% within CatSpeedLimit	100.0%	100.0%	100.0%	100.0%	100.0%
OS-NL2	BrakeType	1ms2During2s	Count	2752a	12848b	5900c	3759d	25259
			% within CatSpeedLimit	90.8%	88.9%	83.8%	86.5%	87.5%
		3ms2During0.5s	Count	272a	1560b	1095c	565d	3492
			% within CatSpeedLimit	9.0%	10.8%	15.6%	13.0%	12.1%
		5ms2During0.3s	Count	7a, b	44b	42a	24a, b	117
			% within CatSpeedLimit	0.2%	0.3%	0.6%	0.6%	0.4%
	Total		Count	3031	14452	7037	4348	28868
			% within CatSpeedLimit	100.0%	100.0%	100.0%	100.0%	100.0%
OS-PL	BrakeType	1ms2During2s	Count	4408a	57846a	8938b	4079a	75271
			% within CatSpeedLimit	89.6%	89.6%	86.9%	89.9%	89.3%
		3ms2During0.5s	Count	503a	6534a	1285b	429a	8751
			% within CatSpeedLimit	10.2%	10.1%	12.5%	9.5%	10.4%
		5ms2During0.3s	Count	10a	200a	59b	31b	300
			% within CatSpeedLimit	0.2%	0.3%	0.6%	0.7%	0.4%
	Total		Count	4921	64580	10282	4539	84322
			% within CatSpeedLimit	100.0%	100.0%	100.0%	100.0%	100.0%
OS-UK	BrakeType	1ms2During2s	Count	827a, b	109463b	55923a	9615c	175828
			% within CatSpeedLimit	89.5%	90.6%	88.6%	85.3%	89.7%
		3ms2During0.5s	Count	96a, b	11140b	6999a	1614c	19849
			% within CatSpeedLimit	10.4%	9.2%	11.1%	14.3%	10.1%
		5ms2During0.3s	Count	1a, b	181b	184a	40a	406
			% within CatSpeedLimit	0.1%	0.1%	0.3%	0.4%	0.2%
	Total		Count	924	120784	63106	11269	196083
			% within CatSpeedLimit	100.0%	100.0%	100.0%	100.0%	100.0%

Each subscript letter denotes a subset of CatSpeedLimit categories whose column proportions do not differ significantly from each other at the ,05 level.

Table B-0-18 : Chi2 tests on speed limit category by operational site

Chi-Square Tests

OperationSite		Value	df	Asymptotic Significance (2-sided)
OS-DE	Pearson Chi-Square	37.267[1]	6	.000
	Likelihood Ratio	35.400	6	.000
	Linear-by-Linear Association	31.299	1	.000
	N of Valid Cases	29739		
OS-FR	Pearson Chi-Square	70.604[2]	6	.000
	Likelihood Ratio	78.666	6	.000
	Linear-by-Linear Association	6.405	1	.011
	N of Valid Cases	225722		
OS-NL2	Pearson Chi-Square	150.282[3]	6	.000
	Likelihood Ratio	147.361	6	.000
	Linear-by-Linear Association	87.244	1	.000
	N of Valid Cases	28868		
OS-PL	Pearson Chi-Square	94.278[4]	6	.000
	Likelihood Ratio	87.405	6	.000
	Linear-by-Linear Association	20.826	1	.000
	N of Valid Cases	84322		
OS-UK	Pearson Chi-Square	448.508[5]	6	.000
	Likelihood Ratio	425.032	6	.000
	Linear-by-Linear Association	419.565	1	.000
	N of Valid Cases	196083		

1. 0 cells (,0%) have expected count less than 5. The minimum expected count is 5,91.
2. 0 cells (,0%) have expected count less than 5. The minimum expected count is 6,12.
3. 0 cells (,0%) have expected count less than 5. The minimum expected count is 12,28.
4. 0 cells (,0%) have expected count less than 5. The minimum expected count is 16,15.
5. 1 cells (8,3%) have expected count less than 5. The minimum expected count is 1,91.

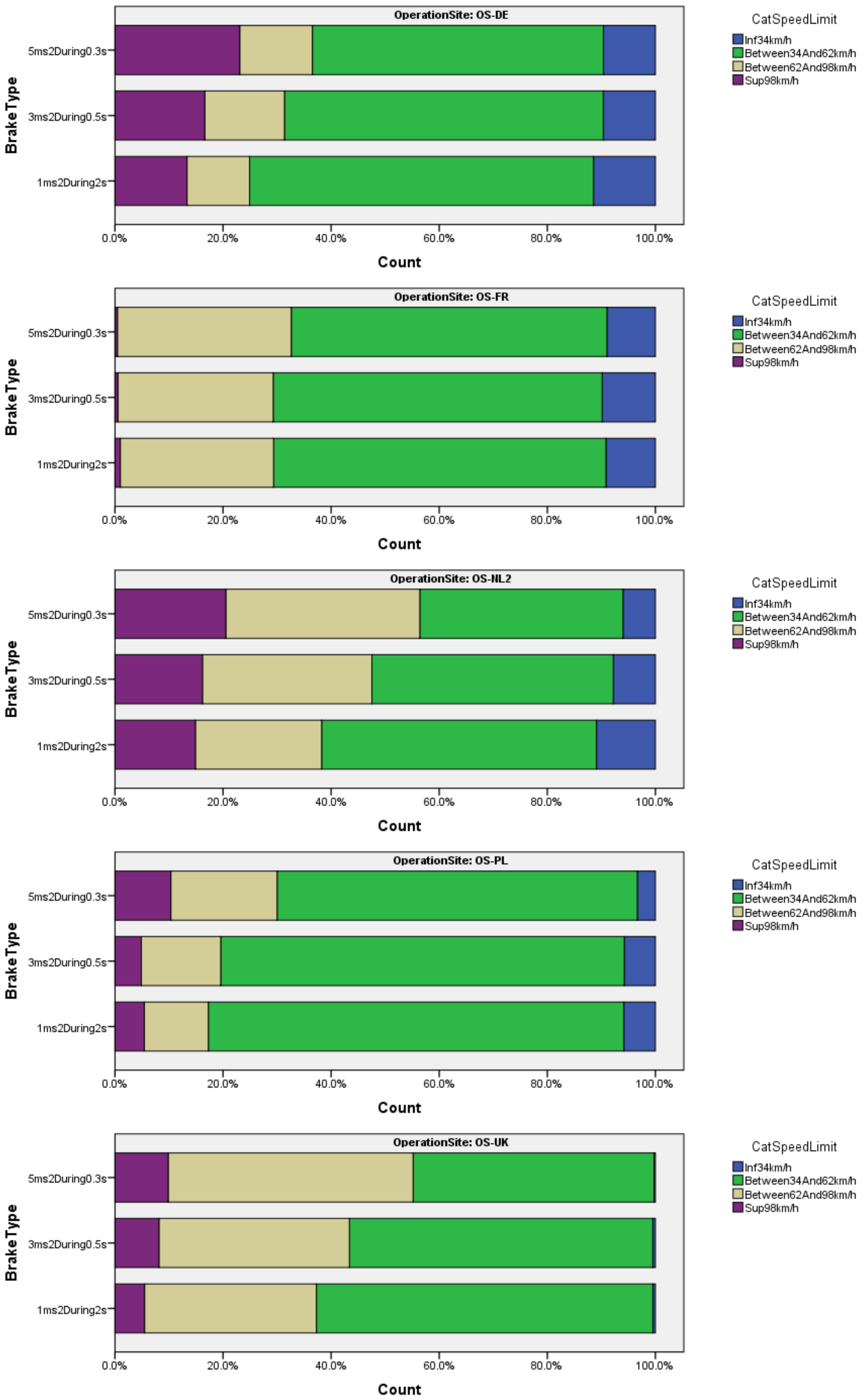


Figure B-6 : Breakdown of speed limit categories for each type of brake for each operational site

B.3 Breakdown of braking by category of period of the day and by operational site

Table B-0-19: Z-test on day category by operational site

BrakeType * DayNight Crosstabulation

OperationSite				DayNight SunRise	Day	Sunset	Night
OS-DE	BrakeType	1ms2During2s	Count	495$_{a, b}$	21789$_{b}$	642$_{a, b}$	5153$_{a}$
			% within DayNight	93.9%	94.6%	94.0%	95.8%
		3ms2During0.5s	Count	31$_{a, b}$	1211$_{b}$	41$_{a, b}$	214$_{a}$
			% within DayNight	5.9%	5.3%	6.0%	4.0%
		5ms2During0.3s	Count	1$_{a}$	40$_{a}$	0$_{a}$	10$_{a}$
			% within DayNight	0.2%	0.2%	0.0%	0.2%
	Total		Count	527	23040	683	5377
			% within DayNight	100.0%	100.0%	100.0%	100.0%
OS-FR	BrakeType	1ms2During2s	Count	7178$_{a}$	123844$_{a}$	8101$_{a}$	59410$_{a}$
			% within DayNight	88.4%	88.1%	88.6%	88.5%
		3ms2During0.5s	Count	921$_{a}$	16336$_{a}$	1007$_{a}$	7585$_{a}$
			% within DayNight	11.3%	11.6%	11.0%	11.3%
		5ms2During0.3s	Count	21$_{a}$	432$_{a}$	34$_{a}$	171$_{a}$
			% within DayNight	0.3%	0.3%	0.4%	0.3%
	Total		Count	8120	140612	9142	67166
			% within DayNight	100.0%	100.0%	100.0%	100.0%
OS-NL2	BrakeType	1ms2During2s	Count	441$_{a}$	19185$_{a}$	813$_{b}$	4751$_{a}$
			% within DayNight	89.6%	87.6%	80.7%	88.2%
		3ms2During0.5s	Count	50$_{a}$	2623$_{a}$	189$_{b}$	618$_{a}$
			% within DayNight	10.2%	12.0%	18.8%	11.5%
		5ms2During0.3s	Count	1$_{a}$	92$_{a}$	6$_{a}$	17$_{a}$
			% within DayNight	0.2%	0.4%	0.6%	0.3%
	Total		Count	492	21900	1008	5386
			% within DayNight	100.0%	100.0%	100.0%	100.0%
OS-PL	BrakeType	1ms2During2s	Count	1162$_{a, b}$	58345$_{b}$	2438$_{a}$	13118$_{a}$
			% within DayNight	90.2%	88.7%	90.9%	91.4%
		3ms2During0.5s	Count	121$_{a, b}$	7169$_{b}$	236$_{a}$	1178$_{a}$
			% within DayNight	9.4%	10.9%	8.8%	8.2%
		5ms2During0.3s	Count	5$_{a}$	233$_{a}$	8$_{a}$	52$_{a}$
			% within DayNight	0.4%	0.4%	0.3%	0.4%
	Total		Count	1288	65747	2682	14348
			% within DayNight	100.0%	100.0%	100.0%	100.0%
OS-UK	BrakeType	1ms2During2s	Count	4982$_{a}$	117351$_{b}$	5921$_{b, c}$	46925$_{c}$
			% within DayNight	88.3%	89.5%	89.9%	90.3%
		3ms2During0.5s	Count	650$_{a}$	13422$_{b}$	653$_{b, c}$	4963$_{c}$
			% within DayNight	11.5%	10.2%	9.9%	9.5%
		5ms2During0.3s	Count	9$_{a}$	290$_{a}$	10$_{a}$	97$_{a}$
			% within DayNight	0.2%	0.2%	0.2%	0.2%
	Total		Count	5641	131063	6584	51985
			% within DayNight	100.0%	100.0%	100.0%	100.0%

Each subscript letter denotes a subset of DayNight categories whose column proportions do not differ significantly from each other at the ,05 level.

Table B-0-20 Chi2 tests on day category by operational site

Chi-Square Tests

OperationSite		Value	df	Asymptotic Significance (2-sided)
OS-DE	Pearson Chi-Square	18.141[1]	6	.006
	Likelihood Ratio	20.155	6	.003
	Linear-by-Linear Association	12.489	1	.000
	N of Valid Cases	29627		
OS-FR	Pearson Chi-Square	13.787[2]	6	.032
	Likelihood Ratio	13.805	6	.032
	Linear-by-Linear Association	6.069	1	.014
	N of Valid Cases	225040		
OS-NL2	Pearson Chi-Square	48.996[3]	6	.000
	Likelihood Ratio	43.951	6	.000
	Linear-by-Linear Association	.003	1	.956
	N of Valid Cases	28786		
OS-PL	Pearson Chi-Square	101.061[4]	6	.000
	Likelihood Ratio	105.976	6	.000
	Linear-by-Linear Association	77.673	1	.000
	N of Valid Cases	84065		
OS-UK	Pearson Chi-Square	37.201[5]	6	.000
	Likelihood Ratio	37.118	6	.000
	Linear-by-Linear Association	30.045	1	.000
	N of Valid Cases	195273		

1. 2 cells (16,7%) have expected count less than 5. The minimum expected count is ,91.
2. 0 cells (,0%) have expected count less than 5. The minimum expected count is 23,74.
3. 2 cells (16,7%) have expected count less than 5. The minimum expected count is 1,98.
4. 1 cells (8,3%) have expected count less than 5. The minimum expected count is 4,57.
5. 0 cells (,0%) have expected count less than 5. The minimum expected count is 11,73.

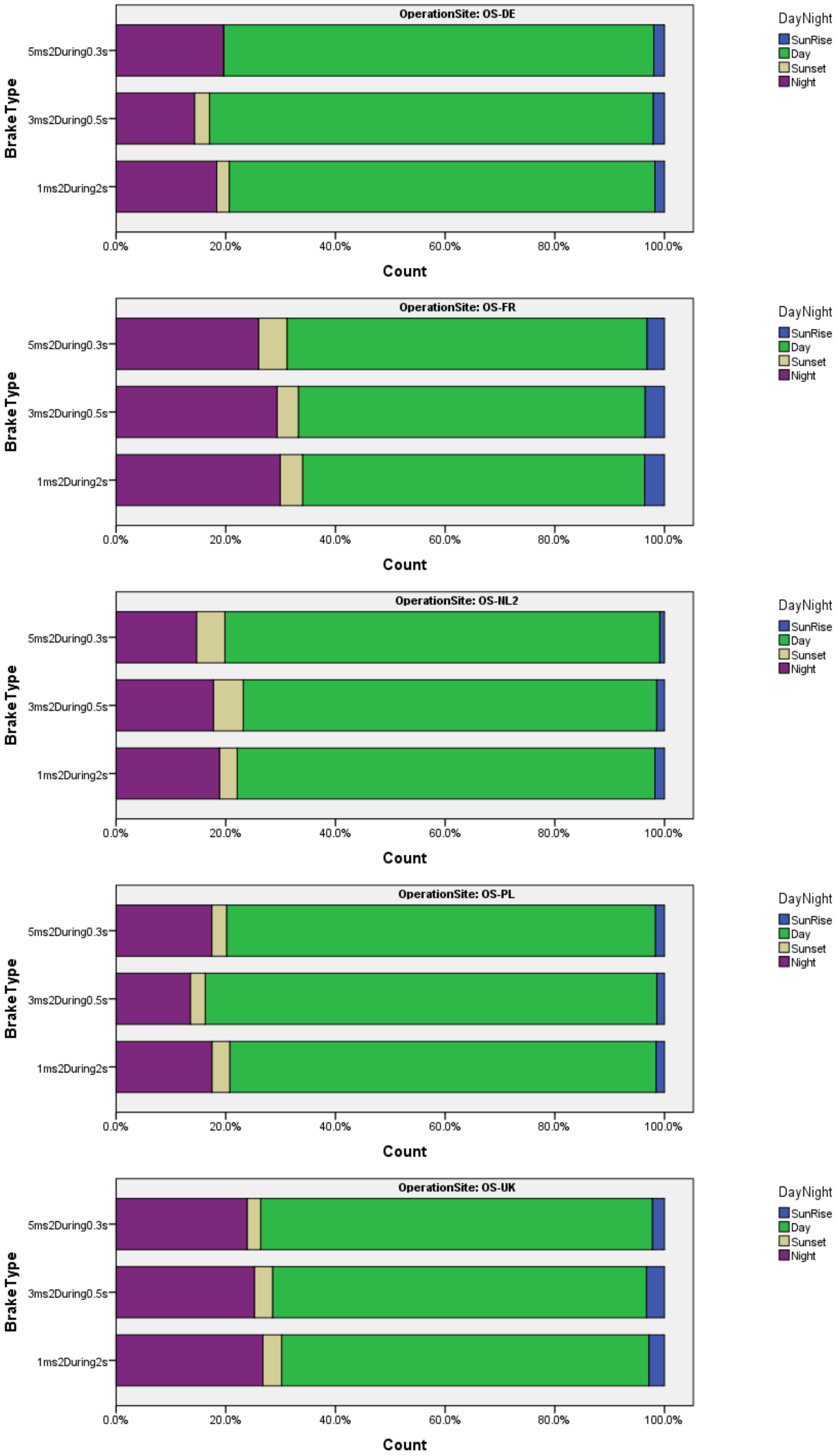


Figure B-7 : Breakdown of day categories for each type of brake and for each operational site

B.4 Breakdown of braking by rain category and by operational site

Table B-0-21 : Z-test on rain category by operational site

BrakeType * RainState Crosstabulation

OperationSite				RainState NoRain	rain	HeavyRain	Total
OS-DE	BrakeType	1ms2During2s	Count	25942a	2158a	82a	28182
			% within RainState	94.7%	95.3%	92.1%	94.8%
		3ms2During0.5s	Count	1398a	100a	7a	1505
			% within RainState	5.1%	4.4%	7.9%	5.1%
		5ms2During0.3s	Count	46a	6a	0a	52
			% within RainState	0.2%	0.3%	0.0%	0.2%
	Total		Count	27386	2264	89	29739
			% within RainState	100.0%	100.0%	100.0%	100.0%
OS-FR	BrakeType	1ms2During2s	Count	191680a	7152b	249b	199081
			% within RainState	88.1%	89.4%	92.9%	88.2%
		3ms2During0.5s	Count	25131a	832b	19a, b	25982
			% within RainState	11.6%	10.4%	7.1%	11.5%
		5ms2During0.3s	Count	641a	18a	0a	659
			% within RainState	0.3%	0.2%	0.0%	0.3%
	Total		Count	217452	8002	268	225722
			% within RainState	100.0%	100.0%	100.0%	100.0%
OS-NL2	BrakeType	1ms2During2s	Count	25259			25259
			% within RainState	87.5%			87.5%
		3ms2During0.5s	Count	3492			3492
			% within RainState	12.1%			12.1%
		5ms2During0.3s	Count	117			117
			% within RainState	0.4%			0.4%
	Total		Count	28868			28868
			% within RainState	100.0%			100.0%
OS-PL	BrakeType	1ms2During2s	Count	68855a	6104b	312a, b	75271
			% within RainState	89.1%	90.9%	93.1%	89.3%
		3ms2During0.5s	Count	8137a	592b	22a, b	8751
			% within RainState	10.5%	8.8%	6.6%	10.4%
		5ms2During0.3s	Count	281a	18a	1a	300
			% within RainState	0.4%	0.3%	0.3%	0.4%
	Total		Count	77273	6714	335	84322
			% within RainState	100.0%	100.0%	100.0%	100.0%
OS-UK	BrakeType	1ms2During2s	Count	154337a	21045a	446a	175828
			% within RainState	89.7%	89.7%	91.4%	89.7%
		3ms2During0.5s	Count	17431a	2377a	41a	19849
			% within RainState	10.1%	10.1%	8.4%	10.1%
		5ms2During0.3s	Count	364a	41a	1a	406
			% within RainState	0.2%	0.2%	0.2%	0.2%
	Total		Count	172132	23463	488	196083
			% within RainState	100.0%	100.0%	100.0%	100.0%

Each subscript letter denotes a subset of RainState categories whose column proportions do not differ significantly from each other at the ,05 level.

Table B-0-22 : Chi-2 tests on rain category by operational site

Chi-Square Tests

OperationSite		Value	df	Asymptotic Significance (2-sided)
OS-DE	Pearson Chi-Square	4.767[1]	4	.312
	Likelihood Ratio	4.661	4	.324
	Linear-by-Linear Association	.283	1	.595
	N of Valid Cases	29739		
OS-FR	Pearson Chi-Square	17.644[2]	4	.001
	Likelihood Ratio	19.537	4	.001
	Linear-by-Linear Association	16.338	1	.000
	N of Valid Cases	225722		
OS-NL2	Pearson Chi-Square	.[3]		
	N of Valid Cases	28868		
OS-PL	Pearson Chi-Square	26.612[4]	4	.000
	Likelihood Ratio	28.286	4	.000
	Linear-by-Linear Association	26.062	1	.000
	N of Valid Cases	84322		
OS-UK	Pearson Chi-Square	2.941[5]	4	.568
	Likelihood Ratio	3.095	4	.542
	Linear-by-Linear Association	.412	1	.521
	N of Valid Cases	196083		

1. 3 cells (33,3%) have expected count less than 5. The minimum expected count is ,16.
2. 1 cells (11,1%) have expected count less than 5. The minimum expected count is ,78.
3. No statistics are computed because RainState is a constant.
4. 1 cells (11,1%) have expected count less than 5. The minimum expected count is 1,19.
5. 1 cells (11,1%) have expected count less than 5. The minimum expected count is 1,01.

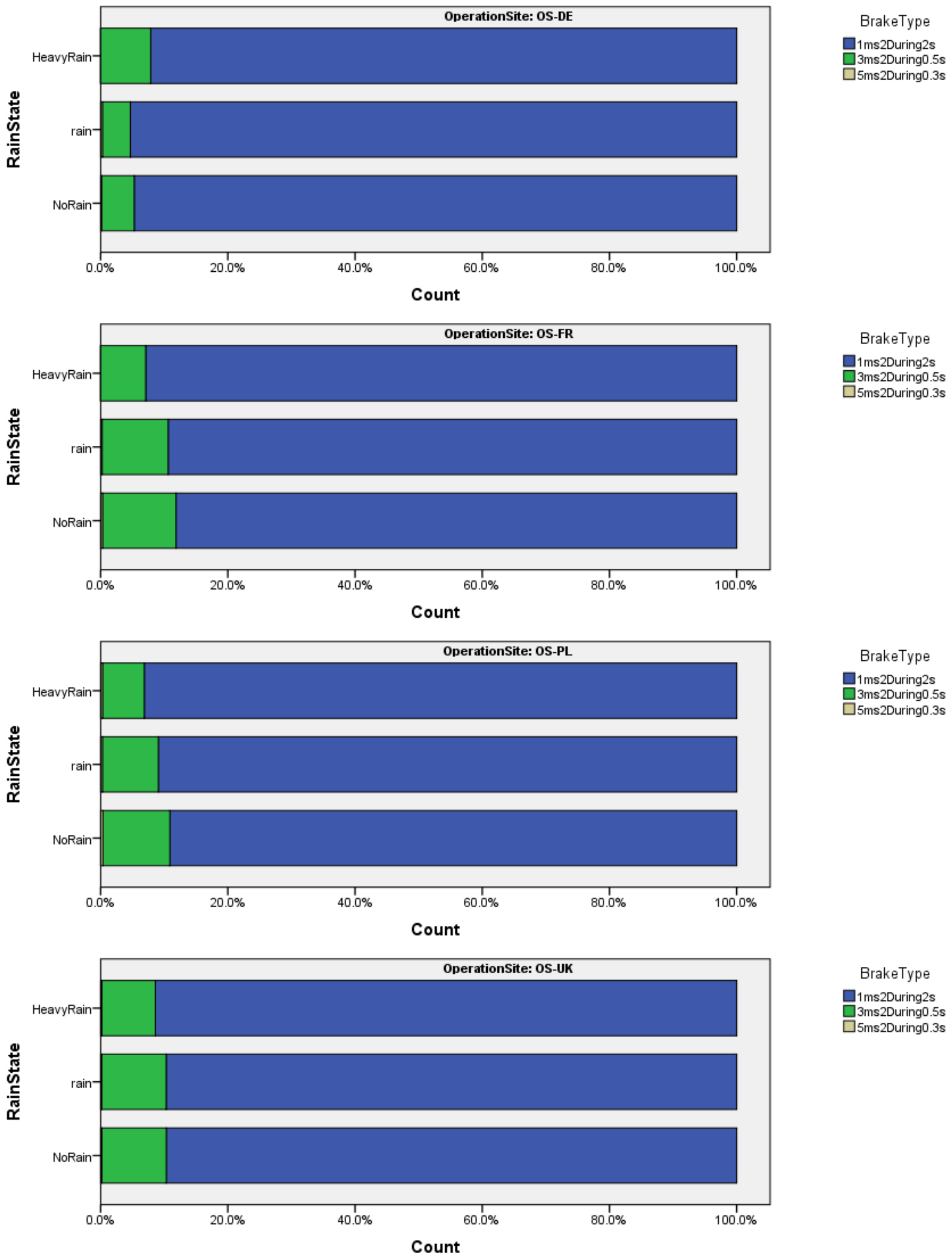


Figure B-8 : Breakdown of brake categories for each type of rain and for each operational site

B.5 Breakdown of braking by category of assistance system and operational site

Table B-0-23 : Z-test on ADAS category by operational site

BrakeType * CCSLState5s Crosstabulation

OperationSite				CCSLState5s: off	CCSLState5s: CruiseControl	CCSLState5s: SpeedLimiteur	Total
OS-DE	BrakeType	1ms2During2s	Count	15001a	458b		15459
			% within CCSLState5s	94.5%	92.2%		94.4%
		3ms2During0.5s	Count	850a	38b		888
			% within CCSLState5s	5.4%	7.6%		5.4%
		5ms2During0.3s	Count	27a	1a		28
			% within CCSLState5s	0.2%	0.2%		0.2%
	Total		Count	15878	497		16375
			% within CCSLState5s	100.0%	100.0%		100.0%
OS-FR	BrakeType	1ms2During2s	Count	176605a	457a	11187a	188249
			% within CCSLState5s	87.9%	90.0%	87.7%	87.9%
		3ms2During0.5s	Count	23694a	49a	1560a	25303
			% within CCSLState5s	11.8%	9.6%	12.2%	11.8%
		5ms2During0.3s	Count	620a	2a, b	13b	635
			% within CCSLState5s	0.3%	0.4%	0.1%	0.3%
	Total		Count	200919	508	12760	214187
			% within CCSLState5s	100.0%	100.0%	100.0%	100.0%
OS-NL2	BrakeType	1ms2During2s	Count	23905a	1081b	273c	25259
			% within CCSLState5s	87.6%	83.7%	97.8%	87.5%
		3ms2During0.5s	Count	3289a	197b	6c	3492
			% within CCSLState5s	12.0%	15.3%	2.2%	12.1%
		5ms2During0.3s	Count	104a	13b	0a, b	117
			% within CCSLState5s	0.4%	1.0%	0.0%	0.4%
	Total		Count	27298	1291	279	28868
			% within CCSLState5s	100.0%	100.0%	100.0%	100.0%
OS-PL	BrakeType	1ms2During2s	Count	38069a	406b		38475
			% within CCSLState5s	88.1%	92.1%		88.1%
		3ms2During0.5s	Count	5006a	32b		5038
			% within CCSLState5s	11.6%	7.3%		11.5%
		5ms2During0.3s	Count	144a	3a		147
			% within CCSLState5s	0.3%	0.7%		0.3%
	Total		Count	43219	441		43660
			% within CCSLState5s	100.0%	100.0%		100.0%
OS-UK	BrakeType	1ms2During2s	Count	80503a	675b	151c	81329
			% within CCSLState5s	85.7%	97.1%	92.6%	85.7%
		3ms2During0.5s	Count	13239a	18b	11c	13268
			% within CCSLState5s	14.1%	2.6%	6.7%	14.0%
		5ms2During0.3s	Count	248a	2a	1a	251
			% within CCSLState5s	0.3%	0.3%	0.6%	0.3%
	Total		Count	93990	695	163	94848
			% within CCSLState5s	100.0%	100.0%	100.0%	100.0%

Each subscript letter denotes a subset of CCSLState5s categories whose column proportions do not differ significantly from each other at the ,05 level.

Table B-0-24 : Chi2 tests on ADAS category by operational site

Chi-Square Tests

OperationSite		Value	df	Asymptotic Significance (2-sided)
OS-DE	Pearson Chi-Square	4.974[1]	2	.083
	Likelihood Ratio	4.458	2	.108
	Linear-by-Linear Association	4.628	1	.031
	N of Valid Cases	16375		
OS-FR	Pearson Chi-Square	21.700[2]	4	.000
	Likelihood Ratio	27.376	4	.000
	Linear-by-Linear Association	.003	1	.956
	N of Valid Cases	214187		
OS-NL2	Pearson Chi-Square	52.130[3]	4	.000
	Likelihood Ratio	60.899	4	.000
	Linear-by-Linear Association	.048	1	.826
	N of Valid Cases	28868		
OS-PL	Pearson Chi-Square	9.428[4]	2	.009
	Likelihood Ratio	10.137	2	.006
	Linear-by-Linear Association	5.055	1	.025
	N of Valid Cases	43660		
OS-UK	Pearson Chi-Square	83.551[5]	4	.000
	Likelihood Ratio	117.514	4	.000
	Linear-by-Linear Association	58.110	1	.000
	N of Valid Cases	94848		

1. 1 cells (16,7%) have expected count less than 5. The minimum expected count is ,85.
2. 1 cells (11,1%) have expected count less than 5. The minimum expected count is 1,51.
3. 1 cells (11,1%) have expected count less than 5. The minimum expected count is 1,13.
4. 1 cells (16,7%) have expected count less than 5. The minimum expected count is 1,48.
5. 2 cells (22,2%) have expected count less than 5. The minimum expected count is ,43.

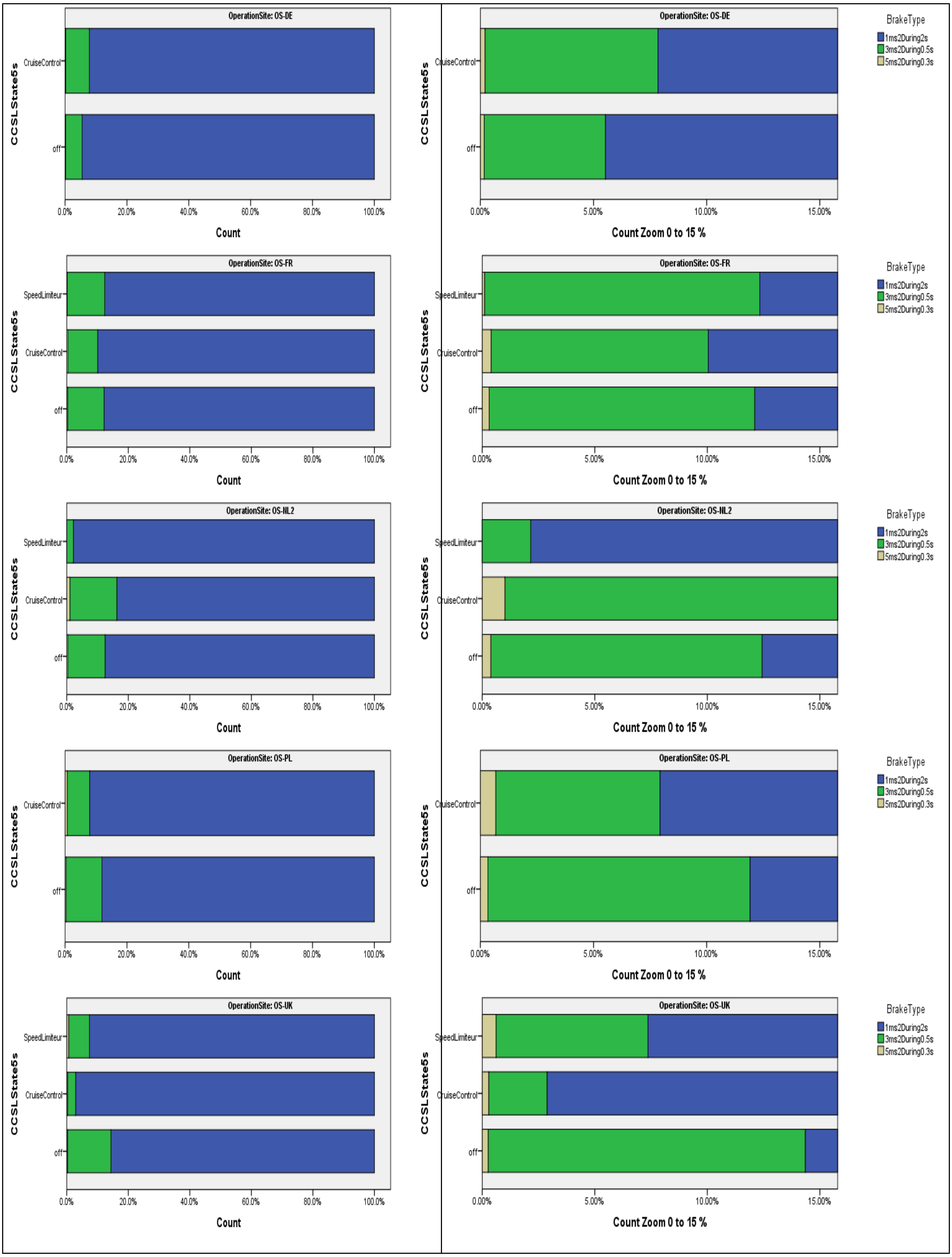


Figure B-9 : Breakdown of brake categories for each type of ADAS and for each operational site

B.6 Breakdown of braking by age group and operational site

Table B-0-25 : Z-test on age category by operational site

BrakeType * CatAge Crosstabulation

OperationSite				Inf30 (CatAge)	30-40 (CatAge)	40-60 (CatAge)	Sup60 (CatAge)	Total
OS-DE	BrakeType	1ms2During2s	Count	7330$_{a}$	3358$_{a, b}$	15092$_{a}$	6058$_{b}$	31838
			% within CatAge	94.9%	94.0%	94.7%	93.8%	94.5%
		3ms2During0.5s	Count	377$_{a}$	206$_{a, b}$	811$_{a}$	388$_{b}$	1782
			% within CatAge	4.9%	5.8%	5.1%	6.0%	5.3%
		5ms2During0.3s	Count	15$_{a}$	8$_{a}$	27$_{a}$	11$_{a}$	61
			% within CatAge	0.2%	0.2%	0.2%	0.2%	0.2%
	Total		Count	7722	3572	15930	6457	33681
			% within CatAge	100.0%	100.0%	100.0%	100.0%	100.0%
OS-FR	BrakeType	1ms2During2s	Count	37040$_{a}$	51215$_{b}$	106800$_{c}$	20864$_{a}$	215919
			% within CatAge	85.2%	87.4%	90.2%	85.6%	88.2%
		3ms2During0.5s	Count	6212$_{a}$	7209$_{b}$	11366$_{c}$	3427$_{a}$	28214
			% within CatAge	14.3%	12.3%	9.6%	14.1%	11.5%
		5ms2During0.3s	Count	217$_{a}$	165$_{b}$	243$_{c}$	92$_{a, b}$	717
			% within CatAge	0.5%	0.3%	0.2%	0.4%	0.3%
	Total		Count	43469	58589	118409	24383	244850
			% within CatAge	100.0%	100.0%	100.0%	100.0%	100.0%
OS-NL2	BrakeType	1ms2During2s	Count	8577$_{a}$	2437$_{b}$	14178$_{b}$	3328$_{a}$	28520
			% within CatAge	93.2%	85.6%	83.8%	92.5%	87.6%
		3ms2During0.5s	Count	604$_{a}$	400$_{b}$	2640$_{b}$	257$_{a}$	3901
			% within CatAge	6.6%	14.1%	15.6%	7.1%	12.0%
		5ms2During0.3s	Count	17$_{a}$	9$_{a, b}$	96$_{b}$	13$_{a, b}$	135
			% within CatAge	0.2%	0.3%	0.6%	0.4%	0.4%
	Total		Count	9198	2846	16914	3598	32556
			% within CatAge	100.0%	100.0%	100.0%	100.0%	100.0%
OS-PL	BrakeType	1ms2During2s	Count	5519$_{a}$	47396$_{b}$	25575$_{b}$	4068$_{c}$	82558
			% within CatAge	92.8%	89.4%	89.2%	87.7%	89.5%
		3ms2During0.5s	Count	408$_{a}$	5440$_{b}$	3002$_{b}$	556$_{c}$	9406
			% within CatAge	6.9%	10.3%	10.5%	12.0%	10.2%
		5ms2During0.3s	Count	17$_{a}$	195$_{a}$	94$_{a}$	16$_{a}$	322
			% within CatAge	0.3%	0.4%	0.3%	0.3%	0.3%
	Total		Count	5944	53031	28671	4640	92286
			% within CatAge	100.0%	100.0%	100.0%	100.0%	100.0%
OS-UK	BrakeType	1ms2During2s	Count	26917$_{a}$	55393$_{b}$	74952$_{c}$	26267$_{d}$	183529
			% within CatAge	93.7%	89.4%	87.8%	91.9%	89.7%
		3ms2During0.5s	Count	1767$_{a}$	6393$_{b}$	10229$_{c}$	2274$_{d}$	20663
			% within CatAge	6.1%	10.3%	12.0%	8.0%	10.1%
		5ms2During0.3s	Count	52$_{a, b}$	158$_{b}$	170$_{a, b}$	44$_{a}$	424
			% within CatAge	0.2%	0.3%	0.2%	0.2%	0.2%
	Total		Count	28736	61944	85351	28585	204616
			% within CatAge	100.0%	100.0%	100.0%	100.0%	100.0%

Each subscript letter denotes a subset of CatAge categories whose column proportions do not differ significantly from each other at the ,05 level.

Table B-0-26 : Chi2 tests on day category by operational site

Chi-Square Tests

OperationSite		Value	df	Asymptotic Significance (2-sided)
OS-DE	Pearson Chi-Square	12.708[a]	6	.048
	Likelihood Ratio	12.454	6	.053
	Linear-by-Linear Association	3.416	1	.065
	N of Valid Cases	33681		
OS-FR	Pearson Chi-Square	1057.324[b]	6	.000
	Likelihood Ratio	1042.434	6	.000
	Linear-by-Linear Association	294.363	1	.000
	N of Valid Cases	244850		
OS-NL2	Pearson Chi-Square	585.566[c]	6	.000
	Likelihood Ratio	627.035	6	.000
	Linear-by-Linear Association	178.592	1	.000
	N of Valid Cases	32556		
OS-PL	Pearson Chi-Square	92.726[d]	6	.000
	Likelihood Ratio	100.347	6	.000
	Linear-by-Linear Association	43.918	1	.000
	N of Valid Cases	92286		
OS-UK	Pearson Chi-Square	989.651[e]	6	.000
	Likelihood Ratio	1054.938	6	.000
	Linear-by-Linear Association	131.395	1	.000
	N of Valid Cases	204616		

a. 0 cells (,0%) have expected count less than 5. The minimum expected count is 6,47.

b. 0 cells (,0%) have expected count less than 5. The minimum expected count is 71,40.

c. 0 cells (,0%) have expected count less than 5. The minimum expected count is 11,80.

d. 0 cells (,0%) have expected count less than 5. The minimum expected count is 16,19.

e. 0 cells (,0%) have expected count less than 5. The minimum expected count is 59,23.

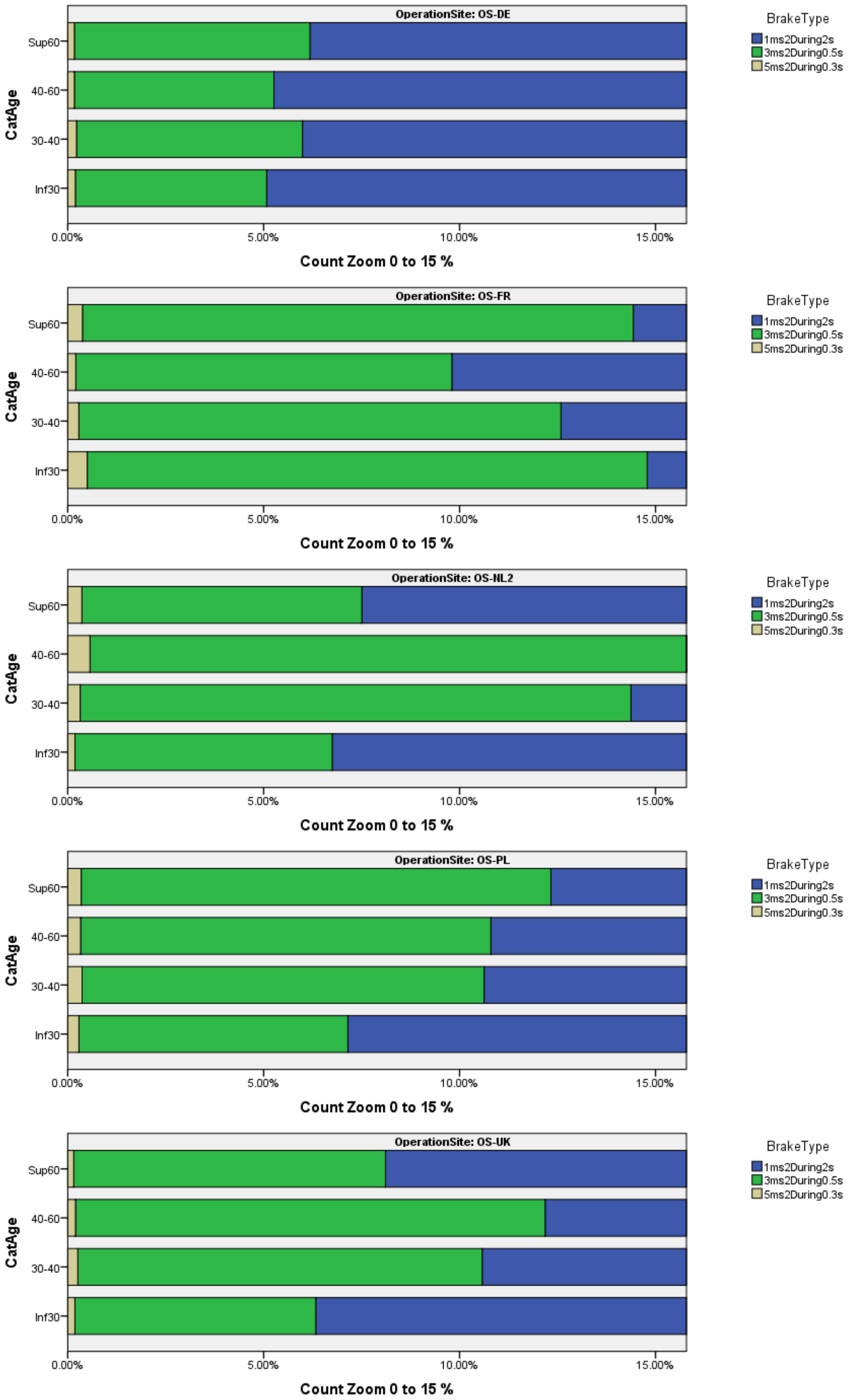


Figure B-10 : Breakdown of brake categories for each type of age each operational site

B.7 Breakdown of braking by genre and operational site

Table B-0-27 : Z-test on genre category by operational site

BrakeType * Sex Crosstabulation

OperationSite				Sex M	Sex F	Total
OS-DE	BrakeType	1ms2During2s	Count	20447a	11391b	31838
			% within Sex	93.6%	96.2%	94.5%
		3ms2During0.5s	Count	1348a	434b	1782
			% within Sex	6.2%	3.7%	5.3%
		5ms2During0.3s	Count	48a	13b	61
			% within Sex	0.2%	0.1%	0.2%
	Total		Count	21843	11838	33681
			% within Sex	100.0%	100.0%	100.0%
OS-FR	BrakeType	1ms2During2s	Count	96174a	119745b	215919
			% within Sex	88.3%	88.1%	88.2%
		3ms2During0.5s	Count	12310a	15904b	28214
			% within Sex	11.3%	11.7%	11.5%
		5ms2During0.3s	Count	374a	343b	717
			% within Sex	0.3%	0.3%	0.3%
	Total		Count	108858	135992	244850
			% within Sex	100.0%	100.0%	100.0%
OS-NL2	BrakeType	1ms2During2s	Count	6708a	21812a	28520
			% within Sex	87.9%	87.5%	87.6%
		3ms2During0.5s	Count	888a	3013a	3901
			% within Sex	11.6%	12.1%	12.0%
		5ms2During0.3s	Count	33a	102a	135
			% within Sex	0.4%	0.4%	0.4%
	Total		Count	7629	24927	32556
			% within Sex	100.0%	100.0%	100.0%
OS-PL	BrakeType	1ms2During2s	Count	59050a	23508a	82558
			% within Sex	89.5%	89.4%	89.5%
		3ms2During0.5s	Count	6729a	2677a	9406
			% within Sex	10.2%	10.2%	10.2%
		5ms2During0.3s	Count	218a	104a	322
			% within Sex	0.3%	0.4%	0.3%
	Total		Count	65997	26289	92286
			% within Sex	100.0%	100.0%	100.0%
OS-UK	BrakeType	1ms2During2s	Count	59937a	123592b	183529
			% within Sex	91.1%	89.0%	89.7%
		3ms2During0.5s	Count	5717a	14946b	20663
			% within Sex	8.7%	10.8%	10.1%
		5ms2During0.3s	Count	114a	310b	424
			% within Sex	0.2%	0.2%	0.2%
	Total		Count	65768	138848	204616
			% within Sex	100.0%	100.0%	100.0%

Each subscript letter denotes a subset of Sex categories whose column proportions do not differ significantly from each other at the ,05 level.

Table B-0-28 Chi2 tests on genre category by operational site

Chi-Square Tests

OperationSite		Value	df	Asymptotic Significance (2-sided)
OS-DE	Pearson Chi-Square	101.742[a]	2	.000
	Likelihood Ratio	107.709	2	.000
	Linear-by-Linear Association	100.089	1	.000
	N of Valid Cases	33681		
OS-FR	Pearson Chi-Square	25.663[b]	2	.000
	Likelihood Ratio	25.545	2	.000
	Linear-by-Linear Association	2.237	1	.135
	N of Valid Cases	244850		
OS-NL2	Pearson Chi-Square	1.173[c]	2	.556
	Likelihood Ratio	1.179	2	.555
	Linear-by-Linear Association	.782	1	.377
	N of Valid Cases	32556		
OS-PL	Pearson Chi-Square	2.305[d]	2	.316
	Likelihood Ratio	2.249	2	.325
	Linear-by-Linear Association	.250	1	.617
	N of Valid Cases	92286		
OS-UK	Pearson Chi-Square	217.425[e]	2	.000
	Likelihood Ratio	222.665	2	.000
	Linear-by-Linear Association	214.248	1	.000
	N of Valid Cases	204616		

a. 0 cells (,0%) have expected count less than 5. The minimum expected count is 21,44.

b. 0 cells (,0%) have expected count less than 5. The minimum expected count is 318,77.

c. 0 cells (,0%) have expected count less than 5. The minimum expected count is 31,64.

d. 0 cells (,0%) have expected count less than 5. The minimum expected count is 91,73.

e. 0 cells (,0%) have expected count less than 5. The minimum expected count is 136,28.

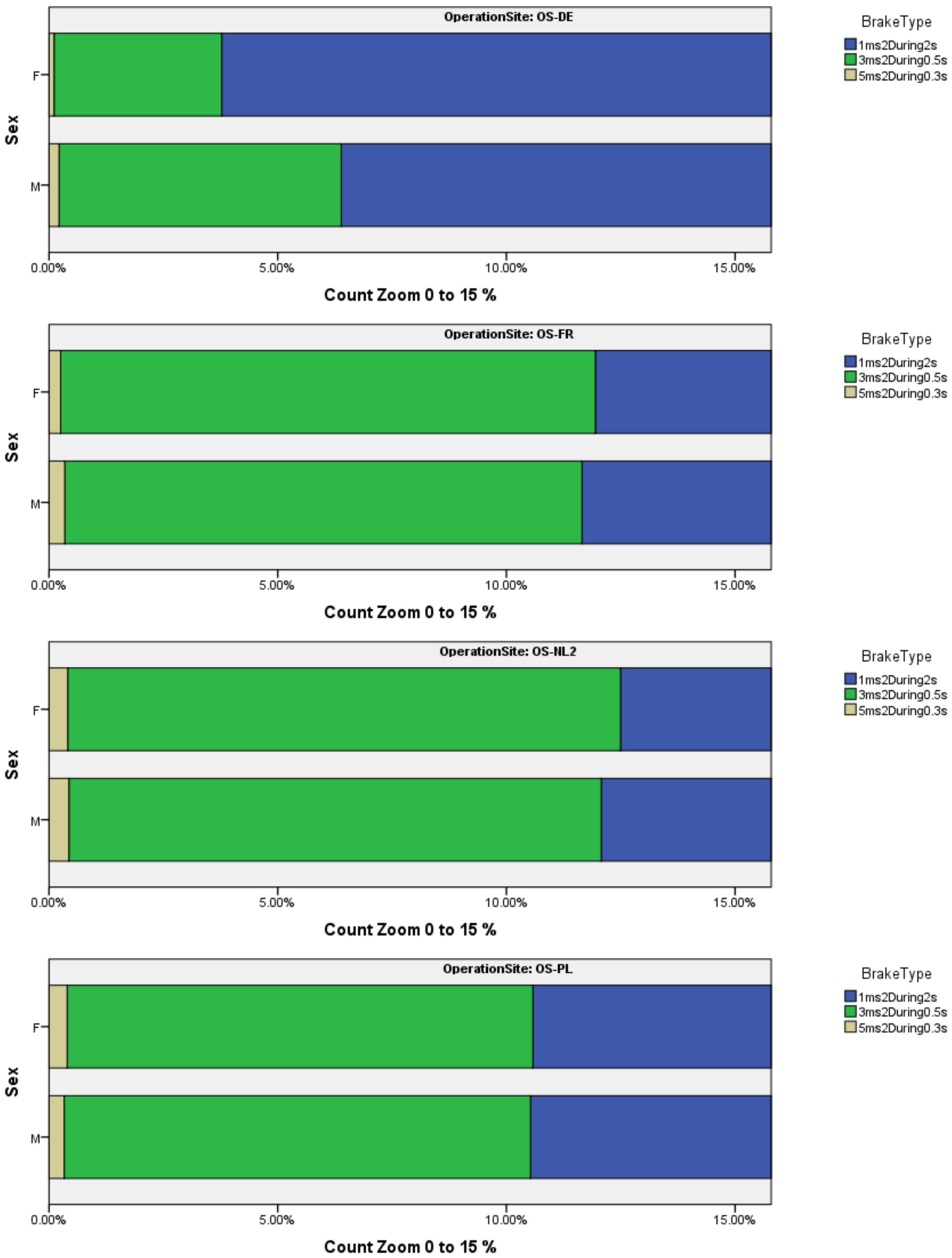


Figure B-11 : Breakdown of brake categories by genre each operational site

B.8 Breakdown of braking by infrastructure type and operational site

Table B-0-29: Z-test on infrastructure category by operational site

BrakeType * InfraType2 Crosstabulation

				InfraType2						
OperationSite				RoundAbout	IntersectionWithSpeedNull5sAfter	Intersection	SpeedReduction	Curve	NoCurve	Total
OS-DE	BrakeType	1ms2During2s	Count	886$_{a, b}$	6052$_{b}$	11285$_{a}$	1842$_{a, b}$	7084$_{a, b}$	1033$_{a, b}$	28182
			% within InfraType2	94.1%	94.2%	95.3%	95.1%	94.5%	94.6%	94.8%
		3ms2During0.5s	Count	53$_{a, b}$	363$_{b}$	536$_{a}$	90$_{a, b}$	407$_{a, b}$	56$_{a, b}$	1505
			% within InfraType2	5.6%	5.6%	4.5%	4.6%	5.4%	5.1%	5.1%
		5ms2During0.3s	Count	3$_{a}$	13$_{a}$	20$_{a}$	5$_{a}$	8$_{a}$	3$_{a}$	52
			% within InfraType2	0.3%	0.2%	0.2%	0.3%	0.1%	0.3%	0.2%
	Total		Count	942	6428	11841	1937	7499	1092	29739
			% within InfraType2	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
OS-FR	BrakeType	1ms2During2s	Count	29032$_{a}$	28130$_{a}$	77658$_{a}$	10726$_{b}$	47400$_{b}$	6135$_{b}$	199081
			% within InfraType2	87.3%	86.8%	86.7%	91.7%	91.0%	91.7%	88.2%
		3ms2During0.5s	Count	4139$_{a}$	4152$_{a}$	11677$_{a}$	950$_{b}$	4524$_{b}$	540$_{b}$	25982
			% within InfraType2	12.4%	12.8%	13.0%	8.1%	8.7%	8.1%	11.5%
		5ms2During0.3s	Count	94$_{a}$	141$_{b}$	218$_{a}$	23$_{a}$	165$_{a, b}$	18$_{a, b}$	659
			% within InfraType2	0.3%	0.4%	0.2%	0.2%	0.3%	0.3%	0.3%
	Total		Count	33265	32423	89553	11699	52089	6693	225722
			% within InfraType2	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
OS-NL2	BrakeType	1ms2During2s	Count	2576$_{a, b}$	3650$_{b}$	9666$_{c}$	992$_{a, c}$	6667$_{c}$	1708$_{c}$	25259
			% within InfraType2	84.9%	83.0%	89.2%	87.9%	88.2%	88.9%	87.5%
		3ms2During0.5s	Count	440$_{a, b}$	727$_{b}$	1139$_{c}$	132$_{a, c}$	849$_{c}$	205$_{c}$	3492
			% within InfraType2	14.5%	16.5%	10.5%	11.7%	11.2%	10.7%	12.1%
		5ms2During0.3s	Count	18$_{a}$	18$_{a}$	30$_{a}$	4$_{a}$	39$_{a}$	8$_{a}$	117
			% within InfraType2	0.6%	0.4%	0.3%	0.4%	0.5%	0.4%	0.4%
	Total		Count	3034	4395	10835	1128	7555	1921	28868
			% within InfraType2	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
OS-PL	BrakeType	1ms2During2s	Count	2345$_{a}$	10705$_{b, c}$	28781$_{c}$	2733$_{b, c, d}$	24520$_{d}$	6186$_{b, d}$	75270
			% within InfraType2	86.3%	88.6%	88.5%	89.8%	90.5%	90.0%	89.3%
		3ms2During0.5s	Count	365$_{a}$	1297$_{b, c}$	3638$_{c}$	298$_{b, c, d}$	2493$_{d}$	660$_{b, d}$	8751
			% within InfraType2	13.4%	10.7%	11.2%	9.8%	9.2%	9.6%	10.4%
		5ms2During0.3s	Count	8$_{a, b}$	76$_{b}$	89$_{a}$	13$_{a, b}$	86$_{a}$	28$_{a, b}$	300
			% within InfraType2	0.3%	0.6%	0.3%	0.4%	0.3%	0.4%	0.4%
	Total		Count	2718	12078	32508	3044	27099	6874	84321
			% within InfraType2	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
OS-UK	BrakeType	1ms2During2s	Count	24789$_{a}$	27370$_{a, b}$	74261$_{b}$	4682$_{c}$	40902$_{c}$	3824$_{c}$	175828
			% within InfraType2	88.0%	88.7%	89.3%	92.7%	91.6%	91.4%	89.7%
		3ms2During0.5s	Count	3305$_{a}$	3390$_{a, b}$	8804$_{b}$	362$_{c}$	3644$_{c}$	343$_{c}$	19848
			% within InfraType2	11.7%	11.0%	10.6%	7.2%	8.2%	8.2%	10.1%
		5ms2During0.3s	Count	76$_{a}$	91$_{a}$	123$_{b}$	9$_{a, b}$	90$_{a, b}$	17$_{a}$	406
			% within InfraType2	0.3%	0.3%	0.1%	0.2%	0.2%	0.4%	0.2%
	Total		Count	28170	30851	83188	5053	44636	4184	196082
			% within InfraType2	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%

Each subscript letter denotes a subset of InfraType2 categories whose column proportions do not differ significantly from each other at the ,05 level.

UDRIVE European Naturalistic Driving Study

Table B-0-30 : Chi2 tests on infrastructure category by operational site

Chi-Square Tests

OperationSite		Value	df	Asymptotic Significance (2-sided)
OS-DE	Pearson Chi-Square	19.858[1]	10	.031
	Likelihood Ratio	19.799	10	.031
	Linear-by-Linear Association	.341	1	.559
	N of Valid Cases	29739		
OS-FR	Pearson Chi-Square	940.608[2]	10	.000
	Likelihood Ratio	983.476	10	.000
	Linear-by-Linear Association	696.199	1	.000
	N of Valid Cases	225722		
OS-NL2	Pearson Chi-Square	143.046[3]	10	.000
	Likelihood Ratio	136.216	10	.000
	Linear-by-Linear Association	16.972	1	.000
	N of Valid Cases	28868		
OS-PL	Pearson Chi-Square	131.963[4]	10	.000
	Likelihood Ratio	126.706	10	.000
	Linear-by-Linear Association	69.234	1	.000
	N of Valid Cases	84321		
OS-UK	Pearson Chi-Square	418.868[5]	10	.000
	Likelihood Ratio	430.480	10	.000
	Linear-by-Linear Association	309.602	1	.000
	N of Valid Cases	196082		

1. 3 cells (16,7%) have expected count less than 5. The minimum expected count is 1,65.
2. 0 cells (,0%) have expected count less than 5. The minimum expected count is 19,54.
3. 1 cells (5,6%) have expected count less than 5. The minimum expected count is 4,57.
4. 0 cells (,0%) have expected count less than 5. The minimum expected count is 9,67.
5. 0 cells (,0%) have expected count less than 5. The minimum expected count is 8,66.

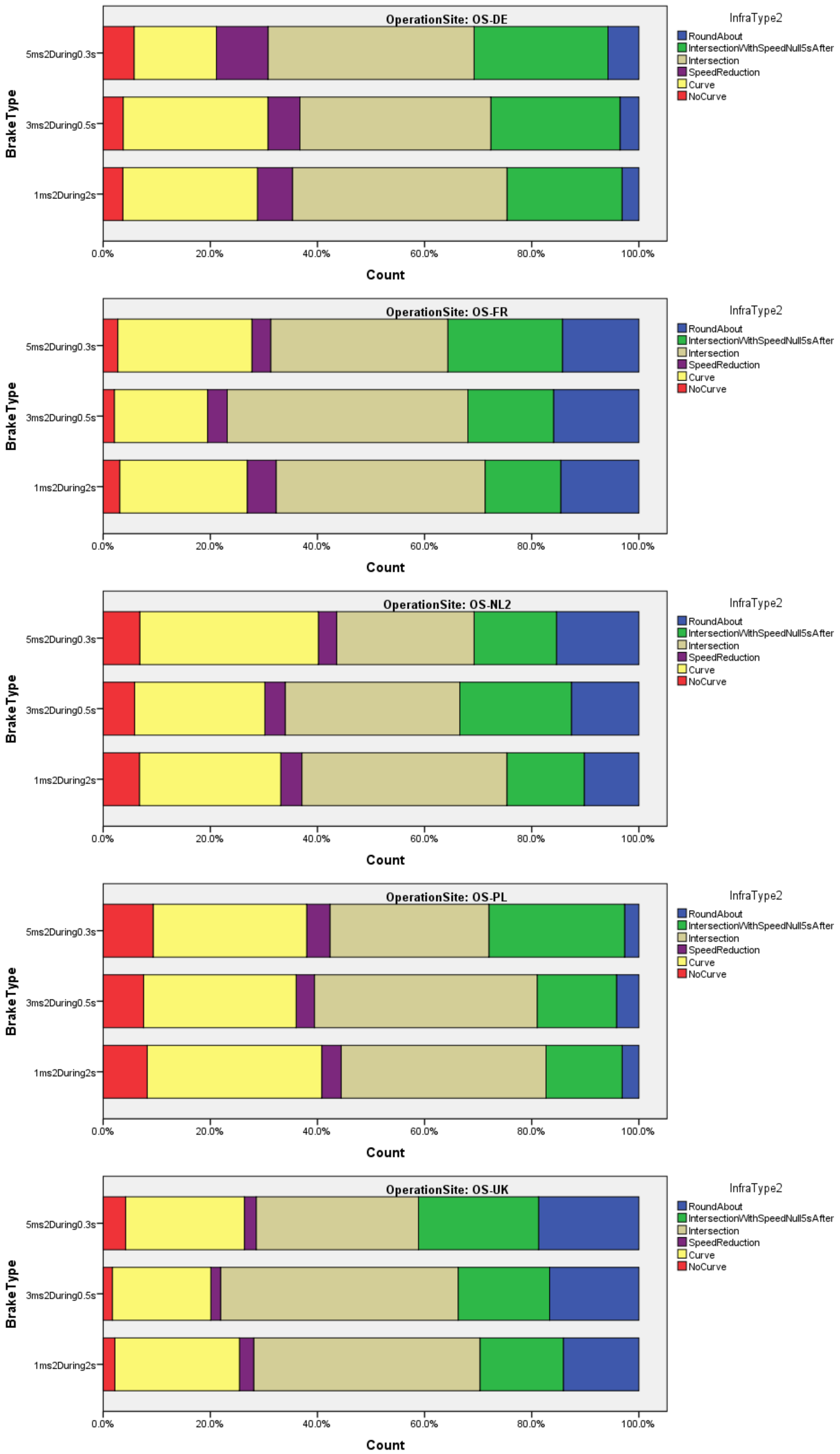


Figure B-12 : Breakdown of infrastructure categories for each type of brake for each operational site

B.9 Post Hoc tests for leading time analyse by infrastructure type and operational site

Table B-0-31 : ANOVA post hoc test for leading time analyse by OS

Multiple Comparisons

Dependent Variable: TIVStartBrakeThreshold

Bonferroni

OperationSite	(I) BrakeType	(J) BrakeType	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
OS-DE	1ms2During2s	3ms2During0.5s	.47275*	.03702	.000	.3841	.5614
		5ms2During0.3s	.51915*	.18057	.012	.0868	.9515
	3ms2During0.5s	1ms2During2s	-.47275*	.03702	.000	-.5614	-.3841
		5ms2During0.3s	.04640	.18396	1.000	-.3941	.4869
	5ms2During0.3s	1ms2During2s	-.51915*	.18057	.012	-.9515	-.0868
		3ms2During0.5s	-.04640	.18396	1.000	-.4869	.3941
OS-FR	1ms2During2s	3ms2During0.5s	.40593*	.01053	.000	.3807	.4311
		5ms2During0.3s	.50703*	.06349	.000	.3550	.6590
	3ms2During0.5s	1ms2During2s	-.40593*	.01053	.000	-.4311	-.3807
		5ms2During0.3s	.10110	.06420	.346	-.0526	.2548
	5ms2During0.3s	1ms2During2s	-.50703*	.06349	.000	-.6590	-.3550
		3ms2During0.5s	-.10110	.06420	.346	-.2548	.0526
OS-NL2	1ms2During2s	3ms2During0.5s	.40499*	.02234	.000	.3515	.4585
		5ms2During0.3s	.62440*	.10983	.000	.3614	.8874
	3ms2During0.5s	1ms2During2s	-.40499*	.02234	.000	-.4585	-.3515
		5ms2During0.3s	.21941	.11147	.147	-.0475	.4863
	5ms2During0.3s	1ms2During2s	-.62440*	.10983	.000	-.8874	-.3614
		3ms2During0.5s	-.21941	.11147	.147	-.4863	.0475
OS-PL	1ms2During2s	3ms2During0.5s	.38023*	.01395	.000	.3468	.4136
		5ms2During0.3s	.47493*	.07441	.000	.2968	.6531
	3ms2During0.5s	1ms2During2s	-.38023*	.01395	.000	-.4136	-.3468
		5ms2During0.3s	.09471	.07547	.629	-.0860	.2754
	5ms2During0.3s	1ms2During2s	-.47493*	.07441	.000	-.6531	-.2968
		3ms2During0.5s	-.09471	.07547	.629	-.2754	.0860
OS-UK	1ms2During2s	3ms2During0.5s	.34723*	.01081	.000	.3214	.3731
		5ms2During0.3s	.59807*	.07319	.000	.4228	.7733
	3ms2During0.5s	1ms2During2s	-.34723*	.01081	.000	-.3731	-.3214
		5ms2During0.3s	.25084*	.07386	.002	.0740	.4277
	5ms2During0.3s	1ms2During2s	-.59807*	.07319	.000	-.7733	-.4228
		3ms2During0.5s	-.25084*	.07386	.002	-.4277	-.0740

*. The mean difference is significant at the 0.05 level.

Table B-0-32 : ANOVA post hoc test for leading time analyse by infrastructure

Multiple Comparisons

Dependent Variable: TIVStartBrakeThreshold

Bonferroni

InfraType2	(I) BrakeType	(J) BrakeType	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval Lower Bound	Upper Bound
RoundAbout	1ms2During2s	3ms2During0.5s	.43190*	.01794	.000	.3890	.4748
		5ms2During0.3s	.79019*	.11777	.000	.5082	1.0722
	3ms2During0.5s	1ms2During2s	-.43190*	.01794	.000	-.4748	-.3890
		5ms2During0.3s	.35829*	.11888	.008	.0737	.6429
	5ms2During0.3s	1ms2During2s	-.79019*	.11777	.000	-1.0722	-.5082
		3ms2During0.5s	-.35829*	.11888	.008	-.6429	-.0737
IntersectionWithSpeedNull5sAfter	1ms2During2s	3ms2During0.5s	.48304*	.01404	.000	.4494	.5167
		5ms2During0.3s	.42708*	.07543	.000	.2465	.6077
	3ms2During0.5s	1ms2During2s	-.48304*	.01404	.000	-.5167	-.4494
		5ms2During0.3s	-.05596	.07649	1.000	-.2391	.1272
	5ms2During0.3s	1ms2During2s	-.42708*	.07543	.000	-.6077	-.2465
		3ms2During0.5s	.05596	.07649	1.000	-.1272	.2391
Intersection	1ms2During2s	3ms2During0.5s	.33028*	.01073	.000	.3046	.3560
		5ms2During0.3s	.50504*	.06963	.000	.3383	.6717
	3ms2During0.5s	1ms2During2s	-.33028*	.01073	.000	-.3560	-.3046
		5ms2During0.3s	.17475*	.07030	.039	.0064	.3431
	5ms2During0.3s	1ms2During2s	-.50504*	.06963	.000	-.6717	-.3383
		3ms2During0.5s	-.17475*	.07030	.039	-.3431	-.0064
SpeedReduction	1ms2During2s	3ms2During0.5s	.19595*	.03415	.000	.1142	.2777
		5ms2During0.3s	.43132	.18166	.053	-.0037	.8663
	3ms2During0.5s	1ms2During2s	-.19595*	.03415	.000	-.2777	-.1142
		5ms2During0.3s	.23537	.18435	.605	-.2060	.6768
	5ms2During0.3s	1ms2During2s	-.43132	.18166	.053	-.8663	.0037
		3ms2During0.5s	-.23537	.18435	.605	-.6768	.2060
Curve	1ms2During2s	3ms2During0.5s	.40455*	.01200	.000	.3758	.4333
		5ms2During0.3s	.63973*	.06958	.000	.4731	.8063
	3ms2During0.5s	1ms2During2s	-.40455*	.01200	.000	-.4333	-.3758
		5ms2During0.3s	.23518*	.07045	.003	.0665	.4038
	5ms2During0.3s	1ms2During2s	-.63973*	.06958	.000	-.8063	-.4731
		3ms2During0.5s	-.23518*	.07045	.003	-.4038	-.0665
NoCurve	1ms2During2s	3ms2During0.5s	.45860*	.02576	.000	.3969	.5203
		5ms2During0.3s	.61626*	.12825	.000	.3092	.9233
	3ms2During0.5s	1ms2During2s	-.45860*	.02576	.000	-.5203	-.3969
		5ms2During0.3s	.15766	.13037	.680	-.1545	.4698
	5ms2During0.3s	1ms2During2s	-.61626*	.12825	.000	-.9233	-.3092
		3ms2During0.5s	-.15766	.13037	.680	-.4698	.1545

*. The mean difference is significant at the 0.05 level.