PHD

Investigating the impact of naturalistic driving behaviour differences on energy consumption and road safety

Malek, Sahand

Award date:
2017

Awarding institution:
University of Bath

Investigating the impact of naturalistic driving behaviour differences on energy consumption and road safety

Sahand Malek

A thesis submitted for the degree of Doctor of Philosophy

University of Bath

Department of Mechanical Engineering

May 2017

To my incredible father, mother and brother, Alireza, Parisa and Sepand.

Table of Contents

List of Tables

List of Figures

List of Equations

Acknowledgements

This thesis was written between October 2015 and March 2016. The 80th anniversary of the compulsory driving test in Great Britain was celebrated on 2015. Moreover, this year is the halfway milestone of the Decade of Action for Road Safety (2011-2020) that was introduced by the United Nations. The important work of those who have dedicated their lives to improving the safety of road users globally has been an inspiration for me to contribute to this field. The aim of this thesis is to motivate all those who care about the safety of road users and researchers working in this area. It is dedicated to the memory of people injured and to those who have lost their lives in traffic accidents.

First and foremost, I would like to express my deepest gratitude to Professor Chris Brace for his exceptional guidance, for sharing his insight and expertise, and for being an excellent supervisor during my time as a postgraduate researcher at the Powertrain and Vehicle Research Centre (PRVC), Department of Mechanical engineering, University of Bath.

Special thanks go to two industry collaborators, the Castel Company and the IPG automotive Company, for their continued support and for providing an opportunity to make conducting this research possible. I would like to thank everyone from the University of Bath and beyond, who supported me in organising and hosting the Eco Safe Driving Challenge events, in order to complete this study. This includes the Researcher Development Unit team, Public Engagement team, Department of Marketing & Communications, Department of Security Services, drivers, and event volunteers.

Finally, my sincere appreciation to my entire family, my parents, my brother, and my cousins, for their never-ending support, understanding, and laughter.

Sahand Malek | 2016

Abstract

The research focus is on two major current challenges facing fleet operators and motor insurance providers globally. For fleet operators, staying profitable means reducing their running cost by increasing their operation efficiency and reducing their running costs by reducing driver fuel consumption and avoiding costly road accidents. For motor insurance providers, having a profitable business requires sale premiums that reflect real-world exposure to risks. This means moving towards metrics beyond traditional pricing parameters, such as age, occupation and gender. Regarding this, there is now substantial interest in understanding drivers driving behaviour, since knowing which ways of driving are resulting in using more fuel and/or being exposed to road accidents can bring huge financial benefits to both the aforementioned industries.

To identify driving behaviours that are costly (both in terms of fuel usage and involvement in road crashes), a naturalistic driving behaviour field study, called Eco Safe Driving Challenge, was developed between 2013and 2015 at University of Bath. The project followed the same format of major real-world driving behaviour test such as 100-cars in the US, PROLOGUE and UDRIVE in the EU by following the FESTA-V approach in designing the experiment, data collection, data management, ethical and legal concerns, data analysis and evaluating the findings. In total, 250 km worth of driving data was collected from nine drivers aged between 25 – 30 using medium size petrol cars. A special route was designed to mimic specific road settings, including downhill driving, uphill driving, traffic lights, a pedestrian crossing and a roundabout. It comprised a 4 km loop (same start and end points), starting and finishing at the University of Bath, with an 11% increase and decrease in road slope. The data were collected

using OBDII dongles fitted with GPS sensors to record location and a sim card for sending driving data to a dedicated server in real-time.

A strategic framework has been developed to analyse the data in two domains, i.e. eco-driving and driving and in three phases. First, identifying specific driving behaviour, secondly classifying and comparing drivers differences and finally, scoring, ranking and model drivers' driving behaviour performances with the aim of assessing eco-driving and safe driving behaviour impacts.

The key contributions of the work are, firstly it has found that drivers' with a tendency to misuse gears, use excessive engine power and/or frequently speed are less fuel efficient than their counterparts To validate this claim, a metric called Vehicle Specific Power – Fuel Consumption (VSP – FC) has been developed, which shows that eco-drivers, on average, have 1.0 – 1.2 points higher than others according to this measure. When evaluating safe driving behaviours, it emerged that there was 75% correlation between historical crash zones (based on public records) and locations (400 metre long road segments) where the nine participants in the driving event undertook harsh breaking and acceleration. This provides evidence that scoring drivers based on the number of harsh braking and acceleration events should be included in metrics aimed at evaluating driving behaviour.

Preface

Evidently, the driver of a vehicle is a major contributor to how efficiently and safely it is controlled. In the case of fuel-based vehicles, efficiency refers to using fuel sparingly for completing a journey, and this helps to reduce vehicle emissions, such as CO_2, thereby contributing to restricting a person's carbon footprint. In terms of safety, there are driving behaviours that are the result of a propensity to drive unsafely and carelessly. This research is aimed at identifying and classifying drivers' behavioural differences that lead to inefficient and unsafe driving.

Studying naturalistic driving behaviour under real-world conditions has gained much attention in recent years. This can provide valuable information to fleet managers and usage based insurance providers about individual driver's habits, styles, and driving patterns. For this research, the principals of conducting naturalistic driving behaviour study were applied, whereby information regarding drivers' driving behaviour in natural driving conditions (familiar roads and car) was gathered unobtrusively. Over the course of three driving events, 240 km worth of driving data were collected from 15 drivers. However, for this thesis, only the data from the final round of the driving event with nine drivers have been analysed to deliver the findings. Specifically, in this final round on 9th February 2014, there were three Vauxhall Corsa drivers, four Nissan Note drivers, one Nissan Juke driver and one Fiat 500L driver. The monitoring devices used in the study were a GPS sensor wired to an OBD dongle with the ability to transmit driving data via a mobile network. In order to analyse the driving data, two categories of methods were employed.

First, there were those aimed at identifying driving variations and second, techniques for classifying (ranking) drivers' differences. The driving data, including vehicle road speed, engine speed, total number of harsh accelerations/decelerations and coasting behaviour were used to identify and classify (rank) eco-drivers and safe drivers.

Here are a number of findings of this project. Firstly, it is shown that drivers who used excessive engine power and frequently sped were less fuel-efficient, which was ascertained by examining drivers' speeding patterns and their reported total fuel consumption amount. Secondly, by calculating the Vehicle Specific Power (VSP) value for all the drivers' trips it was demonstrated that this value allows for meaningful comparisons across the participating drivers. Drivers who tended to break the speed limit frequently and had been scored poorly, based on the fleet management deducting points scoring method, according to the total number of counted alarming events (harsh acceleration and harsh deceleration), were shown to use excessive and unnecessary vehicle power to complete the driving task. Subsequently, the Vehicle Specific Power – Fuel Consumption (VSP – FC) value was introduced as a metric to compare and rank drivers' fuel consumption in regard to road gradient and their vehicle velocity, acceleration and drag coefficient. The results show that drivers with a low VSP – FC value were those with a high number of harsh accelerations / decelerations and they tended to drive with a higher range of vehicle and engine speeds. The final major contribution is the method that is proposed to associate similar drivers based on driver's vehicle speed profiles. It provides a statistical solution for grouping drivers with similar vehicle speed distribution based on the effective sample size value between every pair of drivers in the sample. This value is then used as an indicator of the level of association (similarity of speed distribution) between every pair of drivers, whereby any two with a high sample size value between them are deemed to be driving in a similar manner. The method is shown to be effective, since similarly classed drivers are demonstrated as sharing other similar characteristics, such as driving scores and driving patterns.

Regarding the safety aspects of driving, it is important to highlight that by conducting correlation analysis between the location where drivers made sharp acceleration and deceleration, it has been shown that in 75% of occasion these locations match with segments on the route where previous accidents were recorded. This becomes numerical proof that scoring driving behaviour based on excessive acceleration and braking can be an effective indicator of drivers' differences and it can be used to score drivers attitudes towards safety. Later, drivers are ranked based on the total number of times they made harsh speeding or braking at locations with previous accident records and those without them.

As part of industry collaboration with IPG Automotive, virtual simulation software named **Carmaker** is used, firstly, to demonstrate the capability of the software package to simulate various driving behaviours and secondly, two simulation scenarios are designed, modelled and simulated with the help of the software. In the first study, drivers' driving efficiency has been virtually simulated by using their actual driving speed on virtual driving simulation software. Under the simulation conditions, it has been shown that drivers who scored poorly by the method used by usage-based insurers and fleet management service providers were not necessarily those who drove inefficiently. The result of a study on the impact of road gradient on fuel usage showed that a 16% increase on a computerised road cause on average 40% increase in fuel consumption (based on simulation test conditions).

Abbreviations

3D	A three-dimensional object
3G	Third generation of mobile telecommunications technology
ACEA	European Automobile Manufacturers' Association
ADAS	Advanced driver assistance system
AECC	Association for Emissions Control by Catalyst
AEMP	The Association of Equipment Management Professionals
AFR	Air-Fuel Ration
Artemis	Assessment and Reliability of Transport Emission Models and Inventory Systems
BSFC	Brake Specific Fuel Consumption
CAS	Crash Avoidance Systems
CMEM	Comprehensive Modal Emissions Model
DAS	In-vehicle Data Acquisition System
DV	Dependent Variable
ECU	Engine Control Unit
EEA	European Environment Agency
EEA	The European Economic Area
EOBD	European On-Board Diagnostics
ESDC	The Eco-Safe Driving Challenge
FESTA	Field operational test support action
FOT	Field Operational Tests
GHG	Greenhouse Gas
GLM	The Generalised Linear Model
GPS	Global Positioning System
GSM	Global System for Mobile communication
HMMs	Hidden Markov Models

HSID	Hotspot Identification
IC engine	Internal-Combustion Engine
ICCT	The International Council on Clean Transportation
IEA	International Energy Agency
ITS	Intelligent Transport Systems
IV	Independent Variable
JB test	The Jarque-Bera Normality Test
KML	Keyhole Markup Language
LAMDA	Learning Algorithm for Multivariate Data Analysis
LDCV	Light Duty Commercial Vehicle
MAF	Mass Air Flow
mini MPV	Mini Multi-Purpose Vehicles
MLR	Multivariable Linear Regression model
MOVES	Motor Vehicle and Equipment Emission System software
MPG	Miles per Gallon
NEDC	New European Driving Cycle
NFDA	The National Franchised Dealers Association
NHTSA	US National Highway Traffic Safety Administration
OBD	On-Board Diagnostic
OGL	Open Government Licence
PAYD	Pay as You Drive Plan
PEMS	Portable Emissions Measurement System
PHYD	Pay How You Drive
PREVIEW	Portable Real-Time Emissions Vehicle Integrated Engineering Workstation
PROLOGUE	Promoting real Life Observations for Gaining Understanding of road user behaviour in Europe
RPM	Revolutions per Minute
RTA	Road Traffic Act

SA	Selective Availability
SAE	Society of Automotive Engineers
SARTRE	Social Attitudes to Road Traffic Risk in Europe
SMP	Sustainable Mobility Project
SWOT	Strengths, Weaknesses, Opportunities, and Threats
SWOV	Netherlands Institute for Road Safety Research
TRL	Transport Research Laboratory
TRRL	Transport and Road Research Laboratory
TWC	Three-Way Catalyst
UDC	Urban Driving Cycle
UTC	Universal Time Coordinated
VISSIM	Microscopic Traffic Model software package
VITO	Flemish Institute for Technological Research
VOEM	VITO on- the-road emission and energy measurement system
VSP	Vehicle Specific Power
VSP-SFC	The Vehicle Specific Power –Specific Fuel Consumption
VTI	Swedish Road and Traffic Research Institute
WBCSD	World Business Council for Sustainable Development
WHO	World Health Organisation
WLTC	Worldwide Harmonised Light Duty Driving Test Cycle
WLTP	The Worldwide Harmonised Light Vehicles Test Procedures
WMW	Wilcoxon-Mann-Whitney U test method

Chapter 1

Introduction

"Human behaviour flows from three main sources: desire, emotion, and knowledge."

– Plato

1.1 Impact of road transport on air pollution and the level of CO_2

The decline in available resources of fossil fuels and increasing concerns about global warming and air pollution have prompted many states, institutions and individuals to take effective action. States are introducing new laws and measures to reduce carbon footprints; institutions are conducting research on a variety of topics to help solve current environmental, and urban challenges and individuals who are concerned about these issues are trying to change their way of life so as to reduce their negative contribution.

In Europe, amongst the sectors producing greenhouse gas (GHG hereafter) emissions in 2012, the energy industry and transport were the highest polluters. The transportation sector alone generated 24.3% of all the GHG producers in Europe (Figure 1). Based on a European Commission Climate Action Report, while the GHG emissions for many areas have been falling since 1990, the amount generated by the transport sector actually increased (European Commission, 2014).

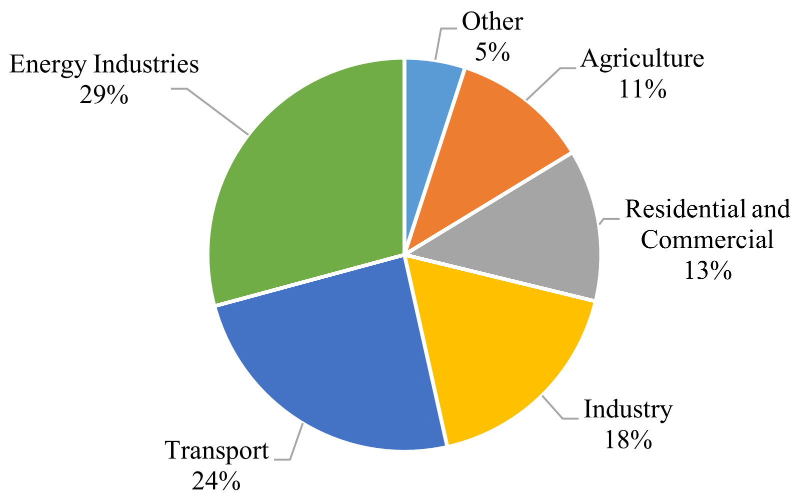


Figure 1. EU greenhouse gas emissions by sector, 2012 (European Commission, 2014)

Specifically, as the historical record of producing GHG emissions from various sectors (agriculture, energy industries, industry, and transport, residential and commercial) has revealed, after 1990 the amount of harmful gases generated by all except transport has continued to decrease. The GHG emissions of all the sectors except transport dropped by around 15% between 1990 and 2007 (Figure 2). However, for the transport sector, there was a 36% increase during the same period. After 2007, due to the wide range of actions that were taken by EU states including the introduction of low emission policies, low emission vehicles, and the promotion of green technologies, the transport sector GHG level decreased remarkably. However, by 2012 it was still 20.5 % higher than the 1990 record (Figure 2) (European Commission, 2014).

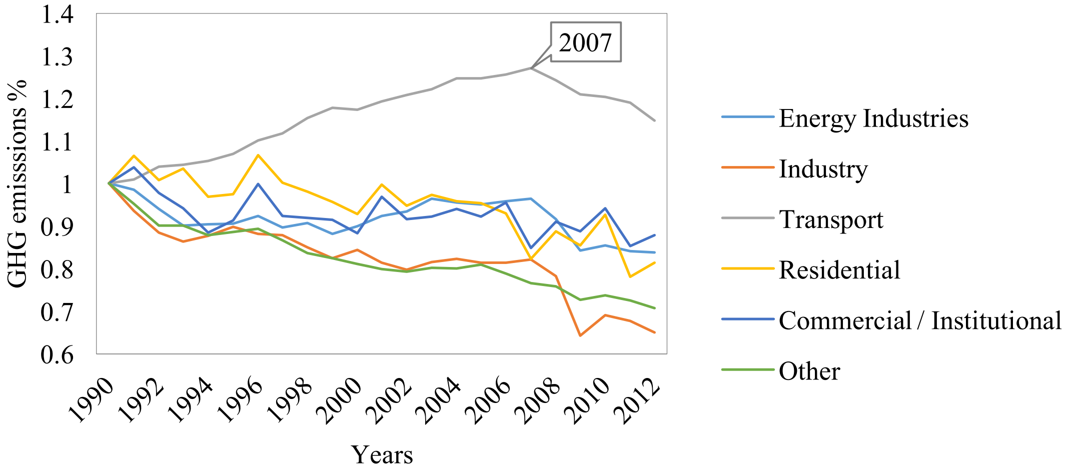


Figure 2. EU percentage of greenhouse gas emissions based on sector[1]

Returning to the 2012 record, among the various forms of transport, the annual increase in personal, and freight transportation means that despite improvements made in vehicle efficiency, the transport sector contribution was the highest compared to other areas. Amongst the transport categories, road transport contributed the most in producing harmful gases. The data show that over 70% of all GHGs were generated by road transport in 2012. In sum, according to the 2012 findings, while other forms of transport such as railways had a very low impact in generating GHGs, combustion based vehicles were major producers of GHGs in EU states (European Commission, 2014). According to the EU 2011 Transport White Paper target, to reduce the level of GHGs to 60% of the 1990 figure by 2050, the transport sector needs to lower its emission levels 67% by 2050 (European Commission, 2014).

[1] "Energy industry sector excludes land use, land – use change and forestry emissions and international bunkers. Transport sector excludes international bunkers (international traffic departing from the EU). Industry sector includes emissions from manufacturing, and construction and industrial processes. Other includes emissions from fuel combustion in agriculture/ forestry/ fisheries, other (not elsewhere specified), fugitive emissions from fuels, solvents and other product use and waste" (European Commission, 2014).

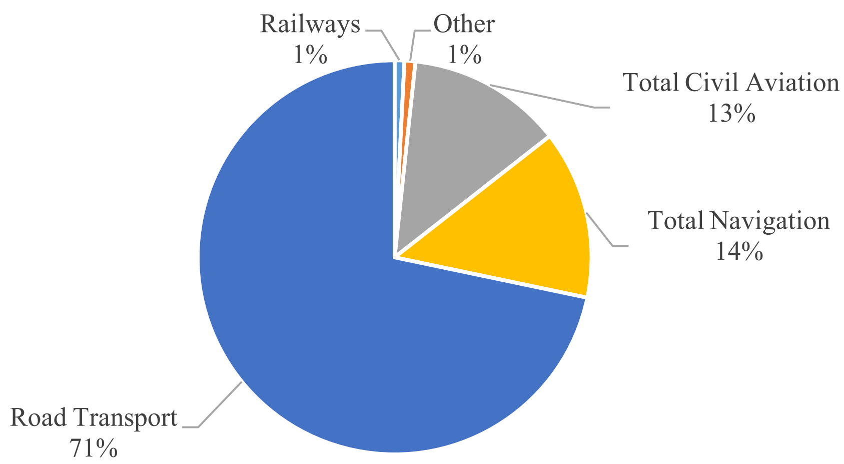


Figure 3. EU greenhouse gas emissions by mode of transport, 2012

The increasing number of passenger cars in the EU has had its consequences. Using a passenger vehicle as a primary means of road transport has exposed EU countries to environmental, and urban traffic management challenges. However, the economic crisis between 2007 and 2009 had an impact on the total number of new passenger cars registered in Europe, which in 2014 was 20% less than before the crisis (The International Council on Clean Transportation Europe, 2015). The EU automotive industry's resurgence during the last three years and significant growth has led to it representing one-third of all EU manufacturing jobs. Owing to this increase, recently there has been an increasing trend of vehicles being introduced to the EU road networks. In fact, in 2014, around 75% of the 12.5 million new passenger cars registered in the EU belonged to the EU's largest markets, which are Germany, France, the United Kingdom, Italy, and Spain. Figure 4 presents the registration of passenger cars in the EU. Data between 2001 and 2007 are from EU-25 only, whereas from 2007 to 2014 there were EU-28 member states (The International Council on Clean Transportation Europe, 2015). Hence, this means a continuing growth in the production of GHG emissions, specifically CO_2.

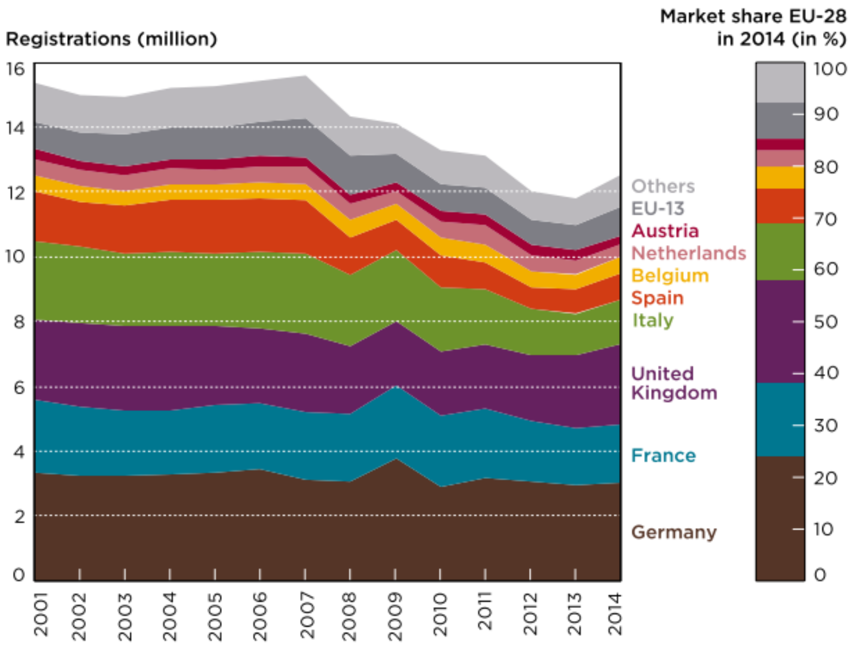


Figure 4. Passenger car registrations by EU member states (ACEA, 2014)

The fuel consumption and CO_2 emissions target for new passenger cars in the EU is legislated by the European Environment Agency (EEA). The EU legislation on such emissions drawn up in 2009 has proven an effective solution, as by 2013, for new passenger cars these had already fallen below the target set for 2015. By meeting the 2015 target in 2013, the annual rate of reducing CO_2 emissions for new vehicles increased to 4% per year. The European Commission target for average CO_2 emissions of new passenger vehicles is 95 g/Km of CO_2 or the equivalent of 3.8 litres/100 km of fuel consumption. Figure 5 presents the EU's CO_2 emission results and targets for new passenger cars (The International Council on Clean Transportation, 2015).

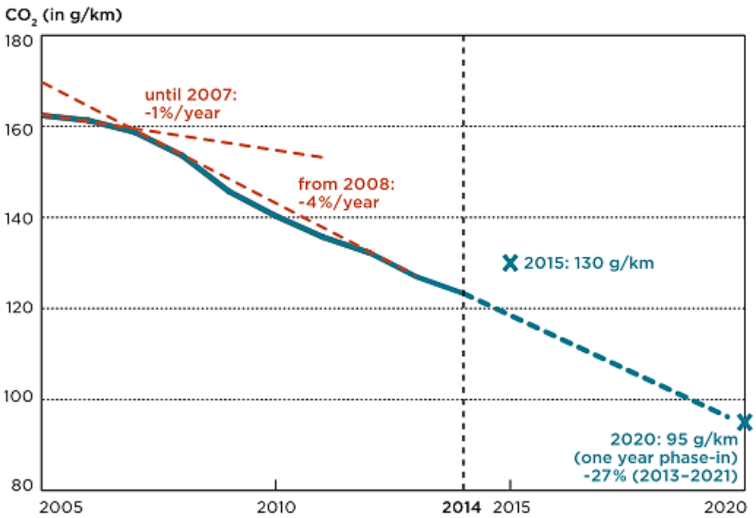


Figure 5. Historical record and future targets for CO_2 emissions levels of new passenger cars in the EU (The International Council on Clean Transportation Europe, 2015)

Whilst the trend in carbon dioxide emissions has been downwards for new vehicles, countries like Germany and England experiencing high sales of new cars, thereby putting pressure on aggregate emissions, which could explain why their levels are still above the EU-28 average. With the set target of 2015 already having been met in 2013 (Figure 5) as mentioned above, EU states and subsequently car manufacturers are committed to reducing their CO_2 emissions between 2015 and 2020 by 27% (The International Council on Clean Transportation, 2013, 2015).

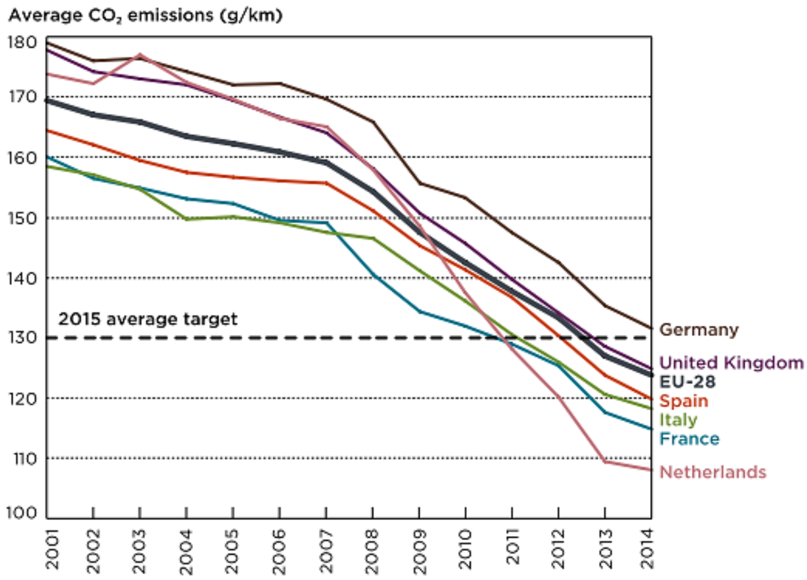


Figure 6. Passenger car CO_2 emissions by member state

The United Kingdom has taken measures over the last two decades to reduce its carbon footprint and its GHG contribution. In 2013, the sectors with the highest contribution of GHG emissions were the energy industries sector and the transport sector, with 33% and 21% respectively (Department for Energy and Climate Change, 2015). According to the Department for Energy and Climate Change report, by 2012, there had been only a 2% reduction in greenhouse gas emissions from the transport sector since 1990. In 2013, there was another 1% decrease in the level of GHG emissions (Figure 7). Road transport and particularly passenger cars are a significant source of these gas emissions in the United Kingdom (Department for Energy and Climate Change, 2015).

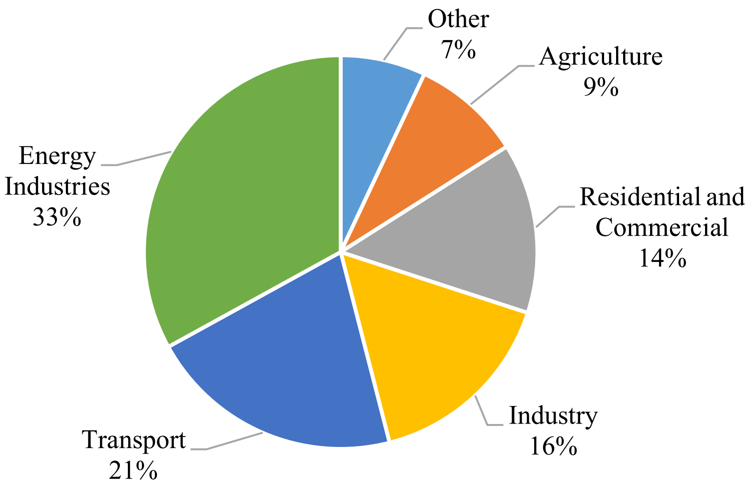


Figure 7. UK greenhouse gas emissions by sectors, 2013

When the GHG emissions are broken down into individual gases, for the UK transport sector, it shows that almost all of the emissions come in the form of carbon dioxide. The Department for Energy and Climate Change confirmed that the level of CO_2 emissions is closely related to the amount of fuel consumed rather than vehicle age or type. Other GHGs, such as nitrous oxide and methane emissions, are closely linked with vehicle age and type. Table 1 presents the GHG emissions of the transport sector by different gases. These data establish the fact that the amount of fuel used by passenger cars has a direct effect on the UK GHG levels since CO_2 is the primary source of GHG emissions of the transport sector.

Table 1. UK transport sector GHG emissions breakdown by gases[2] (1990 - 2013)

	1990	1995	2000	2005	2010	2012	2013
Carbon dioxide	119.7	119.9	124.9	129.3	119.3	116.9	115.7
Methane	0.8	0.6	0.4	0.2	0.1	0.1	0.1
Nitrous oxide	1.3	11.7	1.5	1.2	0.9	1.0	1.0
F-gases	0.0	0.0	0.0	0.0	0.0	0.0	0.0

[2] Millions of tonnes of carbon dioxide equivalent (Mt CO_2e).

The European Commission has proposed measures to be taken by 2020 regarding CO_2 emissions and a fuel consumption reduction target. The action plan includes strategies to reduce emissions from all forms of road transport including cars, vans, and light- and heavy-duty vehicles. Further plans are putting limitations on tyre rolling resistance, passing legislation to label tyres, enforcing mandatory tyre pressure monitors on new vehicles, providing government grants towards the cost of purchasing electric vehicles by EU member states and finally, introducing a new target for a weight-based target system in the EU so as to reduce new vehicle masses (The International Council on Clean Transportation, 2013).

1.2 Road safety facts and figures

The issue of drivers and other road users' safety has been a concern of every country worldwide. In 2010, the United Nations adopted resolution 64/2551, which initiated a global campaign calling for *A Decade of Action for Road Safety*. This will end in 2020, by which time it is estimated that more than five million lives will have been saved. According to a joint WHO and UN report (2013), in 2010 up to 1.24 million people lost their lives due to road accidents, and as many as 50 million suffered non-fatal injuries. This figure includes drivers, passengers, cyclist, and pedestrians. The WHO has estimated that without taking significant action, the fatality rate due to road accidents would increase by about 65% by 2024. These estimates also highlight the fact that among the six global regions the European region has the lowest death rate due to road accidents with 10.3 per 100,000 population. In contrast, the African region has the highest rate of fatalities, standing at 24.1 per 100,000, because of poor road infrastructure (Figure 8) (World Health Organisation, 2013).

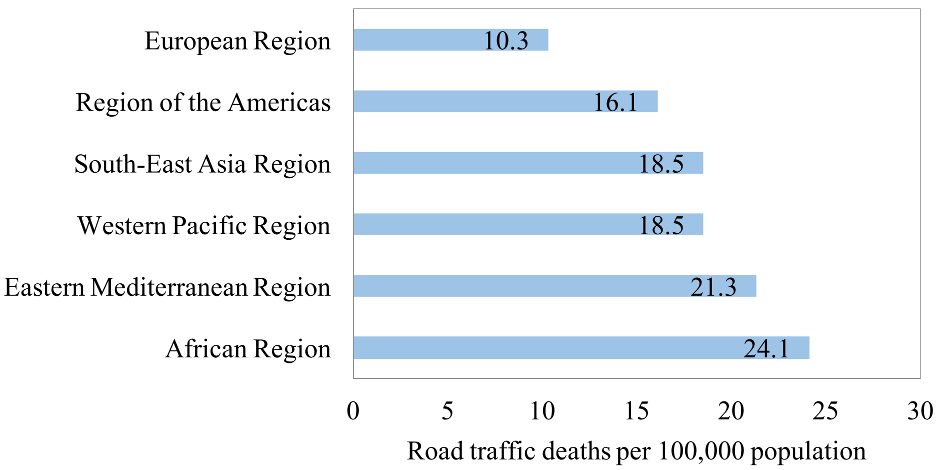


Figure 8. Global road traffic deaths per 100,000 population[3]

[3] Regions are based on the WHO road safety report, 2010.

The total percentage of road fatalities by road users for the five EU-28 member states with the largest market share of new passenger cars, according to this report, shows that, except for Italy, the other four countries have a higher proportion of fatalities among vehicle users than the global average when compared to all road users. The United Kingdom, after Italy, had the lowest percentage vehicle user casualties. In contrast, France, and Spain had the highest proportion of road casualties attributed to vehicle users amongst the five countries in Figure 9 (World Health Organisation, 2013). This figure highlights the vulnerability of drivers and passengers of vehicles compared to other road users' safety.

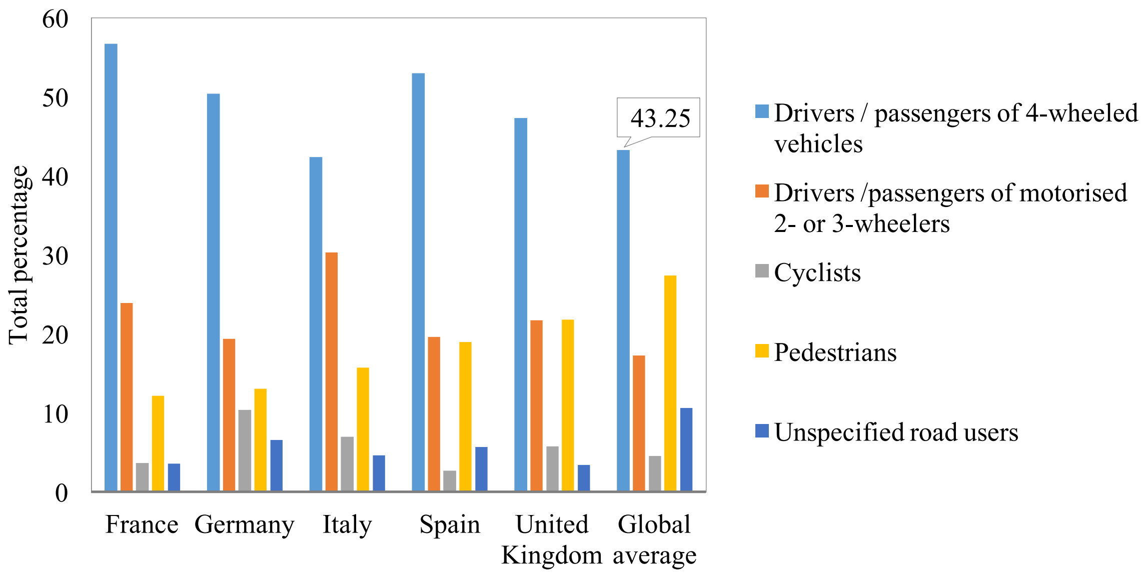


Figure 9. Distribution of road traffic deaths by type of road user as a percentage, 2010

Increasing road safety requires a scientific evidence-based approach. This is usually tackled by studying the driving habits of travellers and identifying hazardous blackspots according to historical collision data. Traveller patterns, including methods of transport and distribution of travel on each road type, are gathered through national surveys and by using on-road traffic flow sensors. In order to identify collision-prone locations, it is important to have collision records. In the UK, these are gathered by police departments in every region. These records are published by the Department for Transport under Open Government Licence (OGL) annually. Previous collision records are commonly investigated at the following levels: on a regional basis, on a number of road networks, regarding the segments of a route, based on road types or in the vicinity of specific points of interest (school, hospital, etc.).

According to the National Travel Survey conducted in 2013, in England, the main mode of travel for social and commuting purposes is by car, which confirms the importance of increasing vehicle users' safety (Department for Transport, 2014). According to the survey, over 65% of all trips in England for the purpose of shopping or commuting are by car. Figure 10 shows the percentage of all trips made in 2013 according to the objective of the journey and method of transport. From the same survey, it has been concluded that up to 70% of all trips to and from work are made by car. The increasing volume of traffic on Britain's roads, as most people prefer to use the vehicle and the fact that the population is rising, is threatening to cause the number of injuries and deaths to start growing again.

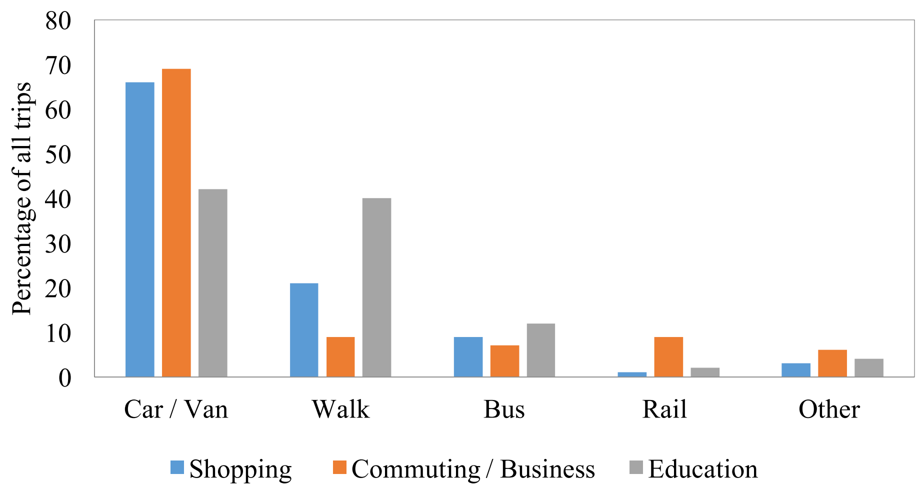


Figure 10. Percentage of trips in England by method of transport, 2013

According to EuroRAP report, the average risk of death or serious injury is 26.7 per billion vehicle km in the UK's ten regions (EuroRAP, 2015). As Figure 11 shows, the risk of death or serious injury on the Wales road networks is the highest (30 per billion vehicle km) and in the West Midlands road networks the lowest (17 per billion vehicle km), while the risk of death or serious injury in the South West is around the average (27 per billion vehicle km).

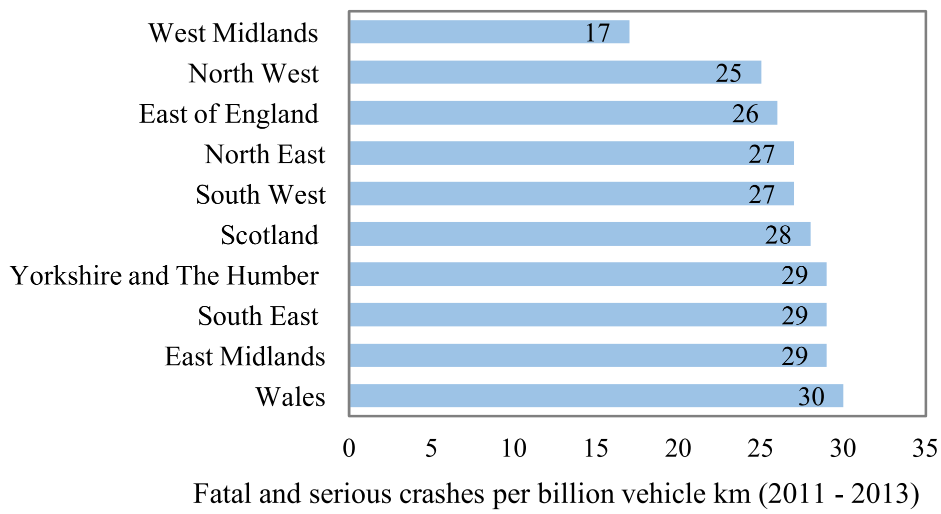


Figure 11. Average risk of death or serious injury for all road networks in the UK by region[4]

Since this study is carried out in Bath, UK, this subsection provides facts and figure about this region. The unitary authority of Bath and North-East Somerset (BANES) is located in the South West of the United Kingdom, the population of which in 2014 was estimated at around 182,000 people. The collision records for this district are provided in Figure 12 for the period of 2005 to 2013.The numbers within the coloured marks show the total number of collisions recorded over the period in question. As can be seen, Bath had by far the highest number of collisions when compared to surrounding towns and villages.

[4] Between 2011 – 2013 based on (EuroRAP, 2015)

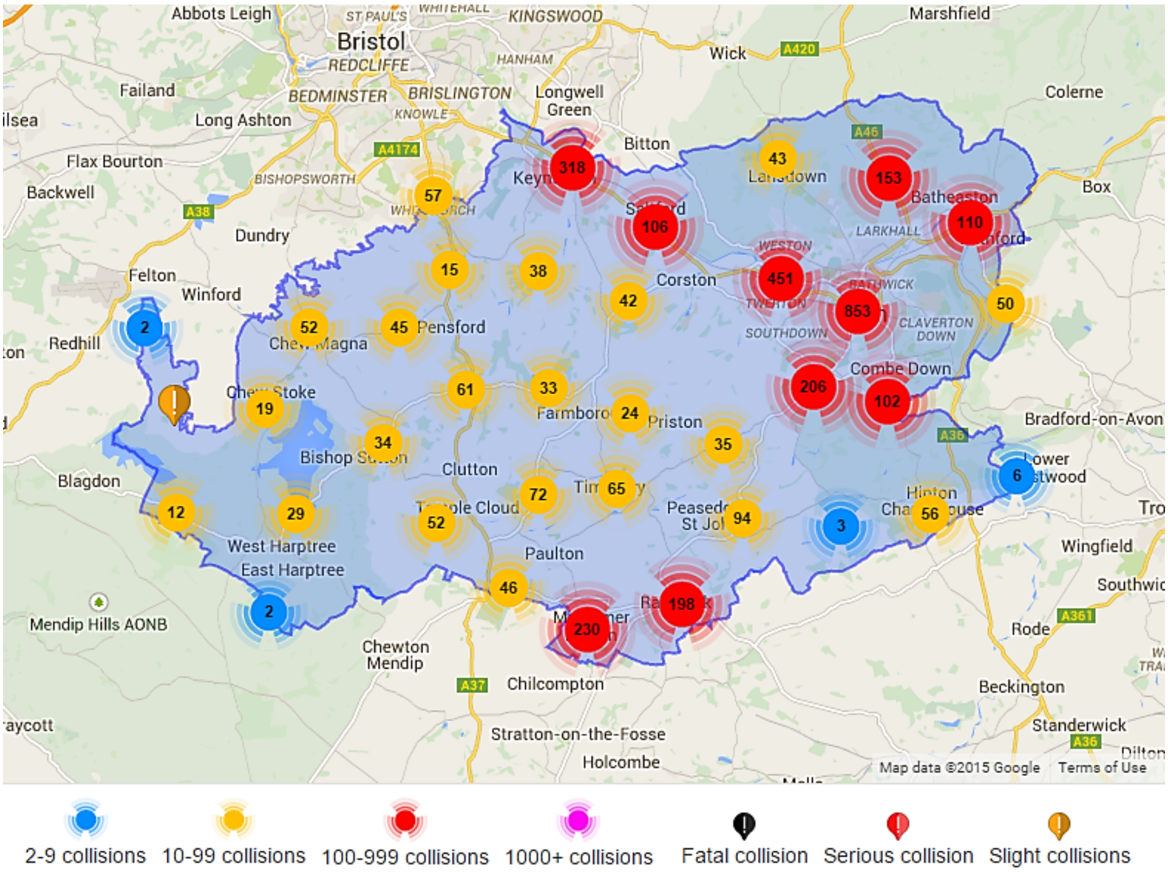


Figure 12. The unitary authority Bath and North East Somerset historical road collisions (2005 – 2013)

The historical collision records published by the Department for Transport show that from 2005 to 2013 the total number of collisions, based on the severity of the crash, decreased steadily nationally. After 2005, the total number collisions with both fatal and slight severity decreased dramatically. That is while total number of collisions with a serious level of damage remained at roughly constant level nationally. The following Table 2 from the aforementioned report compares the trend of three classes of collisions: fatal, serious, and slight between Bath and North East Somerset (BANES) and nationally. Trends are also shown in a plot format for the same periods.

Table 2. The historical collision data for Bath and North-East Somerset (BANES) and the UK

	Collision severity	2005	2006	2007	2008	2009	2010	2011	2012	2013
Unitary authority	Fatal	10	10	5	10	4	9	4	4	6
	Serious	61	54	46	24	39	30	18	29	46
	Slight	634	556	549	506	521	452	469	407	363
National	Fatal	3,201	3,172	2,946	2,538	2,222	1,901	1,857	1,754	1,713
	Serious	29,000	29,000	28,000	26,000	24,690	23,122	20,803	23,039	21,657
	Slight	238,862	226,559	217,060	202,333	195,234	178,927	185,995	170,930	160,300

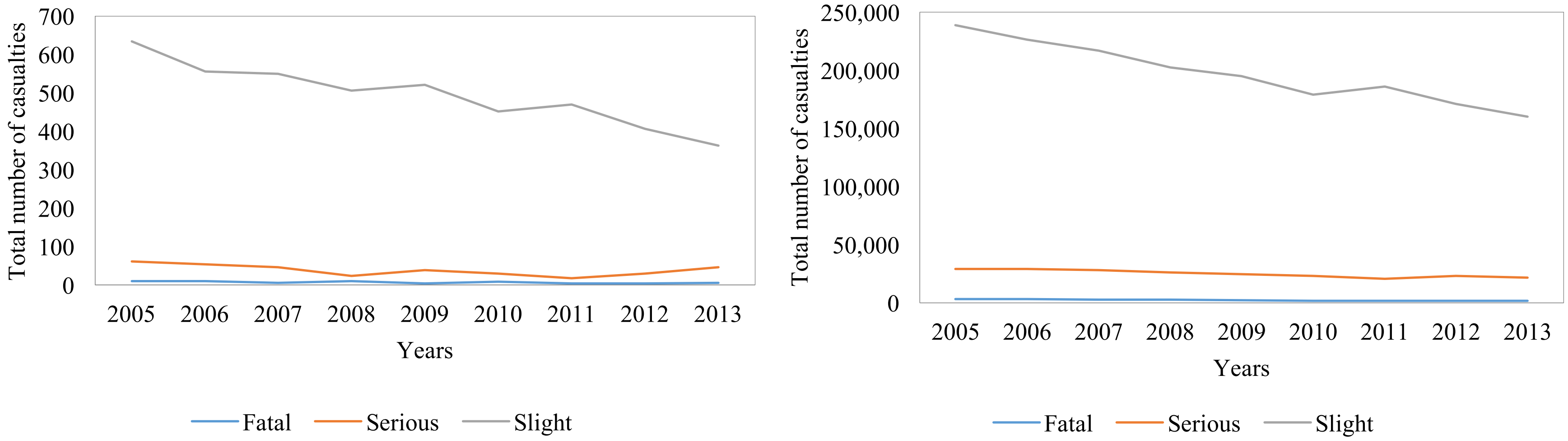


Figure 13. Trends of collision severity (2005 – 2013): BANES (bottom left); nationally (bottom right)

1.3 Motivation

The underlying motivation for this research is that by collecting real-world driving data the following gaps, as found in the relevant studies, will be addressed:

1. Understanding and characterising the impact of real world driving on fuel economy;
2. Identifying the contributing factors of drivers' driving behaviour to road safety;
3. Developing potential analytics approaches for ranking drivers and the impact of drivers' driving behaviour differences;
4. Designing a low-cost project of short duration to investigate drivers' behaviour according to naturalistic driving behaviour studies methodology.

Despite an extensive body of scientific research already having been published in this domain, there are still research gaps that remain to be fulfilled. Whilst some improvements have been made in some respects, such as improving cars' aerodynamics features, less attention has been paid to the impact of external factors, such as road geometry or drivers' driving characteristics (Redsell, Lucas and Ashford, 1993; Saffarzadeh and Arjroody, 2003; van Basshuysen and Schäfer, 2008; Schipper, 2011). There are four areas that have received attention in respect of reducing fuel consumption, which are:

- Modification of the physical properties of the vehicle, such as reducing air resistance, optimising vehicle weight, and minimising wheel resistance;
- Engine modification methods, such as engine downsizing, variable valve timing, ignition improvements, and cylinder shutoff modification;
- Modification of the vehicle powertrain, including selecting and arranging transmission ratios that deliver desirable speeds and vehicle power;

- Reducing the impact of external factors, such as road geometry, traffic flow, environmental parameters, and drivers' driving behaviour, which will all help to reduce fuel consumption;

In relation to the above, improvements have been made regarding many factors related to vehicle body design as well as engines and powertrain characteristics. However, the major factors affecting fuel consumption of current vehicles on roads are beyond the control of car manufacturers (Casadei, Broda and Ricardo Inc., 2008; van Basshuysen and Schäfer, 2008; The International Council on Clean Transportation, 2013). Factors such as driving style, vehicle load, weather, driving terrain, and maintenance are those that are not fully understood and they not been characterised accurately in detail.

In order to put the effect of the factors relating to fuel consumption into perspective, the proportional effect of a range of factors has been collated from the following studies: (Redsell, Lucas and Ashford, 1993; McTavish and BP Castrol Technology Centre, 2008; The Goodyear Tire & Rubber Company, 2012; Detroit Diesel Corporation - Demand Detroit, 2015). After investigating these factors, it is concluded that nine factors, including driving style, weather, road terrain, gearing, aerodynamics of the body, vehicle maintenance, vehicle load, tyres and lubrication are the most influential external factors impacting on vehicle fuel consumption. Individual estimations from the aforementioned studies, as appropriate, have been combined and averaged to generate the following pie chart (Figure 14).

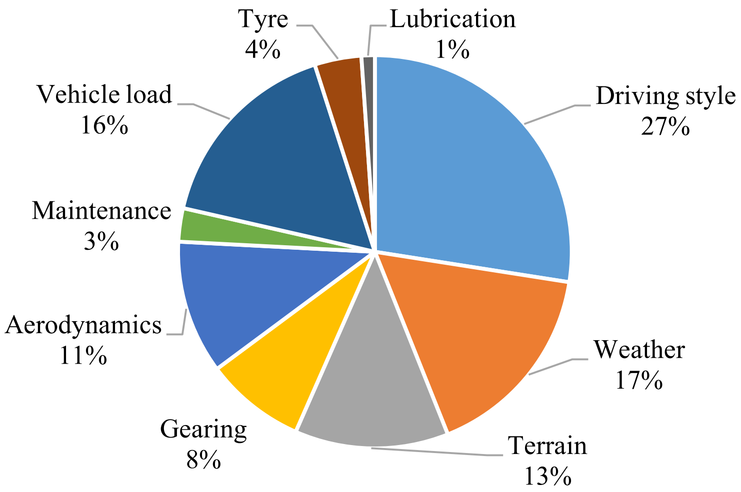


Figure 14. Factors affecting fuel consumption efficiency excluding vehicle load

The driver, as the operator of a vehicle, has a major influence on the amount of fuel consumed while driving. Journal papers and technical reports from academia and industry claim that different driving styles can cause anything from a 3% and 18% increase in fuel consumption (Mierlo *et al.*, 2004; Felstead, McDonald and Fowkes, 2009; Berry, 2010; Mujanovic, 2011; Malek, Brace and Liu, 2012; The Goodyear Tire & Rubber Company, 2012). Hence, it is important to identify which driving behaviours can negatively affect fuel economy. A driver can negatively influence the level of fuel consumption in two ways: firstly, by having inefficient control of a vehicle, exhibiting such as speeding, excessive usage of vehicle power, sudden acceleration/deceleration (braking), and using the wrong gears. Secondly, failing to comply with appropriate vehicle maintenance and vehicle usage, such as not having regular servicing and overloading, can also impact detrimentally on fuel consumption (Fontaras, Zacharof and Ciuffo, 2017). According to Fontaras, Zacharof and Ciuffo (2017), neglecting regular maintenance, a wheel misalignment of 2 mm can potentially increase energy losses by 3%.

The road safety report by the WHO (2013) concluded that four major factors influence the safety of road users: excessive speed (speeding), driving under the influence of alcohol and/or drugs, not wearing seat belts or child restraints, and failing to wear a motorcycle helmet. Since law enforcement was put in place regarding these factors, there has been a reduction in the number of road collisions and fatalities (Wang, Quddus and Ison, 2013; World Health Organisation, 2013). An extensive body of research has been conducted by road safety and transport institutions, highway safety foundations and investigation teams about the effect of these factors as well as other contributors to driver safety, such as age and fatigue (Horberry *et al.*, 2006; Wang, Quddus and Ison, 2013; Zhang *et al.*, 2016).

By reviewing the UK's driving offence codes and list of endorsements (penalty points), it has been concluded that out of the 15 categories, three of them are related to drivers' driving performance and their speeding. These classes are careless driving, reckless/dangerous driving, and exceeding the speed limit. The law states that careless driving offences shown by codes CD10, 20, and 30 are "driving without due care and attention or without reasonable consideration for other road users". Regarding reckless and dangerous driving, under the codes DD10, 40, 80 and 90, this refers to "furious driving, causing serious injury, and causing death by dangerous driving or committing manslaughter or culpable homicide while driving a vehicle" [5] (Driver and Vehicle Standards Agency, 2015). The careless driving endorsement category is aimed at targeting drivers with poor driving behaviour and discipline, for which on-the-spot penalty points and a fine came into effect on 16 August 2013. Since these data are published based on postcode groups, the following table is aggregated based on the postcodes started with BA. Out of 36 areas with their postcodes started with BA, only BA1 and BA2

[5] Law Road Traffic Act (RTA) 1988 sects 2 & 3 as amended by RTA 1991. (Driver and Vehicle Standards Agency, 2015).

cover Bath city. Consequently, the following table has been constructed by extracting endorsements awarded only in these two postcode areas. As it can be observed, exceeding the statutory speed limit on a public road for the BA1 and BA2 postcodes were 521 and 1053, respectively. This is concrete information as exceeding the speed limit is a measurable offence, whereas careless driving is a subjective concept.

Table 3. Total number of convicted non-heavy-duty vehicle drivers with driving endorsements 2013[6]

Bath postcodes	endorsements codes	Endorsement awarded
BA1	Driving without due care and attention	3
BA1	Exceeding statutory speed limit on a public road	521
BA1	Exceeding speed limit on a motorway	106
BA2	Driving without due care and attention	10
BA2	Exceeding statutory speed limit on a public road	1053
BA2	Exceeding speed limit on a motorway	107

The importance of reducing CO_2 emissions of passenger cars in the EU and the necessity of reducing road casualties due to car accidents in the same region initiated a body of scientific research on the effective parameters for reduce both ameliorating both of these matters. Drivers' driving behaviour is one of the primary causes of increasing fuel usage beyond the parameters under the control of the car manufacturers. Apart from dangerous driving habits, such as careless driving (failing to look properly), failing to judge other the person's trajectory or speed and reckless or hurried driving, there are specific driving behaviours that increase a driver's chances of involvement in a road accident, such as harsh acceleration, harsh braking and speeding for a high proportion of a trip. In fact, identifying drivers' patterns of driving

[6] Endorsement related to careless driving and exceeding speed limit, in Bath area (BA1, BA2), 2014.

habits that contribute to excessive usage of fuel and attitudes that are unsafe are increasingly becoming the focus of the automotive industry and research institutions.

1.4 Research aims and objectives

Understanding drivers' driving behaviour without an intrusive monitoring device in their natural driving setting (their own car, on a route that they are familiar with) in a real-world traffic situation, has led to a field of research with the primary goal being to identify driving behaviours that correlate highly with the risk of collision. Furthermore, the naturalistic driving studies are aimed at determining the effect of human factors on vehicle performance (LeBlanc, Sivak and Bogard, 2010), road topography along with road safety (Al-shihabi and Mourant, 2001) and traffic management (Rakha *et al.*, 2011). The EU PROLOGUE project (Promoting real Life Observations for Gaining Understanding of road user behaviour in Europe) and the 100-Car Naturalistic Driving Study in the US are two successful examples of such studies (Klauer *et al.*, 2006; Van Schangen *et al.*, 2011).

The inapplicability of this prior research to the UK setting is because drivers' driving behaviour is influenced markedly by countries' driving culture and attitude towards the road. Despite the extant research the impact of driver behaviour on vehicle fuel economy and road safety still remains under researched, according to engineers and road safety experts. Finally, investigating scoring and ranking methods to rank drivers' differences is gaining huge interest due to its relevance to usage-based insurance and the fleet monitoring industry.

Building on naturalistic driving behaviour studies, this research is aimed at gaining an understanding of the effect of driving performance differences in terms of fuel usage and road safety. Specifically, the goals are to identify driving habits that affect fuel usage, eliciting drivers' attitudes towards driving safely, classifying and ranking drivers with similar driving behaviours and finally, using this information to model and simulate driver variations and their effects on driving performance. To accomplish these aims, the guide to conducting naturalistic driving behaviour study was followed, i.e. the framework from the handbook[7] on the methodology of implementing and operating field operational tests was used throughout this research. The project was divided into three stages: study preparation, data acquisition / data management and finally, post-processing. The general format and method of the ISO standards for testing passenger cars on open loops was employed to develop a research strategy and to conduct a comprehensive study.

[7] Field operational test support action (FESTA) handbook.

1.5 Research topics and questions

In the discussion above, two areas of research on drivers and driving behaviours have been identified for investigation. These are the effects of drivers on using fuel and driver attitudes to driving safely. In naturalistic driving behaviour studies, driving behaviours are identified with the purpose of grouping drivers together or modelling their behaviours. The core research topics are in relation to identifying and classifying driving behaviour differences by using collected driving data from vehicle engines and driver locations using GPS sensors. For each research topic (eco-driving and safe driving), two stages of research interests were defined (identifying and classifying) (see Figure 15). Finally, the recorded data were used to simulate drivers' driving differences.

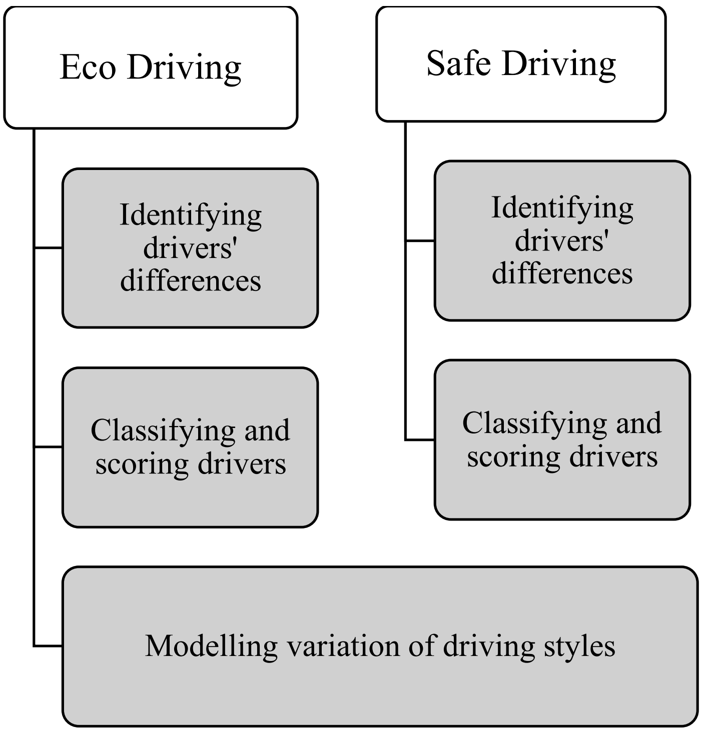


Figure 15. Research topics

The list of research questions is as follows.

1. What driving parameters (e.g. vehicle speed, acceleration, and engine speed, road geometry) can identify drivers' driving differences?
2. What is an effective way to classify, rank, and group drivers based on driver differences and similarities?
3. Is it feasible to model variations of driving behaviour based on collected real driving data?
4. What driving parameters (e.g. vehicle speed, acceleration, and engine speed, road geometry) can identify drivers' attitudes to driving safely?
5. What is an effective way to classify, rank, and group drivers' safe driving based on dangerous driving habits?

1.6 Project overview and structure

To gather real-world naturalistic driving data, a series of driving events was designed and hosted. Specifically, in order to engage with potential drivers, the Eco Safe Driving Challenge (ESDC) was designed and introduced as a driving campaign. The ESDC had two main aims: firstly, it provided a real-world platform for gathering reliable driving data and secondly, the events provided an opportunity to exchange and educate the public about the effect of their driving behaviour on their daily driving and fuel usage. The campaign was very successful[8]. In total, three driving events were hosted by the author and a team of volunteers; 250 miles were safely driven and around 500 minutes' worth of driving data were collected. In this thesis only the data gathered during the final driving event on 9th February 2014 have been used in the analysis. The public engagement and knowledge exchange approach that was adopted in the early stage of this project created an opportunity for the author to work with various departments at the University of Bath and it provided an opening to collaborate with industry leaders. Appendix A is the final report about the driving campaign milestone and the project's public engagement and knowledge exchange outcomes.

1.6.1 Funding approach and ESDC project scope

The Eco Safe Driving Challenge was funded[9] by the Research Development Unit at the University of Bath. The fund covered the cost of running the three driving events and the required expenses related to the project. The project scope included an information session, the driving events, an academic seminar, and an awards ceremony. For collecting the real driving data, unobtrusive-monitoring devices, namely OBD dongles and GPS sensors, were chosen, as

---

[8] ESDC dedicated website: www.esdc2013.com

[9] The total funding was £2500; this covered the cost of the monitoring devices, SIM cards and hosting the driving events.

they were inexpensive and would not distract those driving in any way. The cost per driver including monitoring device, sim card top up, the amount of fuel used at the event and extras to do with hosting the (refreshments, prints, etc.) was around £165.

1.6.2 Collaboration with industrial partners

This project involved collaboration with several transport research institutions and companies. For instance, the Transport Research Laboratory (TRL) was invited to make a presentation at the ESDC seminar. The EE mobile network operator and internet service provider supported the project by providing SIM cards for the monitoring devices. There have been two industrial partners that the researcher had been able to secure and collaborate with; firstly, there was the CASTEL Wireless Telecommunications Co. Ltd (Castel Company), who sourced the monitoring and data acquisition devices. They also provided consultation about their approach to monitoring and scoring driving behaviour; their wireless fleet monitoring technologies have been used in 27 countries and regions globally. The Castel Company has been backed by the China Aerospace Industry, and it receives technical support from the Hong Kong and Shenzhen R&D centres. They collaborated in the project because of a mutual interest in developing methods to identify and classify drivers' driving performance for monitoring and insuring purposes.

The second partner of the project was the IPG Automotive Company[10], a leading virtual driving simulator software provider, with a high level of homologation approvals by car manufacturers as well as a history of extensive collaborative practices. It sponsored the project by supplying their passenger car simulation software package called 'Carmaker', which was because of their

[10] For more information, visit: www.ipg.de

interest in increasing the level of realism of the driver models in their software. That is, as will apparent, this collaboration involved using their the IPG Carmaker software to model drivers based on their actual driving behaviour. Two scenario-based studies were developed, and the software was used to establish the effectiveness of modelling driving behaviour using their software package.

1.7 Thesis outline

This thesis is organised as follows: Chapter 2 is the Literature Review; Chapter 3 presents the Methodology; Chapter 4 gives results and discussion, and Chapter 5 contains the conclusion. Each chapter contains the context of two subjects; firstly, topics related to fuel consumption and the eco-driving performance of drivers and secondly, a section covering road safety and the attitudes to safety by drivers while driving. This structure is kept throughout the thesis and thus, each chapter begins with topics and parts related to eco-driving followed by those pertaining to safe driving.

In Chapter 2, the relevant literature, including previous empirical works by scholars who have studied the impact of drivers' driving behaviour on fuel consumption and road safety is investigated. In this chapter, advancements made in relation to these topics are discussed, in terms of a chronological timeline of the technological advancements in the past 80 years Specifically, the most salient works for purpose of this thesis, dating back to 1938, are reviewed so as to guide the research process. The chronological part focuses on studies between 1938 and 2010, whilst the technological part covers highly cited relevant journals and academic papers published between 2010 and 2015. Subsequently, a second round of literature searching was conducted, which unearthed a number of key studies published between 2016 and 2017 for review.

In Chapter 3, the methodology, the methods employed are divided into three sections: field study preparation, data acquisition, and post-processing analysis of these data. The lattermost consists of two parts, one pertaining to the eco-driving analysis and the other contains studies exploring safety in driving.

In Chapter 4, the results from identifying driving behaviours that affect driving efficiency are presented in the first section, whilst in the second, drivers' classification outcomes are covered. The findings from the Carmaker software virtual modelling of real world driving are presented and discussed in subsection three. The final two subsections cover the outcomes of identification (subsection four) and classification (subsection five) of driving behaviours regarding the road safety aspects of the driving task.

In Chapter 5, there is discussion of the key findings for all the studies, reflection on the methodology deployed and consideration of the relevance of the work. Regarding the lattermost, the significance of findings is examined and the impact of the key contributions are re-examined in the context of previous studies. Finally, the researcher's contributions are provided and this study limitations are discussed.

In Chapter 6, the project outcomes are summarised, and directions for future research in the field of naturalistic driving behaviour studies are proposed.

1.8 Research contributions

The project's initiative and approach, including the results of the pilot study "Effects of driving behaviour on fuel consumption" were published in the fifth volume of the *Driver behaviour and training* in 2012. The paper "Identifying Collision-Prone Locations on Commuting Routes Toward the Workplace and Forecasting their Future Trend" was presented in November 2015 at the *4th International Traffic Accident Conference,* Tehran, Iran. Following are a list of papers that will be published this year:

1. "Establishing the association between collision-prone locations and places where drivers often made harsh accelerations and decelerations" will be submitted to the *Accident Analysis & Prevention*
2. The study that classified drivers by their vehicle speed will be submitted to the *SAE International Journal of Transportation Safety.*
3. Two studies conducted by modelling driver performance with Carmaker software will be jointly published at the *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*.

The researcher continues to maintain the Eco Safe Driving Challenge website as a platform for publishing results to non-academic readers. Since completion of the project; the web page has been transformed into an educational hub providing news regarding issues relating to the subject matter of thesis.

Chapter 2

Literature Review

"Measure what is measurable, and make measurable what is not so."

– Galileo Galilei

2.1 Introduction to the importance of reducing transport emissions

As has been discussed in the previous chapter, the inefficient usage of fuel by UK car drivers has become one of the most influential factors affecting the level of cars' CO_2 emissions contributed by the transport sector in the country. In recent years, cars have had a remarkable reduction in fuel usage and GHG emissions due to efficient engine designs, well-designed aerodynamics body shapes, and improvements in catalyst technology (De Vlieger, De Keukeleere and Kretzschmar, 2000). Whilst in modern vehicles with Internal-Combustion engines (IC) the three-way catalytic converter (TWC)[11] is designed to operate within a tolerance of temperature and air/fuel ratio. However, under certain conditions, such as cold start and high engine load, the catalytic converter can underperform and as a result, vehicles' level of GHG emissions can increase (Yamada, Kayano and Funabiki, 1993; De Vlieger, 1997; De Vlieger, De Keukeleere and Kretzschmar, 2000). Operating a vehicle under these conditions depends on factors, such as initial idle time when starting the vehicle, ambient air temperature,

---

[11] A common device used by car manufactures to convert carbon monoxide (CO), oxides of nitrogen (NO_x), and hydrocarbons (HC) into carbon dioxide (CO_2), water (H_2O) and nitrogen (N_2).

and time between trips, which can increase the cold start effect (May, Bosteels and Favre, 2014; Wyatt, Li and Tate, 2014).

A high engine load effect occurs owing to a combination of high speeds, high acceleration rates and a positive road gradient along with secondary ones, such as operating the air conditioning, e.g. driving uphill with the air conditioning on. Drivers' driving habits have a direct effect on operating a vehicle under these conditions and as a result, they increase vehicle exhaust contributions to GHG emissions (Wyatt, Li and Tate, 2014; Fontaras, Zacharof and Ciuffo, 2017) .

One of the ways to increase drivers' awareness of their level of CO_2 contribution is by including vehicle tailpipe CO_2 emission levels on the vehicle excise duty, commonly known in the UK as road tax. From March 2001 onwards, every car producing carbon dioxide emissions and registered was obliged to pay road tax based on the level of g/km of carbon dioxide emissions it produced. Table 4 presents part of the list of the UK Vehicle Excise Duty rates for cars registered after 1st March 2001. The price for newly registered cars is aligned with current EU CO_2 emissions targets.

Table 4. Current rates of road tax in the UK for cars registered after 1st March 2001[12]

Band	Vehicle CO_2 emissions	Standard rate 2016-2017	First-year rate 2016-2017
A	Up to 100 g/km	£0	£0
B	101-110 g/km	£20	£0
C	111-120 g/km	£30	£0
D	121-130 g/km	£110	£0

[12] For full list and data source visit: https://www.gov.uk/vehicle-tax-rate-tables/rates-for-cars-registered-on-or-after-1-march-2001

Band	**Vehicle CO_2 emissions**	**Standard rate 2016-2017**	**First-year rate 2016-2017**
E	131-140 g/km	£130	£130
F	141-150 g/km	£145	£145
G	151-165 g/km	£185	£185
H	166-175 g/km	£210	£300
I	176-185 g/km	£230	£355

The allowed level of tailpipe CO_2 emissions and vehicle fuel economy is regulated and measured by environmental agencies globally. By testing vehicles under regulatory driving cycles, the standards determine the accepted level of fuel consumption and carbon dioxide emissions of new cars sold in the EU and the European Economic Area (EEA). These laboratory-based vehicle emission tests are conducted under standardised conditions to provide a common basis so as to be able to compare the CO_2 emissions and vehicle fuel economy of vehicles (International Energy Agency, 2012; Olivier, Muntean and Peters, 2015).

Driving cycles are standardised driving patterns, which are defined by velocity against time tables. There are three globally well-known driving cycles: the European driving cycles, the US driving cycles, and the Japanese driving cycles. The EU and Japanese driving cycles are modal ones, which means that for some parts of the driving cycle the speed is constant, while the US driving cycles are transient ones, which involve a continuous change of driving speed (Joumard *et al.*, 2000; Li *et al.*, 2007; Barlow *et al.*, 2009; Fontaras *et al.*, 2014). The vehicle acceleration is assumed to be constant during these driving cycles. Over time, new additions have been introduced to include suburban[13] driving into test cycles. In the EU, these regulatory tests are called Euro cycles. Each successive Euro standard has been aimed at tightening the

[13] Suburban or Extra-Urban Driving Cycle (EUDC).

allowed level of accepted CO_2 emissions and fuel consumption when the vehicle is tested under a standard driving cycle. Table 5 shows the milestones of Euro emission standards and driving cycles (International Energy Agency, 2012; Delphi, 2013).

Table 5. The EU passenger car vehicle emission standards and driving test cycles

<table>
<tr><th>Euro 1</th><th>Euro 2</th><th>Euro 3</th><th>Euro 4</th><th>Euro 5a</th><th>Euro 5b</th><th>Euro 6b</th><th>Euro 6c</th></tr>
<tr><td colspan="4">Urban + Extra-Urban</td><td colspan="3">Revised urban + Extra-Urban</td><td>WLTC[14]
+ RDE</td></tr>
<tr><td colspan="8">1995 2000 2005 2010 2015 2020</td></tr>
</table>

Source: (International Energy Agency, 2012; Delphi, 2013)

The New European Driving Cycle (NEDC) is the current one used as part of the Euro 6b standard. The NEDC pertains to a combination of urban and extra-urban driving conditions. The following references provide details about the US, Japanese and European driving cycles (Barlow *et al.*, 2009; International Energy Agency, 2012; Delphi, 2013; Fontaras *et al.*, 2014).

[14] Worldwide Light Duty Test Cycle (WLTC).

2.2 Shortfalls in the fuel consumption and CO_2 measurements of driving cycles

The main deficit of all driving cycles and laboratory-based fuel consumption and emission measurements is that they are not representative of real-world driving. The issue of lack of representation of real-world emissions has been a point of criticism of the modal driving cycles. The first modal urban driving cycle, UDC, was introduced in 1970 as part of international efforts to reduce CO_2 emissions globally. In 1973, Kruse and Huls pointed out the fact that if the aim of using driving cycles is to measure emissions of real-world traffic conditions, then cycles failed to represent the dynamic conditions of real-world driving (Kruse and Huls, 1973). A random roadside survey conducted in 1985, 1987, and 1989 in California, US, was a confirmation that driving cycles were not representative of present real-world driving conditions (Ashbaugh and Lawson, 1991). The shortfall inaccurate representation of real-world driving by driving cycles led to studies on measuring real-world emissions by using on-board emission instrumentation (Kelly and Groblicki, 1993). Specifically, the necessity to measure vehicle emissions while driving resulted in the development of portable on-road emission and energy measurement devices, which today are called the Portable Emissions Measurement Systems (PEMS). The Flemish Institute for Technological Research (VITO) on-board system[15] was an early attempt to capture the effect of real-world traffic conditions of fuel usage and CO_2 emissions levels by taking tailpipe samples of real-world driving (De Vlieger, 1997).

[15] VOEM: VITO on-the-road emission and energy measurement system.

2.3 Representation of real-world driving conditions in vehicle testing

As explained at the beginning of this chapter, modern cars fitted with three-way catalyst (TWC) devices give less GHG emissions. However, their efficient performance can only be achieved under an ideal minimum theoretical air/fuel ratio. This also known as stoichiometric combustion[16] is achievable under laboratory conditions for which all regulatory driving cycles are designed to perform. New cars performed increasingly well under regulatory driving cycles since they were designed to operate under stoichiometric conditions during tests in labs. Under real-world driving conditions, TWC did not always perform well, e.g. the vehicle operated under a cold start situation or had a high engine load. Consequently, initial attempts to mimic real-world driving conditions in laboratory tests had to take the poor performance of TWCs into account. Over time, other parameters such as road gradient, drivers' driving behaviours and traffic condition, which could not be accurately simulated, gained more attention, and hence are included in the test procedure.

The importance of identifying influential factors that affect fuel consumption in real-world conditions, and measuring their impact, led to their being investigated according to two test scenarios, real road driving and on a laboratory test bed, i.e. chassis dynamometer. Road based studies were focused on external factors affecting vehicle emissions, whilst also investigating topics such as traffic management and road safety. Whilst laboratory tests were aimed at identifying factors affecting inaccuracy in chassis dynamometers and developing methods to closing the gap between laboratory fuel consumption test results and real-world fuel usage (Kelly and Groblicki, 1993; Burke, Brace and Moffa, 2009; Berry, 2010).

[16] The stoichiometric combustion is the ideal combustion process where all oxygen is used and all fuel burned.

Since the objective was comparing the road results with chassis dynamometer ones and adjusting the lab fuel consumption outcomes to reflect the conditions on the road (usually in the form of correction factors), the effect of driving behaviour and road gradient became an important part of these improvements. These studies led to an accurate measurement of fuel consumption by adjusting the test results to match real-world driving (see section 2.5). The aforementioned studies on the previous page provided the basis for the development of driving models as well as the introduction of driving cycles with a better representation of such driving (Tutuianu *et al.*, 2013; Fontaras *et al.*, 2014). Influential advances go back to early 1970s with pioneers such as Kruse and Huls (Kruse and Huls, 1973), were brought about through technological evolution, such as fuel efficient vehicles, accurate measuring and testing devices all driven by revolutionary progress in computing and post-processing software packages. In the following subsections, the milestones of these studies and developments are discussed with a primary focus on road vehicle studies.

2.4 Chronological and technological review of studies investigating the role of drivers

The significant result of an experimental study on the effect of drivers' differences in fuel consumption by Evans (1979), led to research being conducted to investigate the role of drivers on fuel consumption. The study concluded that experienced drivers can save more fuel without increasing their trip time by anticipating the road ahead and by adjusting their speed in accordance with any approaching road setting, such as a set of traffic lights (Evans, 1979). Waters and Laker's (1980) report from the Transport and Road Research Laboratory (TRRL), currently known as the Transport Research Laboratory (TRL) , emphasised that conservation of fuel usage can be done by addressing the following major factors (Waters and Laker, 1980):

- Driver driving performance;
- Vehicle design;
- Road gradient;
- Traffic flow.

They conducted a road study involving nine drivers with each completing 10 traffic driving laps. The report suggested that the least and most fuel efficient drivers had up to 50 percent difference in fuel usage. As part of their study, Waters and Laker (1980) investigated the effect of acceleration rate and poor speed control on fuel consumption and found that by educating the driver about his /her role on fuel usage, it is possible to save of up to 15 percent fuel usage. The study also addressed the effect of changing the aerodynamic drag coefficient and final drive ratio of a vehicle on fuel consumption (ibid). Since instantaneous fuel consumption measurements were not accessible at the time, an indirect measurement would be used, such as the Clarkson estimation model for fuel usage (Clarkson and Hicks, 1982), called the average speed fuel consumption model, shown as Equation 1 below. Despite Clarkson and Hicks'

(1982) model being developed under the specific condition of the average fuel consumption of a passenger car at 30 km/h, in having formulated a fuel consumption model linked to vehicle speed, this meant that a distinction could be made between two drivers and as a result, the more efficient driver could be identified (ibid). This concept was deployed in a study by the Swedish Road and Traffic Research Institute (VTI) on the impact of various driving styles on fuel consumption and established the concept of economical driving (eco driving) to save fuel and fuel cost (Laurell, 1985). To this day, the concept of eco-driving is still relevant and consequently, is widely discussed (Mensing *et al.*, 2014; Ayyildiz *et al.*, 2017).

Equation 1. The average fuel consumption model (Clarkson and Hicks, 1982)

$$Q_m = 16.57 \times P_m \times V^{-1} \times \exp\,(0.0195.V)$$

Where,

Q_m= The average fuel consumption rate for a passenger car (litres/100km)

P_m= The average fuel consumption of passenger car at 30 km/h

V = Average speed (Km/h)

The aim of using the above equation was to compare the influences of drivers' driving speed on fuel consumption. At the time, comparisons were mostly focused on understanding the impact of the driver on fuel consumption. Notably, a study by Redsell et al. (1988) achieved this by conducting a field study using different vehicles and fuel types. The outcomes of the study of three Vauxhall Cavalier cars with different engine sizes and fuel types suggested that depending on traffic conditions, a 1600 cc diesel car could save between 4% to 20% of fuel when compared to a petrol car with a 1300 cc engine capacity (Redsell, Lucas and Ashford, 1988). The authors concluded that out of seven different drivers the best performance was achieved by the professionally trained driver in urban conditions with the diesel version of the car model. It was reported that when compared to regular drivers, the professional drivers, on average, has 6% less speed, 9% less vehicle acceleration and 14% less vehicle deceleration. According to Redsell et al. (1988), these differences meant that adopting this style of driving and operating a car in a moderate manner could significantly reduce running costs for motorists as well as potentially lead to a saving of up to 9% in fuel consumption.

By 1989, a large number of petrol-driven vehicles were equipped with an electrical fuel injection system, and therefore, it became possible for researchers to investigate instantaneous fuel consumption accurately. From the 1990s onwards the positive effect of the aforementioned three-way catalyst could be measured with increasing accuracy, as well as the actual effect of factors that increased the amount of fuel consumption and vehicle emissions. This was due to the advancements made in the portable emissions measurement system (PEMS) by the Warren Spring Laboratory in the UK (Potter and Savage, 1982), the VOEM by the Flemish Institute for Technological Research (Van Mierlo *et al.*, 2004), the Portable Real-Time Emissions Vehicle Integrated Engineering Workstation (PREVIEW) developed by Ford and the US Environmental Protection Agency on-board and on-road High Precision Remote Sensing devices (Jiménez-Palacios, 1999)[17]. Because real-world driving conditions are not repeatable and well-defined, studies with portable emission data acquisitions systems have tended to focus on the design of the study, the demography of the drivers, vehicle types and the topography of the road (Wyatt, Li and Tate, 2013).

[17] Exhaust emissions scanning device that is placed near the road to collect exhaust gas levels.

A comprehensive study by the US Environmental Protection Agency in 1993, compared the effect of drivers' operating a vehicle under a high engine load and concluded that the primary indicators for identifying and measuring the effect of real-world driving is by comparing drivers' vehicle speed and acceleration (US Environmental Protection Agency, 1993). The secondary factors that were identified by the report were vehicle type and age, travel duration, time of day as well day of the week, and trip patterns (trip lengths) (US Environmental Protection Agency, 1993). The EPA's introduced the concept of "specific power" to compare drivers' excessive usage of available vehicle power above that required. The EPA's specific power value is developed based on the Watson Positive kinetic energy model (Alimoradian, 1983), defined as the square change of speed during positive acceleration. The value later became an indicator of a driver's aggressiveness. According to the EPA, the following equation can be used to calculate specific power:

Equation 2. The specific power according to the EPA (US Environmental Protection Agency, 1993)

$$\boldsymbol{EPA\ specific\ power = v_f^2 - v_i^2 \approx 2 \times v \times a}$$

Where,

v = Vehicle speed and $v_f > v_i$

a = Vehicle acceleration

A study in 1993 by Kelly and Groblicki, compared real-world driving data with chassis dynamometer driving data. They concluded that calibrating a portable device in certain ways (based on different calibration gases) led to very different emissions levels for the same car. The researchers found that events, such as hard acceleration while entering freeways or ascending hills, can result in ineffective catalyst performances and therefore, higher CO_2 emission levels when compared to laboratory results (Kelly and Groblicki, 1993). Research into determining the factors affecting fuel consumption in 1993 was a critical study since by on-road measurement of driving parameters; the authors were able to quantify the factors influencing fuel economy under different road types, driving conditions and fuel types (Kelly and Groblicki, 1993).

De Vlieger's (1997) study for the Flemish Institute for Technological Research focused on the advantages of cars with three-way catalysts installed. The findings of a study elicited that cars with a three-way catalyst had 70% lower generation of GHG emissions when compared to non-catalyst ones (De Vlieger, 1997). The researcher made a comparison between normal driving and aggressive driving and concluded that the former produced four times more emission gases (De Vlieger, 1997). While An and his team's (An *et al.*, 1997), study involved comparing the emissions of 300 vehicles in a laboratory, where they developed a predictive tailpipe emissions model for a variety of driving conditions. Other researchers focused on defining aggressive driving and its effect on fuel consumption and emissions (Martens *et al.*, 1998; Johansson, Färnlund and Engström, 1999). The highly cited work of (Jiménez-Palacios, 1999), who introduced the Vehicle Specific Power (VSP) concept and were able to quantify vehicle emission under different conditions, provided the basis for a broad range of new research and emissions software development[18] (software such as US EPA Motor Emission Simulator software). The emissions model concept was simplified by (Hual *et al.*, 2005; Zhai, 2007) into its current format (see Figure 16 and Equation 3).

[18] Motor Vehicle and Equipment Emission System software (MOVES).

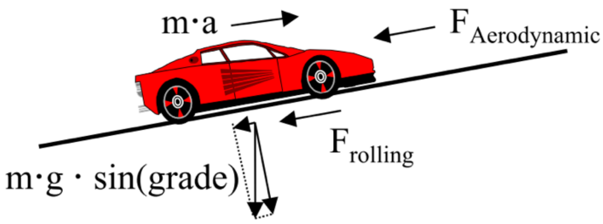


Figure 16. Demonstration of forces applied to a moving car (Jiménez *et al.*, 1999)

Jiménez-Palacios (1999) formulated the impact of road gradient, aerodynamic drag, rolling resistance and other forces (see Equation 3) in order to calculate the true required power by a vehicle based on its mass. The advantage of VSP over the other aforementioned methods is that it provides a one-dimensional value, which depends on the level of emissions. Using VSP makes comaprison of different driving tests easier as it is already includess the imapct of parameters such as each cars aerodynamic drag coefficient and road gradient (Jiménez-Palacios, 1999).

Equation 3. The Vehicle Specific Power by (Jiménez-Palacios, 1999; Hual *et al.*, 2005; Zhai, 2007)

$$Vehicle\ Specific\ Power\ (VSP) = \frac{Power}{Mass} = \frac{\frac{d}{dt}(KE + PE) + F_{rolling}.v + F_{Aerodynamic}.v}{m,}$$

$$= \frac{\frac{d}{dt}\left(\frac{1}{2}mv^2 + mgh\right) + C_R mg.v + \frac{1}{2}\rho_a C_D A(v + v_w)^2.v}{m}$$

$$= v.(a.(1 + \varepsilon_i) + g.grade + g.C_R) + \frac{1}{2}\rho_a \frac{C_D A}{m}(v + v_w)^2.v$$

$$\cong\ v \times (a\ +\ g \times sin(\phi)\ + \psi\ )\ +\ \zeta \times v^3$$

Where,

VSP = Vehicle Specific Power ($W/Kg = m^2/s^3$)

v = Vehicle speed (m/s)

a = Vehicle acceleration (m/s^2)

ϕ = Road gradient

ψ = Rolling resistance coefficient

ζ = Drag coefficient (m^{-1})

m= Vehicle mass

$Grade$= Length of road vertical rise

C_R= Coefficient of rolling resistance

C_D= Drag coefficient

A= Fontal area of the vehicle

ρ_a= Ambient air density

v_w= Headwind into the vehicle

ε_i= Gear dependent "translational mass of the rotating components", mass factor

While in the US, eco-driving training had been introduced by 1970, the concept of eco-driving only emerged in 1998 in the EU as part of driving training courses (Bandivadekar, 2008). The term eco-driving refers to fuel-efficient driving that has environmental and financial benefits in the long-term by saving fuel and having less maintenance costs. The study of the effect of eco-driving on fuel consumption in comparison with normal driving, as conducted by Johansson, has been influential as it shows that drivers who receive eco-driving training are able to achieve a 10% reduction in fuel consumption (Johansson, Färnlund and Engström, 1999). Through conducting a road test with real-world traffic congestion, several research groups investigated the effect of traffic flow on fuel consumption (De Vlieger, De Keukeleere and Kretzschmar, 2000; Rapone *et al.*, 2000). Regarding which, De Vlieger's (De Vlieger, De Keukeleere and Kretzschmar, 2000) medium-size study of five diesel cars and four petrol cars on city, rural and ring roads was not able to establish a significant correlation between traffic jams and fuel consumption. However, it did confirm that aggressive driving increases fuel consumption by 40% (De Vlieger, De Keukeleere and Kretzschmar, 2000).

The concept of On-board diagnostics (OBD), and giving the owner access to the vehicle ECU and subsystem was introduced by Volkswagen first in the form of its scannable on-board computer system in 1968, and since then advancements have been made by car manufacturers in their own in-house versions of such a system. By 1988, the Society of Automotive Engineers (SAE) recommended car makers standardise their connector (connection port) and diagnostic test signals. The OBD-I standard in 1991 was the basis of a mandatory OBDII standard specification in 1996 for every passenger car manufactured in the US (Greening, 1992). The EOBD,[19] the European version of the OBDII standard, became mandatory for every new petrol car with less than 8 passenger seats in 2000. Apart from the primary function and application of accessing the ECU via OBD ports and OBDII protocol, which involves scanning engine faults and logging driving data while driving, accessing engine information in this way gained more attention. The OBD dongles became affordable tools to collect large scale driving data, and to conduct emission tests[20]. Moreover, data collected in this way could be used to monitor fleet drivers' fuel consumption and driving behaviour.

[19] European On Board Diagnostics (EOBD).
[20] For instance, currently it is mandatory in the Netherlands to have an EOBD test annually. In some states of the US it is permitted to do an emissions test with an OBD dongle instead of the tailpipe emissions test.

Coincidentally, with a very similar timeline, the U.S. global positioning system policy in 1996, opened public access. The policy then later expanded in 1998 to provide enhanced GPS data to civilian users worldwide. By 2000, the U.S. had lifted all degrading[21] on GPS signals, which meant an increase in accuracy 10 times more than what was available before the 1st May 2000. The importance of this development was that it led to new era of advancement in mobility and transportation, in general, examples of which can be found in McNally et al.'s comprehensive study on using in-vehicle and GPS technology for transport and traffic management studies (McNally *et al.*, 2003). More importantly, having access to position data and driving data simultaneously, led to a paradigm shift in how driving behaviour is studied. Given the change in relation to how experiments conducted (using new devices), there was great level of concern about driver's privacy and monitoring individual driving behaviour. As early as 2002, telematics technology was adopted to monitor and track drivers, thus raising the issue of privacy and personal data needing to be addressed by industry. The highly cited work of Duri et al. (2002) introduced a detailed framework to address security and privacy concerns related to automotive telematics data (Duri *et al.*, 2002).

Having access to instantaneous fuel consumption through the ECU with OBD loggers, and vehicle location by GPS sensor, provided the opportunity for on-road vehicle real fuel consumption testing and the development of an eco-driving advisory tool for drivers while driving (Frey *et al.*, 2001; Van der Voort, Dougherty and van Maarseveen, 2001; Unal, 2002). The fuel-efficiency advisory tool by Van der Voort helped drivers to achieve a 16% reduction in fuel consumption compared to their normal driving style (Van der Voort, Dougherty and van Maarseveen, 2001). As has been mentioned before, De Vlieger et al. (De Vlieger, De

[21] Selective Availability (SA).

Keukeleere and Kretzschmar, 2000) were not able to find a significant correlation between traffic conditions and fuel consumption. However, the study by Frey et al. (Frey *et al.*, 2001) and later by Frey et al. (2001) showed a reduction in vehicle emissions through better traffic management (Frey *et al.*, 2001; Unal, 2002).

Between 2002 and 2003, an experimental study conducted in South Korea was very influential in relation to using in-vehicle monitoring technologies to identify driving patterns affecting fuel consumption. The team managed to collect 1,300 kilometres of driving data and as a result, they were able to measure, firstly, the percentage of time drivers spent in each driving mode[22] and secondly, they quantified the effect of average speed on fuel consumption (Sa, Chung and Sunwoo, 2003). Estimating the amount of time drivers drive in each driving mode provided the basis for a predictive fuel consumption model based on the modes. The driving mode consumption model helps to estimate the amount of fuel consumed per vehicle in different traffic conditions, but it ignores how a driver's behaviour impacts on fuel consumption in each driving mode (Equation 4). In addition, the model provided the basis for more complex predictive fuel consumption models (Ardekani, Hauer and Jamei, 2001; Montazeri-Gh and Naghizadeh, 2003; Saffarzadeh and Arjroody, 2003; Akçelk, Smit and Besley, 2012; Mathew, 2014).

Equation 4. The drive mode fuel consumption model

$$G = f_1 L + f_2 D + f_3 S$$

Where,

G= Amount of fuel consumed per vehicle (Litres/vehicle-km)

[22] Driving modes are: idle, accelerating, decelerating, and cruising.

L= Total measured travel distance

D= Total idle time (stopped delayed)

S= Total number of stops

f_1= Fuel consumed while cruising

f_2= Fuel consumed while idling

f_3= Excess fuel consumption while accelerating and decelerating

2003 was also a significant year, as Progressive Insurance in the US launched its first Pay-As-You-Drive insurance. Even though the trial was not a very successful, it did provide the basis for the more successful telematics based motor insurance policies currently on offer by Progressive.

Owing to the gap between cars' fuel consumption on roads and in laboratories, as part of a project, the World Business Council for Sustainable Development (WBCSD), the Sustainable Mobility Project (SMP) and the International Energy Agency (IEA) worked together to address the shortfall between vehicles' fuel economy labels and on-road fuel consumption, known as the "on-road gap", by developing reference cases and an adjustment model[23] for driving cycle results (Fulton and Eads, 2004). Moreover, the correction factors method was proposed by standard agencies (SAE J1349, DIN 70020 and ISO 1585) to adjust engine power output in relation to atmospheric conditions. Consequently, the engine power and fuel consumption amount was obtained in laboratory test cycle conditions (standard ambient conditions) matched

[23]The global transport spreadsheet model (Fulton and Eads, 2004).

the road tests (Sodré and Soares, 2003). The Ford Scientific Research Laboratory attempted to improve emission inventory estimates by comparing on-board real-world driving emissions with laboratory test results and the outcomes of a virtual model simulation[24] of traffic emissions (Nam *et al.*, 2003). The study involved developing the EPA specific power (see Equation 2) into a complex metric for measuring drivers' driving aggression. Nam argued that the aggressivity value can be used as an explanatory variable for emissions studies (Nam *et al.*, 2003). The work of (Jiménez-Palacios, 1999), (Nam *et al.*, 2003; Nam, 2004) led to the development of the Physical Emission Rate Estimator by the EPA and the introduction of a comprehensive version of the Motor Vehicle Emissions Simulator (MOVES) (Nam, 2004).

Equation 5. The metric introduced by the Ford Scientific Research Laboratory in 2003 to present drivers' aggressive driving

$$Ford\ Aggressivity\ value = RMS(P\ ) = \sqrt{\frac{1}{N}\sum_{i}^{N} P_i^2}$$

Where,

P= $EPA\ specific\ power \approx 2 \times v \times a$

N= Total number of positive vehicle acceleration profiles

i= Individual vehicle acceleration profile

While the MOVES simulator package provided the basis to estimate the effect of various real-world driving conditions on vehicle emissions and fuel consumption in the US (Hual *et al.*, 2005), in the EU the development of vehicle simulation programme software[25] led to in-depth

[24] The integrated model by combining the microscopic traffic model (VISSIM) and Comprehensive Modal Emissions Model (CMEM).
[25] Van Mierlo J.: Simulation software for comparison and design of electric, hybrid and internal combustion vehicles with respect to energy, emissions and performances", PhD dissertation, Vrije Universiteit Brussel, 2000.

analysis of the effect of vehicle weight, gear shift manner, tyre pressure and different driving styles conducted to establish the effect of driving behaviour and styles (eco-driving and sportive styles[26]) on fuel consumption (Van Mierlo *et al.*, 2004). The Van Mierlo approach to simulating differences in driving behaviour took the speed profile of each driving style corresponding to every traffic setting from the field study and used them on a chassis dynamometer to measure the actual fuel consumption and emissions of corresponding driving behaviours (Van Mierlo *et al.*, 2004).

By 2005, advancements made in wireless and the cellular network combined with real-time spatial information provided by Global Positioning Systems (GPS) signified an opportunity to collect real-time information about drivers, vehicles and traffic conditions (Li *et al.*, 2005; Tong, Merry and Coifman, 2005; Byon, Shalaby and Abdulhai, 2006; Hong *et al.*, 2007; Zhai, 2007). The in-vehicle tracking systems integrated with on-board engine monitoring (OBD dongles) changed scholars' approach to studying the real-world driving factors affecting fuel consumption and vehicle emissions. Work conducted by Alessandrini in 2006, is an example of using OBD dongles and a GPS sensor to capture the effect of everyday driving behaviour on the environment. The study concluded that drivers who accelerate calmly kept the air /fuel ratio around a stoichiometric value, while aggressive usage of the acceleration pedal resulted in higher fuel consumption (Alessandrini *et al.*, 2006) .

The most significant milestone in the field of driving behaviour studies and using in-vehicle monitoring devices on a large scale[27] was the 100-Car Naturalistic Driving Study. The project

[26] Classified according to (Van Mierlo *et al.*, 2004) study.
[27] The project scope in terms of numbers: 100 cars, 241 primary and secondary drivers, 43,000 hours of worth of driving data, 12 to 13 months data collection period for each vehicle (Dingus *et al.*, 2006).

was conducted by the Virginia Tech Transportation Institute and sponsored by the US National Highway Traffic Safety Administration (NHTSA) in 2006. Whilst the aim was to learn about drivers' behaviour and events before the occurrence of collisions, the usage of in-vehicle monitoring devices (OBD dongle, GPS sensors, facing in and out camera recorders) on a large scale initiated the basis to observe drivers in their daily driving settings unobtrusively, and as such, initiated the field of the naturalistic driving behaviour study (Dingus *et al.*, 2006). The 100-Car study was important for guiding the current research in terms of the methods used, regarding both the eco-driving and safe driving investigations (see 2.10).

While collecting real-world driving data became accessible and useful for understanding drivers' driving behaviour, measuring the effect of driver driving differences required laboratory-based tests, i.e. a chassis dynamometer. As part of the Competitive and Sustainable Growth Programme of the fifth framework programme of the European Commission, efforts were made by members of the DECADE project and a team from the former French National Institute for Transport and Safety Research (INRETS[28]) to reduce the gap between test bench and actual fuel consumption on the road (Joumard *et al.*, 2006; Pelkmans and Debal, 2006), i.e. the "on-road gap". Joumard and his team addressed the key issues of the Artemis project "Assessment and Reliability of Transport Emission Models and Inventory Systems", which were to reduce the systematic errors between laboratories' emission results and the necessity of quantifying the impacts of factors affecting fuel consumption (Joumard *et al.*, 2006). The Artemis Driving Cycles (ADC) were developed from the European driving database obtained from private cars in France, the UK, Germany and Greece. The vehicle speed, acceleration,

[28] Institut national de recherche sur les transports et leur sécurité (INRETS).

engine speed and trip information of 77 cars were recorded by in-vehicle monitoring devices[29] (Joumard *et al.*, 2006). The Artemis Driving Cycles aimed to represent European driving behaviour patterns on urban, rural and motorway roads by including various increases and decreases in vehicle speed and gear shifting strategy (Joumard *et al.*, 2000, 2006; Barlow *et al.*, 2009).

Parallel to those aforementioned studies that focused on impact of driving and engine components on fuel consumption; a study conducted by the International Energy Agency to recommend policies on potential savings to be made on fuel consumption from non-engine components, emphasised the impact of elements, such as tyres (20% fuel energy to overcome rolling resistance of tyres), cooling technology (between 15% and 30% of fuel usage) and lighting (consumes up to 32% of vehicle energy), on fuel consumption. The researchers also raised the matters of drivers operating the car in a responsible manner as well as the impact of educating drivers about eco-driving skills (Onoda and Gueret, 2007).

As part of efforts to quantify the impact of eco-driving training on reducing fuel usage, Af Wåhlberg investigated both the short-term and long-term effects of economical (efficient) driving on fuel consumption, accident involvement and changes in driver accelerating habits (Wåhlberg, 2006, 2007). The author discovered that whilst the effect of training about fuel consumption had always been reported as significant immediately after training, over a period of 12 months, only 2% reduction in fuel usage was recorded by drivers (Wåhlberg, 2007). Reed also came to similar conclusions when assessing the impact of fuel efficiency training on the

[29] The project scope in terms of numbers: 77 cars, 2200 hours of worth of driving data and a 2 years data collection period.

total fuel consumption of fleet drivers (Reed, 2007). The comprehensive project on "assessing fuel-efficient driver behaviour through tachograph information" conducted by the TRL explored the impact of driving styles on fuel consumption by using a vehicle driving simulator under five sets of experiments (Reed, 2007), the key results of which were as follows. Firstly, 99% of the variation in fuel consumption can be explained by a change in accelerating behaviour. Secondly, despite the fact that the participants drove on different routes (on the simulator), the correlation between acceleration and fuel usage remained a significant indicator as to who was the most fuel efficient. Finally, Reed ascertained that the standard deviation value of the vehicle acceleration can be used to identify fuel efficient drivers, regardless of their vehicle type and load (Reed, 2007).

Probing drivers' influences on fuel usage, a PEMS based study by Frey et al. (Frey, Zhang and Rouphail, 2008) investigated the impact of driver, road gradient and time of day on fuel consumption. The authors elicited that different drivers had different driving speed profiles for the same route and as a result, their average emissions differed by 4 to 5% for CO_2, 9 to 11% for HC, 16 to 18% for NOx, and 102 to 114% for CO, when the result was stratified by vehicle type, route and travel direction, and time of day (Frey, Zhang and Rouphail, 2008). Felstead et al. (Felstead, McDonald and Fowkes, 2009) conducted a similar study with the aim of classifying drivers' driving profiles by comparing driving data collected in a field study and from a chassis dynamometer. They identified two driving profiles: aggressive driving and passive driving. During data collection, the following instructions were given to drivers for each driving profile: for aggressive driving, they were told to undertake harsh acceleration and deceleration, keeping pace with the car in front at a safe distance; and for passive driving, they were asked to accelerate and decelerate with moderation, following the speed limit at all times. The results showed that aggressive driving was identifiable when comparing drivers' instantaneous acceleration. Moreover, aggressive drivers spent more time operating their vehicle in an inefficient part of the engine consumption map[30] (Felstead, McDonald and Fowkes, 2009).

Table 6. A summary of Felstead et al.'s (2009) study on the effect of driving style on emissions

Parameters	Passive	Aggressive	NEDC	WSL[31]
Duration (s)	2110	1799	1185	3977
Average speed (km/h)	32.97	38.81	33.2	38.28
Maximum speed (km/h)	79/69	116.81	120	118.59
Idle period (s), %	372, 17.6%	316, 17.6%	298, 25.1%	654, 16.4%
Acceleration standard deviation (m/s)	0.4463	1.007	0.4243	0.5252
Average acceleration (m/s^2)	0.2751	0.8052	0.5412	0.4277
Average deceleration (m/s^2)	-0.3251	-0.6685	-0.7885	-0.4141

Source: (Felstead, McDonald and Fowkes, 2009)

The concept of the "Pedal Busyness" was introduced by Burke (Burke, Brace and Moffa, 2009) as one of the test factors[32] in chassis dynamometer testing that can affect the accuracy and repeatability of fuel consumption measurement. The pedal busyness or oscillatory behaviour of the acceleration pedal exhibited by drivers in laboratory testing led to studies, such as those by (Bonnington, 2009; Daniel, Brooks and Pates, 2009), which investigated the pedal busyness value of data collected from real-world driving. The MAHLE Powertrain Ltd research outcomes, using the OBD monitoring device to monitor the driving styles of fleet drivers in the

[30] The brake specific fuel consumption (BSFC) map is the contour plot to show where engine operates efficiently (Goering and Cho, 1988).
[31] UK Department for Transport (DfT), Warren Spring Laboratory driving cycle representing real-world driving.
[32] Other factors investigated by Burke were: Battery discharge (V), Engine start temperature, Engine oil level, Pedal busyness, Speed error, Road speed fan, Vehicle alignment, Tie-down straps, Tyre type, Tyre pressure, Vehicle mass, and PAS pump (Burke, Brace and Moffa, 2009).

US and the UK, led to the terming of a new metric called "Total Aggressivity"[33] (Daniel, Brooks and Pates, 2009). The proposed metric (Equation 6) combined the concept of pedal busyness, i.e. pedal aggressivity (Daniel, Brooks and Pates, 2009), with the Ford aggressivity value (See Equation 5).

Equation 6. The MAHLE Powertrain Total Aggressivity value (Daniel, Brooks and Pates, 2009)

$$MAHLE\ Powertrain\ Total\ Aggressivity$$
$$= \sqrt{Ford\ Aggressivity\ value} \times \sqrt[4]{Pedal\ Aggressivity}$$

Where,

$$Ford\ Aggressivity\ value = RMS(P\ ) = \sqrt{\frac{1}{N}\sum_{i}^{N} P_i^2}$$

$$Pedal\ Aggressivity, (\%\ /\ s) = \sqrt{\frac{1}{N}\sum_{i}^{N} Q_i^2}$$

Q= $2 \times pedal \times pedal/s$

N= Total number of positive pedal acceleration profiles

i= Individual pedal acceleration profile

$pedal$= Average pedal position for the current profile

$pedal/s$= Individual profile pedal acceleration

The study not only established the proposed metric as a significant identifier of drivers' driving behaviour, but also compared the total aggressivity value of real-world driving data from the

[33] The unit for this metric is km/h.%/s$^{1/4}$.

UK and the US against a number of driving cycles[34]. The results confirmed the lack of representation of real driving behaviours in the selected driving cycles.

Table 7. A comparison between real-world driving and driving cycles maximum and mean values of total aggressivity metric

	Total aggressivity maximum value	Total aggressivity mean value
UK real-world driving data	343.2	55.5
US real-world driving data	332.4	47.1
FTP-72 driving cycle	59.7	30
LA92[35] driving cycle	71.9	39.5
NEDC driving cycle	41.8	19.9
US06 driving cycle	153.5	52.5

Source: (Daniel, Brooks and Pates, 2009)

Collecting real-world driving data led to studies with the aim of classifying and predicting drivers' driving patterns. Dealing with a broad range of real-world driving data, chassis dynamometers test results and data generated by driving simulators required careful attention to the data management process of these studies. Research by Inata (2008) is an example of using a driving database to develop a driver model (Inata, Raksincharoensak and Nagai, 2008). Whilst the model concerned also involved adopting the approach of using of large-scale real-world driving data to classify drivers, the primary aim was to provide personal assistance to drivers. A study by Amata and his team the following year was influential, as when classifying real-world driving, they concluded that there were significant differences between experienced and novice drivers in terms of pedal operation and fewer differences in relation steering wheel operation (Amata *et al.*, 2009). Moreover, after using real-world driving data, they proposed

[34] For details of driving cycles see (Barlow *et al.*, 2009).
[35] The California Unified Cycle.

two predictive models: firstly, a multiple linear regression analysis to predict driver manoeuvres, acceleration and deceleration behaviours and secondly, a Bayesian Network model, which predicts driver decelerating intentions. By using the driving simulator, the team validated their finding, with a 70% and 50% success rate for each model, respectively (Amata *et al.*, 2009).

The first telematics standard[36] was released by the Association of Equipment Management Professionals (AEMP) in 2010. As a result of the AEMP efforts, end-users were able to interact with and integrate key telematics data into their existing fleet management systems. Moreover, as has been explained above, the on-board diagnostic protocol and OBD II standard has become mandatory for every passenger car globally since the early 2000s. As a result of advancements made in mobile and GPS technology, the application of using OBD dongles and GPS sensors received huge attention from fleet operators by 2006 and widely by motor insurers by 2010. Hence, various studies using these technologies emerged (Bandivadekar, 2008; Berry, 2010; LeBlanc, Sivak and Bogard, 2010; Van Schangen *et al.*, 2011; Gonder, Earleywine and Sparks, 2012; Duarte *et al.*, 2015; Sentoff, Aultman-Hall and Holmén, 2015).

LeBlanc et al.'s (LeBlanc, Sivak and Bogard, 2010) use of large-scale real-world data logged via an in-vehicle monitoring system and OBD dongles and the highly cited work of Berry (Berry, 2010) on the effect of drivers' behaviour on the fuel consumption and performance of the US light-duty vehicles, established the effectiveness of this method in assessing the impact of drivers on fuel consumption. The study conducted by LeBlanc et al. investigated the assessed drivers' variation in fuel usage as a function of speed and acceleration by using a vast amount

[36] The standard is currently called the AEM/AEMP Draft Telematics API Standard.

of naturalistic driving data[37]. The study outcomes are presented in the table below (LeBlanc, Sivak and Bogard, 2010).

Table 8. The impact of speed and acceleration on fuel consumption

Mode	Acceleration (m/s^2)	Speed bins (km/h) <2.5	2.5 to 3.0	31 to 60	61 to 90	90 to 120	Over 120	All Speed
Significant acceleration	More than 1.05	0.1%	4.6%	5.2%	1.0%	0.2%	0.1%	11%
Notable acceleration	0.55 to 1.05	0.1%	1.6%	4.0%	2.3%	0.8%	0.2%	9%
Speed almost constant	-0.55 to 0.55	5.7%	3.3%	10.1%	20.0%	27.9%	10.7%	78%
Notable deceleration	Less than -0.55	0.2%	1.3%	0.5%	0.1%	0.0%	0.0%	2%
All modes		6%	11%	20%	23%	29%	11%	100%

Source: (LeBlanc, Sivak and Bogard, 2010)

The aim of the study conducted by Berry was to calculate and characterise driver aggressiveness using real-world driving parameters[38]. The authors proposed that the aggressiveness factor reflects driving behaviour, being positively correlated with fuel consumption and vehicle mass (Berry, 2010).

Equation 7. Driver aggressiveness factor value based on wheel work and vehicle mass, (Berry, 2010)

$$Berry\ aggressiveness\ factor = \frac{Wheel\ Work - Steady\ speed\ wheel\ work\ at\ average\ speed}{Vehicle\ mass}$$

[37] The project scope in terms of numbers: 117 cars, 342,941 kilometres (km) worth of driving data and there was a 36 to 42 days data collection period for 103 of these vehicles (LeBlanc, Sivak and Bogard, 2010).

[38] The project scope in terms of numbers: 15 cars and 12,620 kilometres (km) worth of driving data (Berry, 2010).

Where,

$$W_{wheel} = \left(\frac{E}{x}\right) = \frac{\int (Av^3 + Bv^2 + Cv + Mav)dt}{\int vdt}$$

$$Road\ load\ power = Av^3 + Bv^2 + Cv$$

In above equation, wheel work is the energy required at the vehicle wheels; the measure is always positive. Moreover, to calculate the road load power of a vehicle, the A, B, and C coefficient values can be measured based on the "SAE J 1263: Road Load Measurement and Dynamometer Simulation Using Coast down Techniques" procedure.

Another excellent example of using logged real-world driving data is work done by McGordon et al. (McGordon *et al.*, 2011) to develop a driver model. The study involved using dongles to collect real-world data to draw up a controller model, which was subsequently employed to investigate the effect of driver behaviour on fuel usage. The authors argued that existing driver models, such as the decision-making process template (Kawashima, Kobayashi and Watanabe, 2001) and manoeuvre-based driver models (Kiencke, Majjad and Kramer, 1999), are not designed to study the effect of individual driving behaviour on vehicle fuel consumption and emissions. They concluded that the proposed driver model combined with the vehicle emission model was able to produce a similar speed profile to real-world driving data (McGordon *et al.*, 2011).

While affordable monitoring devices, i.e. the OBD dongle and the GPS sensor, became widely accessible and have been used to collect real-world driving data in the last five years (Berry, 2010; LeBlanc, Sivak and Bogard, 2010; McGordon *et al.*, 2011; Van Schangen *et al.*, 2011; Malek, Brace and Liu, 2012), studies that use PEMS devices have remained an important part of quantifying the impact of factors affecting fuel consumption (Bokare and Maurya, 2013; Wyatt, Li and Tate, 2013; Ma *et al.*, 2014; Duarte *et al.*, 2015). A study on the effect of drivers' speeding and acceleration on vehicle emissions confirmed that whilst decelerating has such a subtle effect on tailpipe emission that it was hardly observed by the research team, speeding and acceleration have a high impact on emissions (Bokare and Maurya, 2013). These findings agreed with previous studies (De Vlieger, 1997; Frey *et al.*, 2001; Unal, 2002; Joumard *et al.*, 2006).

Table 9 The average tailpipe emission rate for different speed and acceleration ranges

Tailpipe emissions		**CO (%)**		**HC (ppm)**		**NOx (%)**	
Acceleration		a=1.0 (m/s^2)	a=1.6 (m/s^2)	a=.1.0 (m/s^2)	a=1.6 (m/s^2)	a=1.0 (m/s^2)	a=1.6 (m/s^2)
Speed range (m/s)	0-3	0.043	0.4	2.4	3.92	15.66	27.53
	3-8	0.006	0.008	1	1.06	2.00	2.46
	Above 8	0.29	0.865	5.29	10.49	31.08	44.77

Source: (Bokare and Maurya, 2013)

Since excessive speeding and acceleration increases fuel usage and driving emissions, classifying and modelling a driver's behaviour by using real driving data or through a driving simulator has become of particular interest (Han, Yao and Liu, 2014; Jaramillo and Narvaez, 2014; Ma *et al.*, 2014). Research conducted by Ma et al. logged the driving data of city buses in China[39] (Ma *et al.*, 2014). The authors argued that since the accelerating process follows a logical operation (shifting to higher gear by order), it is possible to develop a driving

[39] The project scope in terms of numbers: 3 cars (city buses), 6 drivers, and 100,000 kilometres (km) worth of driving data.

classification model based on driver gear shifting and acceleration habits (pedal position). The pattern classification method chosen by them is the C4.5 decision tree, which works based on input parameters, such as duration of the accelerating process (s), final vehicle speed after accelerating (km/h), average depth of acceleration pedal during the gear shifting process from first to fourth gear (%), and vehicle velocity during each shifting of gear. The model has 85% accuracy in classifying drivers based on their accelerating habits, and it can be used to train drivers to become fuel efficient (Ma *et al.*, 2014). Classifying real-world driving data by using pattern classification methods is not only used to detect abnormal driving behaviour, for it can also be utilised to detect engine faults and diagnose faults (Hasan *et al.*, 2011) in vehicle electronic parts, such as those in the fuel injection system(Yong He and Lei Feng, no date), or an uncalibrated oxygen sensor (Jaramillo and Narvaez, 2014). Jaramillo and Narvaez used a fuzzy classifier called the Learning Algorithm for Multivariate Data Analysis (LAMDA), to develop a diagnosis system that is able to detect driving behaviour (by how a car is operated) and to identify faults in the vehicle (Jaramillo and Narvaez, 2014).

The most recent approach to quantifying the effect of real-world driving on fuel consumption and vehicle emission is by integrating the effect of the road gradient and driving style. To accomplish this aim, the most relevant metric introduced in this subsection has been the Vehicle Specific Power (VSP), which presents a driver's vehicle speed and acceleration as well as road gradient and vehicle aerodynamics properties (Jiménez-Palacios, 1999). Since the adverse effect of speeding and excessive accelerating on fuel usage was established, the effect of driving style and road topography has been investigated (Wyatt, Li and Tate, 2013; Sentoff, Aultman-Hall and Holmén, 2015) in order to make a better estimation of vehicle emissions. Moreover, the VSP has proven to be useful, as demonstrated in a study conducted by Duarte in 2015, for establishing the association between vehicle certification data and real-world vehicle fuel consumption (Duarte *et al.*, 2015). The following table lists a number of methods that are highly relevant to this study.

Table 10. A brief list of the driving models aimed at assessing drivers eco driving performance and aggressive driving [40]

Study Approach	Author(s)	Commentary
The Vehicle Specific Power (VSP)	(Jiménez-Palacios, 1999)	The highly cited work by Jiménez-Palacios formalised the instantaneous power demand of the car related to its mass (as a kilowatts per tonne)
The drive mode fuel consumption model	Highlighted by (Saffarzadeh and Arjroody, 2003) and others	An early attempt to link driving behaviour and mode of driving including: idle, accelerating, decelerating, and cruising to fuel consumption
Ford Aggressivity value	(Nam *et al.*, 2003)	As part of the development of the Motor Vehicle Emissions Simulator (MOVES) software, the Ford team introduced an aggressive value based on EPA specific power (US Environmental Protection Agency, 1993), total number of acceleration events
MAHLE Powertrain Total Aggressivity	(Daniel, Brooks and Pates, 2009)	The work was aimed at developing a quantifiable model to measure drivers aggressivity. The model's primary goal was to link driver aggressivity to intensity of pedal usage. However, it would appear that they over simplified the formula they applied
Berry's aggressiveness factor	(Berry, 2010)	This comprehensive study by Berry reviewed all recent developments in the effects of driving style and vehicle performance on the real-world fuel consumption. Berry linked driver aggressiveness to the concept called wheel work and vehicle mass

[40] The following is not exhaustive and does not cover all methods, it is only based on the relevancy of the work to this project's aims and objectives.

2.5 The future of vehicle emissions testing under real-world conditions

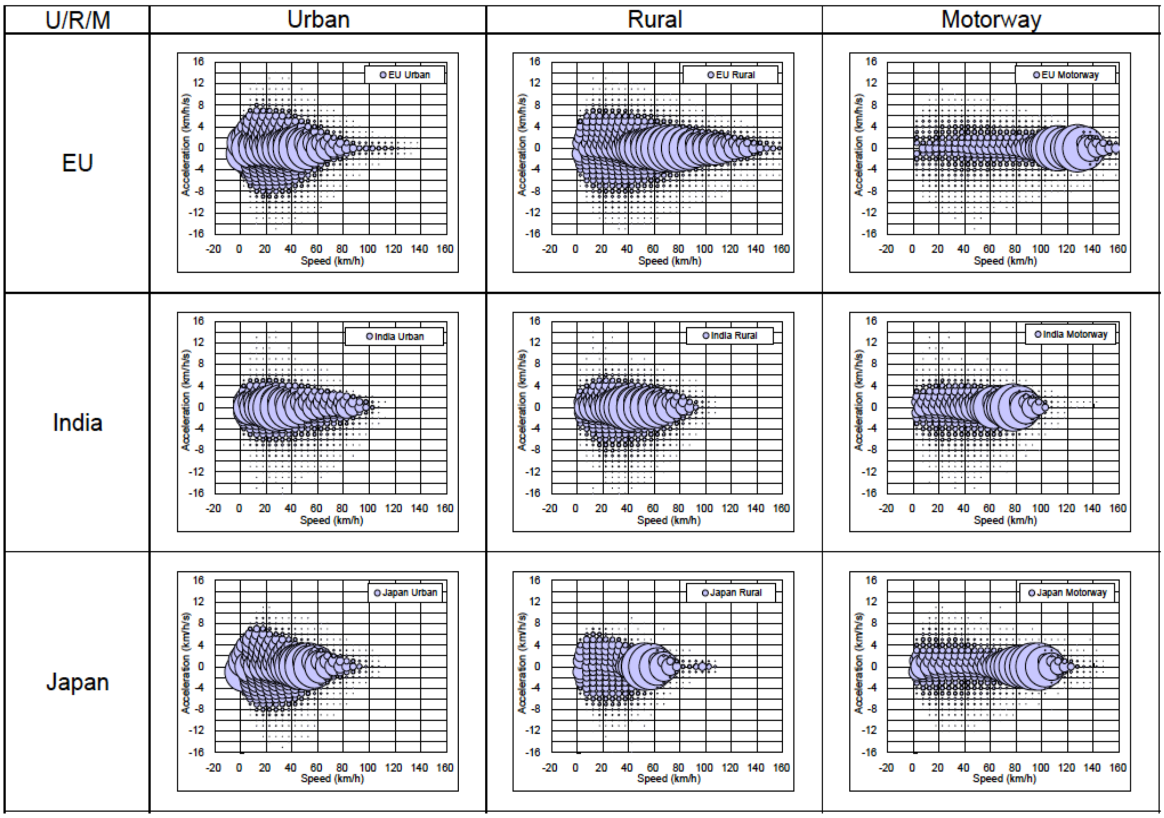
The current EU plan to improve the NEDC driving cycle is by modifying the driving cycle as well as adding new test conditions to reflect the impact of real-world driving conditions, such as testing a vehicle under the following conditions: AC on/off; daylights on/off and audio system on/off. These form part of a comprehensive proposal that is scheduled to be introduced with the Euro 6c regulation in September 2017 (Delphi, 2013). To provide a basis for the future modification, the Worldwide Harmonised Light Vehicles Test Procedures (WLTP) has been introduced by the EU. The WLTP is the global project that aims to standardise vehicle emissions test procedures in the Europe, India, Japan, Korea and the US. As part of this project, the Worldwide Harmonised Light Duty Driving Test Cycle (WLTC) was developed by collecting real-world driving data in all the aforementioned regions[41]. This is one of the largest global naturalistic driving study (NDS) with the purpose of capturing real-world driving conditions, to this day. A significant amount of information, such as vehicle types, road types, and driving conditions (rush hours, off-peak, weekend), were gathered globally (Tutuianu *et al.*, 2013). The project led to the development of the harmonised global test cycle. As part of the process of developing the aforementioned driving cycle, Tutuianu *et al.* (2013) gathered driving data from the EU, India, Japan, Korea and the USA. From the UK, 48,934 km worth of driving data were collected by 10 passenger cars and 12 Light Duty Commercial Vehicles (LDCV) (Tutuianu *et al.*, 2013).

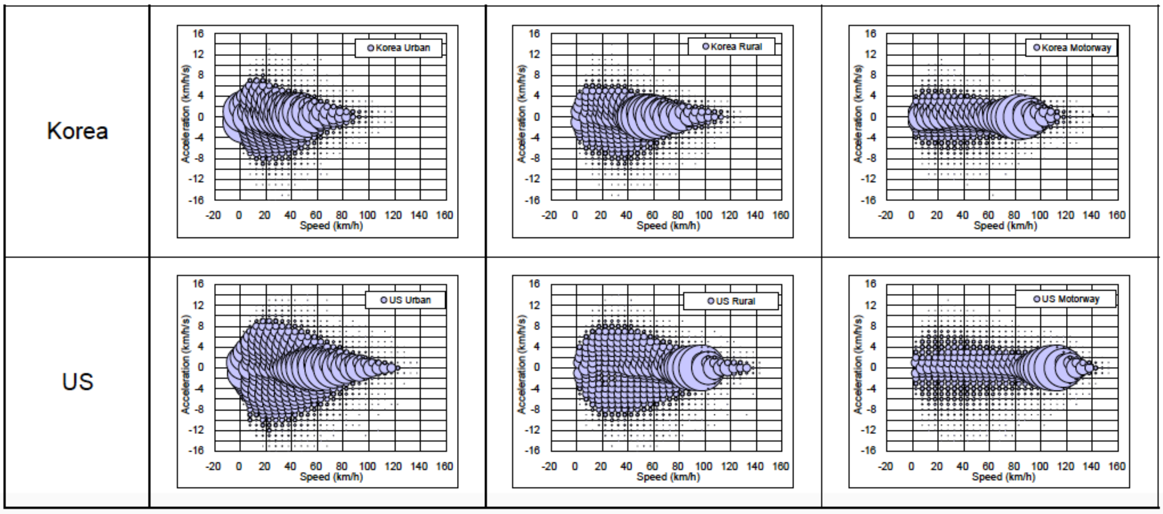
According to the technical report by Tutuianu et al., one of the most distinguishable driving differences observed from the collected driving data was drivers' gear shifting strategy and

[41] The project scope in terms of numbers: 394 cars and 765,000 km worth of driving data.

behaviour. The report suggested that gear shifting behaviour is more influenced by individual driving habits rather than vehicle transmission design (Tutuianu *et al.*, 2013). Another important conclusion from the report is that it establishes the fact that driver behaviour is widely diverse among drivers globally (see Figure 17), which is important for global fleet or insurance companies when they want to compare drivers. It highlights the fact that driving models and methods to score driers have to be adjusted to local settings and conditions. Figure 17 below illustrates the analysis conducted by Tutuianu *et al.* (2013) to present the significance of driving pattern differnces in the following regions: the EU, India, Japan, Korea and the USA.

Figure 17. Distribution Speed (km/h) and acceleration (km/h/s) (Tutuianu *et al.*, 2013)





As shown in charts above, driving characteristics are highly influenced by the region, highway regulations and infrastructure differences. Moreover, driving behaviour varies between three road types urban, rural and motorway even within one region or a country. Despite the speed limit on urban roads being universally below 100 km/h, there are notable cross-regional differences. For example, whilst US drivers' acceleration range is between ±12 km/h/s, this is not the case in the EU or Japan (±8 km/h/s) and it is even narrower in India (±4 km/h/s) due to high traffic density.

The NEDC is being superseded by the WLTC driving cycle in that the latter has longer duration (total distance of driving on a chassis dynamometer) (Table 11). In addition, it is a better representation of real-world driving as there is a continuous change in driving speed. Moreover, the maximum acceleration and speed has been modified from 1 m/s^2 and 120 km/h to 1.6 m/s^2 and 131.6 km/h in order reflect observed behaviour (Tutuianu *et al.*, 2013). However, the credibility of WLTC has been questioned since it was first introduced. There are a number of valid concerns about the cycle, such as the slow rate of increase in vehicle speed, no representation of hill driving and according to the Association for Emissions Control by Catalyst (AECC) study, there are substantial differences between emissions recorded from the

WLTC test and those measured by the PEMS device on the road (May, Bosteels and Favre, 2014). Currently, the Euro 6c test standard and the WLTC driving cycle are planned to come into effect in 2017.

Table 11. A comparison between the NEDC and WLTC driving cycles

	NEDC	WLTC
Duration (s)	1,220	1,800
Length (km)	1,106	2,326
Idle time (%)	33	13.4
V_{max} **(km/h)**	120	131.6
$V_{average}$ **(km/h)**	31.6	46.3
Max acceleration (m/s^2)	1	1.6

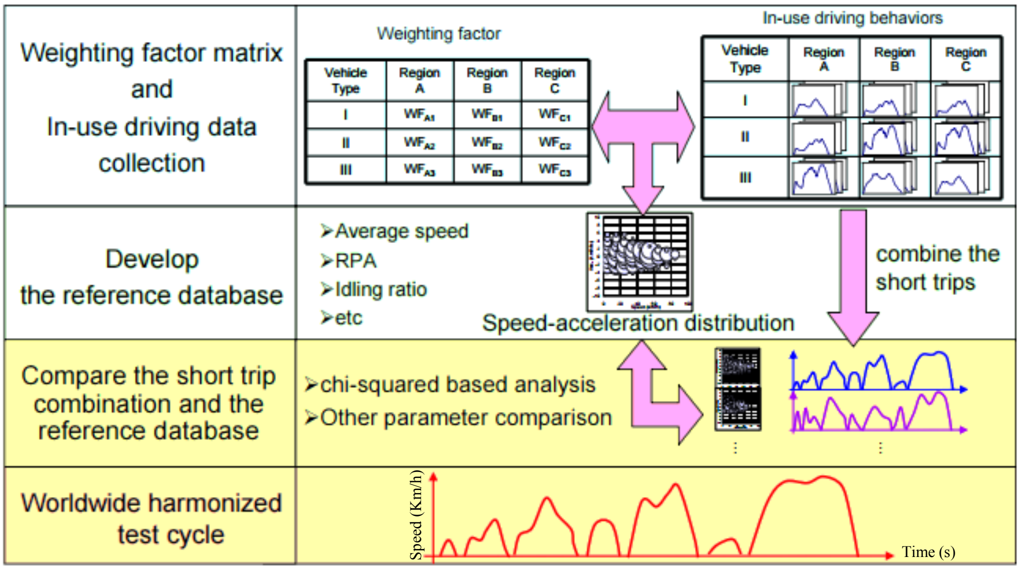


Figure 18. Development of the worldwide harmonised test cycle[42]

[42] The WLTC in red is the final speed – time driving test cycle (directly from the source (Tutuianu *et al.*, 2013)]).

Figure below is an example of the Worldwide Harmonised Light Vehicles Test Procedure cycle for class 3 vehicles, which are vehicles driven in the EU and Japan.

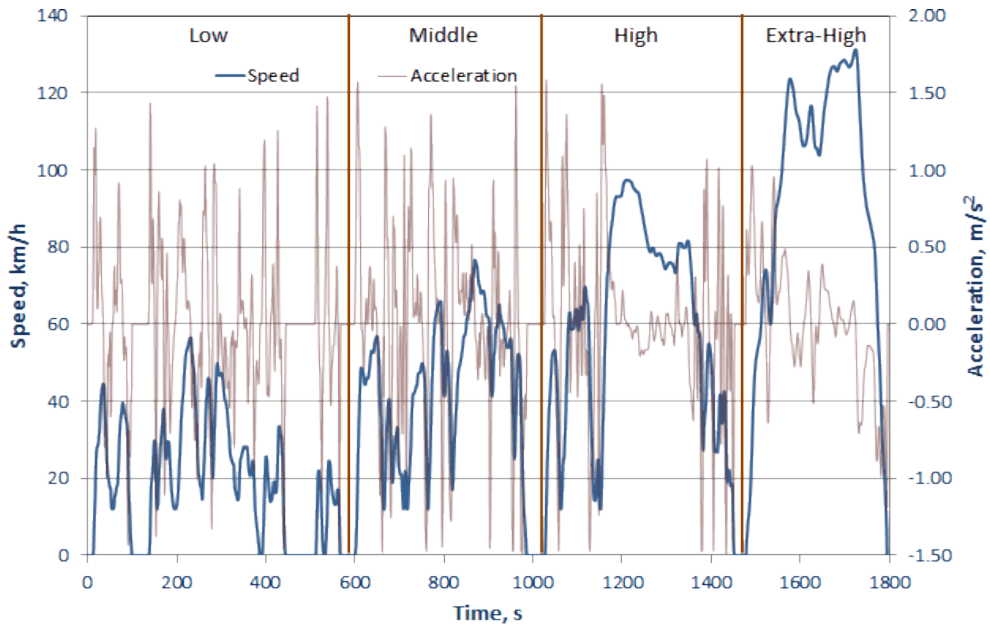


Figure 19. WLTP Cycle for cars in the EU and Japan (May, Bosteels and Favre, 2014)

In addition to the WLTP project, the EcoDriver project[43] was an EU-based one that was initiated with the purpose of reducing CO_2 in the transport sector by providing drivers with their driving performance information. The four-year long project was aimed at promoting eco-driving behaviour among drivers by providing them with real-time eco-driving recommendations based on their driving behaviour (relaxed or aggressive) and their vehicle characteristics (vehicle types and powertrains). In this project, a range of vehicle types, powertrains and driving profiles were tested to finalise the most customised and appropriate feedback for any given conditions. Through various analyses conducted by multiple teams on the impact of providing feedback to drivers the following conclusions stand out: these systems reduce cruising behaviour on highways by 4% and up to 15% on other roads. Moreover, they

[43] The project held its last event in March 2016. Details can be found at: http://www.ecodriver-project.eu/ [last accessed March 2016].

helped to reduce number of harsh accelerations and decelerations by 5% to 10%, respectively (Mejuto, 2016).

Table 12. Summary of all the partners participating in the EcoDriver project

Test sites	Partner	Route type	Car types	Mileage covered
Sweden	VTI	Urban, rural, motorway	-	90 km
England	University of Leeds	Urban environment	10 hybrid buses	11 km
Germany	BMW	Urban, rural, motorway	BMW passenger car	-
Germany	TomTom	Free route	Light Commercial Vehicle (LCV)	-
Germany	ika	Urban, rural, motorway		-
Germany	Daimler	Urban, rural, motorway, hilly routes	-	-
Netherlands	TomTom	Free route	LCV's: Fiat, MB, VW Truck: MB	-
France	IFSTTAR	Urban, rural, motorway	-	25 km
Italy	CRF	Urban, rural, motorway	-	52 km
Spain	CTAG	Urban, rural, motorway, hilly routes	-	-

2.6 The impact of the human factor on the causation of road collisions

In the field of road safety, the priority has always been to identify risk factors through investigating the causation of road collisions. Collision-prone locations were the basis of conducting these studies and finding stable parameters to address the accident causation were usually the conclusions of this approach. Influential factors can be classified into two groups: human factors (i.e. impaired driver, driving under the influence of alcohol and/or drugs) and non-human factors (external reasons: poor road lighting, inclement weather, etc.).

The early interpretation of the role of driver in travelling safely was based on his/her skillfulness in the act of controlling a vehicle under dangerous circumstances. That is, by steering effectively (steering away to avoid collisions) , accelerating and decelerating (speeding to drive away from a hazardous situation or making an emergency stop brake) (Gibson and Crooks, 1938). By 1949, the influential work of Tillmann and Hobbs had led to the concept of accident-prone individuals and the influences of psychiatric and social background in determining individual involvement in accidents (Tillmann and Hobbs, 1949). By considering driver background and demographic, the authors concluded that drivers who are often involved in car crashes have a higher degree of aggression and sociopathic behavioural traits (Tillmann and Hobbs, 1949; Shinar, 1998).

The concept of exhibiting aggression towards other road users while driving was discussed as early as 1968 in a book called "Aggression on the road" by Parry (Parry, 1968). Whitlock (Whitlock, 1971), drawing on Parry's (1968) approach, investigated the relationship between social violence with road accidents in Great Britain. He concluded that 85% of all crashes in Great Britain were due to driver aggression while driving. A reviewed study conducted by

McKenna in 1982 on the methods used to quantify the impact of human factors on road accidents concluded that including psychological tests as part of statistical analysis to identify collision-prone drivers is useful because it helps to characterise traits of behaviours that lead to car crashes.

The domination of behavioural psychology studies in the latter half twentieth century as part of a systematic approach to understanding human behaviour in relation to its environment, provided a basis for investigating the cognitive stance of drivers (Altman, Wohlwill and Everett, 1981). This also led to the development of driving behaviour models with an emphasis on the cognitive status of driving tasks. By the end of the twentieth century, a substantial number of driver and driving behaviour models, frameworks and theories had been proposed by scholars with different academic backgrounds and research interests. These models were developed largely in an attempt to explain the mechanism of how a driver's behaviour (driving performance) can be a reason for car crash involvement. Model here refers to the mechanism or flow chart developed to explain driver driving behaviour rather than computer models.

To signify the work of early researchers in this field (1959 – 2000), the most relevant and highly cited studies are investigated and tabulated to be part of a background study about the importance of studying driving behaviour. For the period 2000 to 2009, the significant studies are discussed in detail. Table 13 contains a summary of the relevant driving behaviour models and theories from 1959 to 1999. These studies reveal the evolution of the research community's perspective towards understanding the role of drivers' driving behaviour (Michon, 1985; Ranney, 1994; Shinar, 1998; Carsten, 2007).

Table 13. A summary of the important driving models introduced from 1959 to 1999[44]

Study Approach	Author(s)	Commentary
Test-based model	(Conger *et al.*, 1959)	The study compared two groups of drivers with (group 1) and without (group 2) road collision history, with no intelligence or psychophysiological differences between the two groups. It emerged that drivers who had been involved in car accidents had problems in terms of lack of control in hostile situations and lower tension-tolerance.
Risk-speed compensation model	(Taylor, 1964)	The study investigated drivers' level of anxiety while driving by using Galvanic skin responses of 20 drivers while driving. It suggested that emotional tension or anxiety can affect driving task performance.
Information flow model	(Kidd and Laughery, 1964)	One of the early computer-based driving models developed was based on information available to drivers for different driving scenarios, such as at junctions.
Skill-based model of driving	(Näätänen and Summala, 1976)	Defining a driver based on his / her personality, motivation, perception, caution, skilfulness and travel reasoning.
The utility maximisation model (A decision theory model of danger compensation)	(O'Neill, 1977)	The model was built based on the assumption that a driver always has a goal while driving, such as time taken to reach the speed limit, showing off, etc., which governs his / her risk-taking behaviour.

[44] The classification is inspired by the work of (Michon, 1985; Ranney, 1994; Shinar, 1998; Carsten, 2007).

Study Approach	**Author(s)**	**Commentary**
Driver steering control models	(McRuer *et al.*, 1977)	Feedback control model based on a driver's steering operation.
DRIVEM event detection control model	(Wolf and Barrett, 1978)	Driver-Vehicle Effectiveness Model designed to control vehicle. The flowchart-based model shows drivers' driving tasks.
Motivational model of driving	(Johnston and Perry, 1980)	The model aimed to relate accident causation to a driver's personality traits.
Risk compensation model or the Risk Homoeostasis model.	(Wilde, 1982)	The servo-control model based on the idea that a driver's risk taking behaviour is only controllable by his / her perception of acceptable risk (risk target level).
Risk- avoidance model	(Fuller, 1984)	The model provided a framework where a driver has to deal with a hazardous situation based on his / her capability and awareness.
The hierarchical control structure model	(Michon, 1985)	The model encapsulates three levels of road users' tasks (in a hierarchy format): strategic level, manoeuvring level and controlling level.
Theory of planned behaviour	(Ajzen, 1985)	Planned behaviour questionnaires have been constructed based on this theory. The work is highly cited and followed due to clarity of the model. The theory suggests that a driver's intention is a predictor of his / her behaviour.

Study Approach	**Author(s)**	**Commentary**
Zero-risk model based on visual attributes and reaction times	(Summala, 1985, 1988)	The model suggests that for most situations drivers learn how to react to avoid collisions and therefore, driving risk behaviour needs to be considered in terms of what the driver perceives as being within the margins of safety.
Developed hierarchical risk model	(Molen and Botticher, 1988)	The hierarchical risk model constructed based on different levels of controlling the vehicle.
Theory of driver error in road safety	(Brown and Groeger, 1990)	The article suggested that driver errors are related to the level of expertise of the driver.
Generic error-modelling system (GEMS)	(Reason *et al.*, 1990)	The study showed that inattention (distraction) and over attention (preoccupation) are part of driver errors.
Visual attention research based on driver useful field of view	(Owsley *et al.*, 1991)	The study established a strong association between a driver's useful field of view (UFV) and culpability for accidents occurring at junctions.
Generalisability of four classes of predictors of accident involvement (a meta-analysis)	(Arthur, Barrett and Alexander, 1992)	Results of a meta-analysis concluded that the following classes of variables were the causation of accident involvement: information processing, cognitive ability, personality and demographic factors.

Study Approach	**Author(s)**	**Commentary**
The frustration-aggression model	(Shinar, 1998)	The study focused on the characteristics of aggressive drivers and examined responses of aggressive drivers in five scenarios, such as honking, passing through red lights, and delays in travel time.
Modelling and predicting driver behaviour	(Pentland and Liu, 1999)	The study used a dynamic controller and driving simulator to model and predict drivers' behaviour through the Markov chain model.
Evaluate the impact of behaviour change techniques	(Goldenbeld, Levelt and Heidstra, 2000)	The study investigated the impact of psychological incentives in positively changing drivers' behaviour

In a critical review paper by Michon (Michon, 1985), the author argued that many driving behaviour models and accident involvement models up to then had failed to incorporate new ideas. He suggested that there was a need for driving models that present a driver's cognitive status in an efficient manner (Michon, 1985). In another influential work, Ranney (Ranney, 1994) reviewed all accident involvement driving behaviour models from the 1930s to 1994 and examined them in terms of their credibility, coming to the conclusion that none of the driving behaviour models were comprehensive or accurate. He argued that the low correlation between model and real-world results, low sample sizes and lack of identifying predictors of safe driving led to models that relied on post hoc explanations. The author concluded that an integrated driving behaviour model containing a task-based model, as well as motivational and cognitive driving behaviour, is an effective way for modelling driving behaviour (Ranney, 1994).

Towards the end of the twentieth century, research conducted by Shinar (1998) had a significant impact, since the study investigated the concept of aggressive driving and the characteristics of an aggressive driver. The researcher made a distinction between aggressive driving and the aggressive driver, arguing that while the latter can be identifiable by personality, and his/her behaviour, the former pertains to road rage while driving and hence, has traffic offences associated with it in many countries. The resultant model proposed by Shinar (1998), namely, the "frustration-aggression model", provides useful predictive knowledge about a driver's aggression on the road (Shinar, 1998).

In his influential work on methodological problems in (psychological) research about traffic accidents, Wåhlberg (2003) argued that most studies have been subject to methodological deficiencies, in relation to three fundamental aspects: the reliability of accident predictor models, the sample time period, and whether driver behaviour is the primary cause of crashes (Wåhlberg, 2003). By 2005, the final version of the Fuller driving model called the "task–capability interface model", had been developed based on driving task demand and difficulties. Under this model, it was contended that there is a correlation between the difficulty of a driving task and the capacity of the driver to accomplish the task safely (Fuller, 2005). In order to address the need for an integrative model, many researchers have suggested a conceptual framework in relation to safe driving behaviour (Toledo, 2003; Strecher *et al.*, 2006; Toledo, Shiftan and Hakkert, 2007).

The advancements made in the field of transport, i.e. Intelligent Transport Systems (ITS) and driving safety technologies, such as the Advanced Driver Assistance System (ADAS), pose new challenges regarding driver behaviour and human-machine interaction in relation to safety. While the aim of these systems is to help the driver with the driving task, they are yet to be installed in every car class coming onto the market. Moreover, using new technologies, such as car stereos and mobile phones by drivers has led to the initiation of new studies on the impact of driver distraction and awareness of road safety (Klauer, 2005; McEvoy, Stevenson and Woodward, 2006; Amditis *et al.*, 2007; Cacciabue, Re and Macchi, 2007; Chen, 2007).

Despite decades of progress in the field of studying driver/driving behaviour and accident research, unresolved problems remain in this area. Wåhlberg (2009) reviewed every critical study that was published prior to 2009 and highlighted a number of issues that were repeated in a significant number of studies conducted in the field of driving behaviour and road safety. According to the author, the poor outcomes of the proposed driving models in predicting and quantifying the role of the driver in accident involvement, are mainly due to the unfounded speculative assumptions about driver behaviour and the cause of accidents. He added that the lack of consistency with studies is because of researchers' blind faith in certain methods and incorrect perceptions about the meaning of road safety facts and figures (Wåhlberg, 2009). Understanding the impact of the human being on road safety is evolving with technology. Evidence of this new breadth of research regarding the studying of driving behaviour is the mimicking of human habits in new autonomous vehicles, work that has been develped by Naranjo et al. (2008). Such research can lead to the identification of good and bad human behaviours, with the former being introduced to these autonomous vehicles. For instance, prior to changing lane, a good driver with safe driving habits positions his or her car close to the side of the current lane and when the situation is safe he or she moves tino the next lane (Naranjo *et al.*, 2008). Another example is how cautious safe drivers are when driving on narrow roads in relation to oncoming traffic (Amata *et al.*, 2009). In sum, adding these details to autonmous vehicle driving models will help make passengers feel safe when using this new technology.

Between 2006 to 2010 computer processing power expanded massively, for instance, intel releeased its first 6 core processor for desktops and i3 in 2010. These advancements made it easier for resaerchers to analyse, model and compute bigger data and hence, be able to process large amounts of real-world driving data for various purposes. This also meant that the field of investigating driving and driver behaviour expanded dramatically, in particular, the last seven years. The most recent published works is this period have been highly speicalised and focused on specific parts of this topic and despite the vast number of publications, only a few tens of them are directly relevant to this work. This is the reason why in following section the focus is on the relevant literature for the current research.

2.7 Driving behaviour and road safety

The loose definitions of aggressive driver and/or driving behaviour has led to a tremendous amount of miscommunication among the research community. This issue was raised by Shinar (1998), whereas aforementioned he distinguished aggressive drivers from aggressive driving. According to his definition, the aggressive driver is someone who exhibits sequences of aggressive behaviours towards others as a result of frustration, whereas aggressive driving is defined as the manner in which the vehicle is operated endangers people and properties (Shinar, 1998). On a similar basis, based on a comprehensive review of the literature addressing aggressive driving, Tasca (2000) proposed the following description of aggressive driving: "A driving behaviour is aggressive if it is deliberate, likely to increase the risk of collision and is motivated by impatience, annoyance, hostility and/or an attempt to save time" (p.9) (Tasca, 2000).

The importance of consistent and unambiguous usage of a single definition was highlighted by Dula and Geller (Dula and Geller, 2003). Their study covered other misused terminologies, such as risk, emotions, and behaviour. As a result, the authors suggested dangerous driving should be defined as pertaining to the following traits: "intentional acts of aggression toward others, negative emotions experienced while driving, and risk-taking" (Dula and Geller, 2003). One of the most recent interpretations of the term 'aggressive driving' is that of the National Highway Traffic Safety Administration (NHTSA, 2011), which states that aggressive driving "is generally understood to mean driving actions that markedly exceed the norms of safe driving behaviour and that directly affect other road users by placing them in unnecessary danger" (American Automobile Association (AAA), 2009).

Concerns about the rising number of aggressive driving incidents in the EU led to a global survey study[45] conducted by the EOS Gallup Europe to investigate the viewpoints of 13,673 drivers from 23 countries[46] in relation to aggressive driving behaviours (EOS Gallup Europe, 2003). It was reported that many drivers had been victims of other drivers' aggression. Moreover, the study suggested that cultural differences between countries have a strong impact on what is acceptable as aggressive driving or normal driving behaviour. That is, the perception of people who were interviewed about what they perceived as aggressive driving behaviour in their countries varied substantially (EOS Gallup Europe, 2003). The study involved using tolerance (yes or no) and irritation (on a scale of 1 to 5) in order to establish what behaviours, constitute aggressive driving. Based on the survey results, the following graphs (Figure 20 and Figure 21) were plotted for the EU-15, the USA and Australia. The data was collected between 2002 and 2003 (EOS Gallup Europe, 2003).

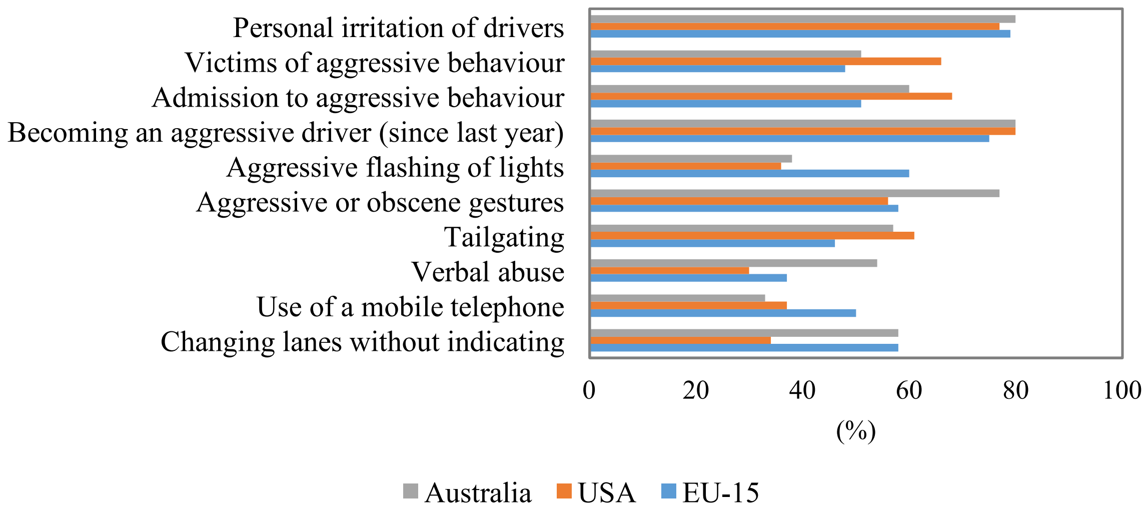


Figure 20. Responses to ten driving behaviour questions associated with aggressive driving[47]

[45] Telephone polling and face to face interviews.
[46] The EU-15 member states plus the Czech Republic, Cyprus, Slovenia, Argentina, Russia, Japan, Australia and the United States (EOS Gallup Europe, 2003).
[47] Data collected from the EU-15, Australia and the US drivers (data source: (EOS Gallup Europe, 2003).

The authors of the work concluded that whilst the results of the aforementioned survey exhibited a high correlation (56%) between drivers who had had at least one accident in the past three years (prior to the date of the study) and perceived aggressive behaviour from other road users (EOS Gallup Europe, 2003). Arguably, the conclusion that others' aggressive driving behaviour and accident involvement are directly linked is rather tenuous. The results of the UK drivers in the above survey show that they had a higher tendency to exhibit aggressive gestures, such as being aggressively pursued or verbally attacked (77%) and became more aggressive while driving (80%) over the year prior to when the survey was carried out in November 2002 (EOS Gallup Europe, 2003).

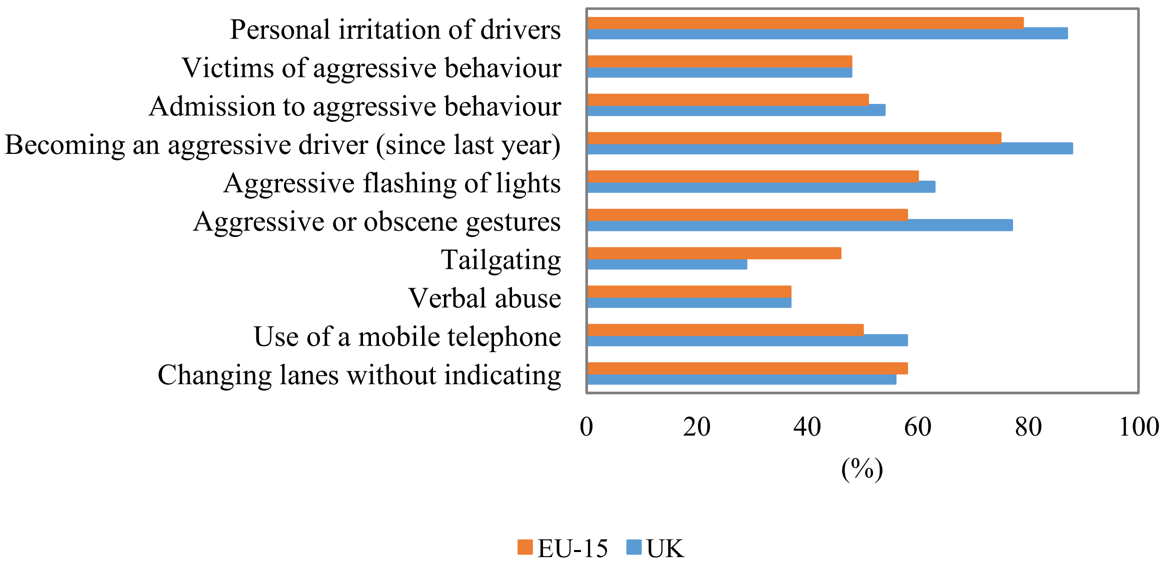


Figure 21. Responses to ten driving behaviours questions associated with aggressive driving[48]

[48] Data collected from the EU-15, Australia and the US drivers (data source: (EOS Gallup Europe, 2003).

As Wåhlberg (2009) concluded, it is not always the case that drivers' driving behaviour (aggressive driving, driver impairment,error, etc.) has been the only reason that the accident occurred (Wåhlberg, 2009). In fact, according to Reason (Reason *et al.*, 1990) and Aarts et al. (Aarts, Wegman and SWOV Institute for Road Safety Research, 2008), crashes occur based on a dangerous action and driver errors over a number of stages. In Figure 22, the thick red arrow represents the occurrence of a road collision. In order for a crash to happen, driver errors have to continue at every level of the driving task, Hence, at every stage, there is a possibility to stop the crash from occurring. In sum, the occurrence of an accident is often related to the continuation of latent errors by drivers (Reason *et al.*, 1990).

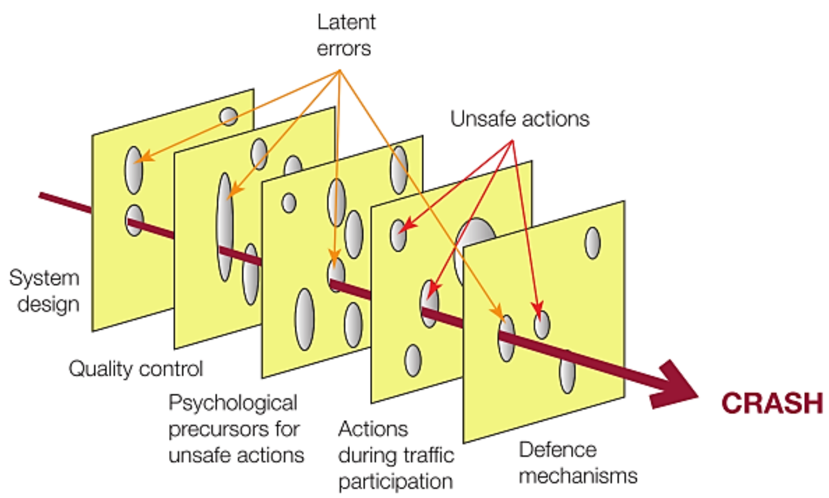


Figure 22. The schematic demonstration of the development of a crash, based on Reason's model[49]

According to the Department of Transport in the UK, the most common reason in the top ten leading causes of a road collision is the inattentiveness of the driver. The driver's inattention in the field of traffic safety and accident prevention has an extensive range of meanings, such as restricted and prioritised attention as well as driver distraction or diverted attention (Regan,

[49] By (Reason *et al.*, 1990; Aarts, Wegman and SWOV Institute for Road Safety Research, 2008).

Hallett and Gordon, 2011). Consequently, driver inattentiveness has been interpreted variously in terms of driver actions / reactions while driving (Klauer, 2005; Regan, Hallett and Gordon, 2011). The following graph (Figure 23) is plotted based on the published records of collision history in the UK between 2013 and 2015 (Department for Transport, 2015). Factors, such as failing to look properly (32%), failing to judge other drivers' driving path and vehicle speed (27%), and losing control over a car (22%), have the highest percentage for causing collisions on UK motorways. The causations of collisions for other road types[50] in the UK are failing to look properly (42%), failing to judge other drivers' driving path and vehicle speed (21%), and losing control over a car (14%). Additionally, careless and reckless driving has a higher percentage for other road categories (16.5%) than highways (10%) (Department for Transport, 2015). The following are not mutually exclusive since more than one factor might contribute to an accident and therefore, the percentages do not add up to 100% (Graves *et al.*, 2014; Department for Transport, 2015).

[50] In the UK, roads are categorised into four groups other than motorways: A roads (large scales of road connecting two areas), B roads (roads that feed other roads to the A roads), classified unnumbered (short roads connecting A/B roads to unclassified roads) and unclassified (city or local roads) (Department for Transport, 2012).

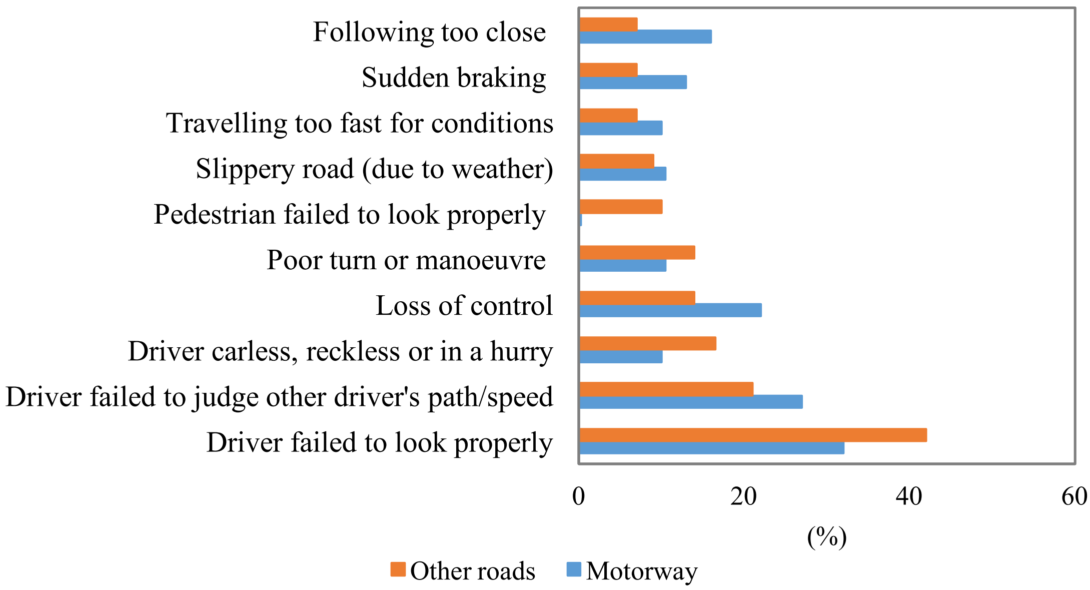


Figure 23. Major factors causing road accidents in the UK (2013 – 2015)

To conclude, factors such as aggressive driving, a driver's psychological status, the inattentiveness of a driver and latent errors of a driver are partially able to explain the reasons why some drivers might be involved more often in car crashes than others. However, quantifying drivers' driving performance in relation to these parameters requires a better understanding regarding driving conditions (whether a driver is distracted, tired or shows a lack of attention regarding the road ahead) and how the vehicle is physically controlled, i.e. in terms of speeding, harsh acceleration and deceleration events, cornering, gear shifting, manner of steering and pedal usage. In the next subsection, studies that associate road collisions with vehicle speed and acceleration are reviewed.

2.8 Vehicle speed and acceleration in conjunction with road safety

The main focus of this section to review studies that have investigated the impact of a driver's vehicle speed, change of speed (acceleration) and/or breaking the speed limit as part of his or her accident involvement. The Department for Transport states that for an aware and fit driver the reaction time is approximately 0.67 seconds and depending on the speed a driver is travelling at, the stopping distance is calculated according to this reaction time (Driver and Vehicle Standards Agency, 2015).

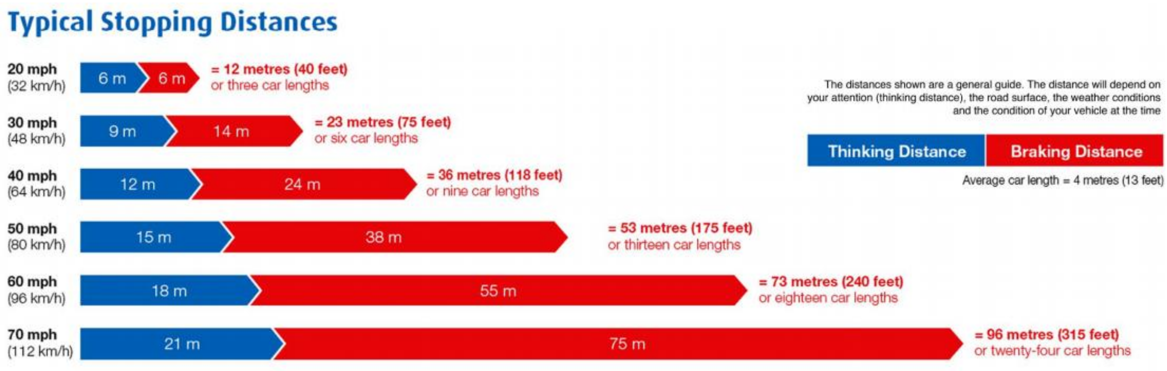


Figure 24. Typical stopping distances based on 6 initial driving speeds (Driver and Vehicle Standards Agency, 2015)[51]

The relationship between driving speed and the severity of injury has been established by many studies. For instance, Fildes, Rumbold and Leening (1991) conducted a multivariate analysis to determine the significate of parameters such as driver age, gender, purpose of the journey, number of occupants, weather, accident involvement, etc. The team established that gender difference does not have a sigfificant impact in relation to speeding attutide (Fildes, Rumbold and Leening, 1991). The highly cited work of Kloeden et al. from the University of Adelaide used a case control study to establish the relationship between being involved in a road collision and driving in a 60 km/h speed limit zone (Kloeden *et al.*, 1997). They conducted 604 cases

[51] From Brake the road safety charity, details at: www.brake.org

where drivers and cars were matched according to “location, direction of travel, time of day, and day of week”, after measuring their speed by speed guns they matched cases with historical records, which led to their seminal conclusion as follows:

“In a 60 km/h speed limit area, the risk of involvement in a casualty crash doubles with each 5 km/h increase in travelling speed above 60 km/h” (Kloeden *et al.*, 1997)

A questionnaire study by TRL led by Quimby et al. involved a thorough analysis of the factors that influence a driver's decision in choosing his or her driving speeds (Quimby *et al.*, 1999). In this work, instead of the case matching approach taken by Kloeden et al. (1997), they develped a questionnaire to measure the relationship between factors influencing driving behaviour and choice of speed, subsequently identifying relationship between speed and accidents (Quimby *et al.*, 1999). In their work, the speed of just cars was recorded at 24 sites exluding motorways. To those who passed the obsevered sites were asked 31driving behaviour related questions. This approach was adopted by the Highways Agency in the same year, where Maycock, Brocklebank and Hall (1999) investigated the impact of choice of speed, overtaking and emergency braking, by measuring these values on the road through postal questionaires. While Maycock et al. (1999) used the data to design better roads, Quimby’s team’s objective was to investigate the association of driving behaviour with crashes. By developing linear regression models, his concluded that younger derivrs, high yearly mileage drivers in large cars, commuters and single occupied cars tend to driver faster than others. Furthermore, based on their comprehensive list of factors that impact on driving behaviour (Table 14), the team examined and ranked the most influential factors regarding choice of speed. Quimby concluded that site effect (road characteristics), driver age, exposure and psychological parameters were all highly influential on driving behaviour (see table 14) (Quimby *et al.*, 1999). This study has been used as a reference work in this project for two reasons, first, it was conducted in the UK, which makes the work relevant to this project setting and secondly, the work methodology and findings had been agreed with by other researchers since then (Elliott, Armitage and Baughan, 2003, 2007; Aarts and van Schagen, 2006; Elliott and Thomson, 2010; Wåhlberg and de Winter, 2012; Zhao *et al.*, 2012).

Table 14. influential factors on driving behaviour

Factors	Class	Sub-group
Driver factors	Demographics	Age/driving experience Sex Exposure (annual mileage, type of road, light/dark etc.) Occupational group
	Visual ability	Static and dynamic acuity Visual field Field dependence
	Driving skill	Car handling ability Hazard perception Judgemental skills
	Psychological factors	Risk tolerance Social/driving deviance Thrill/sensation seeking
	Temporary states	Mood Fatigue Impairment due to drink or drugs Illness Speed adaptation
Other factors	Trip characteristics	Length, purpose and urgency
	Car characteristics	Performance and comfort
	Road environment	Road type Design speed Speed limit Enforcement levels Maintenance Presence

	Environmental factors	Presence of passengers Presence of pedestrians Time of day Signs/warnings Local knowledge (familiarity) Weather

Source: (Quimby *et al.*, 1999)

The work of Taylor, Lynam and Baruya has become an industry reference despite in being published in 2000. In this work, Taylor's team focused two methods, namely, road-based studies and driver-based studies, to assess the effect of drivers' speed choice on the frequency of road accidents (Taylor, Lynam and Baruya, 2000). In their road-side based studies, the research team investigated the speed of all drivers, traffic and pedestrian flows as well as road layouts. Their objective was to predict the number of injury accidents on both urban and rural segments. Regarding the driver-based side, their responses to questionnaires were matched with their accident history and their actual recorded speed on observed road, The authors managed to quantify the association between choice of speed and drivers' personality (Taylor, Lynam and Baruya, 2000). According them, the results of their extensive studies were in line with the following two important statements:

1. Reduction of speed by drivers travelling faster than the traffic average speed leads to potential reduction of deaths and injuries;
2. Road characteristics is a determining factor in reducing accidents

This confirmed that speed limit reduction policies would provide higher safety for road users (Taylor, Lynam and Baruya, 2000).

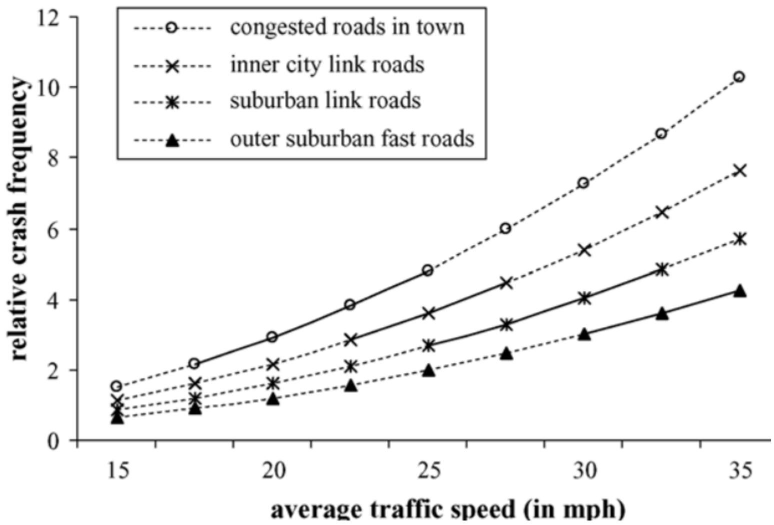


Figure 25. Established and forecasted relationship between average traffic speed and relative crash frequency (Taylor, Lynam and Baruya, 2000)

Following Taylor's team findings many institutions and road safety operators began to build upon and validate these findings. For example The Institute of Transport Economics (TOI) developed a model containing six equations that relates the change of average speed of traffic to the number of road accidents (Elvik *et al.*, 2004). They concluded that the severity of the accident exponentially increases with vehicle speed, based on collision data. This led to them calling for speed limit reductions, particularly in urban areas, so as to reduce accident levels.

The third Social Attitudes to Road Traffic Risk in Europe (SARTRE) survey reported the views of 1,000 drivers in 23 EU states about speed and speeding. Quimby (2005) summarised the questionnaire results with the remarks that, firstly, UK drivers had a negative attitude towards speeding as, 87% of the respondents associated it with accident involvement. Secondly, only 5% of UK drivers enjoyed driving fast compared to their counterparts in France, 8%, Germany, 11%, and Denmark, 15% (Quimby and Department for Transport, 2005).

A study by Aarts and van Schagen (2006) reviewed all the existing studies that had investigated the relationships between driving speed and the risk of getting involved in road crashes. The authors concluded that, without exception, the higher the vehicle speed, the higher the chance of getting involve in a road collision. However, they reported an inconsistency in the results of studies investigating the rate of accident reduction due to a decrease in vehicle speed (Aarts and van Schagen, 2006).

As aforementioned, in Fuller's (2005) driving model (see section 2.6), it is assumed that drivers will take risks to the extent that they believe that they can maintain control whilst driving (controlling the vehicle) (Fuller, 2005). In 2006, Fuller et al. combined the Musselwhite (2006)[52] drivers' risk-taking classification, which involved introducing four types of driving according to speed and risk-taking behaviour. Driver groups were developed based on these two models as follows: 1) the high-risk threshold drivers, 2) the low-risk threshold drivers 3) opportunistic drivers, and 4) reactive drivers (Fuller *et al.*, 2006). The following table was developed by Fuller and colleagues based on the two studies' conclusions.

Table 15. Four classes of drivers based on risk taking and speeding behaviour,

Drivers groups	Demographic	Characteristics
The high-risk threshold drivers	14.4% of the total drivers, 90% young male drivers (average age 26.4)	▪ Don't reduce their speed if they realise they are travelling faster than the speed limit ▪ Driving faster when in a hurry ▪ Lots of sharp accelerations and decelerations ▪ Exceed the 30 mph speed limit even when it feels unsafe

[52] Musselwhite, C. (2006) Attitudes towards vehicle driving behaviour: categorising and contextualising risk. Accident Analysis and Prevention, 38, 324–334.

The low-risk threshold drivers	38.7% of the total drivers, 53% female drivers	▪ Would reduce their speed if they realised they are travelling faster than speed limit ▪ They don't drive faster when in a hurry
Opportunistic drivers	23% of the total drivers, 69% male drivers	▪ Exceeding the 30 mph speed limit when feeling safe ▪ They don't drive faster when in a hurry ▪ They mostly like to use a faster lane to avoid queuing
Reactive drivers	23% of the total drivers, 73% female drivers	▪ Driving faster when in a hurry ▪ Exceeding the 30 mph speed limit when feeling angry, annoyed or irritated ▪ Don't make dangerous overtaking ▪ Don't exceed the 30 mph speed limit when it feels unsafe

Source: (Fuller *et al.*, 2006)

Following the classification made by Fuller et al. (2006), as part of the UK-based project the High UnSafe Speed Accident Reduction HUSSAR, Fuller et al. (Fuller *et al.*, 2007) conducted a qualitative study on 36 drivers' responses to evaluating the Fuller driving task difficulty model. Additionally, for the same project (HUSSAR) Stradling et al. (Stradling *et al.*, 2007) interviewed 1,005 drivers in the UK to gain a better understanding of the four typologies of drivers described in Table 15. All the findings were consistent with those from previous studies. However, Stradling et al. (2007) found that for the high-risk threshold drivers the speed / acceleration relationship was not as strong as the earlier studies had suggested.

The concept of driver "celeration" behaviour was introduced by Wåhlberg in 2008. He suggested that the celeration is the total sum of all accelerations and decelerations of the vehicle. Wåhlberg suggested that drivers' accelerating, decelerating, skidding, lane changing and overtaking behaviours are linked and can be represented by the celeration value. He proposed a theoretical method for predicting the total number of culpable road collisions based on driver acceleration behaviour over a chosen time period. The author claimed that the celeration value can be useful in classifying and ranking drivers' performance and safe driving (Wåhlberg, 2008).

Identifying important driving parameters that differentiate drivers' driving behaviours was the subject of a study conducted by Takano et al. (2008). The researchers used Hidden Markov Models (HMMs) to distinguish these differences. The team used a driving simulator and real-world driving data in order to construct their driving model. The outcomes suggested that driver speed and acceleration are two parameters that are useful for distinguishing driving behaviour (Takano *et al.*, 2008). Later, Amata et al. (Amata *et al.*, 2009) took this finding and constructed a predictive driving behaviour model for road safety purposes. The predictive model of driving behaviour that was developed by them is salient as it was built to provide feedback to drivers about their safe driving behaviour in order to reduce collisions. The authors argued that based on their research, between 20% and 80% of road accidents could be prevented if a driver knew about his / her performance. The method they used to conduct the study is also of interest; the team managed to collect real-world driving data and then by using pattern classification methods (multiple linear regression analysis and Bayesian network analysis) it was able to classify drivers based on their safety attitudes while driving. The continiousion of this work led to comperhsive work by Halim and his team that developed a vehicle crash prediction model by using real-world driving data and artificial intelligence techniques (Halim *et al.*, 2016).

2.9 Identifying collision-prone sites by analysis of road collision records

By nature, road collision records are fraught with uncertainties, inaccuracies and errors. However, the geographical information system (GIS) technology and universal recording system are aiding studies that are conducting spatial analysis on road collision records. In the UK, reported traffic accidents are documented using the STATS19 system. The form with the same name is used to record a comprehensive set of data about each collision[53]. Every year, collision data is gathered from all local authorities approved by the National Statistics Authority and published by the Department for Transport. The published record is open for access by members of the general public. The website www.crashmap.co.uk is a map-based website that post processes these records and presents them on an online map. The provided map, CrashMap, is the core level of spatial analysis of the collisions records. Advanced spatial analysis of road accident historical data involves using this information for the following purposes: traffic safety study, traffic management and enforcing new traffic settings (speed limits, traffic lights, etc.).

Road safety engineering is the field that audits and investigates the causation of road accidents in specific locations, with the aim being to reduce road collisions at that location. After studying collision records at any given location, a road safety engineer often proposes new measures (new speed limit, providing better road lighting, instaling speed camera, adding a roundabout, etc.) to reduce the road collisions at that particular site. Owing to the sheer volume of collision

[53] Road collisions have required this documentation since 1979.

records and road segments with frequent accident problems, it is vital to identify collision-prone locations prior to conducting any field observation activities or auditing collision records.

Identifying collision-prone places, also referred to as crash hotspots, black spots, and hazardous and high-risk locations, is aimed at prioritising sites with significant collision problems that need to be improved. Basic methods, such as ranking collision locations based on the total number of accidents and their severity, are widely used by professionals to this end. However, given the increasing importance of reducing road injuries, identifying collision-prone sites more accurately based on historical data has come to the fore. In recent years, various methods have been introduced by researchers to overcome problems relating to collisions records. Issues, such as small recorded data size, low sample mean, and having a historical record with zero observation of accidents for a period of time, are common problems that have affected the results of the identification process (Cheng and Washington, 2005; Loo and Anderson, 2015). Table 16 has been developed based on various hotspot identification methods in the literature (Cheng and Washington, 2005; Lord, Mannering and Pankow, 2010; Montella, 2010). As explained in the following chapter, the confidence intervals method was deemed most appropriate for the current research.

Table 16. Common collision-prone (hotspot) identification methods

Method types	Method name	Description
Crash frequency methods	Ranking	Ranking road segments with the highest collision records in descending order
	Equivalent property damage only	Weighted ranking method based on the severity of the crash and the property damage cost associated with each type (slight, serious, and fatal)
	Crash rate	Ranking road segments with the highest collision records divided by measured traffic volume
Proportion methods	Confidence intervals	Based on the statistical confidence intervals method, segment 'A' is a hazardous location if the total number of recorded collisions at location 'A' is equal to or greater than the sum of the average and the standard deviation value of neighbouring segments
	The threshold proportion	Based on the Bernoulli trials[54], finding crash occurrence patterns by crash types (night time, poor visibility, wet road, etc.)

[54] The probability of an accident based on the Bernoulli trials: k is a type of crash is $p(k) = \binom{n}{k} p^k q^{n-k}$. k is the comparison group, n is total number of trials.

2.10 Naturalistic driving behaviour studies

As has been discussed in section 2.4, by 1996 in the US and by 2000 in the EU all passenger cars were equipped with an OBD-II port. Hence, accessing vehicle parameters, such as vehicle speed, engine speed and acceleration provided opportunities for new ways of studying driving behaviour in terms of road safety. Additionally, in-vehicle technologies, such as video cameras, front radar, GPS sensors and eye tracking cameras, have provided new countermeasures to reduce crash involvement. Earlier driving behaviour models were developed based on collision history records, self-reporting accident involvement and driving behaviour questionnaires. Gathering real-world driving data provided a reliable basis for studying pre-crash driving behaviours and therefore, revolutionised the way driving behaviour is investigated and modelled (Olsen and Wierwille, 2001; Klauer, 2005; Dingus *et al.*, 2006). By 2006, a combination of in-vehicle monitoring systems, cellular technology and GPS sensors allowed the possibility of evaluating driving behaviour and his or her safety. A concept that is now widely used by fleet management service providers and motor insurers in the form of the Pay as You Drive Plan (PAYD) and Pay How You Drive (PHYD) vehicle insurance policies.

Naturalistic driving behaviour studies are those in which driving behaviour is observed and recorded in great detail using in-vehicle driving monitoring without interfering with drivers' driving routines. In 2012, 40 large-scale naturalistic driving behaviour studies with the purpose of looking at road safety were conducted by transport and road safety intuitions globally (Regan *et al.*, 2012). Boyce and Gelle's study in 2001 is an example of an early naturalistic driving study. In their blind experiment, 61 drivers aged between 18 and 82 were asked to complete 22 miles of normal driving without prior knowledge that their driving was being monitored. The outcomes of the study suggested that the in-vehicle monitoring method had the potential to become the primary method of collecting driver's driving behaviour data. Moreover, these researchers nominated five measures for comparing drivers' driving behaviours as follows:

1. Percentage of safe speed;
2. Percentage of safe following distance;
3. Percentage of on-task behaviour;
4. Percentage maintaining lane position;
5. Percentage of turn signal use (Boyce and Geller, 2001).

Following up the study results, they contended that risk-taking behaviour was a function of age and not gender (as other self-reporting and survey studies had suggested) (Boyce and Geller, 2002).

An early work by Olsen and Wierwille (Olsen and Wierwille, 2001), for which they conducted a field study with a vehicle equipped with cameras facing inside and outside along with a front radar, provided the basis for a highly cited study carried out by the Virginia Tech Transportation Institute and National Highway Traffic Safety Administration (NHTSA) that started in 2005 (Neale *et al.*, 2005). Initially, the 100-car project was designed to provide information about the characteristics of crash, near-crash and incident driving events. Figure 26 demonstrates the in-vehicle data question system used in the 100-car project. The study established the fact that driver inattention is the primary reason for incidents, followed by the use of a mobile phone (Dingus *et al.*, 2006)[55]. In the US, an ongoing transport research programme, called the second Strategic Highway Research Program (SHRP2)[56], has completed

[55] Further: (Dingus *et al.*, 2006; Klauer *et al.*, 2006).
[56] Details can be found at: http://www.trb.org/StrategicHighwayResearchProgram2SHRP2/Blank2.aspx [last accessed March 2016].

the largest in-vehicle field study with 2,900 participants, but the results are yet to be published (Dingus *et al.*, 2006).

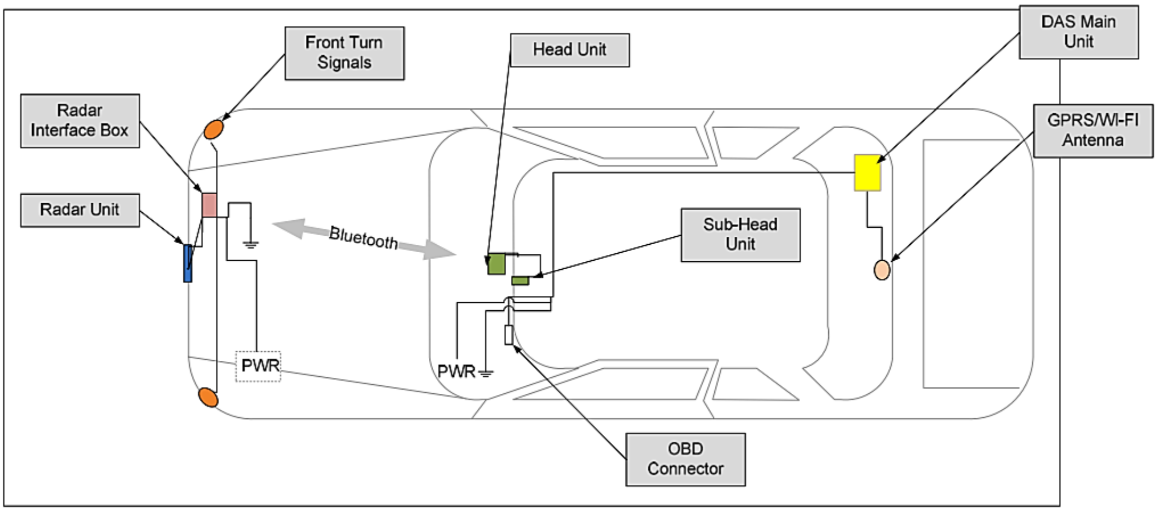


Figure 26. The in-vehicle monitoring system used in the 100-car project

Studies by Toledo and Lotan (2006) and Toledo et al. (2007) provided a framework for analysing, evaluating and modelling drivers' driving behaviour using an in-vehicle naturalistic driving study method. Moreover, collaborative projects in the EU member states as part of PROLOGUE, Promoting real Life Observations for Gaining Understanding of road user behaviour in Europe, have provided comprehensive documentation about the concept of naturalistic driving studies. The project provided a framework for conducting a large-scale European naturalistic driving study (Van Schangen *et al.*, 2011). For the first step, a series of small-scale projects were undertaken in Israel, Austria, the Netherlands, Spain, and Greece (Backer-Grøndahl, Lotan and Schagen, 2011). The success of these led to the UDRIVE[57] project (Eenink *et al.*, 2014; Eenink and European Commission, 2016), which is the EU's first large-scale naturalistic driving study project, is being coordinated by the Netherlands Institute for Road Safety Research (SWOV). The project started on October 2012, and it is planned that

[57] Details to be found at: http://www.udrive.eu/ [last accessed March 2016].

during the four-year period 150 lorries, 80 motorcycles, and 240 passenger cars will be monitored.

The UDRIVE project funded by European Commission was the EU first driving and riding safety study. In this project, the driving behaviour of drivers using passenger cars, trucks and scooter riders was collected using five to six different camera views in six countries across Europe (Eenink *et al.*, 2014; Eenink and European Commission, 2016). The UDRIVE objective can be summarised into the following three points:

1. Examine and quantify drivers and riders' behaviour in different regions of EU in regular and near-) crashes conditions;
2. Investigate and quantify participants behaviour regarding emission levels and fuel consumption;
3. Increase traffic systems safety in a sustainable manner.

The final naturalistic driving that is discussed in this subsection is one from Australia. Regan et al., in their 2011 article on driver distraction, suggested that by conducting in-depth crash investigations from an in-vehicle recorded video, it is possible to understand the causation of attention failures for drivers involved in car accidents. This conclusion led him and his team to study the feasibility of conducting a large-scale naturalistic driving project in that country. In 2013, four institutions in Australia namely Transport and Road Safety (TARS) Research, University of NSW, Monash University Accident Research Centre, Centre for Accident Research and Road Safety, Queensland University of Technology, and Centre for Automotive Safety Research, University of Adelaide collaborated on the first Australian large scale (400-car) naturalistic driving study and according to the schedule, the results of the study will be published by the end of 2016 (ibid).

2.11 A critical review of previous studies

As has been discussed, there are two types of fields of studies in relation to understanding the impacts of drivers' driving behaviour. The first group comprises those aimed at reducing fuel usage and CO_2 emissions, whilst the second relates to those that aim to reduce road casualties and increase driver safety. A summary of the methods of driving behaviour studies is summarised in Table 17, there are various ways to study drivers' driving differences.

Table 17. A summary of the methods of driving behaviour studies

Method	**Measures**
Surveys, self-reporting, driving behaviour questionnaires and/or historical collision data	Impacts of drivers' demographics, personality, aggression, errors, violation (violating the Highway Code) and lapses in driving performance
Laboratory testing, chassis dynamometer testing	Engine performance under different conditions and scenarios (cold start, hot start, controlled weather simulation), fuel consumption, and accurate measurement of emission gases
Simulation (computer based, driving simulator)	▪ Monitoring vehicle dynamics ▪ Monitoring drivers' perceptions and attitudes based on task-based studies ▪ Measuring hypothetical fuel consumption and emissions

Method		**Measures**
On-Road	Test track	▪ Monitoring vehicle dynamics ▪ On-road and on-track emissions ▪ Drivers' eye direction whilst driving ▪ Monitoring specific driving tasks (turning left or entering a motorway) ▪ Drivers' speeding and braking behaviour ▪ The distance from the car in front ▪ Recording near crash, crash and incident events ▪ Steering wheel performance and gear shifting strategy ▪ Lane keeping and in-vehicle distractions (e.g. using a mobile phone, changing music/radio channel)
	Field Operational Trial (FOT)	
	Naturalistic Driving Study (NDS)	

Source: adapted from the works of (McLaughlin, Hankey and Dingus, 2009; Van Schangen *et al.*, 2011; Regan *et al.*, 2012)

As has been explained in the previous subsection, drivers' driving differences can be discussed based on various assumptions. However, there are commonly shared views about the parameters that can affect a driver's fuel efficiency performance and his / her safety. Nevertheless, the approaches applied in the existing studies and the research questions raised have been inconsistent in terms of the terminology used and data acquisition methods as well as being subject to bias in relation to the findings.

To make a critical assessment of the existing studies in the field of driving behaviour, in particular, naturalistic or real-world driving behaviour studies, four dimensions have been developed to provide a clear perspective on some of the strengths and weakness of the prior work carried out in this field, these being:

- Agreed parameters: a list of parameters that gained a consensus among scholars in studies that were published between 1980 and 2016 and that have been reviewed in this research;
- Terminology differences: highlighting the fact that despite the vast number of publications in this domain, a number of misused words led to weak conclusions;
- Data collection methodologies: variation of the data collection strategies makes it difficult to evalaute and compare study conclusions;
- Biassed conclusions and confounding: lack of clear understanding of the complexity of naturalistic driving behaviour studies has led to poor, irrational and illogical conclusions.

Agreed parameters

It is generally agreed that drivers' driving behaviour is governed by driver demographics, emotions, driving conditions (road condition, weather conditions, road lighting, unfamiliar routes) and travel habits (always using A/C, travelling with extra load, always driving with more than one passenger). These parameters can affect a driver's driving performance. Regardless of which of these factors triggers a change in someone's driving, their impact determines how the vehicle is controlled, which can be measured by such phenomena as how the pedals are being used, the gears are shifted, and whether excessive vehicle power is engaged in. To some extent it can be said that harsh braking can influence excessive fuel usage due to the fact that the driver has to speed up (burn more fuel) to reach to required speed after his or her harsh braking. If a driver controls a vehicle in such a way that it operates under inefficient conditions (driving in the wrong gear, high engine speed), then his or her driving affects fuel consumption. Regarding safe driving, it is not always clear which part of a driving task can cause a driver to get involved in a car crash. However, indicators such as exceeding speed limits, harsh acceleration and deceleration, losing control of the vehicle, driving in a way that harms others and public property, distraction, not driving inside the lanes, and changing lane aggressively, can all be signs of a driver who is driving unsafely.

Terminology differences

According to Shinar, Tasca, Dual and Geller, loose usage of terminologies in this field, e.g. aggressive driver and aggressive driving, have led to many misinterpretations in road safety studies (Shinar, 1998; Tasca, 2000; Dula and Geller, 2003). This is also true when it comes to classifying driver behaviour into different categories, because a whole range of unclearly defined terms and their attributes, such as calm driver, aggressive driver, sportive driver, risk-taking driver have been used. Regarding emissions studies, drivers' driving behaviour and aggressive driving are inappropriate concepts when researching fuel usage and vehicle emissions as they are purely mechanical matters, where the driver should only be considered in terms of being the operator of a vehicle, i.e. using the pedals, shifting gears, and using a steering wheel and not in terms of his or her psychological state. However, a review of emissions studies shows that these psychological terms have been used, which has led to some confusion as to what is actually being measured. This is clearly not satisfactory when the aim is to study drivers as individuals with their unique set of behaviours, emotions and demographics. It needs be clarified that in this study an aggressive driver / driving aggressively refers to an unusual set of driving habits that situate a particular driver away from other drivers in the group. The method through which outlier driving behaviours are detected is based on their driving parameters, including: driving above the speed limit, excessive usage of vehicle power (high engine speed) and sharp acceleration/deceleration.

Data acquisition methods

There are a broad range of data acquisition methods for quantifying drivers' driving behaviour. It is important to acknowledge that the field of studying driving (driver) behaviour is a multidisciplinary subject and a large number of questionnaires and driving monitoring methods have been developed and employed by researchers over time to study drivers' differences. Regardless of the study approach in studying driver behaviour, it is crucial to avoid issues such as poor and bias sampling methods and lack of randomisation in the sample. For example, studies that use only any one or two cars or just one or two drivers to conduct a driving behaviour field test are not effective for investigating the phenomenon. Additionally, studies that use only one round of surveys with a low sample size, or those that only use one set of questionnaires are subject to sample bias. That is, such studies do not represent any population and therefore, their results are not sufficiently statistically significant to draw any scientific conclusions.

As has already been addressed by the PROLOGUE project, it is important to have an integrated data acquisition method, where both the field test and survey test results of a vast number of participants are reviewed together. A similar approach was used by the EcoDrive team in their data collection strategy. In their work, the team developed a clear and integrated data acquisition method strategy to collect, use and compare driving data from multiple devices from multiple providers. It is important to mention that the data acquisition device is not a critical issue as long as the study is designed with a naturalistic driving behaviour mind-set and with the help of frameworks such as FESTA-V approach.

Biassed conclusions and confounding

There have been many studies which, despite their efforts to examine the differences in drivers' driving behaviour, have conducted their research in the following ways (the following wording is based on a number of studies that have been reviewed and to keep the confidentially of their authors no references have been made to their work):

"*…drivers asked to drive aggressively...*" or "*… First-day driver invited to drive calmly. On the second day, he was asked to drive aggressively…*" or "*…one female driver and two male drivers are asked to drive normally and then aggressively...*"

Whilst these types of studies are useful as pilot studies or for proving a concept, unfortunately, they show a lack of understanding of the nature of the driving task. They suffer from a lack of representation of different driver types, vehicle models, and road conditions and have too few trials, hence resulting in biassed conclusions. This also provides evidence that undertaking naturalistic driving behaviour and a FESTA-V approach will allow for these issues to be addressed during the design of the study.

Confounding refers to when during a study the impact of a factor that is of interest to the researcher (in this case aggressive driving) is not distinguishable from the effect of other factors that might contribute in the same measure, when, for instance, "*...by comparing driving time, it shows aggressive drivers finish faster than others…*".

In this example, the researcher omitted the impact of factors, such as road traffic or the amount of time each driver waited at red lights during a trial. This is a problematic because the paucity of experimental description makes it impossible for other researchers to replicate the research and thus, test the validity of the outcomes. Moreover, the driving parameters are interconnected. Therefore, it is difficult to quantify one parameter as the only major factor of an accident, or of an excessive amount of fuel usage. To avoid this type of mistake, it is important to be careful that these issues are clearly addressed at an early stage of the research, such as during the designing of the experiment phase.

2.12 The eco-driving and safe driving studies

As has been discussed, drivers, as the operators of vehicles, have a significant impact on how efficiently the car is operated (eco-driving) and how safely a journey is completed (safe driving). However, studying the effect of these two driving aspects together has only been undertaken during the last decade. In both the fields of study (eco-driving and safe driving), several highly cited scientific works have been conducted to understand the impact of driver driving differences. Research carried out by institutions, research teams and individuals led to projects, such as the WLTP and EcoDriver, with the primary goal of understanding the effect of driver behaviour on fuel usage and vehicle emissions. Similar efforts were made by road safety scholars to understand driver behaviour in regard to them driving safely. These studies have led to projects, such as the 100-car and SHRP2 programmes in the US and UDRIVE and EcoDriver in the EU. While both concepts are still under investigations and debate (Mensing *et al.*, 2014) the new concepts emerged which aims to identify the overlapping between being safe and being eco-friendly (Young, Birrell and Stanton, 2011).

The work of Haworth and Symmons (2001) has been an influential milestone towards linking the two concepts of eco driving and safe driving together (Haworth and Symmons, 2001). Harworth and Symmons' reasoning for this link is that since the widely accepted driving manner to use less fuel is to driver smoothly with less harsh acceleration and braking, this can be directed towards driving with higher anticipation and hence, lead to reduced driving risk (Haworth and Symmons, 2001). However, these authors agreed that this link was complex to prove empirically and even provided examples that appeared to refute it. Their work highlighted the importance of driving training, more specifically, eco-safe driving training (ibid).

Studying the relationship between drivers' eco-driving and safe driving gained more attention after Wåhlberg's (Wåhlberg, 2007) study on the long-term effect of eco-driving training on drivers' safe driving attitudes.

In a follow-up study in 2007, Wåhlberg reviewed the effect of his initial work a year earlier. (Wåhlberg, 2006). In his first study, he had examined the short-term effects of eco-driving training on bus drivers' fuel consumption in terms of fuel usage, accident involvement and acceleration behaviour. Despite the study conclusions not being significant one year after the training (2% reduction in fuel usage as well as no effect on accident involvement and acceleration behaviour), it is highly cited and has become very influential, because it was pioneering in its aim to bridge the gap between the effects of eco-driving behaviour and safe driving (Wåhlberg, 2006, 2007).

Moreover, advancement in technology has greatly facilitated researching these two phenomena in tandem. Regarding which, feedback devices for advising drivers on their fuel consumption were developed as early as 2001 (Van der Voort, Dougherty and van Maarseveen, 2001) and a new generation of these was introduced in 2007. This initiative was expanded to intelligent transport systems and to using these technologies to reduce fuel consumption by avoiding traffic congestion. For example, Kono et al. (2008) used traffic information, route characteristics and vehicles parameters to develop the "Eco Route Search" predictive model that gave routes that required the least amount of fuel (Kono *et al.*, 2008; Ben Dhaou, 2011). Similar initiatives were taken by Ichihara et al. (2009) by developing a driver assistance system that measured drivers' eco-driving skills and their motivation to drive in an eco-friendly manner (Ichihara *et al.*, 2009).

Extensive research has been conducted to understand the psychological and behavioural forces behind drivers' decision making in relation to operating cars in eco-manner. For example, Boriboonsomsin, Barth and Vu (2011) examined the relationship between drivers beahviour and attitude towards eco-driving. With a sample of 20 drivers, they evaluated the impact of on-board eco-driving devices on drivers and managed to show an improvment of 6% on fuel econmoy on urban roads (Boriboonsomsin, Barth and Vu, 2011). Emphasising drivers' driving behaviour attitude led to extensive works on eco-driving training and its impact on saving fuel and reducing gas emissions (Barić, Zovak and Periša, 2013; Rolim *et al.*, 2014; Ho, Wong and Chang, 2015)

While many scholars have measured eco-driving performance of drivers by only assessing their driving behaviour, Sivak and Schoettle (2011) have argued that eco-driving should be discussed even during an individual's car purchasing thought process (Sivak and Schoettle, 2011). Accordingly, they examined drivers vehicle selection and maintenance decisions (strategic decisions), route selection and car load (tactical decisions) and finally driving behaviour, which they categorised as operational decisions. The classification highlighted the fact that, in total, about 45% is the maximum reduction that can be made per driver, if he/she drives in an eco-friendly manner (Sivak and Schoettle, 2011).

One of the projects that advanced the single concept of eco-driving and safe driving was the UK-based project called Foot-LITE[58]. The idea was to provide innovative driver/vehicle interface systems to promote both safe driving and eco-driving, work which then led to the development of an in-vehicle advisory system and the notion of 'smart driving' (Fairchild *et al.*, 2009; Young, Birrell and Stanton, 2009). The concept of smart driving was promoted by Zarkadoula, Zoidis and Tritopoulou (2007) in their pilot study on training urban bus drivers in Greece. However, according to them, smart driving only rested in the domain of eco-friendly driving. (Zarkadoula, Zoidis and Tritopoulou, 2007). In recent years, the concept has changed to being more about both eco and safe driving. According to Young et al. (2011), smart driving pertains to driving that is both eco-driving and safe, which they defined as follows: "(1) Plan ahead to avoid stopping and minimise sharp braking, (2) Use smooth but positive acceleration to reach high gears sooner, and use engine braking for smooth deceleration, (3) Use moderate engine speeds and a uniform throttle for steady speeds, (4) Obey speed limits" (p. 537).

A similar concept was also developed in Japan at the Toyota Transportation Research Institute. The team used a driver's locations (the GPS sensor) to predict his or her performance. Based on the vehicle speed, acceleration and its location, the model was able to promote eco-driving and safe driving to drivers (Ando, Nishihori and Ochi, 2010). Ando's work has been significant as it attempted to quantify the eco and safe driving performance of drivers by using a threshold rules for evaluating drivers' eco-driving and safe driving (Ando, Nishihori and Ochi, 2010). In 2012, Yun et al. developed a smartphone-based, eco-driving and safe driving advisory app by using information gathered from an OBD dongle and GPS locations (Yun *et al.*, 2012). Their model works based on threshold rules for scoring drivers' eco-driving performance and safe

[58] Details to be found at: www.foot-lite.net [last accessed March 2016].

driving by a simple comparison of fuel consumption so as to be able to classify drivers. As identifying eco-driving behaviour without in-vehicle device has its limitations (e.g. inaccurate assessment and poor data quality), most recent studies have focused on classifying drivers' behaviour based on their aggressive driving/safe driving behaviours by using smartphones (Meseguer *et al.*, 2013; Saiprasert and Pattara-Atikom, 2013; Zhao *et al.*, 2013; Hong, Margines and Dey, 2014; Castignani *et al.*, 2015; Koh and Kang, 2015)

In fact, many driving parameters can be used to identify and classify eco-drivers and safe drivers. The most common method is based on using linear equations with a different coefficient ratio being assigned to each driving parameter, according to their impacts on eco-driving and safe driving. The following table has been developed based on the parameters used in the aforementioned studies to identify and classify eco-drivers and safe drivers (Table 18).

Table 18. A summary of the driving parameters used to classify eco-drivers and safe drivers

Eco-driving	**Safe driving**
▪ Exceeding vehicle speed ▪ Exceeding the engine speed to above 4000 rpm ▪ Number of sharp accelerations and decelerations ▪ Idling time ▪ Fuel consumption	▪ Quick start and take-off ▪ Quick stops and turning off the vehicle without considering safety measures ▪ Waiting time in the driveway ▪ Exceeding vehicle speed ▪ Speeding for a long duration of time ▪ Number of sharp accelerations and decelerations ▪ Bad handling of the vehicle, including rash overtaking, inappropriate turning, and unsafe lane changing/handling ▪ Failing to signal and to obey traffic signals ▪ Following too closely

Developed based on the following studies: (Amditis *et al.*, 2007; Inata, Raksincharoensak and Nagai, 2008; LeBlanc, Sivak and Bogard, 2010; Handel *et al.*, 2014; Luke and Heyns, 2014; Jamson, Hibberd and Jamson, 2015)

In order to have scoring methods for differentiating drivers in terms of eco driving and safety it is necessary to have key driving performance indictors capable of providing that information. Combining parameters shown in the table above led to conclusion that the following represent the minimum requirements for building a viable driving scoring model. That is, listed below is a summary of the minimum set of driving parameters for building a scoring model and the threshold rules that need to be violated by drivers in order to count the event as an alarming one as defined by the CASTEL research and development team (CASTEL Company, 2014).

- Speeding events: Number of times above the speed limit and percentage of journey spent speeding
- Exceeding PRM (revolutions per minute): Based on excessive engine speed, typically above 3500 rpm
- Harsh acceleration: Based on g force value, typical threshold rules from 0.2 g to 0.6 g
- Harsh braking: Based on g force value, typical threshold rules from - 0.2 g to - 0.6 g
- Sharp cornering: Based on speed and cornering force, if the value exceeds 0.2 g
- Unsafe lane changing: Based on driving speed, location and a g force value exceeding 0.2 g

The most recent study linking eco-driving and safe driving is that conducted on a driving simulator at the Institute for Transport Studies, University of Leeds (Jamson, Hibberd and Jamson, 2015). The researchers examined the impact of real-time visual feedback on the eco-driving and safe driving performances of drivers. Two hypotheses were examined using a driving simulator. Firstly, the researchers checked whether drivers learn to drive in an eco-friendly manner at different rates and secondly, they probed whether drivers prioritise between eco-driving and safe driving (Jamson, Hibberd and Jamson, 2015). They did not observe any improvement in terms of eco-driving performance among the participants, who were asked to drive a driving scenario three times with eco-driving feedback being provided. The second part of the study was interesting in that the researchers introduced a high traffic density scenario in the simulator and the participant drivers prioritised safe driving over eco-driving (Jamson, Hibberd and Jamson, 2015). The findings of this study suggest that in the field of studying eco-driving and safe driving there are unknown parameters that require in-depth investigation, if they are to be uncovered.

Jamson, Hibberd and Jamson, 2015 have suggested that drivers tend to piroritse safe driving over eco-driving. However, the two related challenges remain, i.e. reducing emission gases and to having safe drivers. To date, both topics (eco and safety,) ,separately or jointly, have been investiaged. For example, work by Ayyildiz' (2017) looked at the impact of fleet (10 light commercial and 15 heavy-duty vehicles) eco-driving trainng in the reduction in fuel consumption and carbon emissions and reportedly managed to reduce the former by 5.5% (Ayyildiz *et al.*, 2017).

2.13 This research in context

In recent years, using affordable technologies, such as an OBD dongle and GPS sensor, have provided the opportunity to track and monitor vehicles and drivers remotely. Two sectors currently benefit from this technology: firstly, the vehicle fleet management providers and secondly, the vehicle insurers who provide usage-based insurance policies (Bruneteau *et al.*, 2013; Husnjak *et al.*, 2015; Karapiperis *et al.*, 2015). The pay-as-you-go concept from the telecommunications industry has been adapted by insurers and fleet management service providers as the most successful pricing model. The Pay as You Drive Plan (PAYD) and Pay How You Drive (PHYD) motor insurance policies are two examples of this adaptation. As a tracking device the basic OBD dongle is utilised with an accelerometer and GPS sensor to detect crashes and careless driving, a concept that Ippisch (2010) has put forward is one of the added value aspects of telematics technology to the insurance industry (Ippisch, 2010). Apart from OBD + GPS models, a standalone blackbox is also a method for tracking drivers and capturing dangerous events. Regarding which, Insure The Box[59] in the UK is primarily using a black box or the in-tele-box. These tracking boxes come fitted with a GPS sensor, accelerometer and transmission unit, which sends the data using the GPRS network.

Figure 27. Example of how telematics insurance works illustrated by Confused.com

[59] Details can be found at: www.insurethebox.com

To clarify the two aforementioned concepts, i.e. PAYD and PHYD insurance products, the following tables has been developed based on Tselentis et. al.'s (2016) idea of risk indicators classification (Tselentis, Yannis and Vlahogianni, 2016).

Table 19. The list of suggested parameters to be used in PAYD and PHYD insurance premiums

PAYD	**PHYD**
▪ Time and date (higher risk in rush hours and specific seasons) ▪ Total distance driven by the user (high mileage high risk exposure) ▪ Road types ▪ Trip frequency (total number of trips per month) ▪ Vehicle brand and model ▪ Previous claims locations (danger zones)	▪ Time and date ▪ Road types ▪ Familiarity (number of times driver has passed down the same road in the last 30 days) ▪ Weather conditions ▪ Previous claims locations (danger zones) ▪ Speeding as a proportion of kilometres driven or time travelled ▪ Harsh acceleration and braking ▪ Harsh cornering ▪ Combined metrics: speeding at night on a highway, in different weather conditions ▪ Telematics claims history ▪ Idle time ▪ Fatigue (as a measure of making a number of stops as part of long trips)

Developed based on the following studies and whitepapers: (A. T. Kearney *et al.*, 2010; Paefgen, Staake and Thiesse, 2013; Jubraj and Moneta, 2014; Husnjak *et al.*, 2015; Intel, 2015; Tselentis, Yannis and Vlahogianni, 2016)

The results of an investigation using telematics technology to reduce accident risk conducted by the International Transport Forum in 2011, suggest that providing financial incentives and feedback to drivers will make them motivated to engage in safe driving behaviour. However, there needs to be clear understanding of how to identify and score drivers who drive poorly (Bolderdijk and Steg, 2011). One recent discovery regarding telematics based solutions is the fact that it can have a positive behavioural impact on drivers' safety. Moreover, the feedback can lead to more efficient control of the vehicle, in particular, in terms emissions, which is why 14 US states have incorporated telematics into their green initiatives and climate projections, According to a TRL report in 2015, there have been at least 30 experimental studies aimed at studying the effect of telematics systems and their feedback on (young drivers) drivers (Tong *et al.*, 2015). It needs to be acknowledged that some initial behaviour changes are due to the effect of drivers knowing that they are being monitored. The impact of providing feedback to drivers is that it has helped a number of fleet companies in reducing their drivers crash rate. For instance, in 2011, GreenRoad (one of the leaders of driver performance and fleet safety management providers) reported that 70,000 fleet drivers who were using their solutions were experiencing, on average, 50% reduction in crashes and 10% reduction in fuel consumption (Ron, 2011).

Only a small number of parameters so far are used for scoring drivers and providing them with feedback by fleet operators and insurers who are offering telematics-enabled policies, these being things that are within the drivers' power to change. These parameters are:

- Speeding events
- Exceeding PRM
- Harsh acceleration
- Harsh braking
- Sharp cornering
- Unsafe lane changing

In addition to these core parameters, insurers are constantly defining and using new metrics to increase their capabilities to score drivers fairly and accurately. Examples of additional metrics are familiarity (driving similar route in one-month period) and driving in rush hour, at night or early mornings.

In recent years, smartphone driving monitoring mobile applications have been accommodated using a Bluetooth enabled OBD dongle to engage more with drivers and to reduce operation cost. However, to insurers, the main benefit of this technology is to be able to predict risk and develop pricing models closely associated with drivers' driving differences. Weiss and Smollik (2012) addressed this in their work geared towards the insurance market and actuaries' profession. These authors argued that telematics data provide a unique opportunity for insurers to customise their underwriting process by using individuals' driving data (Weiss and Smollik, 2012).

Azzopardi and Corti (2013) investigated the benefits of using the aforementioned technologies for the fleet management industry. They concluded that the key strength of using this system is because of its ability to monitor driver fuel consumption and behaviour, while other benefits, such as recording mileage and driver locations can be controlled by cheaper methods (regular mileage recording, geo-fencing)(Azzopardi and Cortis, 2013). According to the Association of British Insurers' (ABI) report, in the UK, the rising market of usage-based technology is due to the young driver (under 25) sector. They are willing to have a device installed in their car and to be monitored in return for cheaper car insurance (The Association of British Insurers (ABI), 2015). However, the report points out that according to a survey[60] conducted by the price comparison website uswitch.com, among UK drivers, 62% were worried about their privacy, and 37% did not want their driving to be monitored (The Association of British Insurers (ABI), 2015). Nevertheless, the market for using OBD dongles and GPS sensors to monitor drivers for fleet management and insurance purposes is emerging, and it is expected to increase in the EU from 5 million policyholders in 2014 to 27 million in 2019 (Bruneteau *et*

[60] Survey conducted on April 2015 with 1,146 UK car insurance owners.

al., 2013; Husnjak *et al.*, 2015; Karapiperis *et al.*, 2015; The Association of British Insurers (ABI), 2015). In sum, as shown in Figure 28, it is predicted that the total number of telematics-based policyholders in the EU and US will increase substantially.

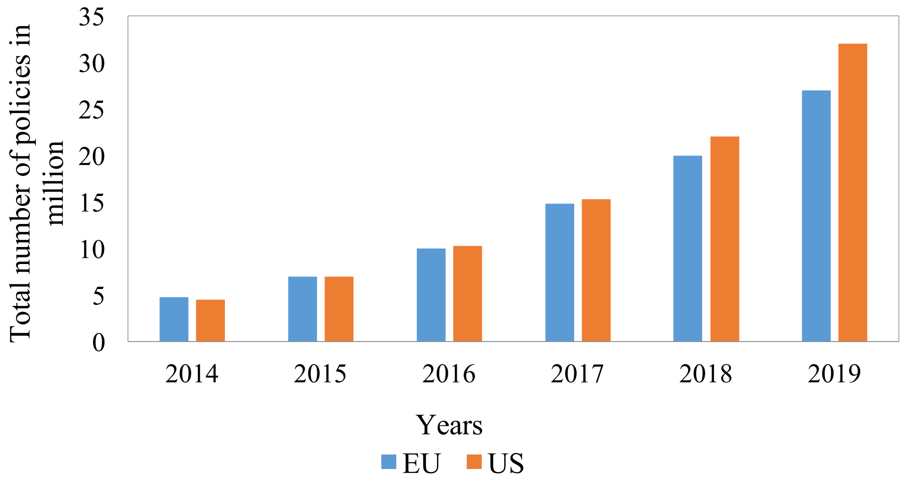


Figure 28. The predicted total number of telematics-based vehicle insurance holders

Whilst telematics-enabled premiums or connected premiums is relatively new, with the first piloting of the concept being conducted by Progressive Insurance in 2004, the concept of usage based premiums is not a new one to the insurance industry. For instance, Allstate in 1939 introduced a premium based on mileage and car usage, whilst in1991 Progressive Insurance tested a credit score concept to underwrite premiums. However, the uniqueness of telematics-enabled premiums is having real-world and real-time driving data streaming towards insurance servers. The motor insurance industry has evolved a lot in the past decade. According to multiple sources, UBI progrmmes are now active in countires such as Italy, the US, Canada, Spain, Austria, South Africa, the Netherland, Japan and the UK (based on public records between 2011 and 2014) (Troncoso *et al.*, 2011; Friedman and Canaan, 2014; Husnjak *et al.*, 2015).

Given the confidentiality of scoring methods that have been used by insurers worldwide, such as Progressive in the US and Generali and Unipol in Italy, publicly available the extensive work conducted by Handel is crucial for those wanting to investigate the impact of driving scoring in telematics-enabled insurance (Handel *et al.*, 2014; Wahlstrom, Skog and Handel, 2015). Handel et al. made a thorough analysis the feasibility of using smartphones as a method to collect and score drivers. The team concluded that whilst smartphone only usage based insurance products can be a scalable option, the huge variation between hardware devices and poor data quality makes it difficult to use this with a high level of confidence (Handel *et al.*, 2014). As a result of their work, (Wahlstrom, Skog and Handel, 2015) investigated various approaches to clean and filter phone' sensor signals, which they then evaluated the usage and effectiveness of using smartphone GPS data to detect the cornering behaviours of drivers.

From the early stages of the current project, it was decided that it should be conducted as if the participatory drivers were part of a group of fleet drivers or as if they were usage-based insurance policyholders. With this approach, the choice of monitoring devices was narrowed down to those that are commonly used by both of the groups. The following table was developed to summarising the current state of widely used technologies for monitoring and tracking drivers, as of December 2015 .

Table 20. A summary of commonly used technologies to track and/or to offer telematics enabled premiums

Industry	Typical usage			
	Blackbox	OBD dongles	Smartphone	Combined
Motor insurance market	Used mostly for PAYD premiums	Used for both PAYD and PHYD premiums	Used mostly for PAYD and on few cases as PHYD premiums, highly used as a communication platform to provide feedback	Blackbox and smartphone or OBD and smartphone as way to collect accurate data and provide feedback
Fleet tracking and managing market	Highly used as a tracking device	Highly used as tracking device and for service and maintenance purposes including fuel usage	Just as a communication & navigation device, rarely used to provide feedback	Highly used as tracking device and for service and maintenance purposes including fuel usage

Source: public information available from the following insurance companies: Allstate, Generali, Progressive, Insurethebox, and Unipol.

Moreover, two of the recent studies in the EU, the EcoDriver and UDRIVE, have been a great inspiration in terms of developing a small-scale eco-safe driving platform to conduct a study, with the aims of investigating drivers' driving behaviour with regard to fuel consumption and their safer driving abilities (Eenink *et al.*, 2014). With this in mind, this study was designed and conducted according to the naturalistic driving study guidelines provided by the FESTA Handbook[61] on conducting field studies. According to this handbook, a field study has to be completed in three stages: preparation (design of the experiment), data acquisition (data management) and post-processing (Van Schangen *et al.*, 2011; FESTA Project Consortium, 2014). This study was designed and conducted according to these guidelines. That is, driving data were analysed in three steps: (1) identification of a trait of behaviour or abnormal driving

[61] Field operational test support action (FESTA).

performance; (2) grouping of those drivers with similar driving habits; and (3) using this information to model drivers' driving behaviours. In the next chapter (methodology) the three stages of conducting the field study, acquiring the data and the post-processing analysis are explained and justified in detail.

Chapter 3

Research Methodology

"What we observe is not nature itself, but nature exposed to our method of questioning".

– Werner Heisenberg

3.1 Introduction

From the critical review of existing studies conducted in the research field of driving behaviour, it is concluded that the most relevant applications of research similar to this project focus on two industries: the fleet management industry and usage based insurance market (see 2.13). The key challenge for these industries is that they want to monitor drivers in their daily routines and score them based on their eco-driving and risk driving performances. In this setting, the monitoring device has to be economical and cost effective as well as the method being robust. Moreover, scores have to reflect reality since business decisions will be made based on them.

As it has been highlighted in the final section of the chapter 2, the general methodological approach to accomplishing the objectives of this project was to conduct a study according to the Naturalistic Driving Study (NDS) guidelines. Specifically, the aim was to observe drivers in their natural setting while they were completing an ordinary driving task by using in-vehicle technology unobtrusively. In this case, the biases discussed in chapter 2 were kept to a minimum, whereby the observed participants were not influenced by allocated interventions from the observers. An unsupervised approach can affect the sampling procedure and make effective comparisons between driver performances very difficult. However, fully supervised Field Operational Tests (FOTs) (e.g. proving ground testing) ignore important information regarding the natural driving setting, such as daily routine. Consequently, this project involved combining both ways of collecting actual driving data. That is, whilst there were controls on some aspects of the driving experiment, there were still opportunities for unexpected phenomena or behaviour to occur. In this way, the occurrence of different traces of driving behaviour and habits could be captured. As a result, this research lies in the overlapping area between the fully controlled and completely unfettered studies, as illustrated in the figure below.

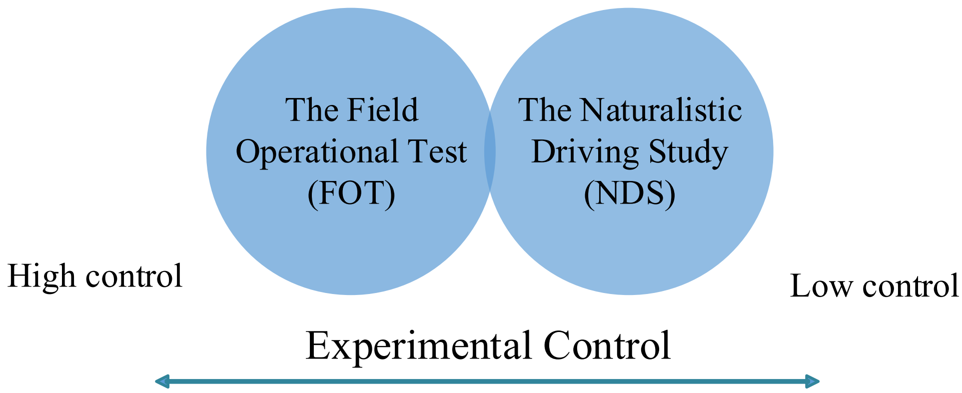


Figure 29. Experimental control continuum of field studies

This means that in this project despite having a number of predefined and predesigned parameters (driving route, vehicle types, fuel, etc), the drivers received very little input on how to behave and how to operate the car. Various methodological considerations need to be taken into account when creating a field test to examine the naturalistic behaviour of drivers (Dingus *et al.*, 2006; Klauer *et al.*, 2006; Van Schangen *et al.*, 2011; Mejuto, 2016).

As discussed in chapter 2, there are three main areas to consider associated with such a study: first, there is the concern about the experimental design not interfering with the drivers' performance; secondly, there is the data acquisition, and monitoring method; and finally, there is the data analysis method as well as the post-processing approach towards the collected data. To address these areas, a framework was developed based on the Field opErational teSt supporT Action (FESTA) handbook on the methodology of implementing and operating FOTs, in accordance with Van Schangen *et al* (2011) (Van Schangen *et al.*, 2011), so as to take into account naturalistic driving behaviour study requirements. The modified version of the framework by Van Schangen et al. (As it is shown graphically in Figure 31) ensured that naturalistic driving research was conducted while addressing every issue of concern in a well-regulated manner. This strategic framework was designed based on the V-model approach that is broadly used by industry for system design, hardware and software development as well as conducting tests (Osborne *et al.*, 2005; Van Schangen *et al.*, 2011). The framework allows for every step to be completed in a scientific manner as well as in a sustainable and repeatable way, as far as possible. In regard to conducting field studies, the V-model places emphasis on two methodological concepts: validation and verification. As achieving the objectives of the tests involves breaking the process down into detailed steps, this helps to ensure test validation. Secondly, since it is visually arranged in the shape of the letter V, every step from the right side of the framework can be validated by the left side. These two concepts (validation and verification) have been disused by Osborne et al. (2005) as being a successful approach to both managing and executing a multi-dimension and multi-task process, including software development, benchmark testing and field studies (Osborne *et al.*, 2005). Hence, all naturalistic driving behaviour studies have adopted this concept in order to design, manage, execute the project (Van Schangen *et al.*, 2011; Sadigh *et al.*, 2013; Eenink and European Commission, 2016). The general version of V-model is illustrated below. Here, the project starts from the left hand side, with project definition phase, then it is the implementation phase (at the bottom of the V) and finally, there is the outcome phase. The usefulness of the model is that if an error is observed during the integration, test and verification phase (right hand side), the cause can be found on the opposite side in the detailed design phase.

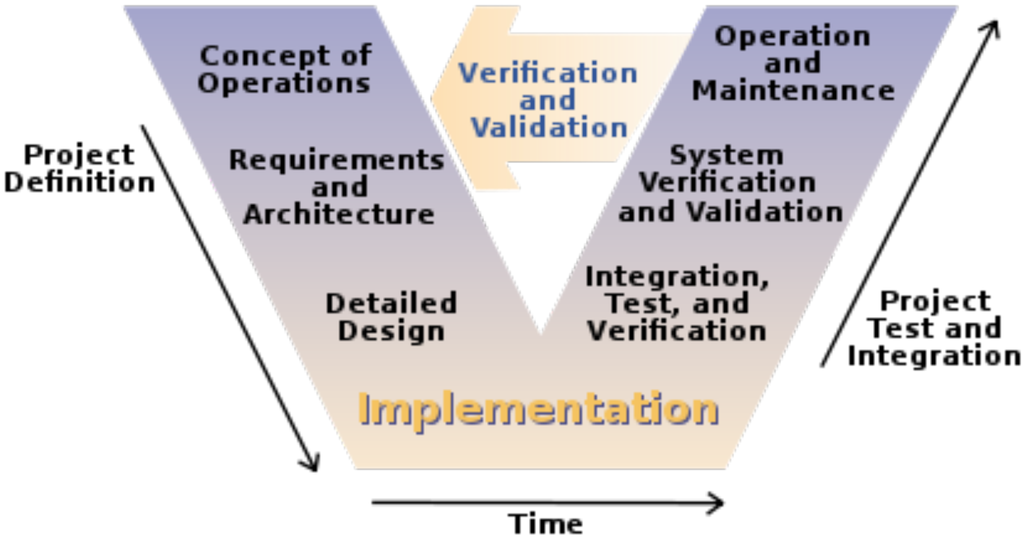


Figure 30. General V-model or V-approach framework (Osborne *et al.*, 2005)

Adapting the research plan to this framework and executing its methodology[62], brought substantial benefits to this project and the following list summarises what these are:

- Providing a structured, and coherent stepwise project plan with common vocabulary to the other scholars working in this domain;
- Adapting the model that has been used by OEMs to conduct field studies such as BMW;
- Providing a platform for validating the project that is in accord with similar studies;
- Assuring issues that otherwise might not be addressed. e.g. ethical issues and the data management plan are covered.

[62] More details available at: http://wiki.fot-net.eu/index.php/FESTA_Handbook

The framework for conducting and implementing a naturalistic driving field study (Figure 31) has three main phases: (1) preparation, (2) data acquisition and management and (3) post-processing. In this case, the field study preparation required identification of the aims and objectives when conducting such an experiment as well as clarifying the procedure and requirements. The data acquisition and management involved providing details about the measurement sensors, the data acquisition procedure, data handling and the data storage method. The post-processing stage of the framework pertained to the approaches to analysing the data and reporting the findings (see Figure 31).

Based on the framework, at each stage of implementing this research study, it was necessary to address issues considering ethical, legal, and personal privacy regarding studies that involved human participants as test subjects (see 3.5). The research presented, at its core, has involved being committed to the Belmont Report's[63] basic principles when conducting an investigation that includes human participants, which is "to acknowledge the autonomy of test subjects and to protect those with diminished autonomy" (if applicable).

This chapter's aim is to present the methodological concerns regarding each stage of implementing and conducting a real world driving study. It includes detailed information about the steps taken to design the aforementioned experiment and the study procedure, as well as a review of the monitoring devices used and the approach chosen in terms of acquiring and managing the data. The chapter is structured as follows. Section 3.2 contains the experimental design and section 3.3 covers data acquisition, whilst section 3.4 explains and justifies the post-processing and analysis of the data. Finally, in section 3.5 ethical and privacy issues regarding

[63] The Belmont Report, Ethical Principles and Guidelines for the Protection of Human Subjects of Research, 1979.

this research are discussed in detail. The table below provides a detailed view of this project's research plan, according to the FESTA V-model guidelines.

Table 21. The detailed research plan based on FESTA V-model steps

Preparation	Function identification and description	Investigating the impact of naturalistic driving behaviour differences on energy consumption and road safety
	Use cases	Telematics-enabled motor insurance premiums and fleet tracking and scoring systems
	Research questions and hypotheses	1. What driving parameters (e.g. vehicle speed, acceleration, and engine speed, road geometry) can identify drivers' driving differences? 2. What is an effective way to classify, rank, and group drivers based on driver differences and similarities? 3. Is it feasible to model variations of driving behaviour based on collected real driving data? 4. What driving parameters (e.g. vehicle speed, acceleration, and engine speed, road geometry) can identify drivers' attitudes to driving safely? 5. What is an effective way to classify, rank, and group drivers' safe driving based on dangerous driving habits?
	Performance indicators	1. Understanding and characterising the impact of real world driving on fuel economy 2. Eliciting the contributing factors of drivers' driving behaviour to road safety 3. Investigating potential analytics approaches to rank drivers and the effect of drivers driving behaviour differences

	Study design	Road: pre-defined 4 km route with uphill, downhill and flat sections Vehicle: Less than 5 years old medium size car Driver: Public announcement to recruit drivers, with the target age being between 18 and 34
	Measures and sensors	OBD II dongle with wired connected GPS sensors and a dedicated sim card
Data acquisition and management	Data acquisition	Speed, engine speed, total fuel consumption, and locations
	Database	Secure MySQL database developed
	Data analysis	Exploratory data analysis, descriptive statistics, and Geo-analysis of driving data
Post processing	Research question and hypotheses	Examine 5 research questions
	System and function analysis	Evaluating the system and function strengths and weaknesses
	Socio-economic impact assessment	Identify drivers who are eco and safe driving
Ethical and legal issues		Anonymised personal information Anonymised driving data

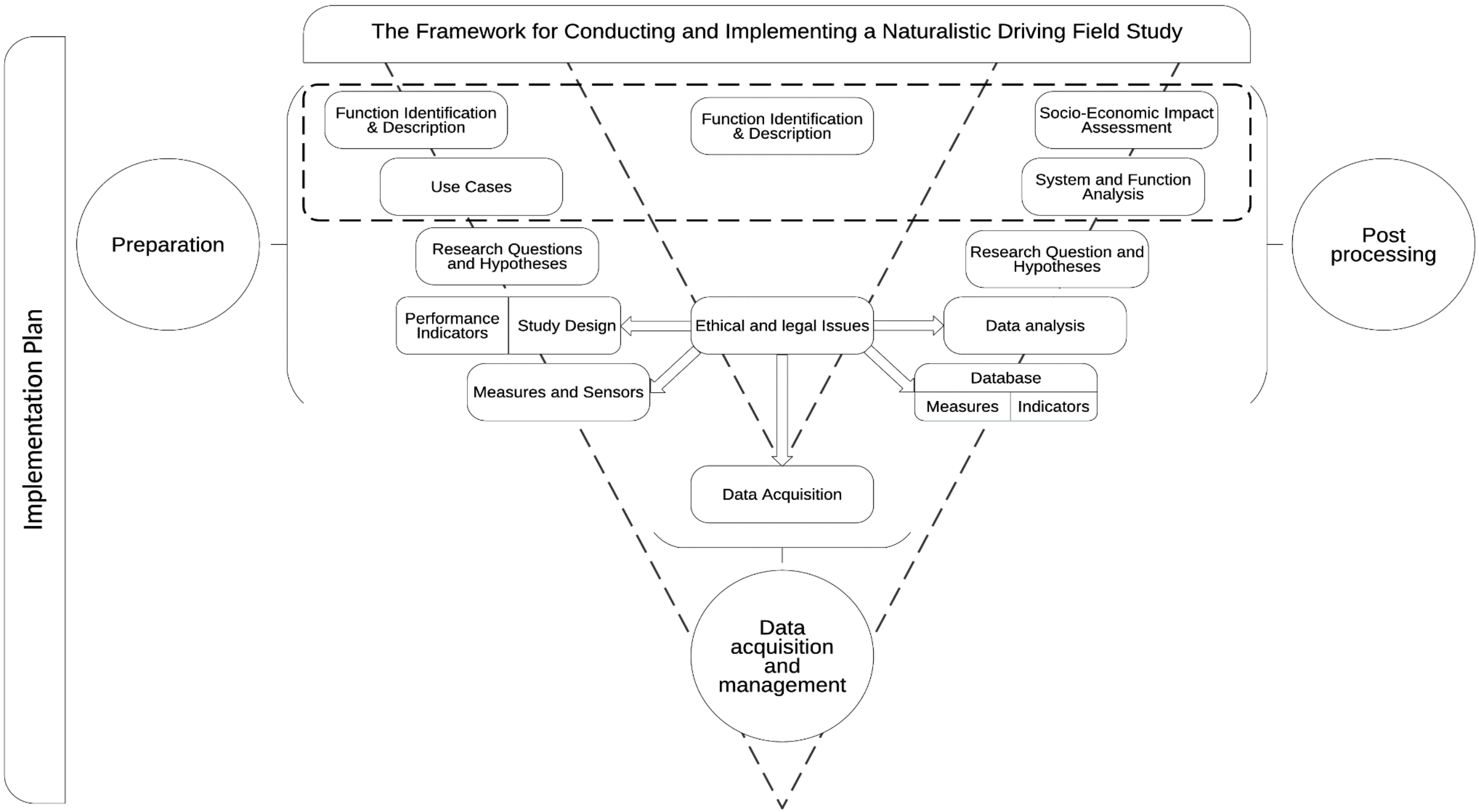


Figure 31. The framework for conducting and implementing a naturalistic driving field study

3.2 Experimental design

Organising a driving event, where the effects of selection bias (inadequate population with insufficiently randomised representation) and performance bias (allocating interventions by the participants and researcher during the test) are kept at their minimum, requires an accessible and equal opportunities communication platform. The **Eco-Safe Driving Challenge** (ESDC) was developed as a communication interface to connect potential drivers to the research team. By being provided with a website and an online signup form, potential drivers were able to apply to participate at driving events. The ESDC web pages also worked as a campaign to raise awareness about the effect of drivers on fuel economy and road safety in addition to providing accessible information about event regulations and criteria. On the web pages, one entire section is dedicated to providing a clear statement of ethical and legal issues and well as answering any concerns regarding data protection and privacy. The general approach towards designing and experimentation in this study is based on the principles of the following standards (listed below) on driving behaviour[64] (vehicle characteristics) and vehicle dynamic field studies, as commonly employed, so as to achieve repeatable and reliable test results. There are eight principles that the following standards are based upon, and that need to be addressed: test scope, test method, testing variables, measurement equipment(s), test conditions, test procedure, data analysis method, and finally, the data evaluation and presentation approach.

1. ISO 4138:2012 Passenger Cars: Steady-state circular driving behaviour, Open-loop test method (BS ISO 4138:2012, 2012).
2. ISO 7975:2006 Passenger Cars: Braking in a turn, Open-loop test method (BS ISO 7975, 2006).

[64] In this case, this does not refer to drivers' driving behaviour, but rather, to vehicle dynamic characteristics.

3. ISO 9816:2006 Passenger Cars: Power-off reaction of a vehicle in a turn, Open-loop test method (BS ISO 9816, 2006).

Since one aspect of this field study was the decision to conduct the trials under control conditions, reviewing these standards helped to establish a test protocol, as follows:

1. Engine cold start (Roberts, Brooks and Shipway, 2014): to avoid the cold start effect on the test, the first round of driving was not included in the study;
2. Review each car's MOT report and previous service records;
3. Make a physical check of tyres, engine oil level, fuel level, lights and diagnostics alarms'
3. Unwanted traffic: to avoid train like traffic caused by the participants, each driver asked to wait 3 to 5 minutes, if another driver left the parking prior to him/her;
4. Record and report weather including humidity, temperature and wind speed as part of the data collection.

Due to the nature of a driving task, there are many possible variables available to measure and monitor, which can be classified into three categories: variables relevant to the road, variables that present drivers' information and finally, those variables that provide information about the vehicle. Designing a successful naturalistic driving study requires consideration of these variables, selecting useful parameters and installing a monitoring system so as to be able to collect data throughout the driving event. This section presents the planning steps taken to host the main driving event, which all the findings and results of this project are based upon. The event took place on 9Th February 2014. Before this, two pilot studies were conducted in May and November 2013 to confirm the feasibility of organising a driving event as well as testing monitoring devices with a large number of drivers.

3.2.1 Driver selection

Drivers and driving enthusiasts were invited to register their interest in the driving event using the online form. The event was an equal opportunities occasion and voluntary in nature, in which all drivers qualified according to certain guidelines were able to participate (Appendix B). Three Sundays were offered for them to choose from and they were asked to note any special requirements they had. Appendix B shows the driver eligibility criteria designed for the event. All drivers had to nominate a co-driver to accompany them during the event. Having met these criteria, a further condition was set for eligibility to attend was being between 18 and 34 years old. This demographic was set for the event, due to the fact that according to Association of British Insurers (ABI), the single biggest cause of accidental death of young people aged 15-24 is being involved in a fatal car accident[65]. Moreover, other campaigns in the UK, such as ABI Campaign for Safe Young Drivers[66], RoSPA's Young Drivers Campaign[67] and Young Driver Focus[68], pointed out that young drivers (18 - 34) are more at risk owing to their being over-confident, having poor assessment of hazards and having a higher tendency to take serious risks while driving (Bates *et al.*, 2014). While the were no limitations to signing up to become a driver in this project, after this period, it became clear that those most interested were between 25 and 34 years old.

3.2.2 Vehicle selection and allocation

Vehicle selection was decided based on three factors: firstly, by reviewing statistics concerning the top brands of vehicles that can be insured cheaply by young drivers (18 to 34 years old); secondly, by ascertaining which had low maintenance and running costs; and thirdly, by

[65] Details available at https://www.abi.org.uk/Insurance-and-savings/Products/Motor-insurance/Young-drivers/ABI-campaign-for-safe-young-drivers
[66] Details available at https://www.abi.org.uk/Insurance-and-savings/Products/Motor-insurance/Young-drivers
[67] Details available at http://www.rospa.com/campaigns-fundraising/current/young-drivers/
[68] Details available at http://youngdriverfocus.org.uk/youngdriverfocus/

ranking the sales figures of the UK's best-selling new cars. Subsequently, the Castel Wireless Telecommunications Co. Ltd Company (the monitoring device provider) and the Hertz Corporation Car Rental Company were asked whether the chosen monitoring device would work in the preferred vehicle and in the case of Hertz, they had been asked whether they would be able to provide sufficient numbers of the selected cars. Firstly, regarding the different brands of vehicles that can be insured cheaply by drivers, taking into account two groups of vehicles (group 1 and 2) classified by insurance groups[69], the results of the UK's largest ownership survey[70] completed by 61,000 UK motorists and car owners, were scrutinised to find the cheapest to insure for young drivers and amongst these were: the Dacia Sandero, Skoda Citigo and the Vauxhall Corsa.

Secondly, regarding car brands sales figures, according to The National Franchised Dealers Association (NFDA) report[71] about new car sales figures, the top three sellers are Ford, Vauxhall, and Volkswagen. Amongst the models that these manufacturers offer, the best-selling cars in 2013 – 2014 were the Ford Fiesta, Vauxhall Corsa, Ford Focus, and Volkswagen Golf. As part of the aforementioned survey, they looked at the running costs, which were based on data provided by drivers about the cost of their new owned vehicle for a three-year period over 36,000 miles of ownership duration. Table 22 shows a summary of the cheapest cars in the UK regarding running cost and insurance group. As the shading shows, the Vauxhall Corsa is the only car that is listed in all three categories.

---

[69] Groups 1 and 2, according to the Group Rating Panel and the insurers, are four or five door small cars with a small engine capacity. According to Thatcham Research, there are eight factors that are used for group ratings: Damage and Parts Costs, Repair Times, New Car Values, Parts Prices, Performance, Safety, Bumper Compatibility, and Car Security (Thatcham Research, 2013).

[70] The survey is organised by Auto Express annually and the results can be seen at: http://www.autoexpress.co.uk/driver-power

[71] Details available at http://www.nfda-uk.co.uk/reporting/new-car-figures/ [Accessed: January-2015].

Table 22. A comparison of top-selling vehicles and cheapest cars to insure and maintain, 2013 – 14

The top 10 best-selling vehicles (ranked)	Cheapest cars to insure (alphabetically ordered)	Cheapest cars and models to run (ranked)
Ford Fiesta	Dacia Sandero 1.2	Toyota Aygo – 1.0 VVT-i 3dr
Vauxhall Corsa	Hyundai i10 1.0	Peugeot 108 – 1.0 Access 3dr
Ford Focus	Peugeot Bipper Tepee	Citroen C1 – 1.0 VTi Touch 3dr
Volkswagen Golf	Renault Twingo	Dacia Sandero – 1.2 Access 5dr
Nissan Qashqai	SEAT Mii	Kia Picanto – 1.1 1 5dr
Vauxhall Astra	Skoda Citigo	Hyundai – i10 1.0 S 5dr
Volkswagen Polo	Skoda Fabia	Vauxhall Corsa – 1.4 3dr
Audi A3	Smart ForFour	Skoda Citigo – 1.0 S 3dr
Mercedes C-Class	Vauxhall Corsa	Volkswagen up! – 1.0 3dr
MINI	Volkswagen Up!	SEAT Mii – 1.0 S 3dr

The Vauxhall Corsa was the only car that statistically showed the highest sales of newly registered vehicles, while providing low running cost and insurance. Consequently, it was selected as the primary vehicle to study. However, as mentioned above, it was necessary to consult with both the Castel Company (project collaborator and monitoring device provider) and the Hertz Rental Company (provider of rental vehicles for the event) to get agreement that this would be our primary choice of vehicle, that it was compatible with the monitoring device and that they would be able to supply sufficient numbers.

The secondary choices were decided upon based on the following criteria. Firstly, it was agreed that instead of drawing on the table above (Table 22), the selection would be made based on the following: the vehicle's similarity to the Vauxhall Corsa (in terms of engine size, build and insurance cost); its availability in the week prior to the event; and finally, that a range of vehicle classes should make up the study sample. Consequently, the other vehicles selected were the Nissan Note (classified as B, B+, or MPV depending on the model), as well as the Nissan Juke and the Fiat 500L, the last two of which are both mini multi-purpose vehicles (mini MPV).

Hertz provided, in total, three of the Vauxhall Corsa VXs, four of the Nissan Notes, one Fiat 500 L and one Nissan Juke. Hence, the Vauxhall Corsa VX and the Nissan Note were the two primary vehicles for the event. Cars were allocated to drivers on a random basis. In total, seven out of nine drivers drove Class B vehicles and as a result, they became the main test subjects The other two drivers were given Class C cars, the Fiat 500L (Driver 8) and the Nissan Juke (Driver 15), and their results were kept separately to those of the Class B drivers, only being used in studies where drivers are compared in terms of fuel consumptions and safety attitudes. With the exception one of the Vauxhall Corsa vehicles, which was manufactured in 2008, the rest were manufactured in or after 2013, and all had manual transmission. Out of the nine participating vehicles, only one of the Nissan Notes (car number 12) and the Fiat 500L (car number 8) were run on diesel fuel. To avoid any misconceptions regarding the drivers' ID (for example the driver's ID that is 1 being considered to be the best driver) drivers ID this started from number 7. The table below presents a summary of the drivers and vehicles participating in the driving event on 9th February 2014.

Table 23. Drivers' details, and vehicle information

Driver ID	Driver Gender	Age	Vehicle Mark	Model	Engine class	Year of manufacture	Drag Coefficient	Fuel type	Transmission types
7	Male	25 – 34	Vauxhall	Corsa VX	B	2008	0.33	Petrol	Manual
8	Male	25 – 34	Fiat	500 L	C	2013	0.31	Diesel	Manual
9	Male	25 – 34	Nissan	Note	B (MPV)	2014	0.29	Petrol	Manual
10	Female	25 – 34	Nissan	Note	B (MPV)	2014	0.29	Petrol	Manual
11	Male	25 – 34	Vauxhall	Corsa VX	B	2013	0.33	Petrol	Manual
12	Male	25 – 34	Nissan	Note	B (MPV)	2013	0.29	Diesel	Manual
13	Female	25 – 34	Nissan	Note	B (MPV)	2014	0.29	Petrol	Manual
14	Male	25 - 34	Vauxhall	Corsa VX	B	2013	0.33	Petrol	Manual
15	Male	25 - 34	Nissan	Juke	c	2014	0.35	Petrol	Manual

3.2.3 The route design and criteria of selection

The main objective when designing the test route is to have a significant amount of control over drivers' safety and to ensure the accuracy of the collected driving data. To ensure repetition and precision of the data collection, a closed driving loop was designed for this study. The route has been developed based on three main objectives: Firstly, it had to be in an urban area so as to represent city driving with everyday traffic settings (e.g. urban traffic flows and traffic obstacles) and road configuration (e.g. traffic lights, a roundabout, and a pedestrian crossing). Secondly, it had to cover a sufficient distance so as to represent a common inner city driving experience as well as such that it provided sufficient data. Finally, for the convenience of drivers and the organising team, it had to begin and end at a location in which drivers' continuous starting and stopping did not violate the Highway Code and put other road users in danger. As a result, a route was selected that was 4 km long and which took approximately 8 minutes to drive.

The University of Bath West Car Park was chosen as the event's primary stop and start location and allocated parking spaces were assigned for the day by the University Security Department. The event organisers' stand, refreshments and toilet facilities were all at the same location for drivers' comfort. The driving route started at the University of Bath West car park entrance to North Road. It then followed the A36 towards Bath city centre towards Bathwick Hill roundabout. After taking the roundabout's first exit, the route went along Bathwick Hill Road back towards the University of Bath. Finally, it went left along North Road and turning right finished at the same starting point. The following tables and figure present driving direction, route properties and location.

Table 24. The step by step direction table

Sign	Descriptions	Distance
Start	Head west on North Rd towards Golf Course Rd	1.40 km
	Turn left onto Warminster Rd/A36	378 m
	Turn left onto Sydney Gardens	306 m
	At the roundabout, take the 1st exit left onto Bathwick Hill	1.66 km
	Turn left onto North Rd	309 m
End	North Road, Bath, Bath and North East Somerset BA2, UK	4.1 km

Table 25. Road properties including the route elevation data

Road properties	Values
Start altitude	164 m
End altitude	164 m
Start and end longitude	-2.335130
Start and end latitude	51.378410
Maximum altitude	173 m
Minimum altitude	29 m
Distance	4.1 km
Total ascent	145 m
Total descent	145 m

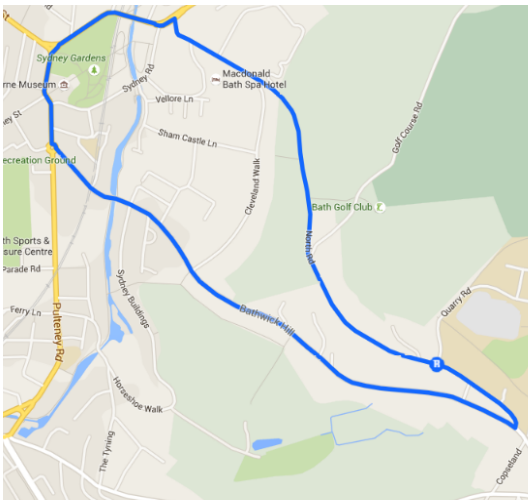


Figure 32. The driving route map

Two secondary characteristics were purposefully included in the driving route. Firstly, it covered the two main roads leading to the two main entrances to the University of Bath coming from the city centre, carrying the highest commuter volume from that direction. Hence, it was anticipated that the findings of this research would be of interest to those driving to and from the university. The second important feature about this route is the fact that it contains an equal amount of downhill driving and uphill driving, that is, around 1.4 km of each. Introducing a steep incline to the route meant that driver behaviour in both the uphill and downhill directions could be analysed and compared.

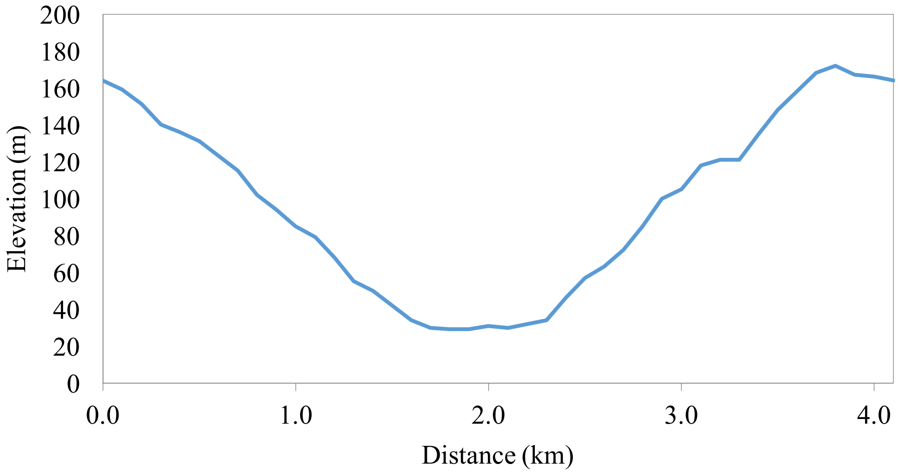


Figure 33. The road profile; the route elevation against distance

As has been mentioned at the beginning of this section, the traffic settings of the route are critical for a better understanding of drivers' behaviour. The route needed to present a sufficient number of traffic settings to become a nonbiased representation of everyday driving scenarios. The selected route contained three sets of traffic lights, one roundabout, and one zebra pedestrian crossing. The table below summarises the traffic settings, road markings, and manoeuvres of the driving course.

Table 26. The traffic configuration, road markings, and manoeuvres

Distance (km)	**Road details**
0-1.4 km	Start point, downhill driving
1.4 km	Entering main road, STOP, turn left
1.5 km, 1.8 km, and 2 km	First, second and third traffic lights positions from the start point
2.1 km	Bathwick Hill roundabout, first exit off the roundabout

2.4 km	Zebra pedestrian crossing with a Belisha beacon lighting
2.4 -3.8 km	Uphill driving
3.8 km	Sharp left turn, short downhill section towards car park entrance

The information above is demonstrated graphically in the two figures below. As is evident, the drivers had to pass through a series of junctions (traffic lights) over the length of the driving route (top map). The topographic contour map[72] in the second figure provides details of ground relief as well as both the natural and man-made surroundings.

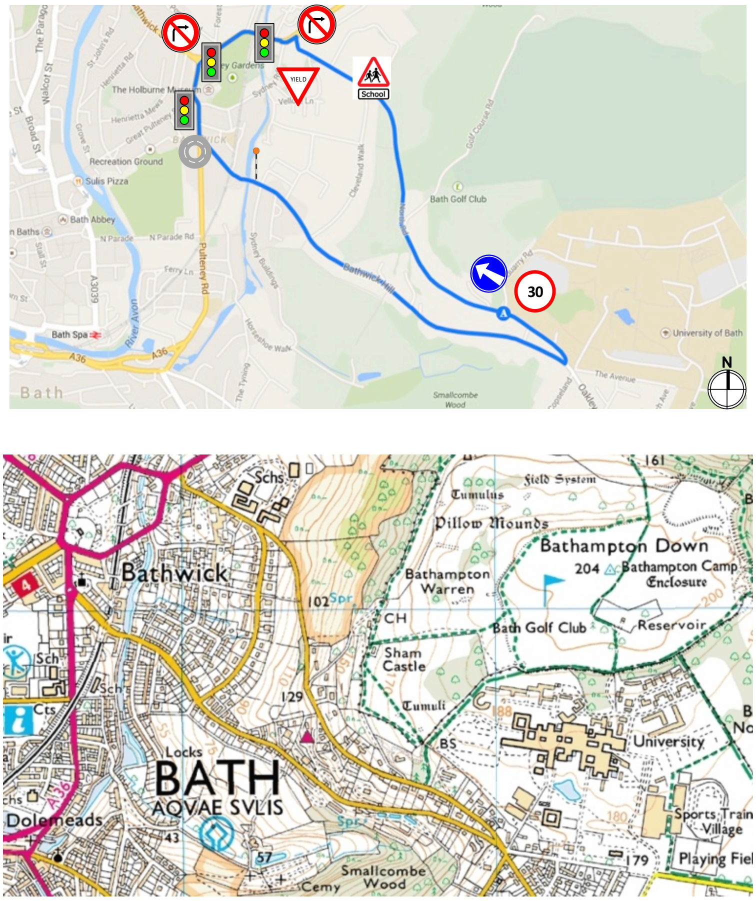


Figure 34. The driving route traffic settings (top), the topographic contour map of the road (bottom)

[72] Also refers to as the ordnance survey (OS) map.

3.3 Data acquisition

The methods of collecting data for naturalistic driving studies include collecting three classes of information: background data, static variables and digitally acquired data from monitoring devices and sensors. This separation meant that the fixed background data would not interfere with the data processing as these were not going to be employed in the analysis. Moreover, the static variables were used to access the digitally collected time series data for comparative analysis, as explained. This section covers data acquisition methods of these variables as well as discussing the utilised monitoring devices, technology and its applications (See Table 27).

3.3.1 Background data

The information about the participating drivers and co-drivers, as well as documents about planning along with running the event was classified as background data. This information was recorded for safety, legal and monitoring purposes. The participants' background data were gathered online from the Eco Safe Driving Challenge (esdc2013.com) website.

3.3.2 Static variables

The static variables refer to information that remains constant throughout the field study, which are assumed to have only a minor effect on the outcomes of driver performance (e.g. vehicle dimensions, or the road profile). Such variables are partially used in the analysis of the data, for instance, to classify drivers based on their cars or to make a comparison between two roads. The static variables of a vehicle pertain to the details about its physical appearance as well as the engine and powertrain system specifications. This class of variables for roads includes unchanging information about road type, traffic settings, and road markings (See Table 27).

3.3.3 Digitally acquired (time series) variables

The information successively measured over a time interval (time series variables) by in-vehicle Data Acquisition Systems (DAS), is classified as digitally acquired variables. This includes any of the data attained through using sensors mounted on the vehicle. For this study, this involved collecting data from an Engine Control Unit (ECU) by data loggers and vehicle location by GPS sensors.

The data logger device uses a standard on-board diagnostic (OBD-II) port to access a car's ECU. It is a common method to monitor vehicle performance along with checking any faults in the vehicle control unit. The dongle connects to the OBD II port of the vehicle to read data from the ECU. Based on the number of available streaming channels of data from the ECU, the dongle can record many parameters, for instance, vehicle road speed (km/h), engine speed (rpm), intake manifold absolute pressure (kPa), intake air temperature (℃), engine coolant temperature (^{o}C) and data from the mass air flow sensor (g/s), to name a few. Each car's location was gathered by the Global Positioning System (GPS) sensor device. This includes information acquired from vehicle direction, latitude, and longitude at 10-second intervals. The GPS sensor time was synchronised with UTC to avoid latency between the time that the actual event took place and the time marked in the dongle, which was then saved in the server. It is important to mention that despite the dominant source of error with the GPS positioning method being due to a satellite's position and atmospheric disturbances, it is an excellent source for providing accurate information about speed and distance travelled (see Table 27)

Table 27. Detailed table of acquired information

Type of variable	**Recorded information**	**Primary data source and instrument**	**Classification level**
Background data	Drivers and co-drivers • Age • Gender • Driving licence • Vehicle ownership details • Personal information	The event sign up forms and ESDC web page All documentation related to host and run the event	Highly classified under the Data Protection Act 1998 (DPA).
Static variables	Vehicle details and specifications • Mark and model • Engine class • Year of manufacture • Fuel type • Dimension details • Engine specification and transmission types	Vehicle user's manual and information provided by the manufacturers	Moderately classified, access limited to the research team and educational use; no public access allowed except under RCUK open access conditions.
	Route information • Road types • Traffic settings and road markings • Roads speed limit information • Route elevation profile • Historical collision data of the route	Online map providers and the road collision historical data published by the Department for Transport.	Moderately classified, accessed through open access and open government licence (OGL) to copy, publish, distribute, adapt, and transmit the information.
Time series variables	In-vehicle data from OBD dongle • Vehicle speed and engine speed (rpm) • Second level of priority of vehicle variables[73] GPS sensor (direction, irection, latitude, longitude and altitude)	Real-time diagnostic and tracking system of IDD212G from Castel and its online platform; user access level: management	Moderately classified, access limited to the research team and educational use; no public access allowed except under RCUK open access conditions.

[73] Parameters such as intake manifold absolute pressure (kPa), intake air temperature (℃), engine coolant temperature (ºC) and mass air flow (g/s) were recorded, if available.

3.3.3.1 Tracking system and monitoring instruments

Castel Company was the industrial partner of this project and provider of the tracking system. The monitoring system included an ECU reader, a GPS sensor and an online monitoring server and database. Their real-time diagnostic and tracking dongle, called IDD212G, comes with a mobile SIM card functionality to send recorded real-time data. The GPS sensor collected vehicle location attributes and was connected to the dongle by an external wire. The ECU parameters and the location data were transmitted to the Castel Company live server via the Global System for Mobile communication (GSM). Figure 35 demonstrates the Castel IDD-212G monitoring device function and data collection procedure. Data from the OBD dongle and the GPS sensor was transmitted by the built-in SIM card function to their online server, which could be accessed live for monitoring purposes and later on, stored securely on their server for post-processing purposes.

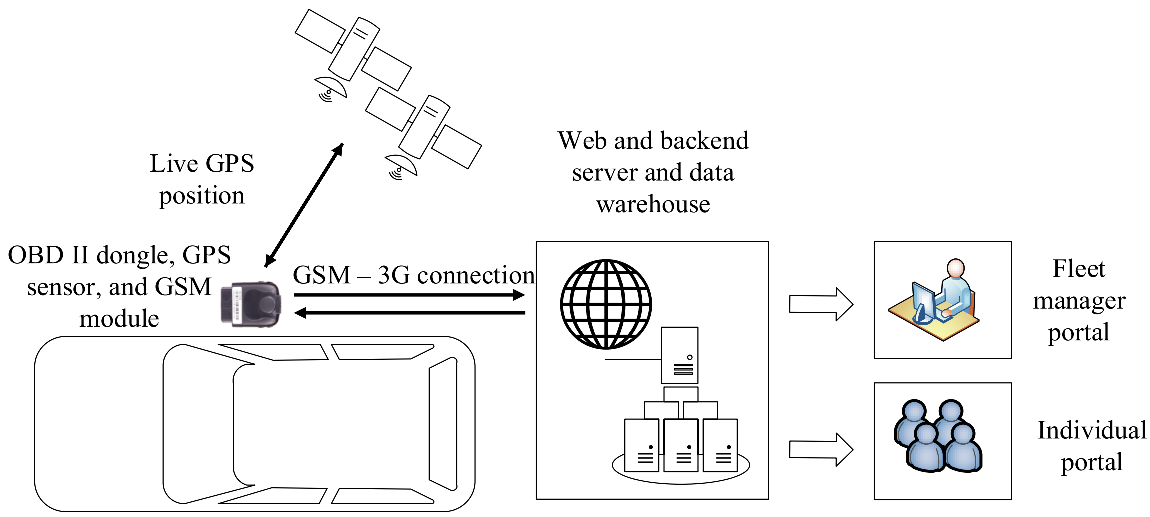


Figure 35. Monitoring device setting

3.3.3.2 Monitoring instrument set-up procedure

The Castel IDD-212G monitoring device is a programmable dongle[74] with GSM functionality. This means that its settings can be customised to record specific vehicle parameters for the ECU, and it can transmit them to the Castel server via a mobile network. The Castel PC-Tool software can program the device and the settings to include: defining the GSM network information, customising threshold rule limits, choosing the uploading time interval, assigning the required parameters to upload, identifying the vehicle engine size and fuel type and finally, synchronising the local time with UTC (Eidson, 2005). All nine monitoring devices used at the event were customised using the provided software. Before the driving event, the mobile network coverage and 3G coverage of all the major mobile network providers were compared[75]. According to the mobile network coverage map provided by Open Signal, the area is strongly covered by all mobile network providers. A special arrangement was made with the EE Limited network provider to supply mobile SIM cards for the event, with 3G coverage as the Castel devices were all configured to support a 3G network including 800, 850, 1700 and LTE bands 1900. Table 28 summarises the general protocol used to configure the parameters of the monitoring devices.

The next step after the required changes were made to the dongles was to connect the monitoring devices to the vehicles' OBD ports. As Figure 36 shows, the GPS sensors were attached to the dongles by cable. This is for two important reasons, firstly, the GPS sensor could use the dongle mobile network to transmit location data to the server. Secondly, the GPS sensor needed to be mounted in the car, on the car dashboard beneath the windscreen, in order to have a clear signal without interference from the vehicle body parts. Each dongle came with

[74] Supported protocols are: J1850-VPW, J1850-PWM, KWP2000, ISO9141, ISO 15765.
[75] Comparison made by using: http://opensignal.com [Accessed: January-2013].

a SIM card tray, in which the full-size SIM card was inserted to transmit both vehicle and location data. The Castel device uploads raw GPS signal to its server second by second. The acquisition frequency of OBD II data is up to the operator to determine. For this project, the uploading data of data was selected to be every 5 seconds. Despite having 1hz location data, Castel only provides the contextual data combined with OBD data to the third parties. Hence, in this study the contextualised data (GPS coordination matched with the map data) has been combined with OBD information, namely, the engine speed for every 5 seconds. As one objective of the project has to been defined as being a low-cost solution, other in-vehicle devices, such as cameras, LIDAR sensors, and steering wheel sensors were not considered for inclusion.

Table 28. Monitoring dongle settings

Configuration parameters	**Setting details**
The GSM network information	The EE Limited mobile network with the 3G SIM card
Threshold rules for driving behaviours categorised as alarming or dangerous driving	The default excessive value setting chosen: ▪ Engine speed over 4500 rpm ▪ Speed over speed limit ▪ Hard acceleration over 0.4 g ▪ Hard deceleration over 0.6 g
The uploading time interval	10 sec intervals for both vehicle engine data and location information
Parameters to monitor	The following parameters were selected: ▪ Engine speed (rpm) – for all vehicles ▪ Vehicle speed (km/h) – for all vehicles ▪ Intake manifold absolute pressure (kPa) – if available ▪ Intake air temperature (℃) – if possible ▪ Engine coolant temperature (°C) – if possible ▪ Mass air flow sensor (g/s) – if available
Engine size and fuel type	For each vehicle, the engine size and fuel type were set
UTC synchronisation	All devices synchronised to UTC for precise time keeping (Eidson, 2005)

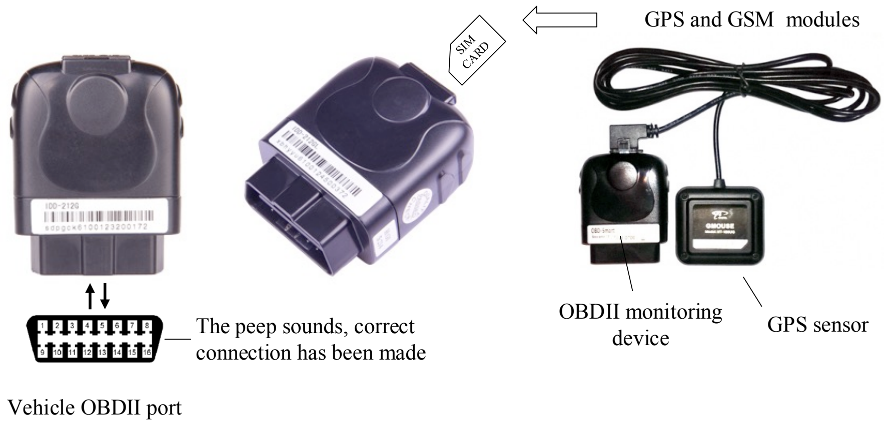


Figure 36. The monitoring device setup

The vehicle OBD port, as suggested in the SAE standards, has to be located inside the vehicle compartment. As is stated in the standard for passenger cars, it needs to be located within 90 cm of the steering wheel, and it must not require any device to access it. The variations in OBD port location depend on the make and models of the car manufactures, but they can only be located in one of the eight locations, as presented in the figure below (source: http://delphiobd.installernet.com/)

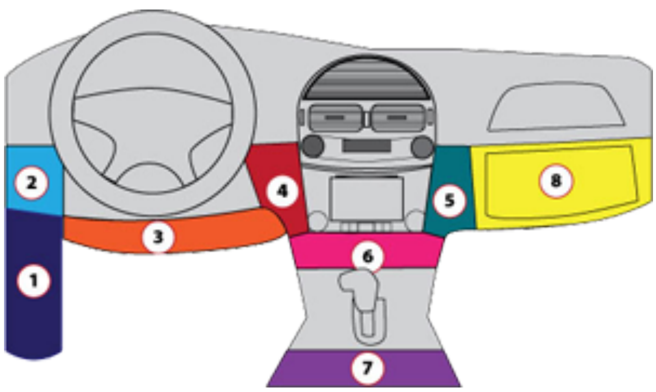


Figure 37. The eight possible locations of an OBD port in a vehicle manufactured after 2000

The OBD port of the Vauxhall Corsa cars is located beneath the dashboard control unit (under the radio and heater control system (location 6). This port for Nissan Note cars is situated on the right side of the steering wheel, and it is covered with a plastic lid (location 1). The correct installation of the monitoring instrument can be checked by the beep sounds the device makes when the right connection has been achieved between the dongle and the vehicle.

3.3.3.3 Live monitoring and the online server

The Castel monitoring package includes an OBD reader, GPS sensor and an online platform for monitoring and storing driving data. The live monitoring feature was extremely helpful during the data collection procedure, as driver safety could be ensured by checking the participants’ performance in real-time. Additionally, live monitoring of the participating vehicles’ locations assured the accuracy of the data acquisition procedures. Each OBD dongle and GPS sensor had a designated ID and password and by using each device’s ID and password, each driver’s performance could be monitored. The Castel online platform provided two levels of access to its users. The first basic access level was for the drivers. The individual drivers could use their device’s ID and password to log into the Castel server. They could obtain their driving history, their driving performance and their driving scores (eco-driving score and safe driving score) according to the Castel scoring method. The second level was access to all drivers’ information, performance and scores. The Castel divides this access level into four main classes: monitoring, tracking, report, and management.

The monitoring section

At the manager access level, the monitoring section was where drivers with active monitoring devices could be monitored in real-time. The primary goal was to check their safety at all times and secondly, in the case of fleet managers, they can assign digital fencing for each driver and to be notified, if one drives outside a specified area. In this project, the monitoring section was used to check whether drivers were keeping to the assigned route during the driving event. Each journey was recorded during the live monitoring for further use.

The tracking section

All recorded journeys could be replayed and checked by providing the date, time and vehicle number. This was particularly useful when the authenticity of the driving data (e.g. the driver following the event route) needed to be checked. The tracking section simply provided access to the recorded data in the form of a video replay.

The report section

After each journey was completed, the server recorded the data and displayed it in a tabulated format. The Castel online platform provided basic reports about vehicle performance including their eco and safety score, fuel usage, total mileage and last recorded location. These reports are available for three months after the date that data are collected and they are then deleted from the server. The Castel reasoning regarding this is, firstly, to protect users' privacy and secondly, the server performance depends on the volume of the saved data. Consequently, regular elimination of past data helps service quality and server response speed. It is important to mention that all recorded data could be accessed and downloaded at any time during the three months subsequent to the event date.

The management section

The final part of the Castel online platform was for changing and modifying settings related to the online platform.

3.3.3.4 SWOT analysis and P-diagram to evaluate the monitoring approach

The Castel company profile has been discussed in the first chapter (see 1.6.2). It's collaboration and support provided an excellent opportunity to conduct this research; however, it is important to assess the chosen monitoring method and monitoring instruments. SWOT (strengths, weaknesses, opportunities and threats) analysis is a useful and a structured method for evaluating the use of an OBD dongle and GPS sensor monitoring system as a means of monitoring driver performance. Since the primary users of these devices are motor insurers and fleet management companies, the SWOT analysis was conducted according to these beneficiaries' interests. This analysis helps to identify the advantages and constraints of the monitoring devices. The SWOT analysis was conducted for both the tracking system and the Castel monitoring devices. The results of the study have been tabulated for comparison purposes and are provided below (see Table 29).

The second part of the monitoring approach is to evaluate the system (i.e. OBDII dongle and GPS sensor) item by item. Through this, potential faults and undesirable outcomes could be identified prior to the case study. To adapt the P-diagram five elements were addressed that were relevant to this project, these being:

- Input signal: system defined input signals;
- Control factors: settings that can be changed or modified to achieve objectives;
- Noise factors: any non-inherent faults that can impact on the design that are not out of the scope of the control factors;
- Desired output: to have a tangible driving and location database useful for addressing the research questions;
- Undesired output: output dataset that facilitates achieving project objectives (see 1.4).

Table 29. The SWOT analysis of the monitoring method and chosen monitoring package

	Monitoring method (OBD dongle and / or GPS sensor)	**Chosen monitoring package from the Castel Company**
Strengths	• Has indirect positive impacts on the environment and fuel consumption • Has an indirect positive effect on road safety due to a positive incentive to drive carefully and be more alert • The motor insurers can individually charge drivers according to their style of driving by using the telematics-based monitoring method	• The device is low cost and requires low maintenance • It is compatible with major mobile network providers • It is easy to use and to set up • The Castel online platform is an excellent tool for monitoring drivers • Secure method of data storage and accessing data
Weakness	• There are some concerns with using telematics as a primary method for insuring vehicles, such as personal data privacy • The costs of purchasing, installing and decommissioning devices are unclear	• There is a limitation regarding the number of parameters that can be monitored at the same time • Data collection interval is 10 secs (0.1 Hz), and this might be an issue in the case of potential studies on collision detection • High-frequency data only available for live monitoring • Continuous fuel consumption is not always easy to monitor
Opportunities	• The cost of telematics monitoring devices is decreasing rapidly • The ability to discriminate based on driving data, rather than a driver's age or gender, is useful for insurers • Online driving feedback to drivers has indirect benefits to be eco-friendly and safe drivers • There is the capacity for added services (real-time feedback, remote diagnostics)	• Many regulations have come into effect since 2014 to track and monitor vehicles • Devices are gaining more attention from the private sector and individual drivers • The online feedback for drivers is an attractive way to help them improve their driving skills and style
Threats	• There are legal risks involved in customer tracking • Laws in some countries preventing insurers from charging for the rental of monitoring devices • New technology and app-based systems might affect the market's future	• The Castel server is sometimes difficult to access, as is the case with all live servers and the three-month data access period can be problematic, especially for longitudinal studies • New types of vehicles (autonomous cars, alternative fuelled and electric vehicles) can change the market future in terms of using these devices

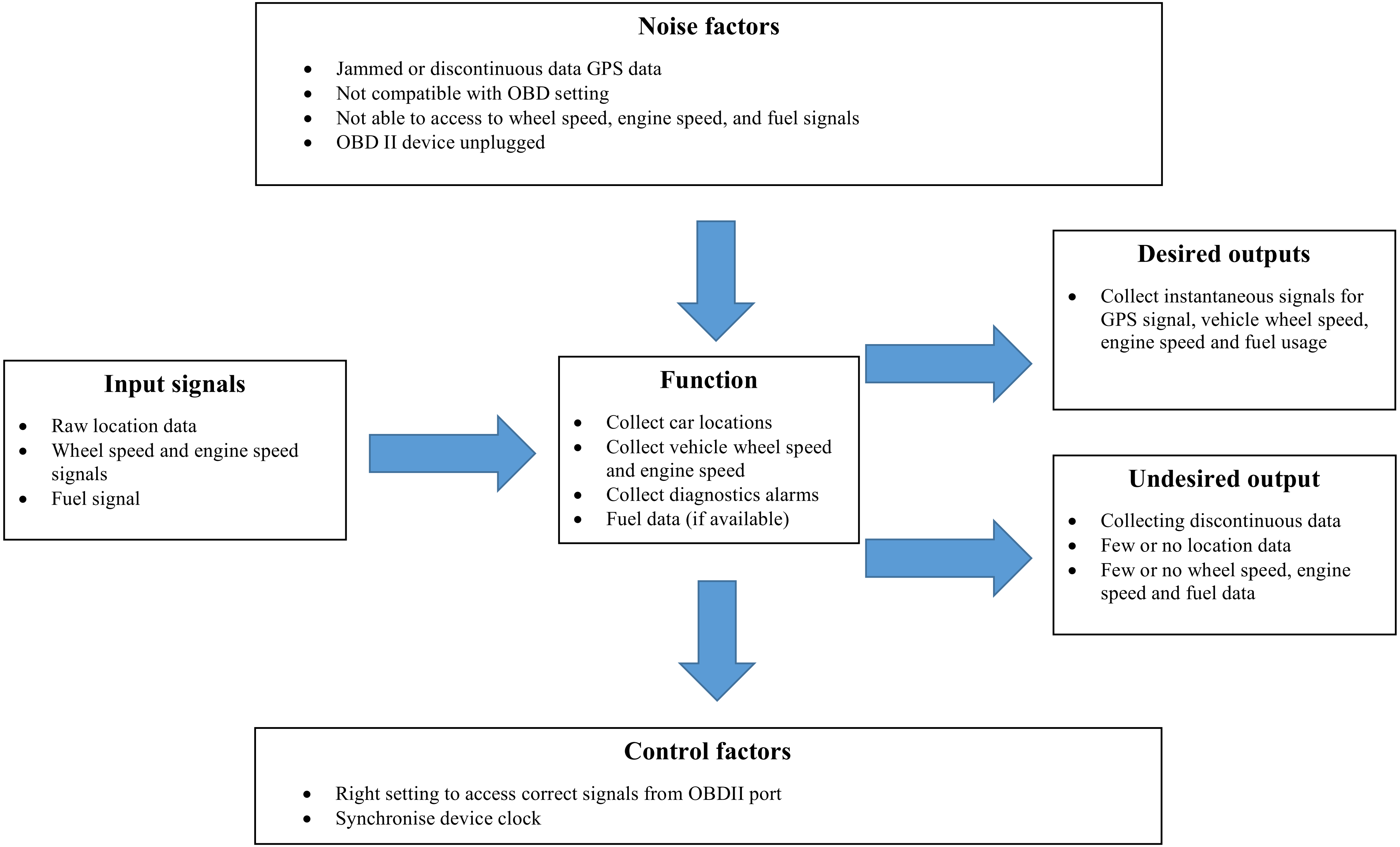


Figure 38. P-diagram assessment of monitoring devices

3.3.4 The study and data acquisition procedure

This subsection covers the data acquisition procedures as well as the driving event day plan.

3.3.4.1 Driving date

To prevent a causation of tail-like traffic flow due to the driving event, over the course of two weeks, the traffic flow of vehicles driving in and out of the West car park was observed on Saturdays and Sundays between 11 am and 2 pm. Finally, it was decided that the event should take place on a Sunday so as to have a low traffic volume. Drivers received three potential dates on a Sunday to participate in the driving event, and the 9th February 2014 was selected as the most convenient day for all drivers to attend. Appendix C presents the recorded weather conditions on the driving day. It was scattered clouds with light rain and visibility of around. 8.7 km (Appendix C).

3.3.4.2 Vehicle arrangement

As has been mentioned, the Hertz car rental company provided vehicles for the study and the drivers were asked to be present at the Hertz branch on Saturday 8th February 2014 with their valid UK driving licence to receive their assigned car. All the drivers drove to the West car park to park their car in the designated area allocated by the university for the driving event. All vehicles were rented with a full tank of fuel, and they were all returned to Hertz in the same way. All drivers, vehicles and, later on, monitoring devices, were assigned to the same numbers between 7 and 15.

3.3.4.3 Monitoring devices

The monitoring devices (OBD dongle, GPS sensor, and SIM card) were numbered and assigned to all nine drivers. Each monitoring device was programmed according to the assigned vehicle's specifications. All were installed in vehicles on Saturday 8th February 2014 and were left in the vehicles for the driving day. The devices were obtained from the drivers after the driving event.

3.3.4.4 The map of the route arrangements

The drivers received maps of the driving route in two formats; firstly, as a printed map with the direction descriptions and secondly, in a digital map format (KML file) to be opened by a driver's mobile phone map and navigation app. Moreover, the drivers familiarised themselves with the route by following the lead car marshal during their first trial.

3.3.4.5 The field study arrangements and tasks

On the day, every driver received the information package which asked them to accomplish five uninterrupted driving rounds on the event route, without interfering with monitoring devices.

The event aims and objectives

The drivers were briefed about the research and its goals and objectives as well as the driving event rules and criteria. They were also informed about the health and safety aspects of participating in the event. Drivers and co-drivers were also assured about the voluntary nature of their attendance at the event. Drivers received an information package including the route map, the event information and consent forms, which they duly signed.

Task one: First driving lap with two lead cars

After the briefing and having completed the consent forms, they were assigned to participate in a driving task to familiarise themselves with the route and driving conditions. The participating vehicles were escorted by two marshal cars to complete a driving lap. These cars (one in front and one at the back) guided the drivers to leave the West car park and followed the route designed for the study. When using the online fleet management platform from Castel, the monitoring devices were checked to make sure they were working appropriately and accurately on an ongoing basis throughout the driving test.

Task two: The driving event

The drivers were then asked to complete five separate driving laps which started and finished in the West car park. They were requested to take a break and to leave at least five minutes between each driving attempt. Moreover, if they saw a driver leave the car park, they had to allow another five minutes to prevent train-like traffic. Special arrangements were made to use the university 8W building foyer as a reception and relaxation area for participants during the event.

The event closure

After completing five laps and having returned to the car park, a final debriefing was carried out, and the monitoring devices were decommissioned from each vehicle. The Castel online server was checked to ascertain whether all the driving data had been uploaded correctly.

Figure 39. The Eco Safe Driving Event, 9th February 2014, the West car park, the University of Bath

Table 30. The driving day and data acquisition procedure and timetable

Date	**Timings**	**Task details**
8th February 2014	To be completed before 5 pm	Collected the vehicle and parked it in the West car park
8th February 2014	-	Monitoring drive installation
9th February 2014	9:00 to 9:30	Set up, preparation and team gathering
	10:00 to 10:30	Drivers arrival, monitoring devices checked (1st attempt)
	10:30 to 11:00	Briefing session consent forms, map of the route
	11:00 to 11:30	Driving trial with two lead cars, monitoring devices checked (2nd attempt)
	11:30 to 12:30	Driving session: cars left the campus with 5 min timing space between each car
	12:30 to 13:30	Lunch break (location: 8W building foyer)
	13:30 to 14:30	Driving session
	14:30 to 15:00	Participants return to the car park
	15:00 to 15:30	Monitoring devices collected, the final check and the end of the event
	15:30 to 16:30	All vehicles refuelled and returned

3.4 Post-processing and data analysis methods

According to the driving field study framework (see Figure 31) (the modified version of the V approach by FESTA), the next stage of the analysis is to post-process the data. This step includes taking required action to prepare the data so as to be able to respond to the research questions and hypotheses. The data preparation step involves managing all recorded data, developing a data cleaning strategy, and finally developing a database. The data analysis contains all methods applied to the data to provide answers to the research questions. The first part of this section is dedicated to data management and the second half explains the methods applied to address the hypotheses and research questions.

3.4.1 Data management

As previously described, the data were stored according to three categories, namely, background information, static variables and time-series data. Each driver was allocated a number, which was duplicated for their car and their monitoring device, thereby anonymising the participants.

3.4.1.1 Data clearing strategy

Because the drivers started their trip from the University of Bath West Car Park, their driving records included driving data while they were driving in and out of the car park. Consequently, a data cleaning process was applied to the data points according to route location to eliminate the driving data inside the car park. The data cleaning strategy involved removing the driving data of drivers driving inside the car park area and knowing where these occurred in the time-series data meant that each individual's driving loop data could be accurately identified. The start of the loop was at the junction between the link road to West Car Park and North Road, which was deemed appropriate because the drivers have to stop according to the Highway Code, as the latter is a main road. That is, it was easy to see where the cars were stationary in the mined data. However, as the engine and GPS data were reported at 10-second intervals on many occasions they came from slightly before or after this particular junction and when this was the case the next sampling point was selected. All the drivers' data were arranged based on the laps they completed during the driving event. Each driver completed five laps and to identify each of them; the following strategy was applied to the dataset. The third trip of driver number 10 was assigned the code 10, 3, thereby linking the lap number to the universal ID for each driver, i.e. their ID number was linked to the vehicle and monitoring device. The significance of this data cleaning approach was to reduce the prime data to a core manageable data set that was easy to access and structured so as to facilitate comparative spatial analysis.

3.4.1.2 The development of the driving database

Having arranged the driving data as explained, the next activity was to develop a database in tabular form so as to allow for multiple enquiries. That is, regarding a naturalistic driving study, it is important to build a responsive and functional database, for this then provides the opportunity to conduct complex analysis and make various comparisons among the different drivers. Several studies have suggested methods for handling driving data and developing a driving database (Inata, Raksincharoensak and Nagai, 2008; Han, Yao and Liu, 2014). Regarding the academic research perspective, such a database can be accessed during future research for comparative purposes, and it can be revisited for further analysis by the lead researcher.

As a result, a data uploading protocol was designed and implemented so as to maintain a reliable database. The data were classified according to sensitivity and the importance of the information in relation to the research objectives. The database was built using SQL (MySQL server) and uploaded to the University of Bath MySQL server. The database was designed according to two main concerns, firstly, as a secure method of data management and secondly, providing dynamic query functionality for analysing data. The database architecture was based on the three core elements of the driving task: the road; the driver and the vehicle. Consequently, three separate tables were added representing these three aspects. Finally, the historical collision data, including crashes and casualties between 2005 and 2013, were taken from published reports by the Department for Transport (Graves *et al.*, 2014). However, due to the size of the recorded data, only those pertaining to the driving route were considered at this stage of database development. The figure below visualises the developed database. As the driving data are kept in the Castel online server for only 3 months, all driving records were downloaded and added to the develop MySQL database on the University of Bath server.

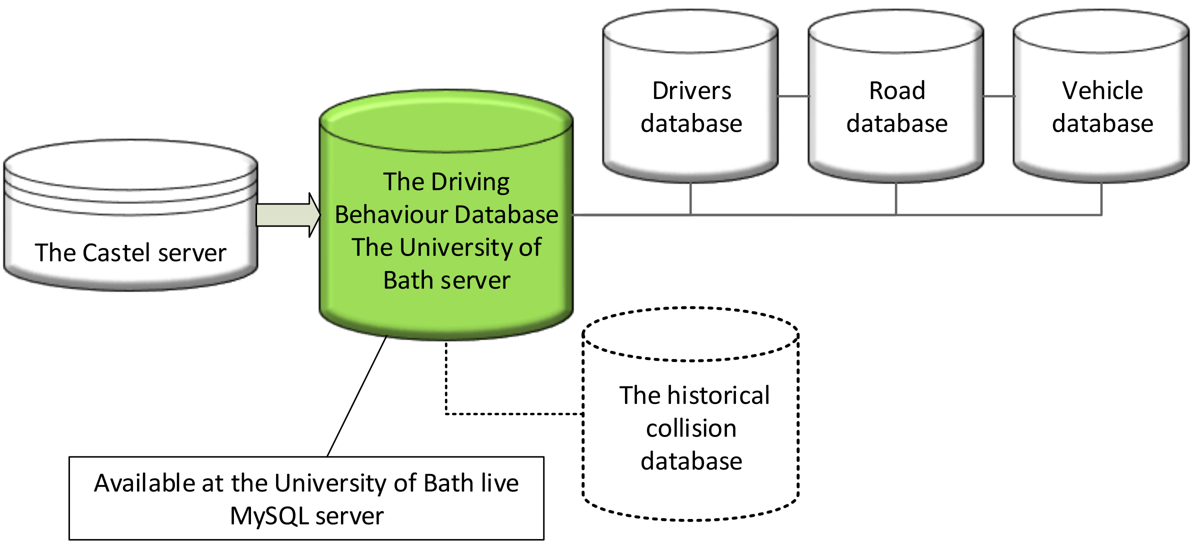


Figure 40. The driving behaviour database current architecture

3.4.2 The Study's Key Goals

From the literature review, it became apparent that prior research into eco-driving or safe driving has focused on identifying, classifying and/or modelling these concepts. Consequently, the aim in what follows is to cover all of these aspects through a series of studies as illustrated in topics below.

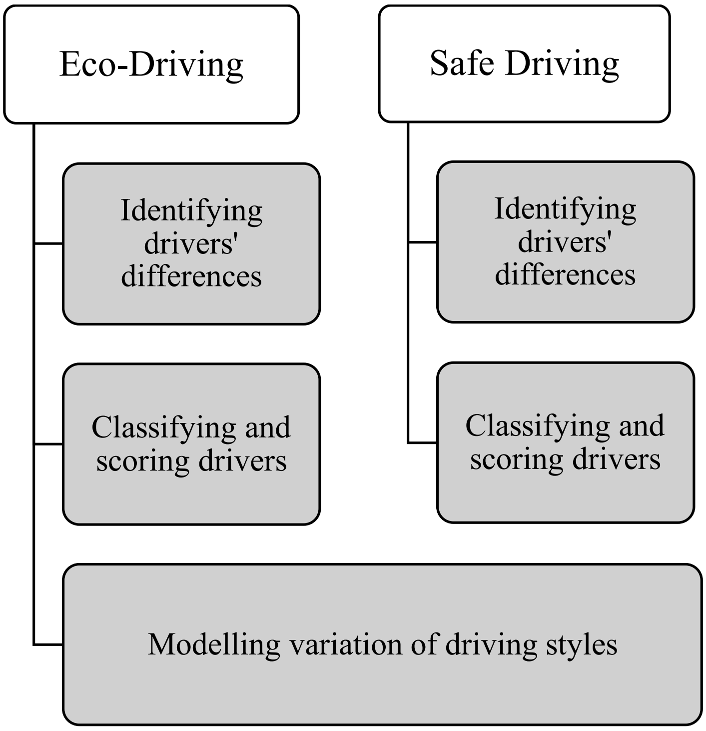


Figure 41. Research topics

Based on the above research topics, the following table (Table 31) was developed to organise the analysis conducted in this research. As the table shows, the post-processing studies have been divided into two topics (eco-driving and safe driving) and three categories (identifying, classifying and modelling). In total, there were 16 studies undertaken to achieve aims of this research.

Table 31. An overview of the post-processing analysis

	Identifying driver differences	**Classifying and scoring drivers**	**Modelling and simulating different driving behaviour**
Eco Driving Studies	1. Study of vehicle speed – distance and engine speed – distance analysis 2. Study of the relationship between drivers' road speed and engine speed 3. Study identifying drivers' aggression by comparing their speed vs. acceleration 4. Geo-analysis of vehicle parameters in conjunction with the route elevation profile 5. Descriptive analysis of vehicles' speed profile	1. Evaluating the similarity of drivers' speed profiles 2. Classifying drives based on their fuel usage and proposing the Vehicle Specific Power – Fuel Consumption metric (VSP-FC) 3. Eco driving classification according to the fleet management scoring system	1. Study of the factors affecting fuel consumption and constructing a fuel consumption forecasting model 2. Modelling different driving styles and road condition effects on fuel usage and car emissions by using IPG Carmaker
Safe Driving Studies	1. Study of drivers' safe speed perception when approaching road settings 2. Study of drivers' hard acceleration and deceleration habits 3. Study for identifying drivers with coasting downhill habits 4. Identifying collision-prone zones	1. Safe driving classification according to the fleet management scoring system 2. Two modifications to the fleet management scoring system	

3.4.3 Identifying drivers' driving performance differences just from vehicle engine data

The first step toward examining driver's differences is to study the primary data that is collected from vehicles via the monitoring devices in the context of eco-driving. The aim of the following analysis is to differentiate drivers' driving performances from one another by observing the vehicle information. This includes investigating such matters as identifying individuals with aggressive acceleration and deceleration behaviour as well as the excessive speed habits of a particular driver. To achieve the aim of distinguishing different driving patterns and driver behaviour five studies were carried out. In the following studies the drivers' vehicle speeds, engine speeds, driven distances and time taken to complete each lap are made as the parameters. Specifically, these studies are as follows:

1. Study of vehicle speed – distance and engine speed – distance analysis (3.4.3.1).
2. Study of the relationship between drivers' road and engine speed (3.4.3.2).
3. Study to identify drivers' level of aggression by comparing their speed vs. acceleration (3.4.3.3).
4. Geo-analysis of vehicle parameters in conjunction with route elevation profile (3.4.3.4).
    1 Geo-analysis of the vehicle parameters
    2 Geo-analysis of acceleration behaviour among drivers
    3 Geo-analysis of the vehicle-specific power among drivers
5. Descriptive analysis of vehicles speeds profile (3.4.3.5).
    1 Analysis of drivers' consistency and the range of chosen speed
    2 Descriptive analysis of drivers' speed distribution

3.4.3.1 The analysis of vehicle speed – distance and engine speed – distance graphs

Two parameters for which data were collected from all the participating drivers were the vehicle and engine speeds. First, for each driver their trips were identified and separated (this is explained in the data cleaning section). One only completed four laps owing to a monitoring device malfunction and two, completed six, with the data for the final lap only being used partially in the analyses recorded correctly. Secondly, the precise distance travelled for each trip was calculated by integrating the speed-time function, with the area under the speed-time curves representing this.

Equation 8. The distance travelled when $v(t)$ is non-negative

$$Distance\ travelled = \int_a^b v(t)dt$$

where, the time between $t = a$ and $t = b$ is the duration of each driving trips.

As a result, the vehicle speed – distance and engine speed – distance graph can be plotted. The results are self-explanatory, showing the variations in their vehicle and their engine speeds, which is a standard procedure when reporting type of investigation.

3.4.3.2 Study the relationship between drivers' road speed and engine speed

In every vehicle powertrain system, the drivetrain characteristics determine the link between the vehicle road speed (km/h) and the engine speed (rpm). Since the relationship is a fixed and linear one, it is possible to examine whether drivers' habits when shifting gears are efficient or not. With a view to find the relationship between the vehicle speed and engine speed, the following equation was used. Drawing on the work of Tutuianu et al. (2013) in relation to determining the gearshift strategy for vehicles with manual transmissions, a simplified version of their method for calculating the WLTP[76] for this aspect is employed (Tutuianu *et al.*, 2013). The calculation requires information on the gear ratios, the final drive ratio and the rolling circumference of tyre, which can be extracted from vehicle specification manuals.

Equation 9. For estimating the engine speed using the known road speed[77]

$$RPM = \frac{(speed \times gear \times final\ drive\ \times 88)}{Cir}$$

where,

RPM = The engine speed

$Speed$ = The vehicle road speed (mph)

$Gear$ = The gear ratio for each gear

$Final\ drive$ = The final drive ratio

The combined conversion factor constant = 88

Cir = The tyre-rolling circumference (feet)

The only uncertain parameter in this equation is the tyre rolling circumference. Since the standard wheel sizes of the Vauxhall Corsa and Nissan Note are known, the rolling circumference values obtained from tyre suppliers' data sheets were used for this calculation. The area of interest when conducting this analysis is to compare the ideal calculated values with the drivers' so as to elicit which of them performed abnormal gear changes and when the

[76] Worldwide Harmonised Light Vehicles Test Procedures.
[77] Simplified equation based on (True, 2003)

drivers' engine speed was higher than usual (one of the signs of a late gear shifting habit). The outcomes of this study have been plotted for all the Vauxhall Corsa and Nissan Note vehicles[78].

3.4.3.3 Drivers' level of aggression by comparing their vehicle speed vs. acceleration

Several studies have investigated the relationship between vehicle speed and acceleration, in order to understand the accuracy of vehicle emission measurements or to ascertain the vehicle performance and fuel consumption efficiency (Jiménez-Palacios, 1999; Johansson, Färnlund and Engström, 1999; Bokare and Maurya, 2013). Other researchers (Mierlo *et al.*, 2004; Joumard *et al.*, 2006; Pelkmans and Debal, 2006; Felstead, McDonald and Fowkes, 2009) have studied the effect of driving styles on vehicle emissions by comparing real world driving data and hardware in loop testing on chassis dynamometer test cycles. It was generally concluded that driver style in terms of calmness or aggression differences can be identified by comparing their speed vs. acceleration diagrams. Some studies have used the vehicle speed vs acceleration diagram in various ways, such as to investigate real-world naturalistic driving behaviour differences (Alessandrini *et al.*, 2006; Felstead, McDonald and Fowkes, 2009) and for improving in-vehicle driving assistance systems (Miyajima *et al.*, 2007; Forbes, 2009). A similar approach to that of Felstead, McDonald and Fowkes (2009) has been taken into account to evaluate the differences among drivers. Appendix D shows an example from Felstead, McDonald and Fowkes (2009). where they compared passive and aggressive driving with laboratory results.

[78] Vehicle number 12 is diesel engine hence it is excluded from this study.

In this research, in order to identify drivers' differences regarding their driving style (passive, normal, aggressive) the speed vs. acceleration data were used in a scatter plot format, to illustrate the drivers' driving variations and their choice of acceleration. It was anticipated that there would be a concentrated distribution of acceleration vs. speed for calm drivers and widely scattered one for aggressive drivers.

3.4.3.4 Geo-analysis of vehicle parameters in conjunction with road profile

Geo-analysis of the vehicle parameters is an analytical method using GPS information to understand the overall distribution of vehicles' road speed, engine speed, and driver acceleration patterns over the route driven. Moreover, it is an excellent way to demonstrate driver's differences in a visual form. This type of analysis has been used in investigations regarding intelligent traffic management methods as well as studies that consider the topology of the road as a factor influencing vehicle emission and performance. Including the spatial characteristics of the route is seen as a robust way for understating differences between driver behaviour (Alessandrini *et al.*, 2006; Van Schangen *et al.*, 2011), determining the accuracy of vehicle emissions performance monitoring techniques (Frey *et al.*, 2001; Noland *et al.*, 2004; Pelkmans and Debal, 2006) or developing driving models (Pentland and Liu, 1999; Amditis *et al.*, 2007; McGordon *et al.*, 2011).

As explained earlier in relation to the route criteria, the selected route contained downhill and uphill driving. Specifically, it comprised 1.8 km of downhill and obviously, 1.8 km of uphill driving, with an average ±16% change in gradient. When taking this into account, it sheds more light on the vehicle engine data in terms of driver differences in relation to uphill and downhill driving situations, which is unavailable when referring simply to the raw data. There were three parts to this analysis, which was aimed at identifying and presenting drivers' different behaviour according road gradient. Firstly, the geo-analysis of the vehicle road speed and engine speed in combination with the route setting was investigated. Secondly, each driver's acceleration profile was compared with the road elevation profile, for the first, third and fifth lap. These were chosen for two reasons, first, so as to reduce the amount of data needing displaying and second because driver differences came over as more distinct by omitting the even laps. This selection of laps only pertains to when graphical illustrations are presented. The final part of this analysis pertained to a comparison between the drivers' vehicle specific power (VSP) against road gradient, for this analysis also the first, third and fifth driving trials were considered.

3.4.3.4.1 Geo-analysis of the vehicle parameters

This study comprised an analysis of vehicle speed and engine speeds in relation to road profile and was divided into two stages. For the first, the drivers' first, third and fifth vehicle and engine speeds were plotted against the driving route elevation data. The second stage involved investigating driver differences using a geoprocessing tool to draw their vehicle and engine speeds according to the GPS location points. The benefits of using this geo-analysis method are, first, its simplicity for exploring the distribution of vehicle parameters within the driving route and second, the ease by which it can identify differences among drivers. The ArcGIS geoprocessing tool provided by the Esri Company was used to build separate layers of data on a map for each driving lap. Every layer contained each driver's location for one of the engine parameters, that is, either the vehicle speed or the engine speed data, at that location. This accumulates across all the laps, thus producing a solid route illustrating the behaviour of that parameter. In total, the Vauxhall Corsa and Nissan Note drivers' data collection led to seven maps representing road speed profiles and seven pertaining to vehicle engine speed profiles.

3.4.3.4.2 Geo-analysis of the acceleration/deceleration behaviour among drivers

A comparison between the drivers' road and engine speed provides insights into their differences regarding speed at different locations (see subsection 3.3.4.1). The first part of the above study involved investigating driver differences with regards to the gradient. The same method can be applied to the instantaneous acceleration/deceleration of the drivers while driving downhill and uphill. To illustrate drivers' differences in acceleration (m/s^2) and deceleration *(-* m/s^2*)* of the first, third and fifth driving laps were calculated at ten second intervals.

3.4.3.4.3 The geo-analysis of the vehicle specific power among the drivers

The final geo-analysis study that was conducted to compare differences among drivers pertained to the use of Vehicle-Specific Power (VSP). As it has been discussed in chapter 2, the VSP was introduced by Jiménez in 1999 as a method to quantify measured vehicle emissions more accurately through the use of remote sensing (Jiménez-Palacios, 1999). The application of VSP has evolved since then, with recent applications involving examination of the effect of road gradient and driving style on real-world vehicle emission measurements (Van Mierlo *et al.*, 2004; Frey, Zhang and Rouphail, 2008; Wyatt, Li and Tate, 2013, 2014). Studies such as those of (Wyatt, Li and Tate, 2013; Sentoff, Aultman-Hall and Holmén, 2015) used the VSP as an instantaneous, second by second, vehicle engine power to examine and model vehicle emissions on the Motor Vehicle and Equipment Emission System (MOVES) software[79]. In contrast, to previous VSP applications that have used it to calculate vehicle emissions, here, it is used to demonstrate differences in driving styles. The VSP is the instantaneous power the vehicle generates based on its speed and acceleration on a certain road

[79] Details available at: http://www3.epa.gov/otaq/models/moves/ [Accessed: January-2014].

with known slope. It can be calculated using the (Jiménez-Palacios, 1999) method, with the simplified version of the VSP equation (see equation 10).

Equation 10. The vehicle specific power (VSP) equation (Frey *et al.*, 2001; Zhai, 2007)

$$VSP = v \times (a + g \times sin(\phi) + \psi) + \zeta \times v^3$$

Where,

VSP = vehicle specific power ($W/Kg = m^2/s^3$)

v = vehicle speed (m/s)

a = vehicle acceleration (m/s^2)

ϕ = road gradient

ψ = rolling resistance coefficient

ζ = drag coefficient (m^{-1})

g = gravitational acceleration (m/s^2)

The road gradient is calculated based on the route elevation record, and a rolling resistance of 0.03 is assumed based on the typical value for car tires on tar or asphalt. The following is the summary of the input variables used to calculate the VSP.

Table 32. Input data to the VSP equation

Equation variables	**Input data**
Vehicle speed	Collected speed data
Vehicle acceleration	Calculated from speed data
Road gradient	Google earth road elevation
Rolling resistance coefficient	0.03
Drag coefficient	Based on each car specification

In contrast to previous studies that suggested using the data binning method to reduce the effect of slightly differing observations, in this study the actual instantaneous value of all the VSPs was calculated (Frey, Zhang and Rouphail, 2008; Wyatt, Li and Tate, 2013; Duarte *et al.*, 2015). Finally, the VSP trajectory of the drivers is compared with the gradient at the different points on the course so as to illustrate drivers' differences as they drove round the undulating route. In this study, the comparison is only made for the drivers first, third and final driving laps, for reasons explained previously.

3.4.3.5 Descriptive analysis of the vehicle speed profile

Drivers' attitudes towards choosing appropriate speeds determine their risk-taking behaviour and hence, their chance of being involved in a car accident. Moreover, a direct link was established by early researchers between vehicle speed and the severity of a crash (Fildes, Rumbold and Leening, 1991; Norris, Matthews and Riad, 2000). Other studies have found drivers with neglectful behaviour, such as disregarding the traffic rules and not following speed limits, are more likely to put themselves and other road users at risk of harm (McCarthy, 2001). Further studies have elicited that there is a positive correlation between measured vehicle emissions and drivers' speed (Mierlo *et al.*, 2004), which implies that reducing the road speed limit will lessen them. From what has been discussed, it can be concluded that drivers with a tendency to ignore the speed limit have a higher risk of involvement in an accident, producing more harmful emission gases (CO_2 and NO_x) and using more fuel (Madireddy *et al.*, 2011).

In the following subsections (3.3.3.5.1 and 3.3.3.5.2), the aim is to investigate the consistency of the vehicle speed for the drivers across the different laps as well as the profile of their chosen speed across all the laps. In order to accomplish this objective, first, they are compared according to their consistency in relation to their chosen vehicle speed (speed cumulative frequency) and in terms of the range of speeds they drove. Secondly, by using descriptive analysis, the drivers' road speed distribution is examined and the normality of their speed distribution investigated. It is only logical to consider that the distribution of speed could not be normally distributed, because over the course of each driving trip, drivers were mostly driving close to the speed limit. That is, for the driving speed to be normally distributed, it would mean that a driver spent an equal amount time at high and low speeds, which would be very unlikely. Hence, the study presented in subsection 3.3.3.5.2 is carried out only to confirm this fact. Here is the summary table of the analysis undertaken in this section:

Table 33. Summary table of the descriptive analysis of the vehicle speed profile

Study objective	Analysis conducted	Approach and visualisation
Driver consistency and the range of chosen speed (3.3.3.5.1)	Drivers speed consistency over the course of five laps of driving	Cumulative frequency is plotted
	Proportion of choice of speed in certain ranges	Stack bar chart
	Comparison between driver's driving speed profiles	Box-and-whisker plot
Descriptive analysis of drivers' speed distribution (3.3.3.5.2)	Speed distribution and normality test	The normality test for all the drivers' speed distribution and the Q-Q plot (probability plot) Individual histogram graphs for all nine drivers so as to compare drivers' speed distributions (Appendix E)

3.4.3.5.1 Analysis of driver consistency and the range of chosen speed

In this subsection, three methods for comparing and demonstrating driver differences in terms of their driving speed are described. Firstly, to illustrate their speed consistency over the course of five laps of driving, their speed cumulative frequency is plotted for all the laps and reproduced in graphic form., with the main aim at this stage being to identify drivers with distinguishable differences in terms of the range of speed they used (Redsell, Lucas and Ashford, 1993; Haworth and Symmons, 2001; De Mol *et al.*, 2009). The work of the Daniel, Brooks, and Pates (2009) has shown that the speed cumulative frequency plots are robust indicators of drivers with a tendency to exceed speed limit as well as drivers' speed range consistency (Daniel, Brooks and Pates, 2009). Analysis. such as comparing drivers speed fluctuations between each trial was considered, however, it was concluded that this would be more beneficial in future works when larger quantities of driving data could be recorded.

The second method for illustrating driver differences in their chosen speed is by calculating the percentages of certain speed ranges that each driver drove within. It is important to mention that the speed limit throughout the driving route is 30 mph, and this process allows for identification of when a particular driver exceeded the speed limit and for what duration. This value was calculated based on the number of times drivers are in a certain speed range. Four speed ranges were defined as a guide, and the outcome is illustrated in a percentage stacked bar chart format, these ranges being:

- Between 0 and 5 mph (time at rest)
- Between 6 and 24 mph
- Between 25 and 30 mph
- Over the speed limit

The relative distribution of speed within the sample reveals useful insights into driver preference regarding the different range of speeds. The final method for identifying driver differences is to compare their speed profile by plotting a box-and-whisker plot (box plot), which demonstrates the patterns of speed range used by the drivers. Specifically, the boxplot illustrates driver speed profile central tendency, interquartile range as well as showing those drivers who exceeded the speed limit. Prior to drawing the plot, the value of the median, the first and third quartiles and the maximum and minimum of the speed profiles were calculated. In both this and the previous study, the outcomes for drivers 8 and 15 are also included.

To summarise, all three methods discussed in this section are aimed at illustrating the differences in the drivers' speed profiles. The findings, as presented in the next chapter, highlight the fact that driver's choice of speed is an identifiable parameter that can reveal useful insights regarding differences in driving behaviour.

3.4.3.5.2 Descriptive analysis of drivers' speed distribution

In this subsection, the distribution of the drivers' speed profiles is compared. This step has been necessary to proceed to compare drivers' driving performance with the proposed power and network analysis.

A Note regarding driving speed distribution

There has been extensive work on investing the overall distribution of driving speed for applications, such as defining the speed limit, or re-examining the road safety measures (Berry and Belmont, 1951; Salter, 1974). As it has been ascertained driving speeds are normally distributed on any stretch of road, such as motorway segments. Recent work by the Iowa State University Institute for Transportation investigated this claim by conducting a exhaustive study on speed distribution on Iowa state highways (Center for Transportation Research and Education, 2001; Souleyrette, Stout and Carriquiry, 2009). In their research, they aimed to establish safe and reasonable speed limits based on real-world driving data rather than predefined speed limits. They worked with the idea that drivers tend to travel at a speed relative to the road conditions, level of congestion, weather conditions etc. Below is the typical driver speed distribution on the road (Center for Transportation Research and Education, 2001; Souleyrette, Stout and Carriquiry, 2009; Hashim, 2011).

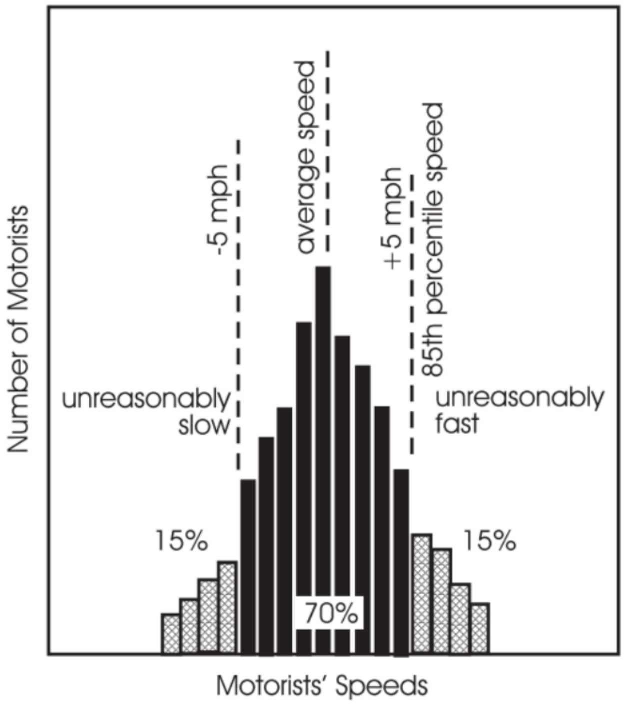


Figure 42. Typical speed distribution on the road (Center for Transportation Research and Education, 2001)

While the overall shape of a driving speed distribution for a large sample takes the form os of symmetric shape around the speed limit, the key to individual differences is not whether they are normally distributed or not, but rather, the degree of the differences between the skewness and kurtosis coefficient. Hence, it is important to emphasise the fact that the nature of speed distribution is not the key, but rather, the interest lies in differences between the peak and the "tailedness" of the distributions.

The significance of conducting the normality test of speed distributions is to verify and present the degree of skewness and non-normally distributed nature of speed profiles, as this needs to be taken into account during subsequent analysis. The following steps were taken to analyse drivers' speed distributions. Firstly, for every driver the standard descriptive statistics parameters. Including the sample mean$s$ ($\bar{x}$), and standard deviation (σ) of the driving speeds are calculated. The degree of symmetry or skewness of the distributions is ascertained, which in this case refers to the measure of speed distribution asymmetry around the average speed and a: positive skew, is where the right tail is longer, whilst a negative one, is where the left tail of the distribution is longer. The skewness value (S) provides a numerical insight into drivers' speed distribution tendency, and it has been calculated by using Equation 12.

Equation 11. The equation to estimate the standard deviation from a sample

$$\sigma(x) = \sqrt{\frac{\sum_{t=1}^{T}(x_t - \bar{x})^2}{(T-1)}}$$

Equation 12. The equation to calculate the skewness of a distribution

$$S(x) = \frac{\sum_{t=1}^{T}(x_t - \bar{x})^3}{(T-1)\times\sigma^3}$$

Where, for both equations above:

$\sigma(x)$= The sample standard deviation

$S(x)$= The sample skew

x_t= The non-missing values in the sample

$\bar{x}$ = The mean of the sample

T= The number of non-missing values in the sample

The sample excess kurtosis value is the final value that is calculated by conducting the normality test and the goodness-of-fit test on the drivers' speed samples. This describes the tailedness and the shape of the sample, which along with the skewness value will help to determine whether the drivers' speed samples are distributed normally. The excess kurtosis (K) is calculated by following equation, where the parameters are the same as in Equations 11 and 12 (Tsay, 2013). 'The importance of communicating the computing method to calculate the skewness and excess kurtosis is the fact that there are different ways to calculate them. The aforementioned formulas are easy to adapt and compute in a timely manner. The calculation has been conducted in Microsoft excel with the help of add-on formulas provided by a software company called NumXL.

Equation 13. The equation to calculate the sample's excess kurtosis

$$K(x) = \frac{\sum_{t=1}^{T}(x_t - \bar{x})^4}{(T-1)\sigma^4} - 3$$

The normality test of the drivers' speed profiles is the final analysis that has been undertaken. There are two methods used to illustrate that how drivers' speeds are distributed and whether they are normal. Firstly, graphically comparing the empirical speed sample distribution against a theoretical Gaussian distribution (normally distributed) and the Q-Q plot (probability plot), shows how closely the distributions are correlated. The diagram is created by plotting the quantiles of all the drivers speed samples against the theoretical sample that is normally distributed. The second method applied to examine the goodness of fit of the drivers speed data is by using the Jarque-Bera test (JB test) (Bera and Jarque, 1982).

Among the widely used methods to test the normality and goodness of fit of driving speed samples including Pearson's chi-squared test, the Kolmogorov-Smirnov test (Hashim, 2011), and Anderson-Darling test (Rakha *et al.*, 2011), the JB test is the most successful in examining the normality of the sample according a skewness coefficient and kurtosis (Bera and Jarque, 1982; Thadewald and Büning, 2007; Kuhn and Johnson, 2016). In Thadewald and Büning's (2007) comparison study of the power of several tests[80], they concluded that "the JB test is superior in power to its competitors for symmetric distributions with medium up to long tails and for slightly skewed distributions with long tails" (Thadewald and Büning, 2007).

The characteristics Thadewald and Büning (2007) described are presented in individual driving speed histograms (see Appendix E). Every histogram shows a symmetric distribation with long tials (in one or both sides) and all are slighly skewed. This confirmed the fact that the JB test would be an excellent choice to test the goodness of the sample. The JB test checks the deviation and departure of the speed sample data from being one that is normally distributed by comparing its skewness and excess kurtosis value against a normally distributed sample. The JB value is found using the following equation, with the parameters being those previously used.

[80] Tests included in the study are the Jarque-Bera, Kuiper, Shapiro-Wilk, Kolmogorov–Smirnov and the Cramér-von Mises tests (Thadewald and Büning, 2007).

Equation 14. The Jarque-Bera equation to test the goodness of fit within the sample (Bera and Jarque, 1982)

$$JB = \frac{T}{6}\left(S^2 + \frac{K^2}{4}\right)$$

The following graphs are presented in the results section.

- The normality test of all drivers' speed distribution and the Q-Q plot (probability plot)
- Individual histograms for all nine drivers to compare speed differences (Appendix E)

3.4.4 Classification of drivers' driving performance from vehicle engine data

Driver profiling, driver classification, driver DNA and driver scoring are terms used in both the literature and industry for grouping drivers based on their driving similarity. The goal of this grouping procedure is to classify drivers in an appropriate manner to distinguish their differences in a measurable order. Given the classification outcome can have repercussions for drivers, the accuracy and fairness of this ranking procedure are crucial. That is, based on the results of the classification, drivers with risky behaviours could face penalty charges and in the case of vehicle insurance providers, it can cost the Pay How You Drive insurance policyholders more money. If this type of classification reflected real driving conditions, which in many cases it does not, then the assessment of driving behaviour would be more accurate and hence, fairer to the motorist. In this subsection, three methods of classifying drivers are covered. Firstly, the drivers are grouped together based on the similarity of their vehicle speed. Secondly, they are ranked based on their fuel consumption and finally, they are scored using a fleet management technique. The following table summarised all the methods covered in this subsection.

Table 34. Summary table of drivers' driving performance from vehicle engine data

<table>
<tr><th>Study objective</th><th>Analysis conducted</th><th>Approach and visualisation</th></tr>
<tr><td rowspan="2">Classifying the drivers based on the similarity of their speed distribution (3.4.4.1)</td><td>The Wilcoxon-Mann-Whitney (WMW) U test method</td><td>Nonparametric test of the null hypothesis</td></tr>
<tr><td>Network analysis</td><td>Visualisation of drivers' relationship</td></tr>
<tr><td rowspan="2">Ranking drivers based on their fuel usage (3.4.4.2)</td><td>Fuel economy and fuel consumption rating</td><td>Metrics including fuel per fixed unit of distance ($L/100km$) and distance per fixed unit of fuel (km/L) or miles per gallon (mpg),</td></tr>
<tr><td>The Vehicle Specific Power – Fuel Consumption metric</td><td>Fuel usage as a ratio of VSP</td></tr>
<tr><td>Drivers' performance classification according to the fleet management scoring system (3.4.4.3)</td><td></td><td>Driving score based on the number events of excessive acceleration, deceleration and engine speed</td></tr>
</table>

3.4.4.1 Classifying drivers based on similarity of their speed distribution

Given the driving speeds of the participants were not normally distributed, but rather, symmetrically distributed with tails and skewed slightly, the parametric techniques, such as ANOVA or a t-test for comparing their driving speeds is not appropriate. Hence, to address this a nonparametric method that first introduced by Mann and Whitney in 1947 is used (Mann and Whitney, 1947). With this method the degree of similarities between each two samples is investigated to measure the effective sample size (Lenth, 2001).

3.4.4.1.1 The definition of statistical power and effective sample size

Power testing of a binary hypothesis refers to finding the probability of whether the test correctly rejects the null hypothesis when the other hypothesis is true (Ellis, 2010). In statistical terms, the power of the test is the probability that the test returns statistically significant differences between two samples (Lenth, 2001; Ellis, 2010; Tabachnick and Fidell, 2012). In the case of the driving trials, the number of laps was five, and each drivers' speed profile was paired off to determine their degree of similarity for this duration. If two drivers have very similar profiles, then the power test will require an effective sample size that is very large before there is a significant difference detected between the drivers' profiles and vice versa (Lenth, 2001; Ellis, 2010). In other words, the test calculates not only the number of times the trial would need to be repeated before two drivers' speed profiles would be distinguished, for it also identifies the degree to which two drivers drive in the same manner. The power test is used to establish the sample size for surveys, field studies and lab tests. For instance Bartlett, Kotrlik and Higgins (2001), used a power test and the effective sample size method to dermine the adequate sample size for survey research (Bartlett, Kotrlik and Higgins, 2001).

This interoperation of the statistical power and the sample size to measure the degree of association (similarity) between two drivers is a novel approach to the field of naturalistic driving behaviour.

3.4.4.1.2 The Wilcoxon-Mann-Whitney (WMW) U test method

The method used to conduct the power analysis so as to determine the effective sample sizes between pairs of drivers' speed profiles and hence, the level of similarity between the two, as explained above, is the Wilcoxon-Mann-Whitney (WMW) U test. The method works under conditions where the samples of studies are not normally distributed. The nonparametric nature of this analysis means that the analysis does not have pre-assumptions about the probability distribution of the assessed samples. This contrasts with other methods, such as the t-test, for which requires a normal distribution. The U test (WMW test) works robustly with small independent samples for calculating effective sample sizes (Mann and Whitney, 1947; Lenth, 2001; Faul *et al.*, 2009; Faul and Erdfelder, 2014).

In order to adapt this analysis for the purposes of the study, every pair of drivers' speed data sets were compared to see whether they were similar or not. The null hypothesis, in this case, is where two samples are equal while the alternative is where they are not. In simple terms, the test computes the differences between each driver by calculating their level of similarity using the "shift model" and subsequently the effective sample size for pair is displayed.

Equation 15. The shift model equation

$\Delta: G(x) = F(x - \Delta)$ for all x

In order to calculate the total sample size, the mean and standard deviation values of each driver were computed. The power error was set to $1 - \beta\ error\ probability$ = 0.95, and it was assumed that every sample had an equal number of trials. The outcome is the effective sample size for these two groups and the results have been tabulated in a matrix format.

3.4.4.1.3 Network analysis of drivers' speed distribution similarity

The network analysis is a visualisation tool to present the relationships and level of association between nodes in the network. For this study, network analysis was conducted to demonstrate, diagrammatically, the associations between the drivers in terms of their vehicle speeds. Link analysis is a form of network analysis where the relationships between members of a sample are investigated, which works based on a node and edge principle. Individual drivers are the nodes in the model, and the links between them (edge) depend on their similarity. In this case, the similarity value is defined as the sample size values from the previous study (i.e. if the sample size value is high, then the two nodes are similar and hence, the link between them is strong). That is, the sample size values determine the edge weight and hence, the strength of the association between two drivers (two nodes).

This relationship is translated into the weight of the connected link between two nodes. The similarity of two drivers is, firstly, demonstrated by the distance between two nodes. Small weights (small effective sample size) means that there is a weak relationship, and hence, the drivers are distant from each other and vice versa. The Gephi software[81] has the capacity to include both the length and the thickness of the line between two nodes to demonstrate the level of association between them, in this case, the drivers' speed profiles. However,

[81] Details at: www.gephi.org

visualisation by distance/thickness is not an effective method of demonstrating the relationship between drivers in this case, because the range of sample size values is from 37 to 22,785. In order to visualise the association between drivers in a homogeneous manner the sample size values are standardised using the following method (Equation 16).

Equation 16. Sample data standardisation method

$$z_i = \frac{x_i - \min(x)}{\max(x) - \min(x)}$$

Where,

z_i = The normalised value of effective sample size value of the i^{th} data point

x_i = The effective sample size value of the i^{th} data point

$\min(x)$ = The smallest sample size value

max(x)= The biggest sample size value

The weighted edge values are changed to the standardised version of the effective sample sizes, thus demonstrating the relationship between all the drivers (nodes). That is, the nodes are arranged in a clockwise manner using NodeXL, where the thickness of the lines connecting them signifies the level of association between the drivers. In sum, whilst the Gephi software was initially used to analyse the network relationships so as to determine the drivers' degree of similarity, after standardisation, the NodeXL package allows for simple representation of the level of association of all the drivers' profiles with all the other participants.

3.4.4.2 Ranking drivers based on their fuel economy and fuel consumption

The second approach towards classifying drivers is to rank them based on the amount of consumed fuel at the driving event. This subsection, firstly, investigates the different methods and metrics for calculating a driver's fuel economy and consumption ratings. Standard methods of reporting and measuring drivers and vehicles' fuel efficiency are then explained. Subsequently, a metric is put forward for investigating fuel consumption in the context of real-world driving, i.e. taking into account the impact of the road topography on fuel consumption.

3.4.4.2.1 Fuel economy and fuel consumption ratio

To rate drivers' driving performances in terms of fuel consumption, it is important to consider the three common ratios employed by the motor industry to rank drivers' (vehicles) fuel economy. The two most important and widely used metrics are:

- The amount of consumed fuel per fixed unit of distance ($l/100km$);
- The distance per fixed unit of fuel (km/l) or (miles per gallon mpg) (which is the exact metric commonly used and is displayed for drivers on modern car dashboards).

As it has been mentioned before, the following parameters are recorded by monitoring devices:

- The total travelled distance (by GPS signals);
- The total travel time (accumulated trip durations);
- The total litre of consumed fuel (total fuel consumption according to the OBD II dongle).

The total fuel economy values were calculated using both measures for all the drivers. The results are presented in bar chart format so as to provide a clear comparison between the participants' amount of fuel consumed.

3.4.4.2.2 The Vehicle Specific Power – Fuel Consumption metric

As it has been discussed in chapter 2, the precise evaluation of the fuel consumption is critical for regulating new measures to increase air quality and reduce gas emissions. The methods covered in the previous sections were aimed at providing a metric measure to evaluate fuel consumption of vehicles, and they are useful because they can help drivers to drive in a more eco-friendly manner (Boriboonsomsin, Barth and Vu, 2011). However, they are not designed to take into account outside influences on fuel consumption, such as road gradient. To this end, first, the standard method of evaluating a vehicle's fuel efficiency is discussed and later a new metric is introduced to compare drivers' performances on level ground by including parameters of vehicle dynamics[4] and road topography. The standard method of reporting and measuring vehicle fuel efficiency is to calculate the Brake Specific Fuel Consumption (BSFC). This makes it possible to compare the fuel efficiency of different engines and hence, different drivers' performances. The BSFC is computed by dividing the fuel consumption ratio (B) by the vehicle power $P\ (kW)$, which becomes $b_e\left({}^{g}/_{kWh}\right)$ (Equation 17).

Equation 17. The (brake) specific fuel consumption equation

$$b_e = {}^{B}/_{P}$$

Power as $P = M.\omega = 2.\pi.M.n$

Where,

b_e= BSFC $(g\ /\ kWh)$

B = Fuel consumption rate (Kg/h)

P = Vehicle power (kW)

M = Torque (Nm)

ω = Angular velocity (rad/s) and n = Engine speed (rpm)

According to two standards of the internal-combustion engine (IC engine) DIN 1940 and DIN70020 (DIN 1940, 1976; DIN 70020-7, 2013), the precise calculation of specific fuel consumption is found by dividing the fuel usage ratio B by the effective power, P_{eff} (kW), instead of vehicle power, P. The effective power is calculated based on the engine pistol displacement value and it represents the net engine horsepower supplied by the IC engine (Robert Bosch GmbH, 2011). Therefore, the brake specific fuel consumption equation from the above section can be modified by replacing the vehicle power with effective power in the denominator, and the specific fuel consumption can be expressed as follows:

Equation 18. The specific fuel consumption based on effective power (Robert Bosch GmbH, 2011)

$$b_e = {}^{B}/_{P_{eff}}$$

$$P_{eff} = V_H . p_e . {}^{n}/_{K}$$

Where,

b_e= BSFC(g / kWh)

B = Fuel consumption rate (Kg/h)

P_{eff} = Vehicle effective power (kW)

V_H= Displacement of the engine (dm^3)

$$V_H = \frac{\pi}{4}. bore^2. stroke. number\ of\ cylinders$$

p_e= Mean piston pressure (bar)

n = Engine speed (rpm)

$$K = \begin{cases} K = 1\ for\ 2 - stroke\ engine \\ K = 2\ for\ 4 - stroke\ engine \end{cases}$$

It is important to note that the aforementioned specific fuel consumption equations are not entirely computable for studies similar to this one, because many of the engine characteristics are inaccessible to researchers. Moreover, without having instantaneous fuel usage data collected with the Portable Emissions Measurement System (PEMS), it is not possible to classify drivers styles' based on previous attempts in terms of such an issue as fuel usage (Zhai, 2007; Wyatt, Li and Tate, 2013). It is also not possible to compare their real-world emissions under the influence of their driver aggressiveness (Nam *et al.*, 2003; Sentoff, Aultman-Hall and Holmén, 2015). However, a solution to this problem is proposed next.

In order to include the effect of road conditions on fuel usage and vehicle dynamic characteristics[82], a new metric is proposed based on the specific fuel consumption model. However, instead of the vehicle effective power being used as the denominator of specific fuel consumption equation, the vehicle specific power is deployed. In this case, the road specific fuel consumption can be calculated, because the effect of the road conditions on fuel usage is included in the model. The Vehicle Specific Power – Fuel Consumption metric (VSP – FC) allows for comparison of drivers on an equal footing since a vehicle's engine capacity, and its characteristics are excluded from the equation. Instead, parameters such as road gradient, rolling resistance and the vehicle drag coefficient are used to calculate the rate of fuel consumption. To demonstrate the effectiveness of this method, the (VSP – FC) values are calculated for all drivers. In this analysis, the total fuel consumption (g/s) has been used to

[82] The vehicle dynamic characteristics includes: vehicle speed and acceleration, aerodynamic drag and rolling resistance (subsection 3.4.3.4.3).

calculate the (VSP – FC) for total fuel consumption per total vehicle specific power. The VSP – FC values can be calculated based on following equation:

Equation 19. The Vehicle Specific Power – Fuel Consumption metric

$$VSP - FC = {}^{B_i}/_{|Total\ VSP|}$$

Where,

$$VSP - FC = \left[\frac{g_{fuel}/_{s}}{W/_{Kg_{vehicle}}}\right]$$

B_i = Fuel consumption (g/s)

VSP = Vehicle specific power ($W/Kg = m^2/s^3$)

3.4.4.3 Classifying driver performance according to the fleet management scoring system

The standard scoring practice among fleet management service providers and the usage-based vehicle insurers, such as the insurance premium called the Pay How You Drive (PHYD), is to rank drivers based on a set of time interval threshold rules. This method of driver profiling is commonly known as the telematics scoring or driving analytics method. This subsection is aimed at examining this approach by applying its techniques to rank the drivers' performance based on the fleet management scoring approach. To achieve this goal, firstly, the telematics driving scoring method is explained. Secondly, the technique is applied to the driving data collected from the driving event so as to classify them according to the fleet management and insurers' methods for calculating premiums, commonly known as driver scoring (driver DNA).

The telematics driving scoring method is a technique that works based on recording the occurrence of dangerous behaviours. A ranking method orders the drivers according to a number of registered alarms. There are four general alarms and risky behaviours that auto insurers and fleet managers are interested in capturing via threshold rules, which are excessive engine speed (rpm), speeding, harsh acceleration, and sudden deceleration. The scoring model is based on counting the number of times the drivers commit these risky behaviours and then ranking them accordingly. Each alarm comes with a coefficient ratio, which presents the severity of the effect it has on the driving score of any given driver. The value assigned to each alarm is set by the service providers and insurers. According to the Castel method of scoring the eco-driving performance of drivers, three alarms affect this, these being high engine speed (rpm), total number of sudden acceleration, and sudden deceleration. The occurrence of these alarms is captured by using the GPS and the OBD monitoring data. The Castel allocates a specific coefficient ratio to these alarms, with the threshold rules for these alarms and their coefficient ratios being presented in the table below.

Table 35. Excessive limit values for eco-performance scoring and eco-driving coefficient ratio

Alarming Behaviour	Excessive limit value	Coefficient ratio
High engine speed events	4500 rpm	4
Hard acceleration events	0.4 g	3
Hard deceleration events	0.6 g	2

Based on the fleet management scoring system, the drivers are ranked according to their eco-driving scores. In order to allocate a numerical value to each driver, every driving attempt is scored out of a maximum 50 marks, where 50 means the driver never made any dangerous mistakes. For each driver, the total number of sudden accelerations, decelerations and times a driver exceeded the engine speed are counted and then the final mark for every driver is calculated using the following equation.

Equation 20. Telematics scoring method to calculate the eco-driving score of drivers

$$\text{Eco} - \text{driving score} = 50 - [(number\ of\ sudden\ accelerations \times eco\ ratio) + (number\ of\ sudden\ decelerations \times eco\ ratio) + (number\ of\ high\ RPMs \times eco\ ratio)]$$

This equation can be rewritten in a shorter format as follows:

$$\text{Eco} - \text{driving score} = 50 - [(SA \times \text{a}) + (SD \times \text{b}) + (RPM \times \text{c})]$$

Where,

SA = Total number of sudden accelerations on all locations

SD = Total number of sudden decelerations on all locations

RPM = Total number exceeding the high engine speed threshold

a = Eco-driving coefficient ratio for sudden acceleration

b = Eco-driving coefficient ratio for sudden deceleration

c = Eco driving coefficient ratio for high engine speed

The following guide was developed to allocate an alphabetic representative for every driving score range, as illustrated in table below. The results of this study are then visualised through graphs using the alphabetic markings and standard energy consumption colouring.

Table 36. An alphabetic representation of drivers' scoring range

Driver scores range	Alphabetic representative
$45 \ll Driver\ score \ll 50$	A
$35 \ll Driver\ score \ll 44$	B
$20 \ll Driver\ score \ll 34$	C
$Driver\ score \ll 19$	D

3.4.5 Developing a fuel consumption forecasting model and modelling drivers' differences by virtual simulation

As explained at the beginning of this section, well-grounded comparative research on drivers' driving includes identifying differences, classification according to similarity and finally, using this information to create models for simulation. In subsections 3.4.3 and 3.4.4, both the identification phase and classification stage were discussed, respectively. This subsection addresses the development of models based on the real-world data gathered from the previous phases. To achieve this objective, there are two methods selected to model drivers' differences. The first involves using the data collected from vehicles to determine the factors that greatly affect fuel usage and subsequently, a regression model is constructed based on these so as to predict the fuel consumption of the tested vehicles. The second study is based on using the IPG Carmaker[83] driving simulation software to model and simulate drivers' driving utilising the collected data from the driving event. Two simulation scenarios are designed and tested to examine the feasibility of using this method to model various driving behaviours in the future. The following list summarises the post-processing studies that are pursued in the next two subsections.

1. The factors affecting fuel consumption and constructing a fuel consumption forecasting model (3.4.6.1).
2. Modelling different driving styles and studying the effect on fuel usage and car emissions by using IPG Carmaker (3.4.6.2).

[83] For more information, visit: www.ipg.de/simulation-software/carmaker/

3.4.5.1 A study on the factors affecting fuel consumption and constructing a fuel consumption forecasting model

The correct evaluation of real world fuel consumption is highly related to factors influencing fuel usage, such as the various vehicle specifications, drivers' behaviour, traffic flow rate and road topography. The measurement of fuel consumption is a complex task, not least because direct measurements are not always available, and consequently, it is necessary to predict the fuel consumption of vehicles in specific testing settings indirectly (Skog and Handel, 2014). This study seeks to predict the fuel consumption of the test vehicles by, firstly, identifying the parameters that are significantly affecting it and secondly, constructing a multivariable (multiple) linear regression model to establish the relationships between the influencing factors and fuel usage (Redsell, Lucas and Ashford, 1993; Sa, Chung and Sunwoo, 2003). The first step involves considering all the recorded data available, as illustrated in the table below, to ascertain which parameters (independent variable [IVs]) have the greatest effect on fuel consumption (dependent variable [DV] and thus, should be considered in the subsequent model building.

Table 37. The available parameters selected to test their level of significance in fuel usage

Available parameters from vehicles engine and drivers driving
Consumed fuel
Average engine speed
Average engine coolant temperature (℃)
Average vehicle speed ($^{Km}/_{\mathrm{h}}$)
Average distance (Km)
Average journey time (s)
Cylinder capacity (cc)
Number of sudden accelerations
Number of sudden decelerations
Number of high engine speeds
Number of high temperatures

3.4.5.1.1 Identifying the factors affecting fuel consumption

For this stage, the association between the variables was investigated by checking the correlation between the parameters. First, the parameters with a high level of correlation (at the 1% level) were eliminated from the study and pertained to those that provided engine data. In the second stage, the correlation between the real values of the consumed fuel, the DV, and the IVs, i.e. the parameters shortlisted from the previous stage, are used to construct a linear regression model. As a result, out of 10 recorded parameters, the most affecting factors on fuel consumption are identified.

3.4.5.1.2 Constructing the fuel consumption forecasting model for the test vehicles

Based on the parameters nominated from a previous study, a Multivariable Linear Regression model (MLR) is built. The goal of the MLR model is to estimate the relationship between multiple independent variables and the dependent one. The regression model was developed in two steps. First, the independent variables with insignificant individual coefficients are eliminated. In the final step, based on the parameters with the highest effect on real fuel consumption the regression model is developed (see equation 21).

Equation 21. The multiple linear regression model general equation

$$Y = \beta_0 + \beta_1.\text{independent variable}_1 + \beta_2.\text{independent variable}_2 + \cdots + \beta_k.\text{independent variable}_k + \in$$

Where,

Y= Model real consumption

β_i= The partial regression coefficient for each independent variable

$\in$= The corresponding random error of the model

k= The number of independent variables in the model

To evaluate whether or not the error is normally distributed within it, the outcomes, including a histogram of the parameters' distribution and the P-P plot[84] of the regression model, are presented in the results chapter.

[84] probability–probability plot.

3.4.5.2 Modelling the variations in driving styles by using virtual driving simulator software

As part of this project, IPG's Carmaker software was offered to the researcher to investigate the effectiveness of their solution in the domain of driving behaviour studies. To pursue this, similar conditios to the real-world driving event were modelled in the software environment and the feasibility of using such a model for further studies was investigated.

This subsection is dedicated to the final phase of conducting the research into naturalistic driving behaviour and involves modelling and simulating different driving behaviours based on the gathered real-world driving data. As discussed in the literature chapter (chapter 2), there is a clear distinction between lab developed driver and driving models that do not consider aspects other than driver behaviour and industrial standard driving simulation software, which whilst containing many more parameters, deals with simulation in a rather simplistic way. This calls for investigation into the best solution for simulating naturalistic driving differences among drivers. Moreover, there is a great desire from car manufacturers and research teams to have an accurate representation of different driving styles within simulation packages so as to be able to simulate virtually drivers' differences with accuracy.

Most of the vehicle virtual test packages available to industry have robust features for modelling and simulating many parameters of real driving, such as various settings to present vehicle dynamics parameters, 3D road construction facilities and options to run countless driving scenarios. However, the capacity to simulate a driving event with different driver categories remains underdeveloped. Applying modifications to existing simulation packages appears to be widely accepted and supported by car manufacturers as well as vehicle research and development groups. This is because any new solutions and developments can be integrated into their existing simulation packages. The IPG Automotive driving simulation packages are amongst the most homologated software and have been widely used by the automotive industry leaders, academia, and formula 1 racing teams. Their Carmaker software is a very well equipped driving simulation software package that focuses on modelling 4-wheel passenger cars. The software's open source functionality and powerful computing capabilities makes it an ideal package for modifying and testing different driving simulation scenarios. The software environment provides an opportunity to design, model and simulate the driver, the vehicle dynamics, the road, the traffic, and environmental conditions in detail.

The aim of this subsection is to investigate the feasibility of using the driving simulation software package (IPG Carmaker) so as to confirm its capability for simulating different driving scenarios, where the interest lies in identifying different driver behaviours. To achieve this aim, the standard procedure of building a driving model has been followed according to the IPG Carmaker user manual. Two driving simulation scenarios have been developed as follows.

1. Scenario I: A study modelling and comparing the Vauxhall Corsa drivers' eco-driving performances (3.4.6.2.3).
2. Scenario II: A study modelling and comparing the effect of road slope on drivers' fuel usage (3.4.6.2.4).

3.4.5.2.1 Virtual driving modelling procedure

The standard procedure for building a virtual driving model in the Carmaker package is based on how the software compiles information to run a virtual test. That is, to provide the setting information of the desired test scenario and then, letting the software calculate and simulate a test run. In order to accomplish a comprehensive driving simulation, the software requires information about the tested road, tested vehicle, driving general manoeuvres, road settings and the route. Having been provided with this information, the package then compiles a driving model and runs a driving simulation. The package can cater for a range of parameters in relation to the road settings and topography as well as the vehicle. All of the data related to the actual driving event[85] were load into the package so as to simulate driver behaviour as realistically as possible. For each scenario, there were a number of test runs, the results of which were saved for further post processing analysis.

The virtual vehicle

From the driving event, the Vauxhall Corsa VX was selected to represent the virtual vehicle in both of the simulation scenarios. The generic Vauxhall Corsa model was built to present the same vehicle characteristics in the software as it had during the driving event. The real Vauxhall Corsa VX, 1.4 L, 5 doors, and hatchback model specifications (see Figure 43) were used in constructing the virtual car for the simulation.

[85] Road elevation, traffic settings and vehicle characteristics.

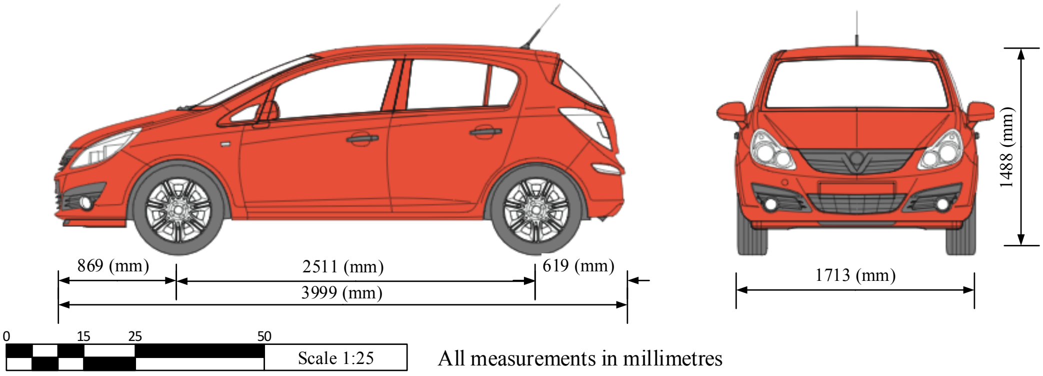


Figure 43. The generic Vauxhall Corsa VX, 1.4 L, 5 doors with dimensions

The virtual route and road settings

In order to simulate the driving tests in realistic conditions, the road settings and the route had to be very similar to the actual driving test. For the road configuration, firstly, the event-driving route was digitalised by converting the GPS longitude and latitude points into a 2D Cartesian format, as shown in Figure 44. Secondly, two important road attributes, track width and road curvature, had to be added to the digitalised route information. Other road attributes, such as surface friction and roadsides profiles, were kept as the Carmaker default settings. Finally, two identical virtual routes were created, one with the real route elevation profile and one without. All the road configurations, such as traffic lights, speed limits and pedestrian crossings, were added to the model as well as environmental aspects that would not have any impact on the outcomes, for aesthetic purposes. The figures on the next page illustrate the steps taken to build the driving route in relation to GPS points, road settings and road topography using the Carmaker simulation software.

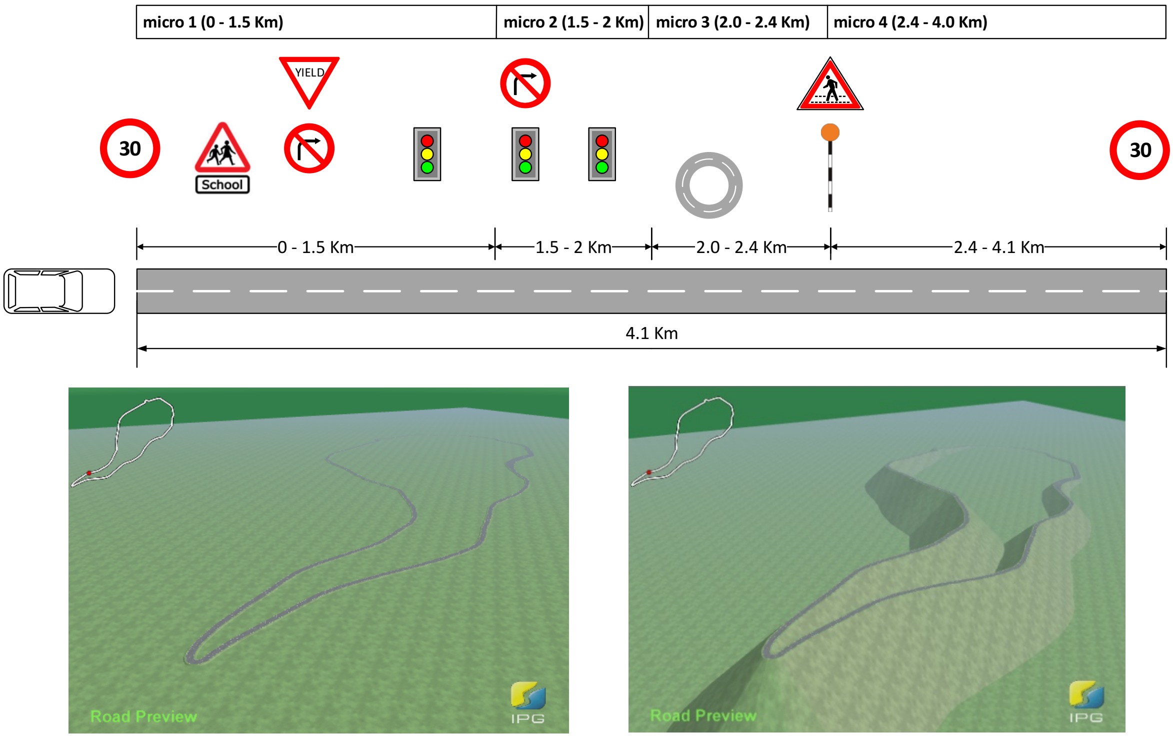


Figure 44. The route details (top), the constructed virtual routes in Carmaker without elevation (bottom left), with elevation (bottom right)

3.4.5.2.2 The simulation approach

The Carmaker model, by default, is integrated with Matlab Simulink, while defining the virtual vehicle and the route. The software itself updates Simulink block sets and modifies the values of every parameter in the model automatically. The core building block of the Carmaker model is summarised in the figure below. As shown in the diagram, by defining vehicle specification and road information, the only parameters the software needs to complete a simulation are the manoeuvre and driver information. Carmaker allows its users to define and test specific manoeuvres, for instance, it is possible to get the program to apply brakes at certain locations or at certain times, while running a simulation.

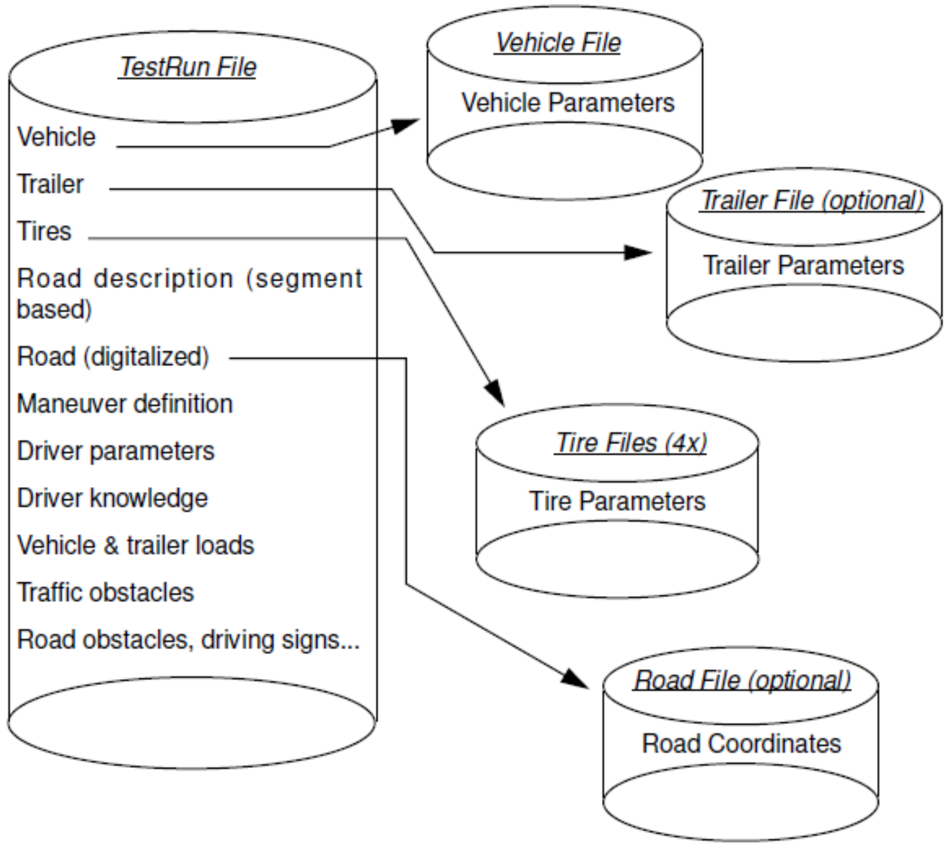


Figure 45. The IPG Carmaker test run parameters

For driver information, IPG developed a driver module called the IPGDriver. The module allows users to add the control actions of a human driver to the model. These actions include braking, steering, accelerator pedal position, gear shifting and clutch operation. In Simulink the IPGDriver module block sets are organised as follows. For each test run, Carmaker's IPGDriver Module first collects user inputs about the test vehicle and the virtual road. Then, it calculates the desired course (driving path) and desired speed (based on the route characteristics and defined manoeuvres). Prior to simulation, Carmaker provides options for users to add the vehicle state and steering wheel torques.

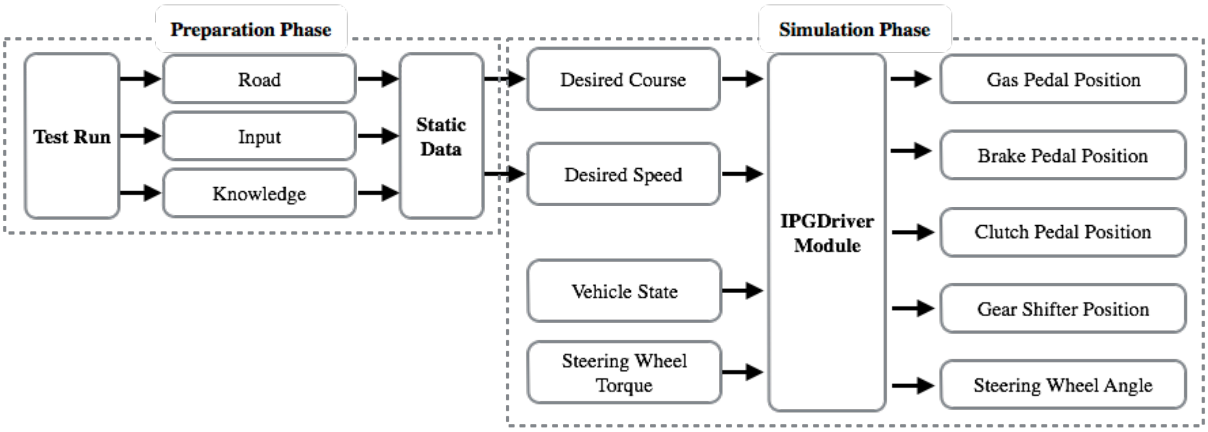


Figure 46. IPGDriver inputs and outputs for every simulation

In following two scenarios (3.4.5.2.2.3 and 3.4.5.2.4), for every simulation (test run) the value of every block is kept the same. To simulate driving performance of the driving event Corsa drivers, their speed profiles (speed vs. time) are introduced into the IPGDriver module as external vehicle speed data. For simulations that resemble the ideal driving scenario, called Ideal Driver, none of the inputs for the IPGDriver model are changed and hence, it is a simulation that is based on the Carmaker calculation of the best driving speed, driving path and steering wheel torque.

3.4.5.2.3 Scenario I: Study modelling and comparing Vauxhall Corsa drivers' eco-driving performances

The study approach

In this study, the simulator was assigned to use the virtual Vauxhall Corsa as the test vehicle, and the constructed 3D road as the test road profile. Once these fixed aspects had been added to the software, the variable for each test was the driving speed (speed vs. time data points) of each driver throughout the route. That is, by keeping all the settings constant throughout the study, it was possible to investigate variations in driving performance among the Vauxhall Corsa drivers, in this case, in relation to their fuel consumption. It should be noted that once the speed had been fed into the program, it was then able to compute the other variables, such as engine speed, emissions level, acceleration and deceleration. The road speeds of the first, third and final rounds of driving of the Vauxhall Corsa drivers were selected. Using the drivers' recorded driving data, nine individual simulations were conducted by Carmaker. In order to check how well the drivers performed during the event, a final simulation was run to evaluate the ideal driving performance based on the road and simulation test settings. The software compiles the best driving performance by calculating the required speed instead of using the real drivers speed as an input. For the final simulation, the road and traffic conditions were kept the same; the only manoeuvre information provided to the model was the road speed limits.

Table 38. List of simulation specifications for Scenario I

ID	Route condition	Vehicle Model	Manoeuvre input to the model
1 to 3	With elevation	Vauxhall Corsa	Corsa drivers' fifth trial road speeds
4	With elevation	Vauxhall Corsa	Ideal driver, 30 mph speed limit max

The study objectives

Carmaker's ability to use real-world measurements as a time series input file in the simulation makes it possible to overwrite the default manoeuvres setting in the model. This provides the opportunity to investigate three important objectives. The first was to study the capability of the Carmaker software package delivering results that identify distinct driver behaviour by only using variations in driving speeds. The second goal was to examine whether the introduced scoring method in the previous sections actually reflected the eco performance of the drivers or not. The final purpose of this study was to explore the feasibility of making the claim that the drivers speed profile is an excellent tool to classify drivers. If these objects were met, then this would show that it is possible to simulate different classified driving behaviour just using drivers' speed profiles. This would prove to be a very useful benchmark for modelling a diverse range of naturalistic driving behaviours.

The post-processing analysis and reporting

It was concluded that to illustrate the outcomes regarding the three objectives of this study, the most suitable method was to compare the drivers' performance by employing Brake Specific Fuel Consumption (BSFC) contour maps. BSFC helps to determine how efficient an engine is being operated at and hence, provides an insight into drivers' driving performance. Figure 47 is the example of an ideal engine BSFC contour map as suggested by (Goering and Cho, 1988), for despite being an old reference the principle still applies. The 'sweet' spot, depicted by (■ in blue) in the figure, is the most efficient spectrum of engine performance. A similar approach by Ma et al. 2014 involved using the same principle to measure the effect of city bus drivers' acceleration behaviour on fuel consumption (Ma *et al.*, 2014).

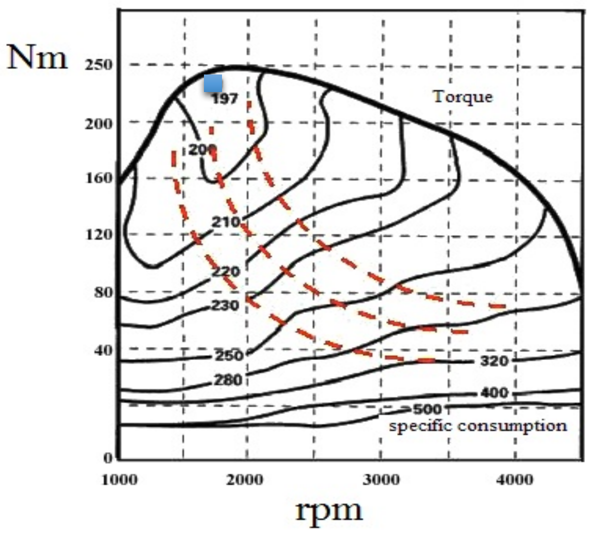


Figure 47. The generic brake specific fuel consumption contour map example (Goering and Cho, 1988)

The BSFC contour maps (fuel island plot) were plotted for all three drivers and from the simulator outputs by using the AVL Concerto post processing package. The contour maps are effective indicators of the amount of time drivers spent in each region of the map and the more they spend around the sweet spot, then the more efficient their driving. Finally, from the BSFC values it is possible to compare the most efficient simulation results. The engine's efficiency is found by calculating the efficiency percentage using the average BSFC value and the Lower Heating Value (LHV) (the LHV for the standard gasoline is 18,679 Btu/lb or 0.0120687 (kWh/g).

Equation 22. The engine's efficiency

$$\eta = \frac{1}{(BSFC \times LHV)}$$

Where,

η = The engine's efficiency

BSFC= The average Brake Specific Fuel Consumption value (g/kWh)

LHV= The Lower Heating Value (kWh/g)

3.4.5.2.4 Scenario II: Study modelling and comparing the effect of road slope on drivers' fuel usage

The study approach

The second scenario was designed as a case study in which the effect of the road elevation profile on drivers' fuel consumption was investigated. For this study, the simulator was assigned to use the virtual Vauxhall Corsa vehicle model as the test vehicle. The simulations were conducted on two modelled roads. The first road was the flat digitalised one without any elevation and the second was the digitalised road with the same elevation profile as the actual route. As with the first scenario, all the settings in the Carmaker software were kept identical, except for the driving manoeuvres, regarding which the fifth lap driving speed profile of the three Vauxhall Corsa drivers was selected. In total, eight driving simulations were run for this study, that is, three with the drivers' fifth lap driving speed data on the flat road and three for the same lap with different road elevation. The last two simulations were the ideal simulation driving behaviours for both road types, with the only variable assigned being the road speed limit.

Table 39. List of the simulation specifications for Scenario II

ID	Route condition	Vehicle type	Manoeuvre input to the model
1	Without elevation	Vauxhall Corsa VX	Driver 7 final lap driving speed
2	Without elevation	Vauxhall Corsa VX	Driver 11 final lap driving speed
3	Without elevation	Vauxhall Corsa VX	Driver 14 final lap driving speed
4	Without elevation	Vauxhall Corsa VX	Simulator, with 30 mph speed limit
5	With elevation	Vauxhall Corsa VX	Driver 7 final lap driving speed
6	With elevation	Vauxhall Corsa VX	Driver 11 final lap driving speed
7	With elevation	Vauxhall Corsa VX	Driver 14 final lap driving speed
8	With elevation	Vauxhall Corsa VX	Simulator, with 30 mph speed limit

The study objectives

This case study is aimed at examining the possibility of using virtual driving simulation packages, such as IPG Carmaker, to investigate the effect of uphill and downhill roads on drivers' fuel usage. The second objective of the study is to investigate the effectiveness of using this package with the actual driving road speeds to compare drivers' fuel consumption. Finally, the study involves examining whether drivers' eco-driving scores are an accurate representation of driver performance.

The post-processing analysis and reporting

A comparison was made between drivers' instantaneous and absolute fuel consumption with those computed as a result of the simulations. Since two road types are designed in Carmaker (without elevation and with elevation) consumption was calculated and compared between these. The ideal simulated results were also included so as to allow for comparisons to be made with actual driver performance. To compare the effect of road slope on fuel consumption, the number of litres of fuel used for the downhill section of the route (distance from the start to 1800 metres) and the uphill section (distance from 2200 metres to the end) was compared with the same sections of the road without elevation (flat road).

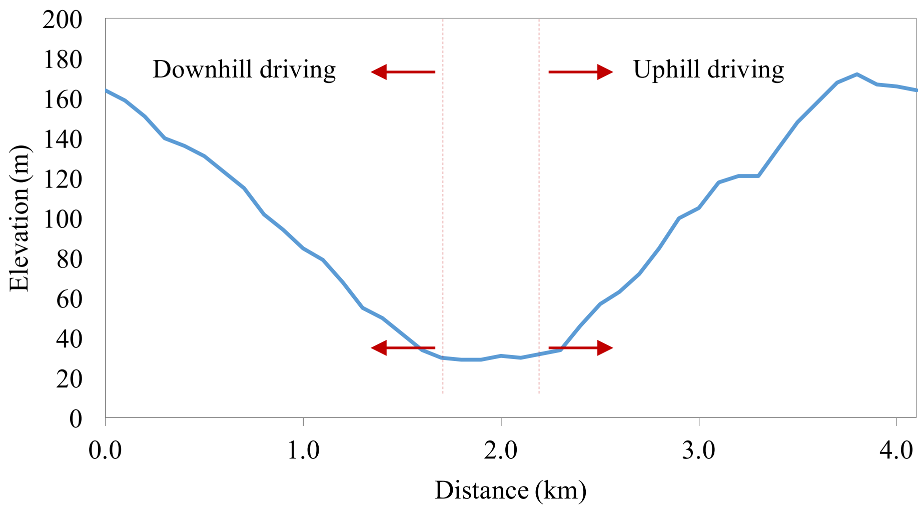


Figure 48. The driving route downhill and uphill driving sections

3.4.6 Identifying drivers' safety perception by using vehicle engine data

The second field of the research for this project is to investigate the naturalistic driving behaviour of drivers in terms their awareness of driving safely. The importance of keeping individual and fleet drivers safe and incident free led to the next aspect of this multidisciplinary research, with the focus shifting away from the drivers' efficiency to investigating their capacity to drive safely. This subsection investigates the possible methods that could help to distinguish driving attitudes that lead to risk-taking behaviours and unsafe driving. To achieve this goal, a number of distinctive analysis techniques are developed, and they are described below.

Four studies are presented here. First, there is the examination of drivers' speed when approaching road settings using vehicle speed data. Second, drivers' total number and locations of hard acceleration and deceleration behaviour are examined so as to establish whether these two metrics are able to identify drivers who were driving in an unsafe manner. The third study investigates whether, on occasion, any driver coasted[86] downhill in neutral gear in an unsafe manner, for which a numerical method is introduced to identify those exhibiting such behaviour. The final study (3.4.6.4) involves adopting a method that is widely used by the road and highway safety engineers and practitioners to identify collision-prone zones and road segments. For this part, police reports about road crashes involving human injury and death in the last 5 years has been used to identify black-spots on the road. The results are used later in subsection 3.4.7.2.2 to modify the drivers' safe driving scores so as to make them more accurate

[86] According to the Highway Code from the Department for Transport: "Rule 122 – Coasting: this term describes a vehicle travelling in neutral or with the clutch pressed down". (Driver and Vehicle Standards Agency, 2015).

in terms of capturing real driving behaviour. The following is the list of the studies that are presented in this section:

1. Identifying drivers with a safe speed perception prior to approaching road settings (3.4.6.1);
2. Total number, repetition, and locations of hard acceleration and hard deceleration made by the drivers (3.4.6.2);
3. Identifying drivers with a coasting downhill habit (3.4.6.3);
4. Identifying historical collision-prone road segments (see 2.9 for more details) on the testing route based on historical data (3.4.6.4).

3.4.6.1 Identifying drivers' safe speed perception prior to approaching road settings

The aim of this study is to ascertain how often the drivers drove too fast when negotiating different road settings, these being: entering a main road, traversing a roundabout and going over a pedestrian crossing. Firstly, as it is mentioned previously, the driving route points of interests were identified and selected. These are locations on the road where the drivers were expected to reduce their speed so as to maintain their vehicle control and hence, drive safely. For each traffic setting, an acceptable speed threshold has been developed, which refers to the speed drivers should have before approaching these road settings in order to have a high level of safety for themselves and other road users. These criteria have been developed according to the UK Highway Code. According to the UK Highway Code rules, drivers entering from a minor road to a main one should assume the stop (speed = 0) position. Therefore, the criteria assigned for their speed was less than 5 mph, for as explained earlier, the GPS data was not always taken at the actual road setting location. When they were approaching the roundabout or the pedestrian crossing, the vehicle speed limit was set to less than 20 mph in order to prevent an unsafe approach. Subsequently, the probability of each driver breaching the set speed limits when negotiating the three road settings was calculated. The table below presents details of the aforementioned points on the route.

Table 40. Locations where drivers' safe driving is compared and each site's testing criteria

Distance	Traffic setting	Algorithm criteria	Location
1.4 km	Entering main road	$0 \leq$ Speed $\leq$ 5mph	The bottom of North Road
2.1 km	The roundabout	Speed $\leq$ 20 mph Speed $>$ 30 mph	Bathwick Hill
2.4 km	The pedestrian crossing	Speed $\leq$ 20 mph Speed $>$ 30 mph	39 Bathwick Hill

3.4.6.2 Study of drivers' hard acceleration and deceleration behaviours

The second method for studying the differences between drivers in terms of safe driving is based on comparing them according to the number of times they harshly accelerated or decelerated. As it has been explained in subsection 3.4.4.3, the occurrences of dangerous behaviours, in this case, harsh acceleration and sudden deceleration, were recorded for every trip made by the participating drivers. This was achieved in two ways. Firstly, it was based on the number of times and on which driving trips, they used the acceleration or brake pedal sharply. Secondly, a comparison was made between the locations of these actions and those of historical vehicle collisions. The following subsections (3.4.6.2.1 and 3.4.6.2.2) explain these two approaches in detail.

3.4.6.2.1 Study of the repetition of harsh acceleration and sudden deceleration behaviour

Similar to previous studies, in this study, the aim is to find the overall trend and each driver's attitude towards road safety. The following criteria, in Table 41 were taken as the thresholds of hard acceleration and deceleration, as drawn from the Castel recommendations. The able below contains the recommended values of the Castel team.

Table 41. Excessive limit value to categorise an acceleration or braking action as dangerous behaviour

Alarming behaviour	Excessive limit value
Harsh acceleration events	0.4 g
Harsh deceleration events	0.6 g

The total number of times the drivers exhibited such excessive behaviour per lap was then presented in logarithmic graphic form to elicit which drivers were driving unsafely and whether their performance improved or deteriorated over time.

3.4.6.2.2 The spatial interpolation of locations with high numbers of alarming behaviours and historical blackspots

In chapter 2 of this thesis, methods for identifying hazardous locations in road networks were discussed. Studies aiming to find the collision-prone spots have widely used historical collision data as their basis for conducting various mathematical and statistical analyses (Lord, Mannering and Pankow, 2010). As it has been clarified in the last chapter, there have been a few published papers pertaining to pre-crash driving behaviour analysis based on data collected from real-world driving (Norris, Matthews and Riad, 2000; Dingus *et al.*, 2006; Klauer *et al.*, 2006; Toledo and Lotan, 2006). While speeding events can be linked to drivers risk taking behaviour and perception of risk (Darby, Murray and Raeside, 2009; Lee *et al.*, 2013), those that are exhibiting more number of harsh acceleration and braking behaviour can be linked to poor driver attitude (Iversen, 2004; Luke and Heyns, 2014), failure to obey traffic signals and driver distraction (Hanowski, Perez and Dingus, 2005; Luke and Heyns, 2014). To examine if there is a meaningful relationship between locations with a previous crash history and drivers harsh braking and acceleration the following study was undertaken.

This study's main purpose was to establish the level of association between locations where drivers performed harsh acceleration and deceleration actions and places with a previous collision history. To this end, the historical collision data between 2005 and 2013 for the driving route were extracted from the Bath and North Somerset District collision database. The information was accessed under open government licence (OGL)[87] from the published database put out by the Department for Transport. The records of all the collisions, including the year of the crash, the number of casualties involved and the severity of the incident were extracted. After extracting these data, two maps from Google map were used to visualise and compare the two types of collected information spatially.

The route was divided into ten segments of approximately 400 metres, and each was assigned the collision data as well as the harsh acceleration/deceleration information. For each segment, the level of correlation (correlation coefficient) for the two types of data was calculated. Three methods were used to examine the correlation between drivers exhibiting sharp acceleration/deceleration and previous collision history along these 10 stretches of road, namely, the: Pearson, Spearman method and Kendall methods. The exploratory data analysis (EDA) approach was chosen to investigate the aforementioned relationship and hence, no pre-assumption was made regarding which method might explain the relationship significantly. Moreover, the correlation was tested through three methods to ensure that that the level of association was calculated consistently across all of them (Hamilton, 1994; Tsay, 2010). The following are the equations used.

Equation 23. The Pearson (1), Spearman rank (2), and the Kendall rank (3) correlation analysis

1. The Pearson correlation, r_{xy} , is defined as follows:

$$r_{xy} = \frac{\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{N}(x_i - \bar{x})^2 \times \sum_{i=1}^{N}(y_i - \bar{y})^2}}$$

2. The Spearman rank correlation r is calculated as follow:

[87] The final visit was made in June 2015: www.data.gov.uk/dataset/road-accidents-safety-data/. The same information is accessible at www.crashmap.com [Accessed: June-2015].

$$r = 1 - \frac{\sum(x_i - y_i)^2}{N\times(N^2 - 1)}$$

3. The Kendall rank correlation τ is calculated as follows:

$$\tau = \frac{N_C - N_D}{\frac{1}{2}N(N - 1)}$$

Where,

$\bar{x}$ = the sample average of the first time series X

$\bar{y}$ = the sample average of the second time series Y

x_i= a value from the first time series data X

y_i = a value from the second time series data Y

N = total number of pairs that don't have missing observation (x_i, y_i)

N_C = total number of concordant pairs of observations[88]

N_D = total number discordant pairs of observations

[88] For every pair of (x_i, y_i) and (x_j, y_j), these are concordant if $x_i > x_j$ and $y_i > y_j$ or $x_i < x_j$ and $y_i < y_j$. The pairs are discordant if $x_i > x_j$ and $y_i < y_j$ or $x_i < x_j$ and $y_i > y_j$.

3.4.6.3 Identifying drivers with a coasting downhill habit

Drivers putting their car in a coasting position by either travelling in neutral gear or with the clutch pressed down, are demonstrating a dangerous driving behaviour, for this causes control of the car to become difficult as it inhibits engine braking[89]. A dramatic increase in vehicle speed with little or no change in engine speed on descending roads indicates coasting, because as explained in subsection 3.4.3.2, a linear relationship should be observed during normal driving. To identify this behaviour, a threshold rule was developed. This involved applying the speed limit to Equation 9, which produced an engine speed of 1900 rpm. As this was for an ideal situation, the threshold was set slightly lower at 1800. This behaviour could only be identified over a distance of 1.8 km immediately after the commencement of a lap, as this was the only downhill section. Consequently, the following logical statement was applied.

Equation 24. Downhill driving algorithm

If route ID is **true**, from distance point 0 km to 1.8 km

If $0 \leq distance \leq 1.8$ **and** $speed \geq speed\ limit$ **while** $RPM < 1800$ **then count the point**.

The total number of hypothetical coasting downhill incidents of all the drivers were tabulated. Regarding the driver with the highest hypothetical coasting incidents, as identified through applying the algorithm, these suspected incidents were assessed by looking at the speed – distance and engine speed – distance graphs from the study in 3.4.3.1, to see if such behaviour most likely occurred.

[89] Engine braking refers to when the driver takes advantage of the vehicle compression and friction from the moving parts of the engine by lowering the gear, which slows down the car.

3.4.6.4 Identifying collision-prone road segments on the testing route

As discussed in chapter 2, choosing the right method to identify collision-prone zones is a data-driven process (Cheng and Washington, 2005; Lord, Mannering and Pankow, 2010). The method for hotspot identification (HSID) is usually selected based on three criteria: firstly, the sample sizes of historical data available and secondly, the presence of road segments with zero collisions (zero observations). Finally, it is based on whether the sample is over-dispersed (sample-mean > sample-variance) or under-dispersed (sample-mean < sample-variance) (Lord, Mannering and Pankow, 2010). For this study, the sample size was comparatively small owing to the short length of road that was driven along by the participants and therefore, the following analysis suffers from low sample size bias. However, whilst the outcomes might not be accurate, they are used in subsection 3.4.7.2.1 along with the results of the pre-collision study in subsection 3.4.7.2.2 to modify the scoring method that is commonly used to score drivers in terms of their safe driving.

Similar to in subsection 3.4.7.2.2, previous collision information on 10 identified segments was gathered and later tabulated, being taken from the historical collision database published by the Department for Transport. Figure 49 shows this arrangement on a map. For the first part of this study, descriptive analysis was conducted on the dataset, which included calculating the sample average, variance, and the standard deviation (σ) (Equation 11). Secondly, prior to the use of the statistical confidence intervals method, the coefficient of variation (C_V) was calculated. This was to determine whether the standard deviation was substantial or not. The statistical confidence intervals method is applicable when the value of C_V is less than 1.0.

Equation 25. Equation to calculate the coefficient of deviation

$$C_V = \frac{\sigma}{\text{Average}}$$

The statistical confidence intervals method is an approach that identifies collision-prone locations based on the following terms: if at the segment (a) the number of observed crashes k_a is equal or greater than the sum of mean (μ) and the standard deviation (σ), then, the segment is unsafe. The algorithm for this is written as follows:

$location\ k_a\ is\ unsafe\ if\, C_V < 1$ and $k_a \geq \mu + \sigma$ (Cheng and Washington, 2005).

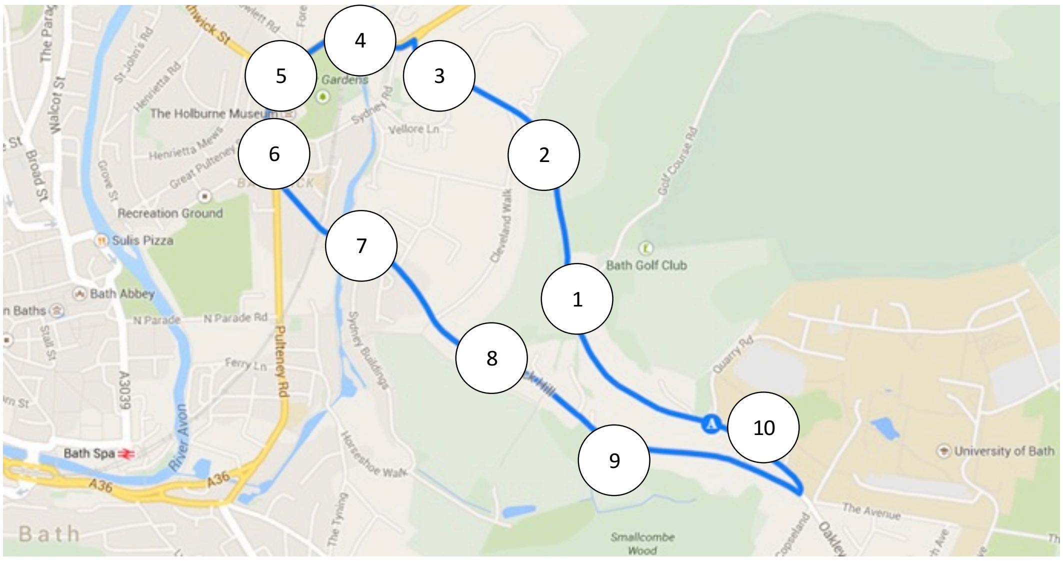


Figure 49. Road segment arrangement for collision-prone analysis

3.4.7 Classifying and scoring drivers' safety differences from vehicle engine primary data

The safety of drivers and other road users is a significant concern for local authorities, motor insurers, and fleet managers. As discussed in subsection 3.4.5, scoring drivers according to different behaviours allows for their classification in a ranked order. In this case, unsafe driving habits are the focus, rather than eco-driving, as in the aforementioned subsection. Any decisions made based on the existing driver scoring methods have their weakness. For example, in the case of drivers who are willing to pay their vehicle insurance premium based on how they drive (Pay How You Drive insurance policy), the use of the insurance provider scoring model has been heavily criticised for being too basic. To this end, the major distinction that has been made between the two proposed scores, namely, the eco-driving score and safe-driving score is the coefficient ratio assigned to each parameter in the scoring function.

In this subsection, the fleet management scoring method that is widely used by industry to classify safety aspects of drivers' driving is employed. Regarding which, the method is explained and then applied to the drivers' data from the event. Subsequently, based on the findings from subsection 3.4.7's studies, two modifications are suggested for improving the scoring model. The structure of the following subsections is as follows.

1. Safe driving classification according to the fleet management scoring system (3.4.7.1).
2. Two modifications to the current fleet management scoring system and subsequent, further classification (3.4.7.2.1 and 3.4.7.2.2).

3.4.7.1 Classifying drivers' safety according to the fleet management scoring system

The standard scoring practice among fleet management service providers and the Pay How You Drive vehicle insurers is to rank drivers' safety score similar to the one used for eco-performance, which has been explained extensively in subsection 3.4.4.3. To recall, the model works based on threshold rules for certain alarming behaviours, i.e. the method ranks drivers according to number of times drivers exceed the engine speed threshold or exhibit sudden acceleration/deceleration. Similar to the eco-performance scoring method, each dangerous behaviour comes with a coefficient ratio set by service providers and insurers. Castel uses the same alarms as the rest of the industry, but with different coefficients and their occurrence is captured by using the GPS and OBD monitoring data. The threshold rules for these alarms and their coefficient ratios are presented in the table below.

Table 42. Excessive limit values for safe driving score and safety driving coefficient ratio

Alarm	Excessive limit value	Safety driving coefficient ratio
High engine speed	4500 rpm	4
Hard acceleration	0.4 g	2
Hard deceleration	0.6 g	2

The same formulation was used to allocate a numerical value to each driver's driving performance in terms of safety, as for the eco-driving scoring equation (see Equation 20). The maximum score allocated for each lap was 50 points and a score near to 50 means that person was a safe driver. The driving score was calculated using following equation.

Equation 26. Calculating the safe driving score

$$\text{Safe driving score} = 50 - [(number\ of\ sudden\ accelerations \times safety\ ratio) + (number\ of\ sudden\ decelerations \times safety\ ratio) + (number\ of\ high\ RPMs \times safety\ ratio)]$$

The Equation 26 can be written in a symbolic format as follows:

Equation 27. Symbolic format of the driving score

$$\text{Safe driving score} = 50 - [(SA \times \mathrm{e}) + (SD \times \mathrm{d}) + (RPM \times \mathrm{f})]$$

Where,

SA = Total number of sudden accelerations for all locations

SD = Total number of sudden deceleration for all locations

RPM = Total number of times exceeding the high engine speed threshold

e = Safety driving coefficient ratio of sudden acceleration

d = Safety driving coefficient ratio of sudden deceleration

f = Safety driving coefficient ratio of high engine speed

The following alphabetic guide was developed to allocate an alphabetic representative for each driving score range. The results of this study are visualised by using alphabetic ranges and graphs.

Table 43. Alphabetic representation of drivers' scoring range

Driver scores range	Alphabetic representative
$45 \ll Driver\ score \ll 50$	A
$35 \ll Driver\ score \ll 44$	B
$20 \ll Driver\ score \ll 34$	C
$Driver\ score \ll 19$	D

3.4.7.2 Two modifications of the fleet management scoring system

3.4.7.2.1 Refining drivers' safety score according to their coasting downhill habits

For the second part of this study, a change in the fleet management safe driving scoring method is introduced, with the aim being to extend the existing threshold rules to create a location, sensitive model. The model, as it stands, is unable to identify where an individual driver is descending a steep road while coasting in neutral gear. The importance of preventing drivers doing so has been discussed in subsection 3.4.7.3., where a method was introduced to identify drivers who were coasting downhill during the driving event. This can be incorporated into the model by deducting the number of times they coasted downhill. Equation 28 shows a modified version of Equation 26 with the deduction of the total number of times drivers were coasting downhill.

Equation 28. Modified version of the safe driving score equation with the deduction of coasting downhill behaviour

$$\text{Modified safe driving score I} = 50 - [(number\ of\ sudden\ accelerations \times safety\ ratio) + (number\ of\ sudden\ decelerations \times safety\ ratio) + (number\ of\ high\ RPMs \times safety\ ratio) + Total\ number\ of\ times\ coasting\ downhill]$$

3.4.7.2.2 Improving drivers' safety score based historical collision data

The study in subsection 3.4.7.2.2, elicited that there was strong positive correlation between locations with a high number of alarming behaviours and collision-prone history (pre-collision study). Thus, it would appear reasonable to penalise those drivers who exhibited such behaviour at incident hotspots. Therefore, for the identified segments of road with a high collision history, the number of times a driver exhibited harsh acceleration/deceleration was doubly penalised as compared to when he/she did so in other segments, as shown in Equation 30. Note that the coasting behaviour is not included in equation so that the incremental changes to the drivers' scores brought about by modifying the scoring model could be assessed independently.

Equation 29. Modified version of the driving score equation based on the effect of collision-prone locations

$$\text{Modified safe driving score II} = 50 - [(number\ of\ sudden\ accelerations\ at\ non\ collision\ prone\ locations \times safety\ ratio) + (number\ of\ sudden\ decelerations\ at\ non\ collision\ prone\ locations \times safety\ ratio) + (number\ of\ high\ RPMs \times safety\ ratio) + (2\times number\ of\ sudden\ accelerations\ at\ collision\ prone\ locations \times safety\ ratio) + (2\times number\ of\ sudden\ decelerations\ at\ collision\ prone\ locations \times safety\ ratio)]$$

To conclude, the two modification equations above were combined to reach to a comprehensive drivers' safety scoring model. Hence, the model presents a safe driving score that includes the adverse effect of coasting downhill as well as making sudden accelerations and decelerations at collision-prone locations in addition to non-prone ones. The final equation is as follows:

Equation 30. Inclusive average safe driving score

$$\text{Inclusive average safe driving score} = 50 - \left[(\overline{SA}\times \mathrm{e}) + (\overline{SD}\times \mathrm{d}) + (RPM\times \mathrm{f}) + \left(2\times\overline{\overline{SA}}\times e\right) + \left(2\times\overline{\overline{SD}}\times d\right) + TCD\right]$$

Where,

$\overline{SA}$= Total number of sudden accelerations for non-collision-prone locations

$\overline{SD}$= Total number of sudden decelerations for non-collision-prone locations

RPM = Total number of times exceeding the high engine speed threshold

$\overline{\overline{SA}}$= Total number of sudden accelerations for collision-prone locations

$\overline{\overline{SD}}$ = Total number of sudden decelerations for collision-prone locations

TCD = Total number of times exhibiting coasting downhill behaviour

e = Safety driving coefficient ratio of sudden acceleration

d = Safety driving coefficient ratio of sudden deceleration

f = Safety driving coefficient ratio for exceeding the engine speed threshold

3.5 Ethical and legal considerations

According to the driving field study framework, ethical and legal issues need to be considered for all three phases of the study, namely, the preparation stage, data acquisition phase and finally, during the post-processing phase. This subsection covers the considerations made for all stages of this research.

3.5.1 Design experiment ethical and legal issues (preparation phase)

Each step of the field study was designed in accordance with the ethical and legal rules applicable to behavioural research involving human participants as the subjects. Specifically, according to the Belmont Report core principles, it is important to take into account the due respect for persons, beneficence, and justice before embarking on any research endeavour and ensuring that these values are maintained throughout the process. Potential participants were recruited via the website set up for the events and initially, anyone was eligible, provided they had a current UK DVLA licence.

All drivers and volunteers received information about the voluntary nature of the study as well as the aims and objectives of the research. This information was provided through three channels, firstly, through an online platform on the event website and a social media page dedicated to the event. The second channel was the information sessions prior to the events, which took place two days before. Finally, on a driving day, everything was explained verbally and in written form, i.e. their driving day information package. All the drivers and co-drivers received this package on a driving day containing all the terms and conditions, event criteria, maps, safety information and the consent form (all signed consent forms were archived[90]). The consent form included each drivers' agreement to have the provided data used for research purposes and permission to use any video footage or photographs for future display. Moreover, special care was taken when sending correspondence to the volunteers not to make them feel pressured into attending the events in any way. All potential participants were informed about their right to withdraw from the study at any time with no sanction being imposed.

Regarding the participants' safety and wellbeing during the events on the university premises and whilst driving, a complete health and safety assessment was conducted before the event. Potential safety hazard was explained to the participants at the beginning of the driving events and relevant documents (university map, driving route, fire assembly points) were given to them for personal use. The Castel monitoring devices made it possible to monitor all drivers (vehicles) while driving in real time, which were used during all driving events to prevent incidents. The university of Bath security department was informed about these events and were on hand to help during their duration if they were needed for any reason.

3.5.2 Confidentiality of the participants' information and acquired driving information (data acquisition phase)

Each contestant's information including their personal details, having been collected, was stored confidentially in a database as explained in subsection 3.4.1.2. All driving data gathered during the events from the drivers were also kept secret, being only accessed by the lead researcher in accordance with Personal Information Protection and privacy act[91]. In sum, no

[90] All documents and signed forms are available upon request.
[91] The Data Protection Act 1998 (DPA).

data were passed on to any third party. As explained in subsection 3.4.1.2, each driver was assigned an ID number for themselves and their vehicle, which was fully anonymised. All the sensitive kept on the Castel online platform was deleted three months after the main event. The offline version of the recorded data, which longer contained any data related to the participants was saved to the University of Bath online server.

3.5.3 Ethical considerations regarding the analysis outcomes (post-processing phase)

None of those displayed in the presentation of the results can be linked back to a particular person owing to the above described anonymisation process. In other words, any driver who exhibited poor driving skills cannot be identified from the graphs or charts provided. Finally, none of the drivers was publically or privately judged on their driving ability.

3.5.4 Company collaboration confidentiality

Castel, the supplier of the monitoring devices and online monitoring platform, was a great collaborator throughout the project. However, at no stage was the research compromised by their involvement. That is, whilst they provided the researcher with technical knowledge, they acknowledged that the outcomes might be critical of their products and procedures. Regarding collaboration with IPG, an agreement was signed in relation to use of their simulation package, and they expressed an interest in being informed about the research outcomes.

Chapter 4

Results

"There are two possible outcomes: if the result confirms the hypothesis, then you have made a measurement. If the result is contrary to the hypothesis, then you have made a discovery."

– Enrico Fermi

4.1 Introduction

This chapter presents the results from the analyses, the methods for which were explained in Chapter 3 subsection 3.4 and the findings are organised in the same order as the post-processing methods in that subsection. Recall from the previous chapter, the main focus of this project is to study driving behaviour differences in terms of two aspects eco-driving performance and controlling the vehicle in a safe manner. As previously discussed, there are three areas of research that are important to address: identifying driving differences, classifying drivers with similar attributes and finally, using those findings to construct models that incorporate the differences between drivers. Table 44 below, sets out the order in which the results are presented, and this has the same format as the corresponding table explaining the different studies in chapter 3. The drivers were asked to complete five driving rounds, with each set of driving data being checked to make sure all had been correctly uploaded to the server. Table 45 presents the participating vehicles and demographics of the drivers.

Table 44. An overview of the results of the post-processing analysis

<table>
<tr><th></th><th>Identifying driver differences</th><th>Classifying and scoring drivers</th><th>Modelling and simulating different driving behaviour</th></tr>
<tr><td>Eco Driving Studies</td><td>1. Results of the study of vehicle speed – distance and engine speed – distance analysis
2. Results of the study of the relationship between drivers' road speed and engine speed
3. Results of the study for identifying drivers' aggression by comparing their speed vs. acceleration
4. Results of the study of Geo-analysis of vehicle parameters in conjunction with the route elevation profile
5. Results of the descriptive analysis of vehicles' speed profile</td><td>1. Evaluating the similarity of drivers' speed profiles
2. Classifying drives based on their fuel usage and proposing the Vehicle Specific Power – Fuel Consumption metric
3. Results of the study of eco-driving classification according to the fleet management scoring system</td><td rowspan="2">1. Results of the study of the factors affecting fuel consumption and constructing a fuel consumption forecasting model
2. Modelling different driving styles and road condition effects on fuel usage and car emissions using IPG Carmaker</td></tr>
<tr><td>Safe Driving Studies</td><td>1. Results of the study of drivers' safe speed perception when approaching road settings
2. Results of the study of drivers' hard acceleration and deceleration habits
3. Results of the study for identifying drivers with coasting downhill habits
4. Results of the study of identifying collision-prone zones</td><td>1. Results of the study of safe driving classification according to the fleet management scoring system
2. Results of two modifications to the fleet management scoring system</td></tr>
</table>

Table 45. Drivers demographics, vehicle information, and number of completed trials

Mark	Model	Car number	Driver	Number of valid trials
Vauxhall	Corsa VX	7	Male	6
Vauxhall	Corsa VX	11	Male	5
Vauxhall	Corsa VX	14	Male	5
Nissan	Note	9	Male	6
Nissan	Note	10	Female	5
Nissan	Note	12	Male	4
Nissan	Note	13	Female	5
Nissan	Juke	15	Male	5
Fiat	500 L	8	Male	5

The valid trials of the driving were those that the data for which were uploaded correctly (for the full length of the route) to the server. Moreover, only driving trials were validated for which at no point the OBD dongle was unplugged. A set of guidelines was given to the drivers at the beginning of the event, which explained all the criteria of a correct driving trial. As aforementioned, the results are based on the data collected during the final eco-safe driving challenge that took place on 9th February 2014, and there were nine cars involved: three Vauxhall Corsa VX and four Nissan Note, one Nissan Juke and one Fiat 500L.

As is evident from the table above, there were two female and seven male drivers attending the driving event. All, except driver 12, were able to complete at least five driving laps of 4km in length. Unfortunately, the data collected on the last driving lap by driver 12 were incomplete owing to a faulty connection between the OBD dongle and the vehicle OBD port. As explained in the previous chapter, time series data collected during the driving event is the primary source for the analysis in this project. Specifically, the vehicle road speed, engine speed and GPS location were these time series data, which were collected for all drivers every 10 seconds.

Other vehicle parameters, such as intake manifold absolute pressure and intake air temperature, were collected from vehicles with available ECU ports for collecting those data. Table 46 shows the time series parameters collected for the participating vehicles. In the following sections, regarding the results where the interest lies in comparing drivers driving the same make of car, then the Nissan Juke and Fiat 500L are excluded; otherwise, they are part of the analysis. The first section (4.2) covers the results for the three research areas in relation to eco-driving, which include identification, classification and modelling, whilst the final section (4.3) presents the outcomes pertaining to the safe driving aspects of the analysis.

Table 46. Available time series variables from the participating vehicles

Car Number	10 (s) intervals	Engine speed (rpm)	Vehicle Speed (km/h)	Intake Manifold Absolute Pressure (kPa)	Intake Air Temperature (℃)	Engine Coolant Temperature (℃)	Mass Air Flow ($^{g}/_{s}$)	GPS Longitude and Latitude
7	☑	☑	☑	☑	☑	☑	—	☑
8	☑	☑	☑	—	—	☑	☑	☑
9	☑	☑	☑	—	☑	☑	—	☑
10	☑	☑	☑	—	—	☑	☑	☑
11	☑	☑	☑	—	—	☑	☑	☑
12	☑	☑	☑	☑	☑	☑	—	☑
13	☑	☑	☑	—	☑	☑	—	☑
14	☑	☑	☑	☑	☑	☑	—	☑
15	☑	☑	☑	—	☑	☑	—	☑

4.2 Result of studies relating to drivers' driving performances

The first step towards understanding individuals' driving differences was to compare the information that had been collected from their driving, which included vehicle data (vehicle speed, engine speed, fuel usage) and geo-location information (longitude and latitude points). This section presents the results of the conducted analysis in relation to identifying, classifying, and modelling the drivers' behaviours based on their vehicle data. The aim of the subsequent comparisons made is to distinguish them in terms of driving performance and their efficiency in relation to handling the vehicle. To start with, drivers with excessive speed and aggressive acceleration behaviours are identified in subsection 4.2.1, with the purpose being to see whether these parameters allow for classification of distinct driver behaviour. Next, the drivers are classified based on the similarity of their vehicle speed and on the scoring method widely used by the vehicle insurance and telematics industries in subsection 4.2.2. The network analysis in subsection 4.2.2.1.2, visualises the closeness of every two drivers according to their driving speed. The section's final study includes the results of employing three metric methods to report classifying drivers' fuel usage and energy consumption differences. The last subsection (subsection 4.2.3) presents the outcomes of two approaches to modelling different driving behaviour. The first involves a predictive model that uses linear regression to model those engine parameters that have a significant effect on drivers' real fuel consumption. The second approach to modelling pertains to using the IPG Carmaker virtual vehicle dynamics software package. The drivers' actual vehicle speeds were used as inputs to the model, and their driving performances were compared based on the two proposed scenarios described in subsection 3.4.6.2. The results of the simulations using the software package are presented in subsection 4.2.3.2.

4.2.1 Results of identifying drivers from vehicle engine primary data

This subsection contains the results of five studies investigating different parameters of a vehicle, with the goal of eliciting driving behaviour differences among the participants and these are:

1. Study of the speed – distance and engine speed – distance;
2. Study of the relationship between drivers' road speed and engine speed;
3. A study identifying drivers' level of aggression by comparing their speed vs. acceleration;
4. Geo-analysis of the vehicle parameters in conjunction with the route elevation profile:
    4.1. Geo-analysis of the vehicle parameters;
    4.2. Geo-analysis of the acceleration behaviour among drivers;
    4.3. Geo-analysis of the vehicle-specific power among drivers;
5. Descriptive analysis of the vehicle speed profiles:
    5.1. Analysis of drivers' consistency and the chosen range of speed;
    5.2. Descriptive analysis of the drivers' speed distribution.

4.2.1.1 Plots of speed – distance and engine speed – distance results

This analysis provides the foundation for all the subsequent studies. In what follows, whilst all the drivers covered the same distance for each lap, that between each GPS reading obviously differed according the speed of the vehicle. Hence, when referring to distance here, it is this that is being referred to. As can be seen below, the graphs representing speed – distance and engine speed – distance, show driving speed (km/h), engine speed (rpm), distance between GPS points and lap number, are all presented together for each of the two multiple used cars, namely, the Vauxhall Corsa and Nissan Note. This allows for a visual comparison of the different profiles for each driver driving the same car for the whole event in terms of vehicle and engine speed. It also shows which drivers broke the speed limit of 48 km/h as well as giving a general impression of where sudden changes in vehicle speed and where excessive usage of a vehicle power, i.e. vehicle engine speed occurred. After each lap in the graphs below, the colours for vehicle speed and engine speed change alternately so as to make the distinction between each lap clear. In the following charts, the driving speed (km/h) is drawn in blue lines for odd numbered trips and the red lines shows the even numbered ones. Moreover, the corresponding engine speeds for odd trips are in green, while those for even trips are in purple. It is important to mention that the speed limit for the entire route was 30 mph or about 48 km/h.

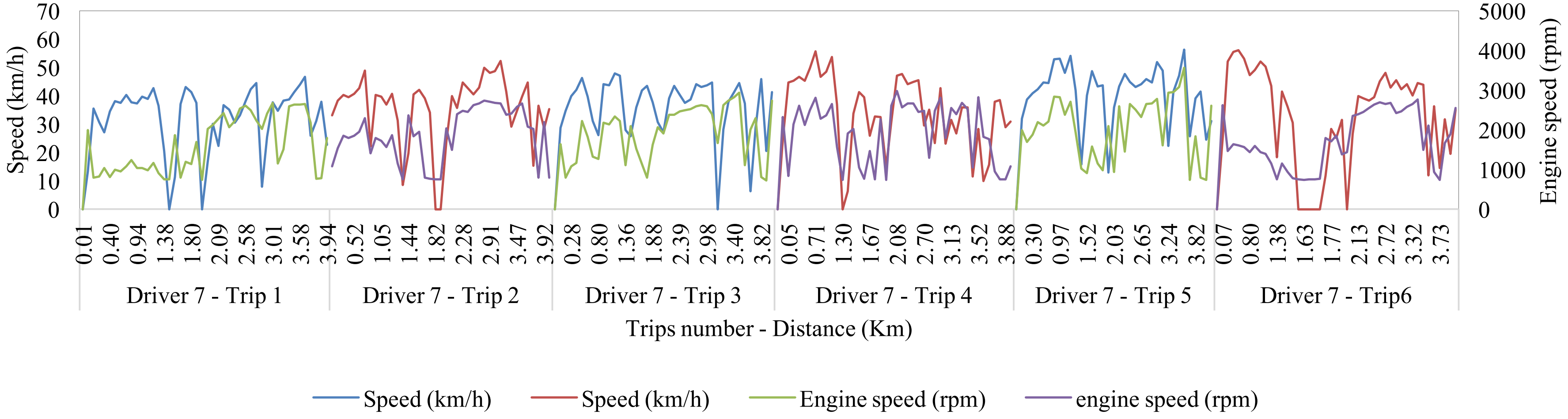


Figure 50. Vauxhall Corsa driver 7: speed - distance and engine speed – distance plots

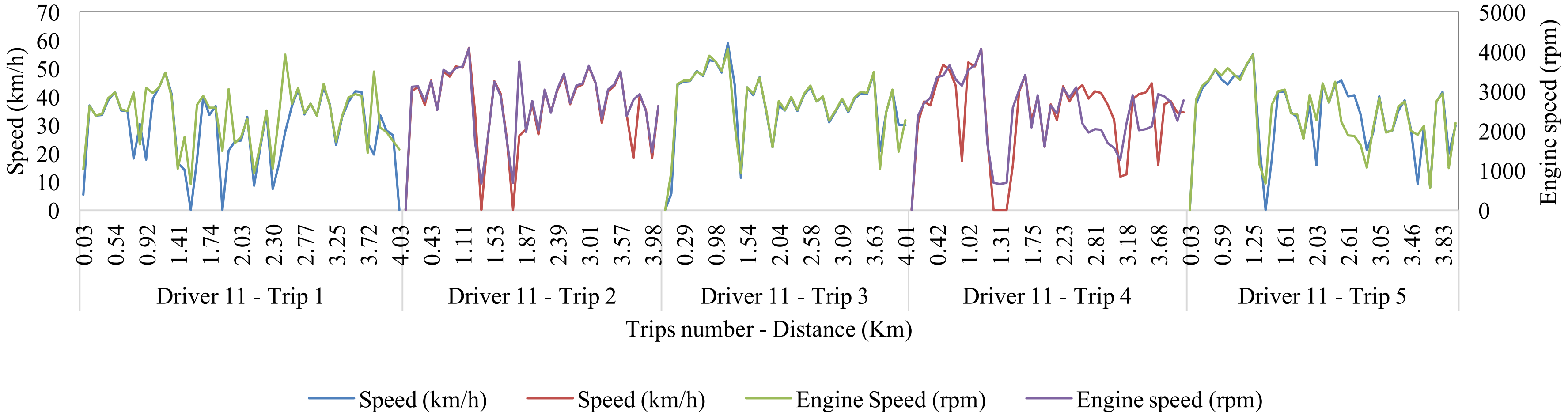


Figure 51. Vauxhall Corsa driver 11: speed - distance and engine speed – distance plots

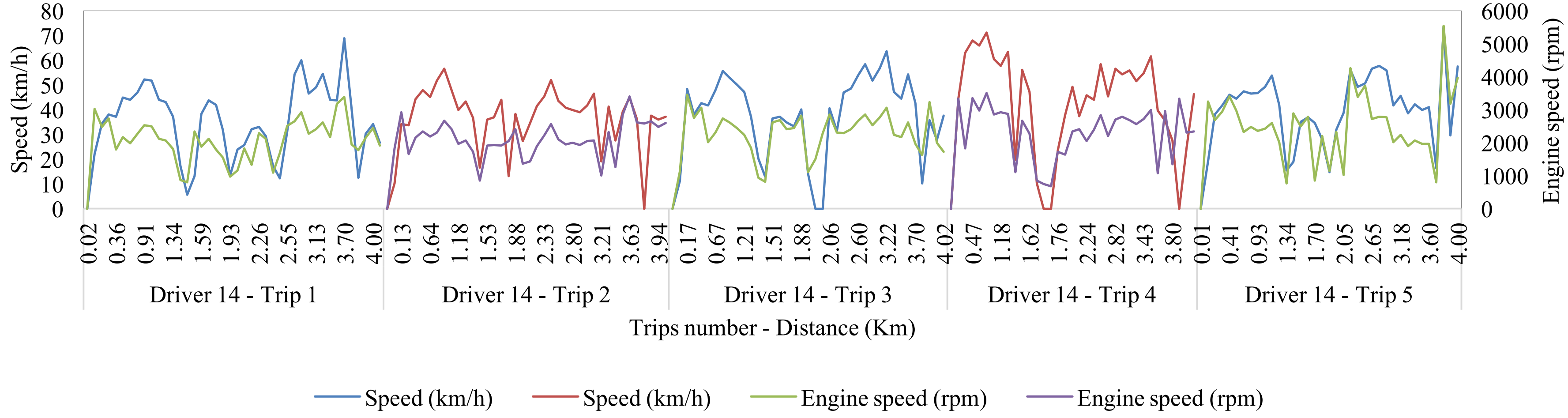


Figure 52. Vauxhall Corsa driver 14: speed - distance and engine speed – distance plot

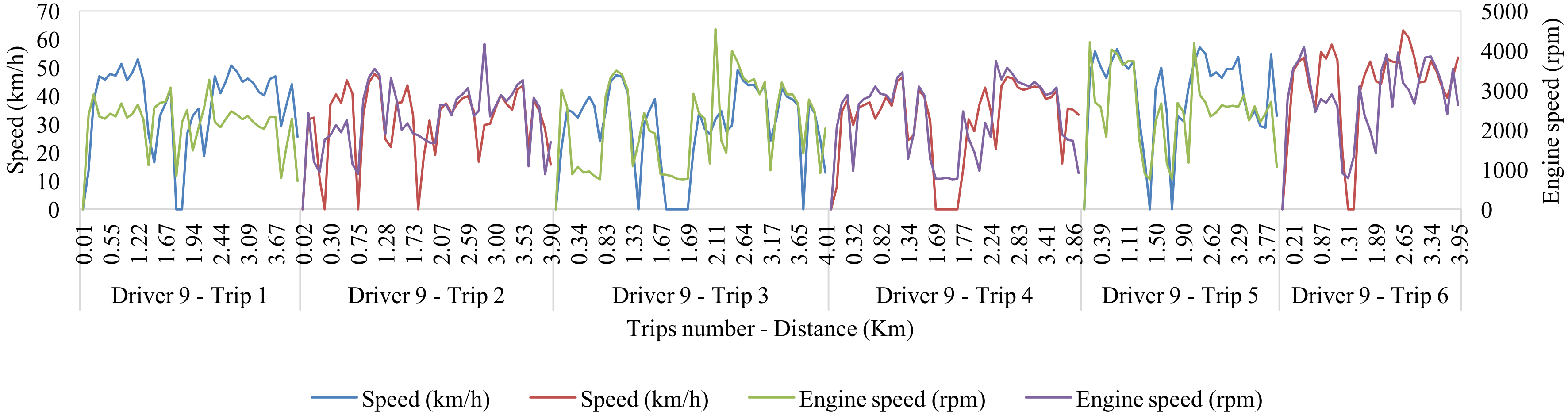


Figure 53. Nissan Note driver 9, speed - distance and engine speed – distance plots

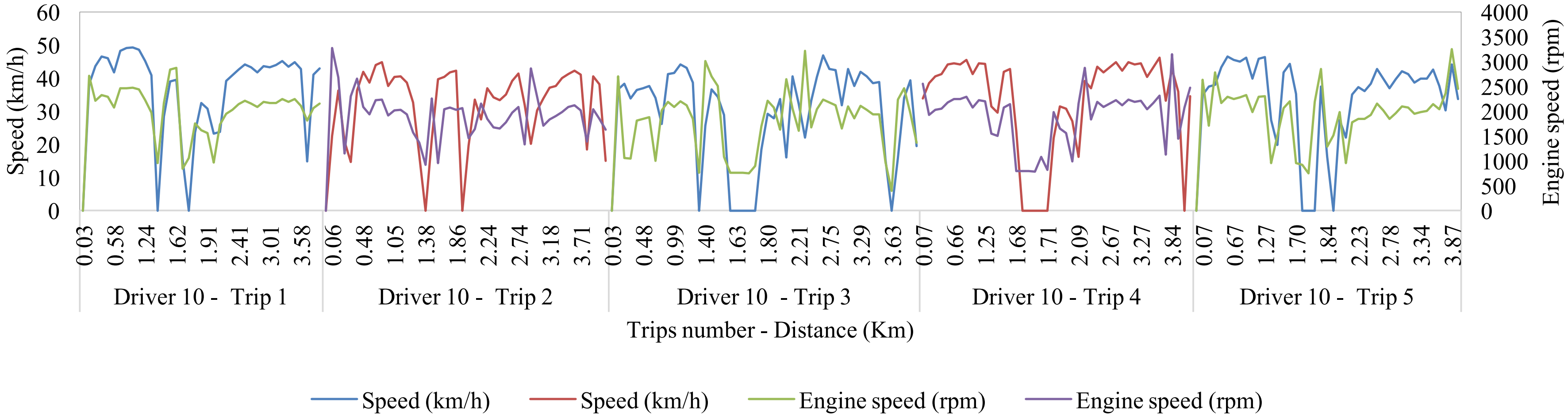


Figure 54. Nissan Note driver 10, speed - distance and engine speed – distance plots

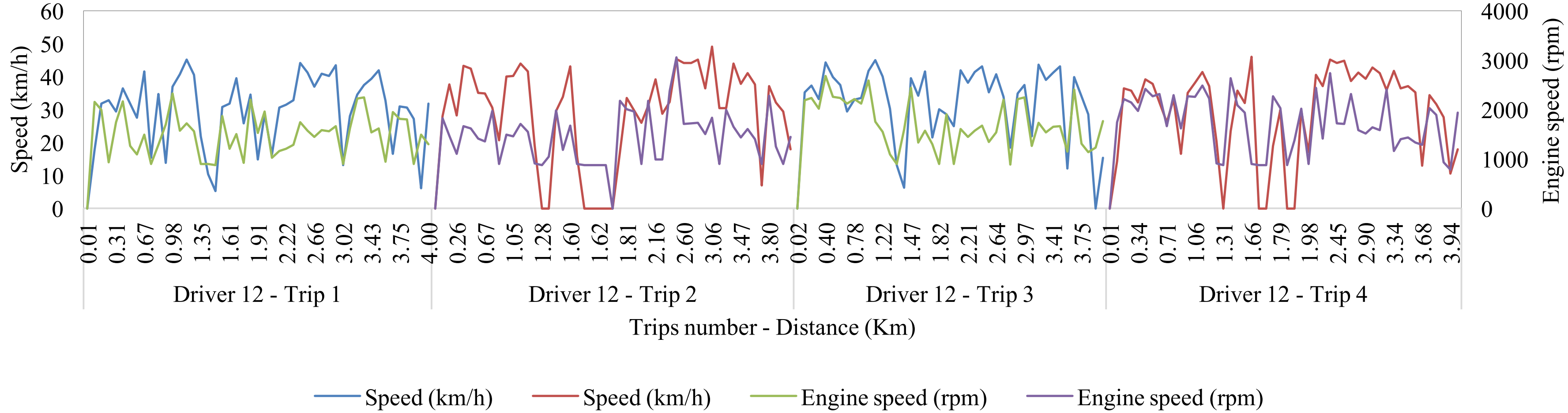


Figure 55. Nissan Note driver 12, speed - distance and engine speed – distance plots

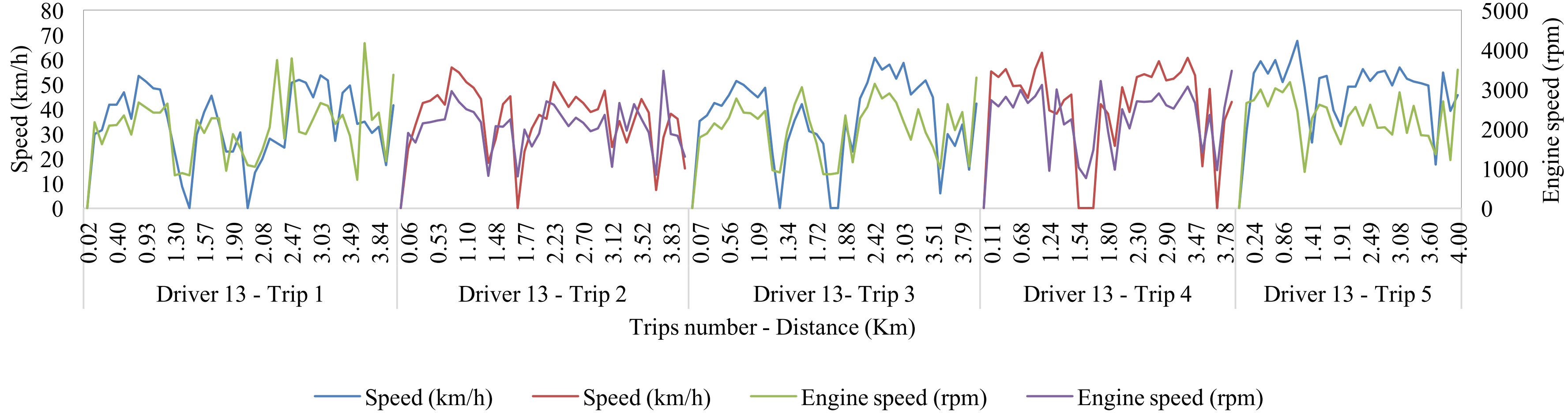


Figure 56. Nissan Note driver 13, speed - distance and engine speed – distance plots

4.2.1.2 Results of study on the relationship between drivers' road speed and engine speed

In order to explore the drivers' gear shifting habits, i.e. whether they did so in an efficient manner or not, the vehicle drivetrain characteristics of both the Vauxhall Corsa and Nissan Note cars were calculated. Using the method in Equation 9 in Chapter 3, the engine speed for known road speed was calculated by using the constant parameters shown in the table below to find the hypothetically most efficient gear shifting strategy pattern. The parameters in Table 47 were used to achieve this (Driver 12 was excluded from the study as the vehicle ran on diesel fuel).

Table 47. Rolling circumference, gear ratios, and the final drive ratio of the tested vehicles

Parameters	Vauxhall Corsa	Nissan Note
The rolling circumference	72.1" (6.00 foot)	72.3" (6.025 foot)
1st gear ratio	3.55	3.727
2nd gear ratio	1.96	2.048
3rd gear ratio	1.30	1.393
4th gear ratio	0.89	1.029
5th gear ratio	0.71	0.821
The final drive ratio	4.18	4.067

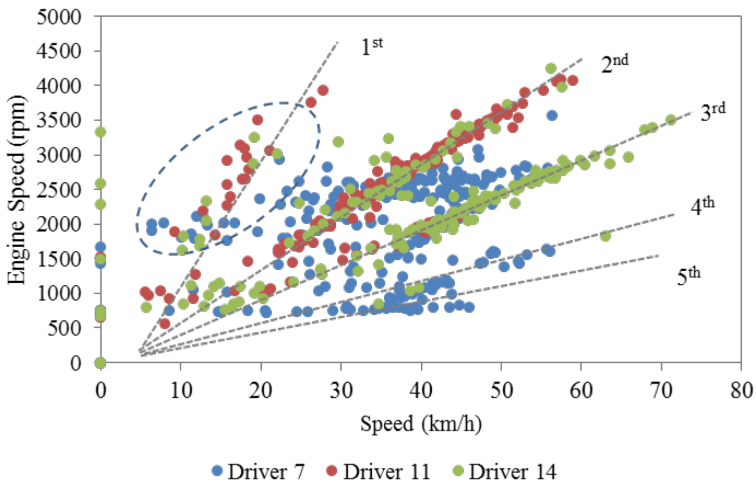
In order to investigate drivers' gear shifting behaviour, first, the linear relationship between vehicle speed and engine speed was calculated, by computing the corresponding values from 10 km/h to 80 km/h at 10 km/h intervals. Then, the drivers' actual engine speeds vs. vehicle speeds were plotted and linear lines representing every gear were added to the graph. The grey dashed lines in the graphs below represent the linear relationships between engine speed and vehicle speed for the ideal scenarios for the five car gears, whilst the dots represent the actual engine and vehicle speeds. There are important observations that can be made from these graphs. Firstly, it is clear that the drivers' gear shifting habits were governed more by individual driving styles rather than vehicle transmission parameters. This agrees with Tutuianu et al. (2013) from their global study on driving behaviours (Tutuianu *et al.*, 2013) . Secondly, these figures can be used to study driver differences at a granular level. As has been discussed in chapter 3, drivers with late gear shifting habits can be identified by comparing actual driving data against the ideal. Those occasions are signified by dashed blue circles on each plot (see

Figure 57 and

Figure 58). Moreover, excessive engine speed while driving over the speed limit is identifiable in each graph's right-hand side. Regarding which, Driver 14 of the Vauxhall Corsa drivers and

Driver 13 of the Nissan Note users were clear examples of this type of driving. The hypothetical lines represent the linear relationship between vehicle speed and engine speed for each gear.

Figure 57. The Vauxhall Corsa drivers' actual and hypothetical engine speeds vs. vehicle speeds



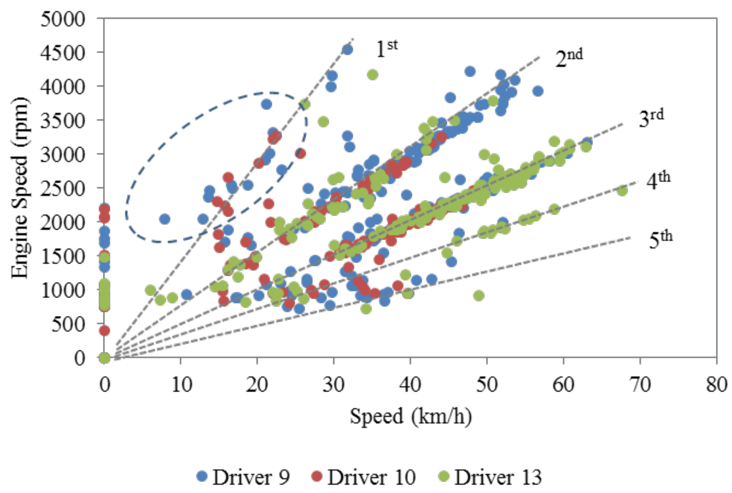


Figure 58. The Nissan Note drivers' actual and hypothetical engine speeds vs. vehicle speeds

4.2.1.3 Results of the study of identifying drivers' aggression by comparing their speed vs. acceleration

In this study, the relationship between the drivers' acceleration and their road speed is investigated, based on Felstead, Mcdonald and Fowkes' study on the potential effect of driving style on fuel consumption and excessive emissions (Felstead, McDonald and Fowkes, 2009). In order to identifying driving styles (calm or aggressive), drivers' instantaneous acceleration values are plotted against speed. The study provides two observations. First, the greater the vertical distribution between acceleration and braking, then the more the driver was driving aggressively. Secondly, if a driver's speed was over the speed limit of 48 km/h (30 mph), then this is another sign that he/she was driving dangerously and ignoring the Highway Code.

As the graphs on figure 59 show, none of the Vauxhall drivers kept below the speed limit all the time. It can be seen that driver 14 had the highest tendency to drive aggressively in terms of acceleration and braking (deceleration) as well spending the most amount of time over the speed limit. Driver 7 was the calmest of the three Vauxhall drivers, with notably very few sudden acceleration incidents and no braking over 40 km/h^2. From figure 60, which shows the acceleration/deceleration/speeding habits of the Nissan Note drivers, it is clear that drivers 10 and 12 always adhered to the speed limit. The latter was a particularly cautious driver when compared to the others, especially driver 11 and driver 13. Moreover, the clear cluster of the dots for driver 12, in particular between 20 and 40 km/h, indicate that his gear changing was the smoothest of all the drivers for both car types. This resembles the result Felstead, Mcdonald and Fowkes' study presented (see Appendix D) (Felstead, McDonald and Fowkes, 2009).

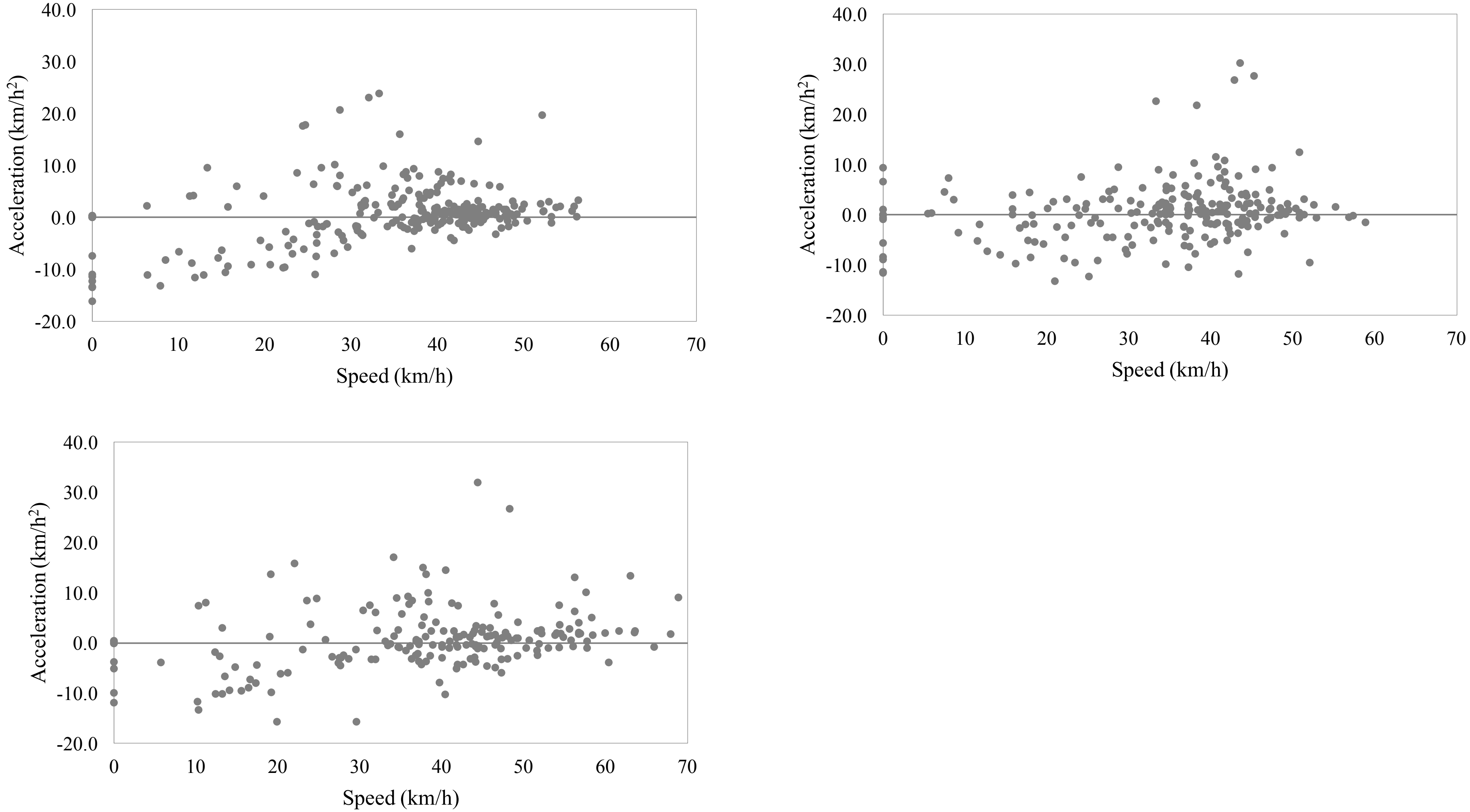


Figure 59. The Vauxhall Corsa drivers' acceleration vs. speed distribution: Driver 7 (top left), Driver 11 (top right), and Driver 14 (bottom left)

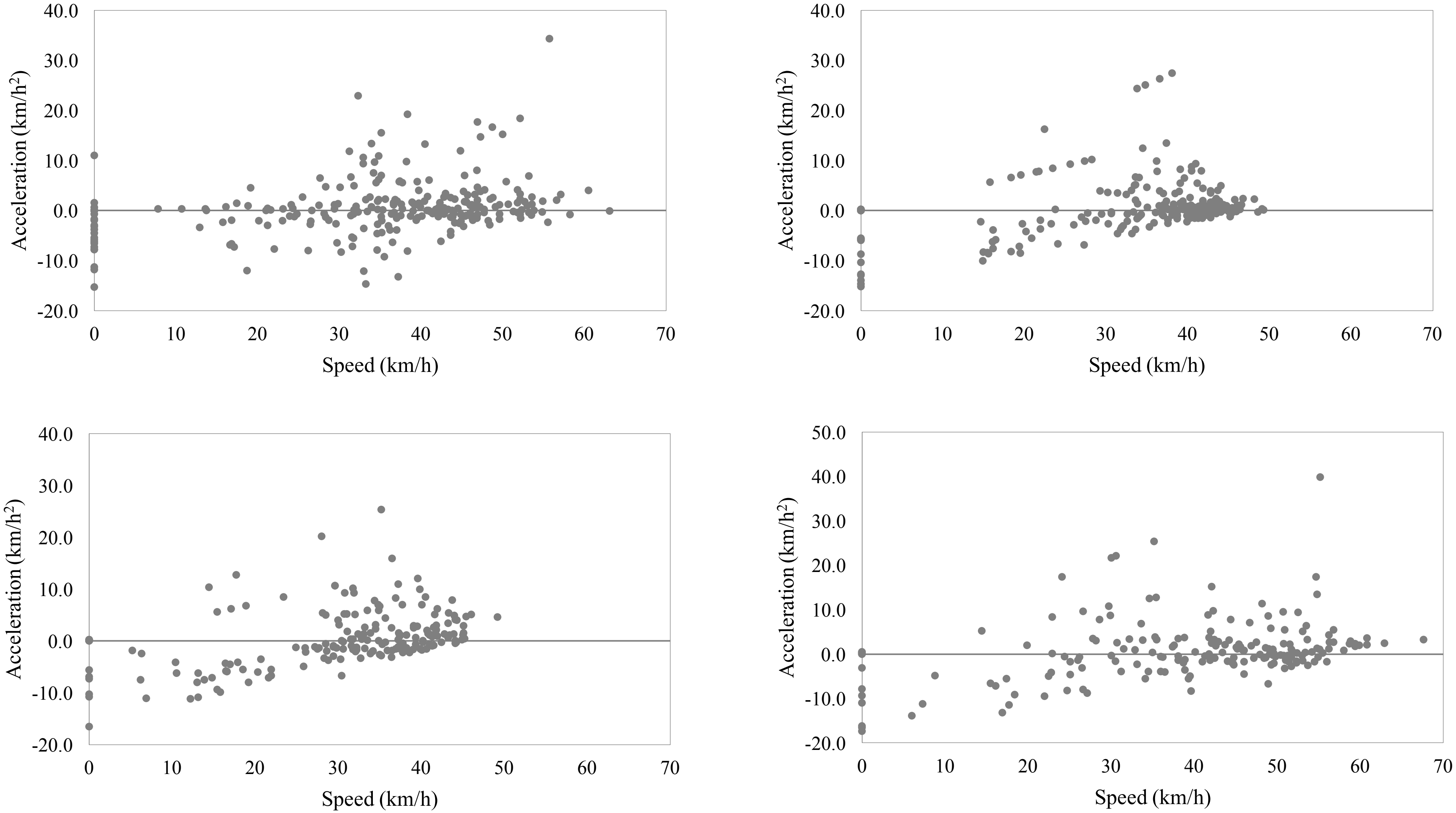


Figure 60. The Nissan Note drivers' acceleration vs. speed distribution: Driver 9 (top left), Driver 10 (top right), Driver 12 (bottom left), and Driver 13 (bottom right)

4.2.1.4 Results of geo-analysis on the vehicle parameters in conjunction with the road profile

This subsection presents the results from geo-analysis of the vehicle parameters. Three distinct analyses were conducted to identify drivers' differences, as follows:

1. Geo-analysis of the vehicle parameters (vehicle speed and engine speed);
2. Geo-analysis of the acceleration behaviour among drivers;
3. Geo-analysis of the vehicle specific power among drivers.

As explained in the previous chapter, in order to present distinguishable differences in the drivers' behaviour and to make the dataset more manageable, only the first, third and fifth driving trial outcomes were used for these analyses.

4.2.1.4.1 The geo-analysis of the vehicle parameters

The geo-analysis of the vehicle parameters is visualised in two formats, with the first being the plotting of the drivers' vehicle and engine speeds in conjunction with road elevation. The second representation is based on the geoprocessing method, where the vehicle parameters according to the GPS points are plotted. Four observations are made from the following graphs. Firstly, from the speed graphs of the drivers' first, third and fifth trials it is clear that there is a big difference between vehicle speeds even with the same driving settings. The second observation is that it is now possible to see where and why acceleration/deceleration occurs, because the road gradient is visible. The third matter is in relation the drivers' engine speeds when driving downhill (first 1.8 km of driving), whereby a low one indicates possible coasting. The fourth observation is that when ascending uphill (from 2 km to 3.8 km), some drivers used substantially greater engine power to climb Bathwick Hill than others.

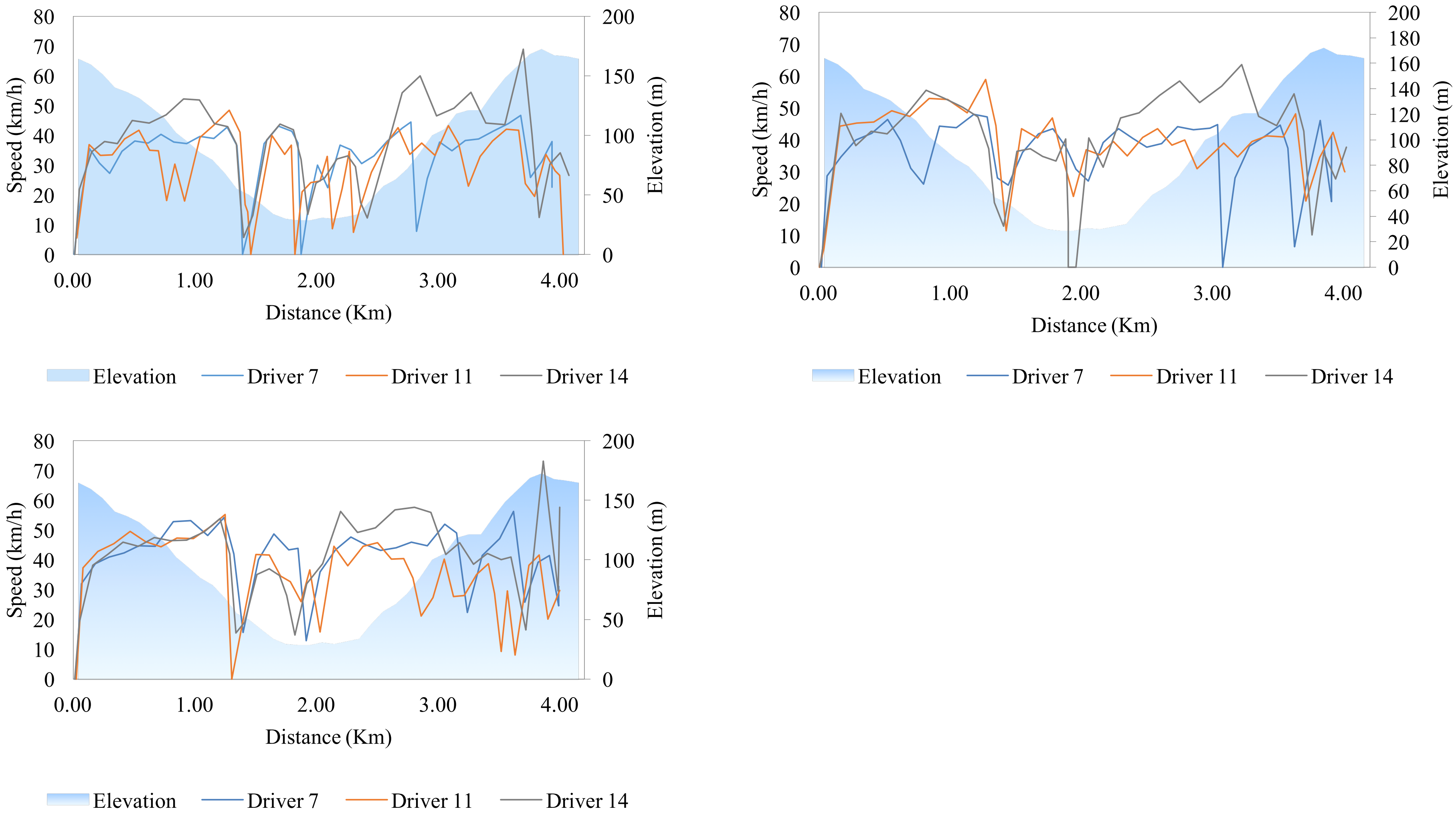


Figure 61. The Vauxhall Corsa drivers' vehicle speed vs. road elevation: first lap (top left), third lap (top right) and fifth lap (bottom left)

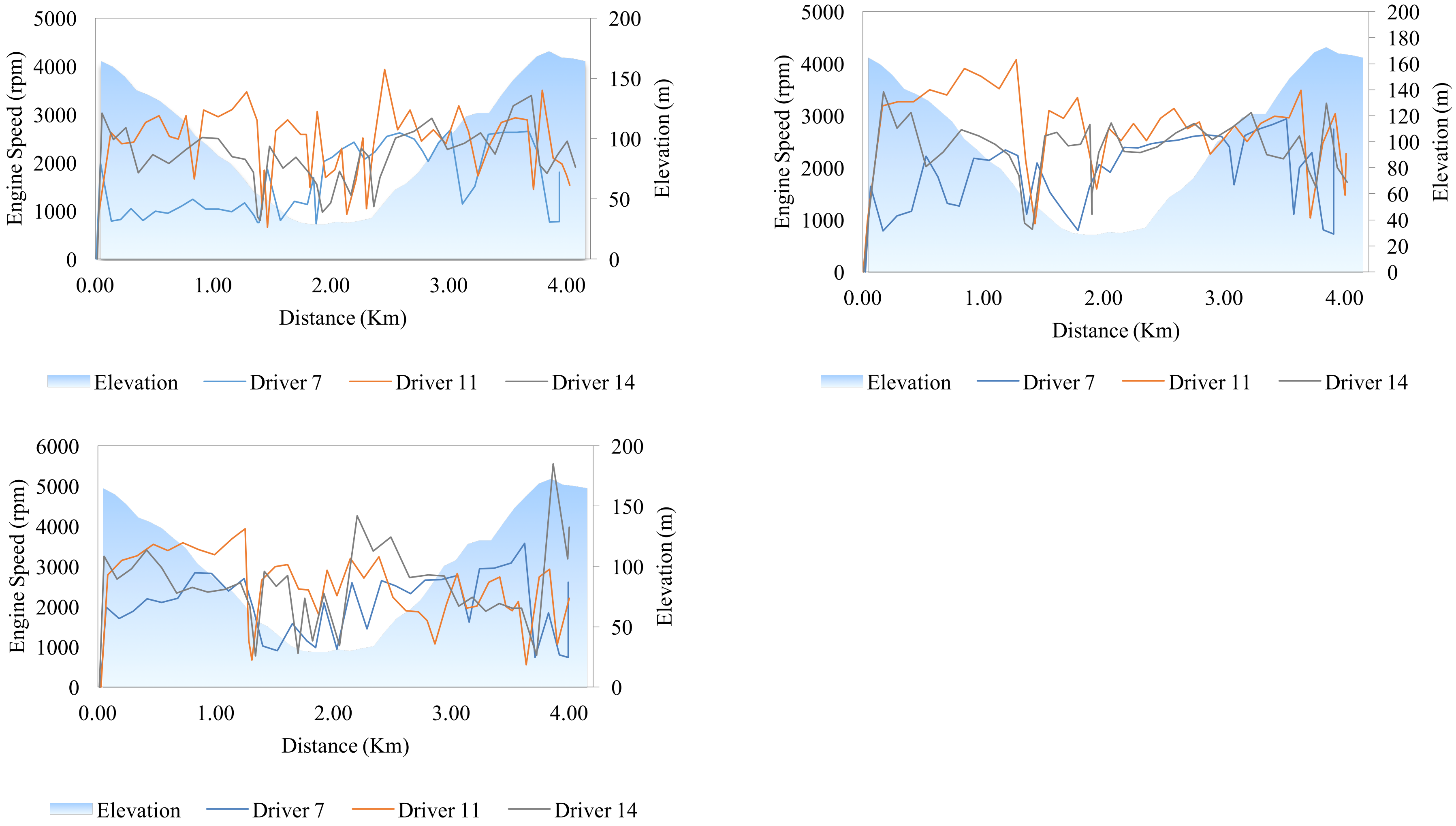


Figure 62. The Vauxhall Corsa drivers' engine speed vs. road elevation: first lap (top left), third lap (top right) and fifth lap (bottom left)

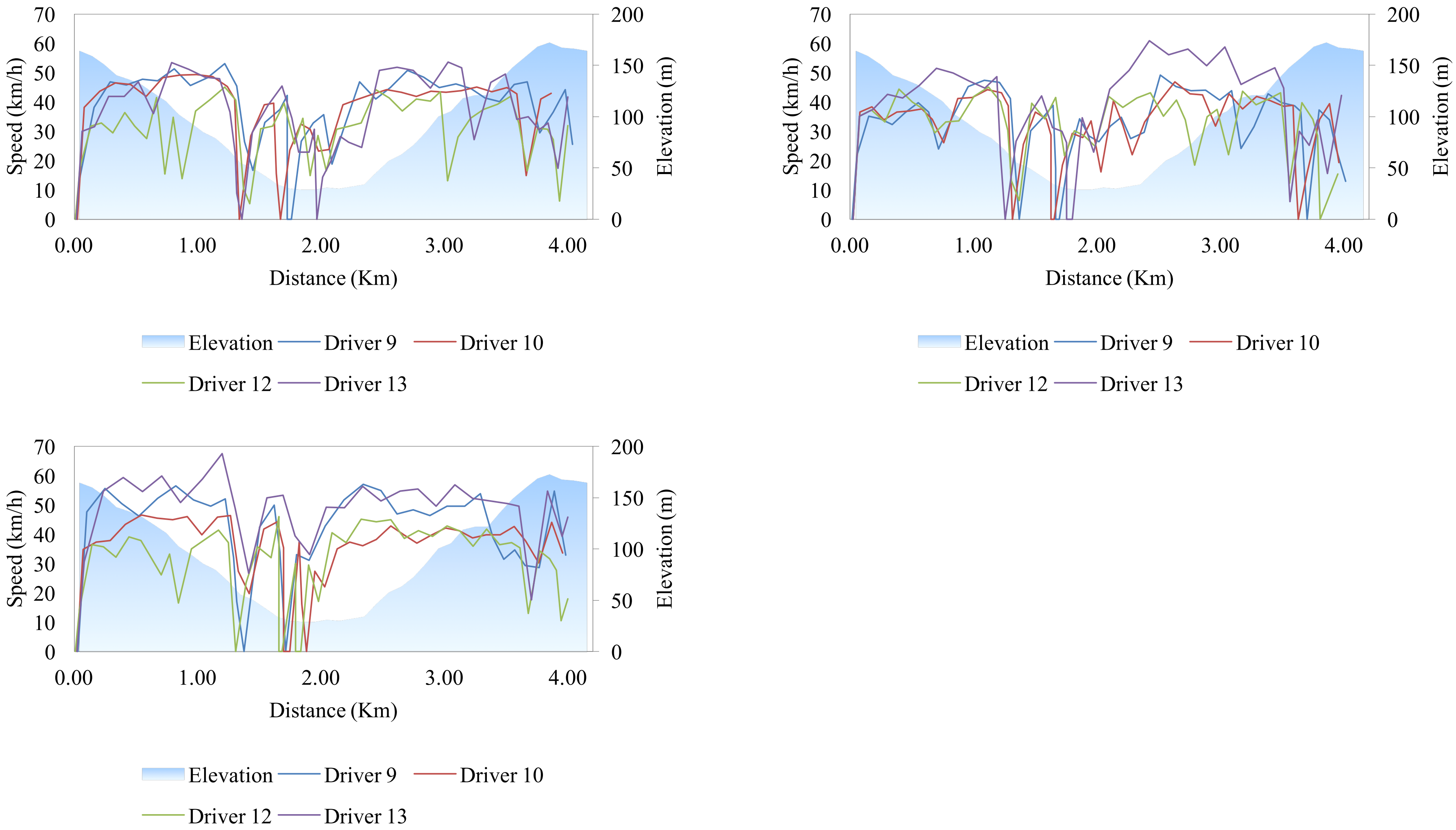


Figure 63. The Nissan Note drivers' vehicle speed vs. road elevation: first lap (top left), third lap (top right) and fifth lap (bottom left)

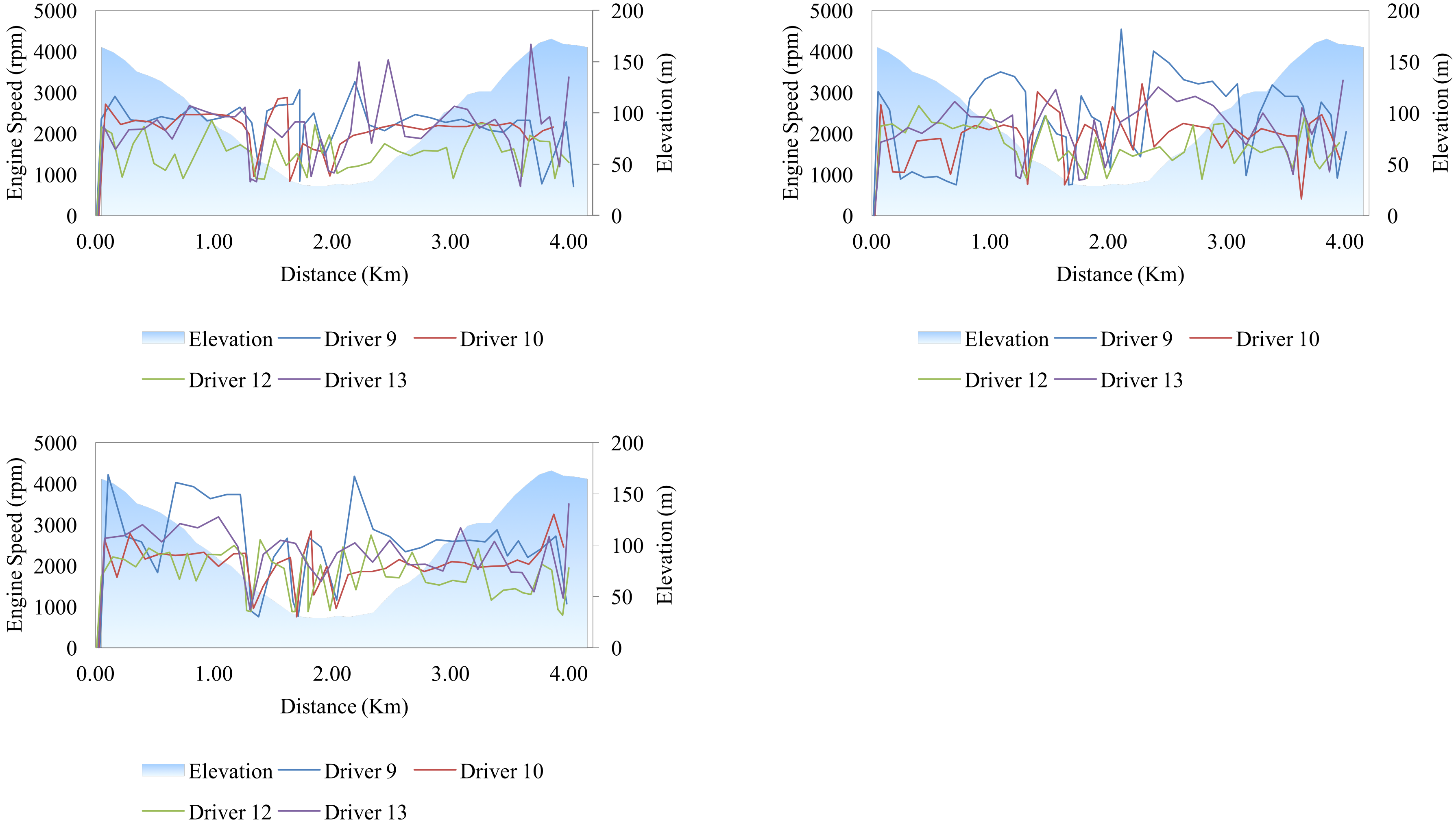


Figure 64. The Nissan Note drivers' engine speed vs. road elevation: first lap (top left), third lap (top right), and fifth lap (bottom left)

It can be seen from the Vauxhall Corsa drivers' speed – distance graphs that driver 7 and driver 11 were very similar in terms of vehicle speed, while driver 14 travelled faster when climbing the hill and generally, had a higher speed throughout the driving event than all the other drivers. Regarding the Nissan Note drivers plotted vehicle speed – distance graphs, it is clearly distinguishable that drivers 10 and 12 had similar styles to each other and drivers 9 and 13 were also alike by the time they traversed the fifth lap, but this was not the case to start with. Overall, driver 13 was the most aggressive Nissan Note driver, with a higher vehicle speed throughout the driving event. The engine speed – distance graphs for both the Vauxhall Corsa and Nissan Note drivers, show occasions where some exhibited a lower engine speed than the other drivers. This can be observed for driver 7's first lap driving data and for driver 9's third. The final observation from the above graphs is in relation to the engine speed differences when the participants approached the uphill part of the route. It was to be expected that they would all have higher engine speed on alighting Bathwick Hill, but clearly, drivers 9, 13 and 14 put their foot down on the accelerator harder than the others did.

The second part of the geo-analysis of the vehicle engine parameters pertained to comparing the drivers' vehicle speed and engine speed geolocation wise. As explained in the previous chapter, for this part of the study, the mapping geoprocessing tool was used to create maps showing the variations in vehicle and engine speeds throughout the route. Layering the data in this format provides a clearer picture of each drivers general driving over different terrain. It is important to recall from the previous chapter that the ESRI software assigns classes to the inputted vehicle and engine speed data that are not uniform. Hence, a comparison between the drivers of the two makes of car is not possible, but these maps are still useful for individual driving behaviour identification. However, one observation that can be made for all the maps is that every driver drove more slowly through the most urban part of the route than when travelling uphill or downhill, perhaps because they were more cautious or there was more traffic. The engine speed maps clearly show how some drivers were over revving when going uphill. Moreover, there is some evidence that drivers 7, 13 and 14 under-revved, i.e. coasted, when going downhill.

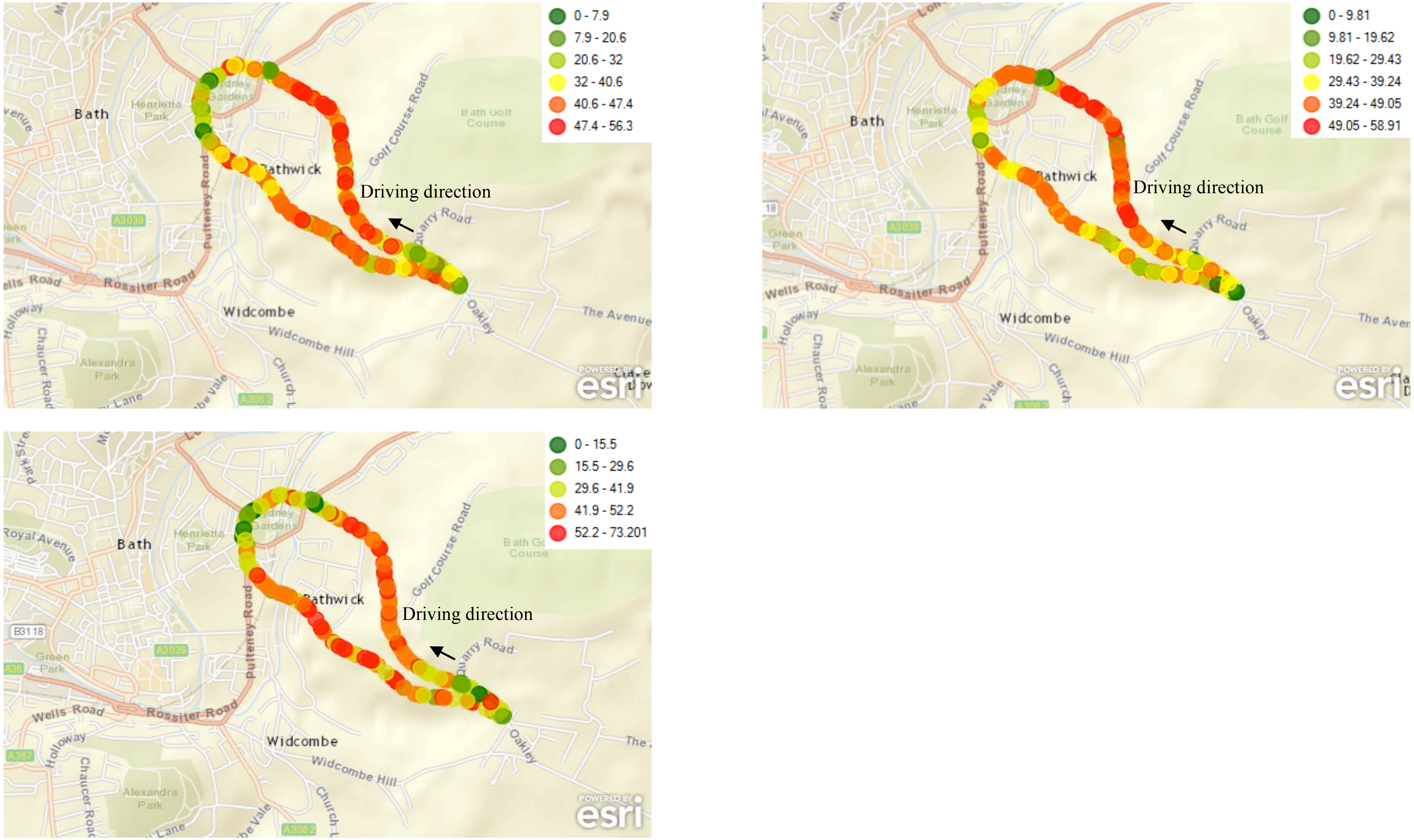


Figure 65. Geo-analysis of the Vauxhall Corsa drivers' vehicle speeds; the maps are based on individual driver's data; the units are in km/h: Driver 7 (top right), Driver 11 (top left), and Driver 14 (bottom left)

Figure 66. Geo-analysis of the Vauxhall Corsa drivers' engine speeds; the maps are based on individual driver's data; the units are in rpm: Driver 7 (top right), Driver 11 (top left), and Driver 14 (bottom left)

Figure 67. Geo-analysis of the Nissan Note drivers' vehicle speeds, the maps are based on individual driver's data; the units are in km/h: Driver 9 (top left), Driver 10 (top right), Driver 12 (bottom left), and Driver 13 (bottom right)

Figure 68. Geo-analysis of the Nissan Note drivers' engine speeds, the maps are based on individual driver's data; the units are in rpm: Driver 9 (top left), Driver 10 (top right), Driver 12 (bottom left), and Driver 13 (bottom right)

4.2.1.4.2 Geo-analysis of the acceleration behaviour amongst the drivers

A comparison between the drivers' vehicle and engine speeds against the road elevation profile provides insights into driving differences with regards to the downhill and uphill sections of the route. In this section, the drivers acceleration behaviour is compared in relation to the gradient throughout the route. In previous studies (4.2.1.4.1), it has been shown that a variation in driving acceleration can lead to the identification of different driving styles, such as calm as opposed to aggressive driving patterns. To illustrate drivers' behaviour differences in relation to their instantaneous acceleration (m/s^2) these data were calculated for the first, third and fifth driving laps and subsequently, plotted in conjunction with the distance and route elevation. As explained in the previous subsection, part of the route involved downhill driving for 1.8 km, when the drivers would be expected to watch their speed. The section between 1.8 km and 3 km from the start entailed urban driving in the centre of the Bath, which has a high volume of traffic, including bicycles as well as a high pedestrian density, thus requiring greater attentiveness of the drivers. The final part of the driving route is between 3 km and the finish, where the drivers' acceleration uphill was to be expected, although not as excessive as in some cases. In the accelerations graphs below, it is possible to observe where sudden changes of speed took place according to the road gradient.

As can be observed from the plots below for the first section of the route (descending downhill), all the drivers became calmer and more confident in handling the vehicle as they completed more rounds. Regarding the second section of the road (urban driving), there were a number of sudden accelerations and decelerations observed to start with, but by the final lap they appear to have been more relaxed. All the drivers accelerated less harshly uphill on the final lap than early on, except for two of the Vauxhall Corsa drivers, namely, drivers 7 and 14. There are two possible reasons for the improved engine efficiency. First, most of the drivers were interested in seeing their progress throughout the event and tried to drive more efficiently on subsequent laps. This was not the primary focus of this study and hence, the level of driving improvement owing to this aspect was not studied in any detail. However, if this was the case, it underlines the importance of educating drivers about how to improve their driving skills. Second, they obviously became increasingly familiar with the course over time and hence, could anticipate the road settings as well as the changes in the terrain. It is also notable that all the Nissan Note drivers exhibited very similar acceleration patterns by the time of the fifth lap. The outcomes of this study support the conclusion that has been made in the previous subsections about the drivers' behaviour and their driving differences.

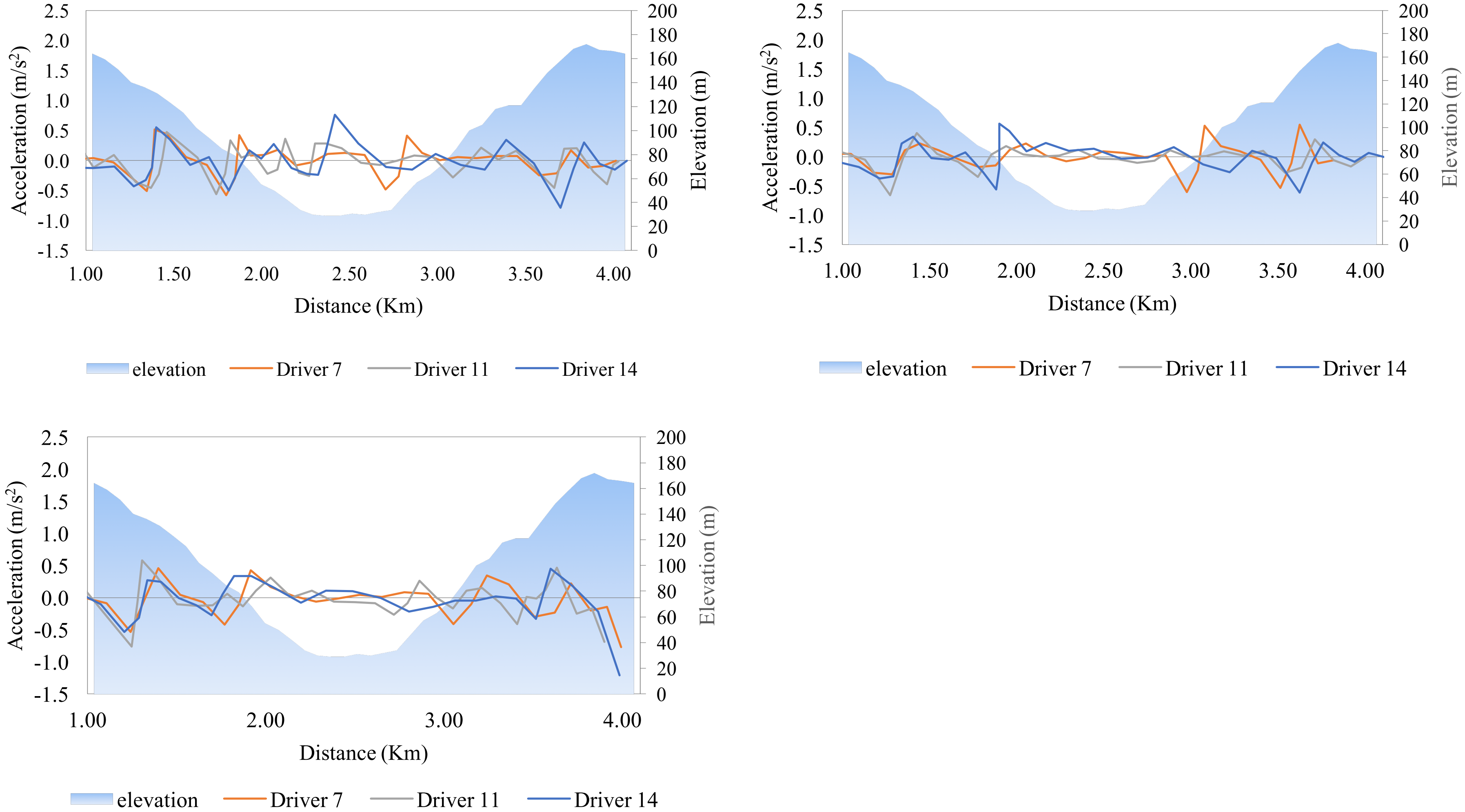


Figure 69. Vauxhall Corsa drivers' acceleration vs. road elevation: first lap (top left), third lap (top right), and fifth lap (bottom left)

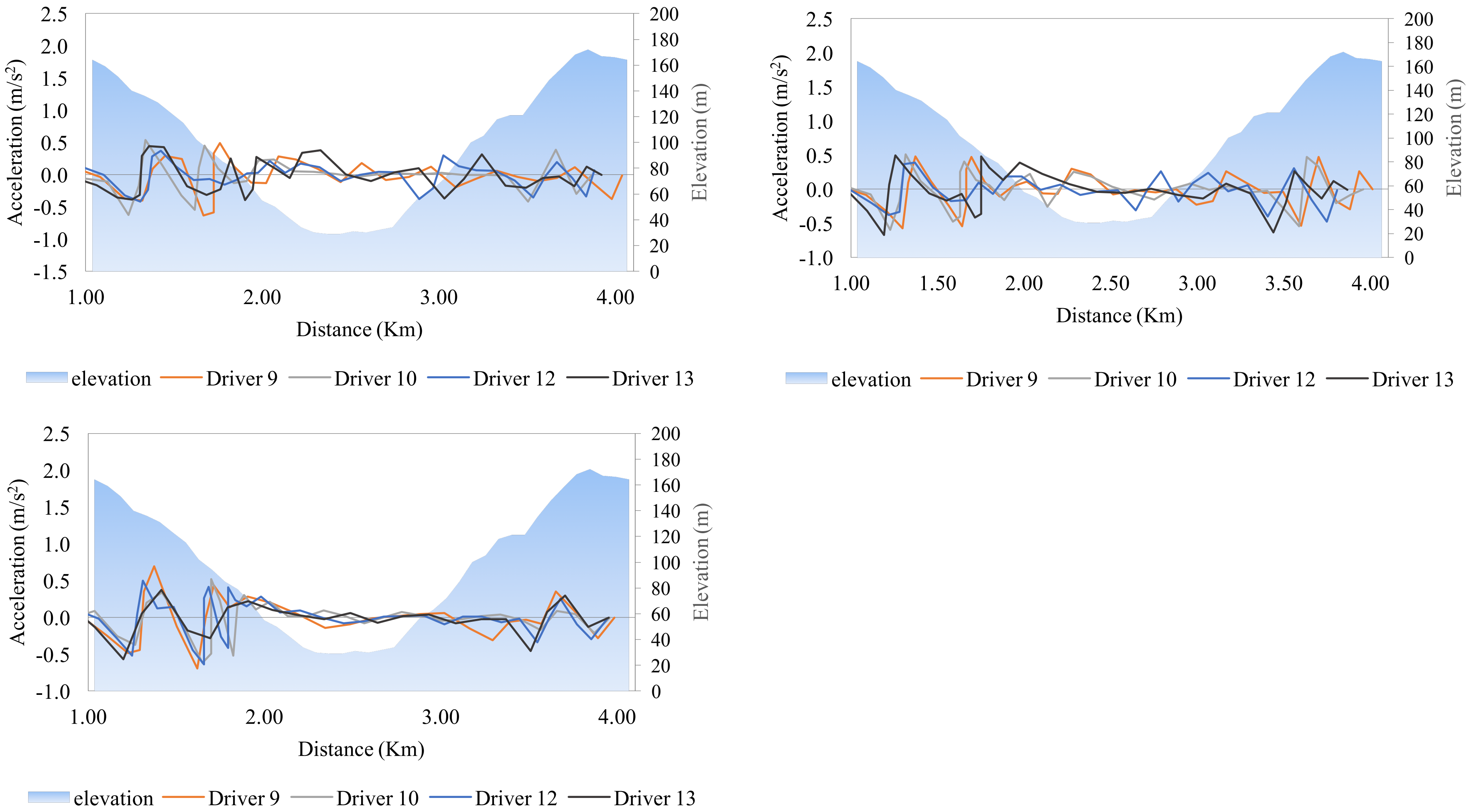


Figure 70. Nissan Note drivers' acceleration vs. road elevation: first lap (top left), third lap (top right), and fifth lap (bottom left)

4.2.1.4.3 The results of geo-analysis of the vehicle specific power among the drivers

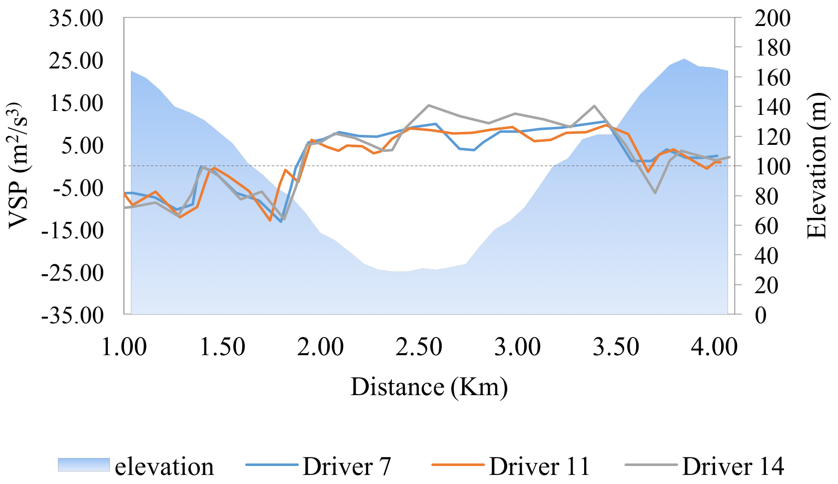
In this subsection, the instantaneous vehicle specific power values of drivers are compared. In order to make a comparison between these values in conjunction with road gradient, the route elevation is included in the following graphs. The VSP values were calculated for the first, third and final driving laps, with the constant parameter values for this process being shown in Table 48 below.

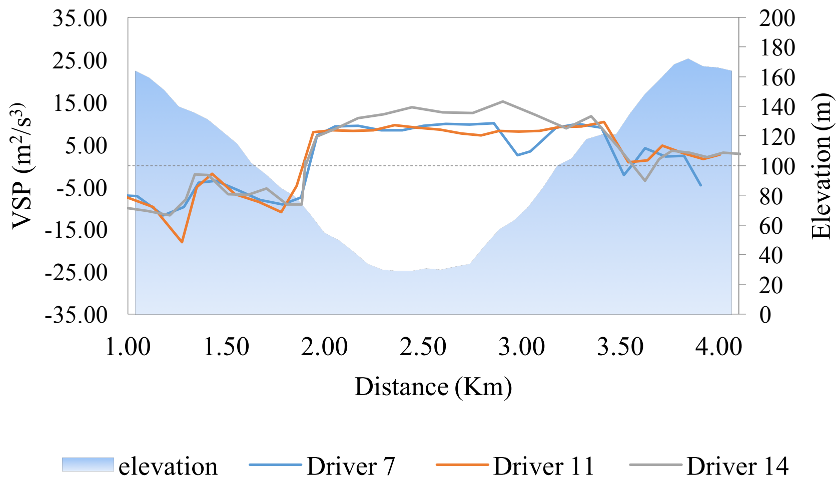
Table 48. Vehicle Specific Power (VSP) constant parameters

Parameter	**Vauxhall Corsa**	**Nissan Note**
Rolling resistance coefficient (ψ)	0.030	0.029
Drag coefficient (ζ)	0.00033 (m^{-1})	0.000298 (m^{-1})
Gravitational acceleration (g)	9.8 (m/s^2)	9.8 (m/s^2)

Differences among drivers' driving performances have already been discussed in terms of their vehicle speed, engine speed and their accelerating/decelerating habits. In this study, the effect of their vehicle speed, acceleration, road gradient and the aerodynamic characteristics of their car are all calculated as one value, namely, the VSP. This allows for more accurate comparisons regarding driving efficiency amongst the drivers as it incorporates real driving behaviour and the road topography with the constant parameters of the vehicle. When including the gradient, the results in the VSP increase when travelling uphill and decrease when going downhill, as can be seen in the graphs below. Most importantly, having the VSP values means that all the cars can now be compared, rather than just those that are the same model.

The plots of the VSP values for the Vauxhall Corsa drivers demonstrate that driver 14 had the highest level in both the urban area and when beginning to drive uphill throughout the laps, when compared to drivers 7 and 11. This indicates that he was a poorer driver than the other two. Regarding the Nissan Note drivers, they had similar VSP values, which was expected given their speed – distance plots and acceleration – distance plots were also similar. Driver 13 occasionally had a higher VSP than the other Nissan Note drivers. In terms of comparisons for both makes of car, driver 14 exhibited the highest VSP values of all throughout the event.





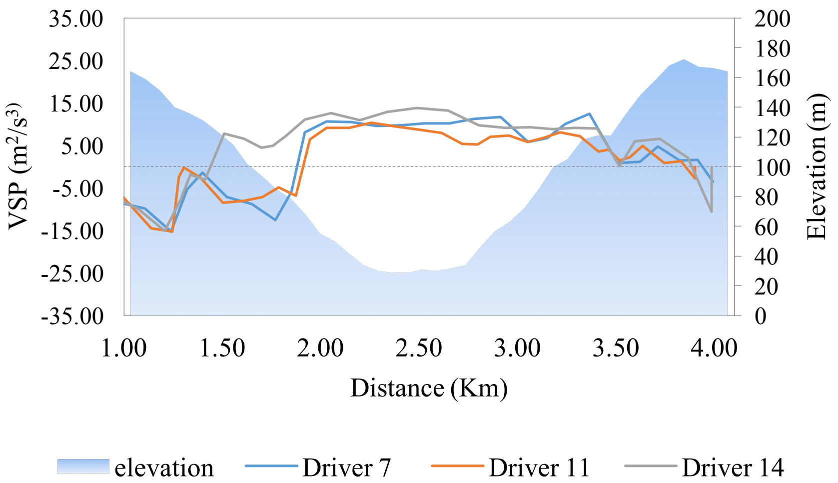


Figure 71. Vauxhall Corsa drivers' VSP values vs. road elevation: first lap (top left), third lap (top right), and fifth lap (bottom left)

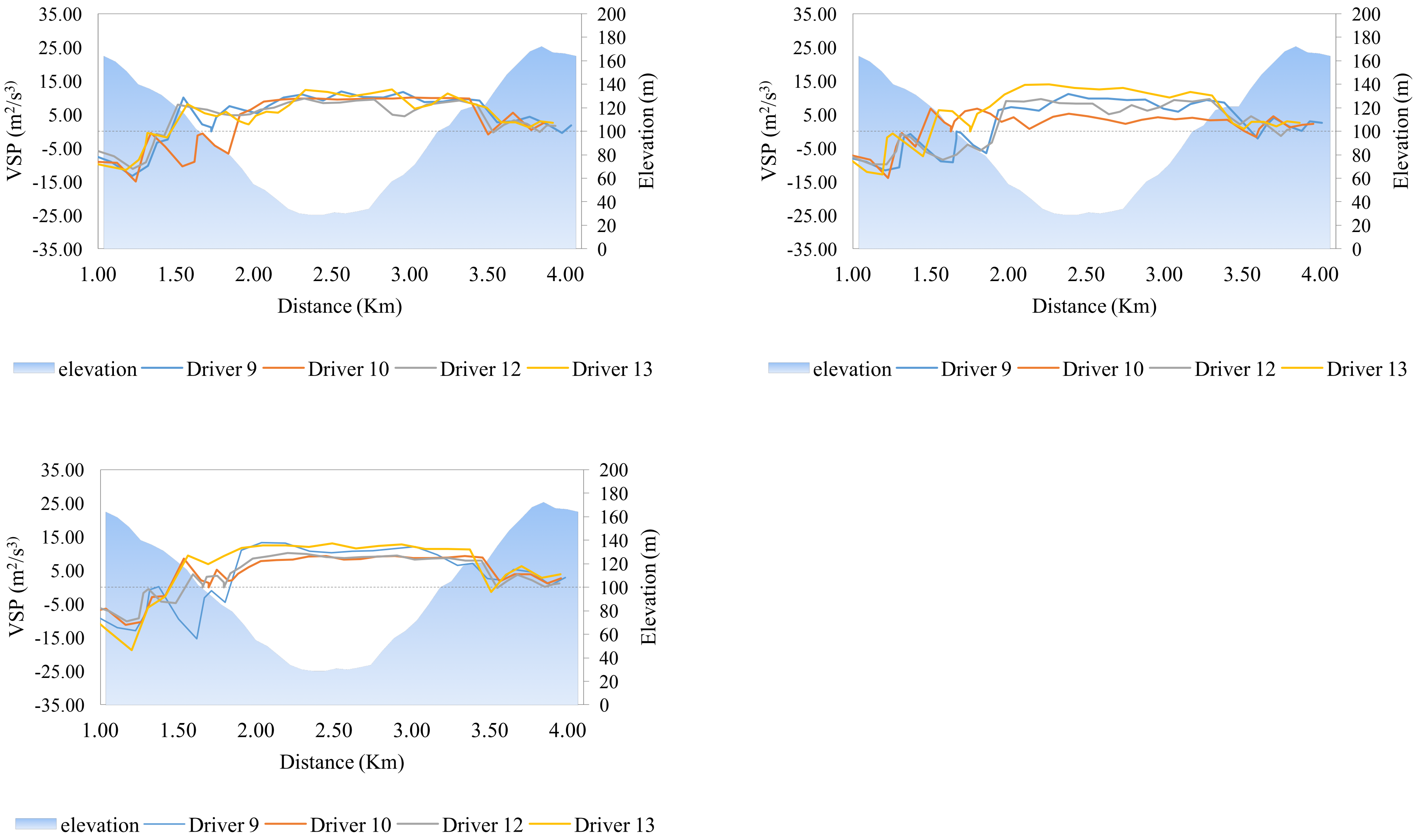


Figure 72. Nissan Note drivers' VSP values vs. road elevation: first lap (top left), third lap (top right), and fifth lap (bottom left)

4.2.1.5 Results of descriptive analysis of vehicles speed profile

This subsection presents the outcomes regarding the descriptive analysis of the drivers' variations in the use of vehicle speed. Firstly, in subsection 4.2.1.5.1, the findings regarding speed consistency by comparing vehicle speed with cumulative frequency plots are presented. Secondly, drivers' speed based on four speed ranges (driving modes) is investigated, which allows for the determination of the range of vehicle speeds that the drivers most preferred and their general level of consistency in their driving. Finally, the results of box-and-whisker plots of the drivers' vehicle speeds are presented, which provide a comprehensive view of the speed variations and ranges. The second part (4.2.1.5.2) of this is dedicated to descriptive analysis of the drivers' speed distribution, which is so as to verify the assumption that their speed profiles were not normally distributed.

4.2.1.5.1 Results of the analysis of the drivers' consistency and the range of chosen speed

The drivers' vehicle speed cumulative frequency plots provide insights into the consistency of the driving habits and the range of speed they primarily opted to drive within. It allows for identification as to whether any drivers exhibited similar vehicle speed trajectories and/or preferred driving speed ranges. Similar to the work of Daniel, Brooks, & Pates (Daniel, Brooks and Pates, 2009), the cumulative frequency plots were created using the monitoring data for all the Vauxhall Corsa drivers, all the Nissan Note drivers, and both groups combined. On each plot, the 'All Drivers' line presents the speed data of all nine drivers participating in the event and including these data allows for comparison of each driver with the overall population. Hence, in each chart, two comparisons can be made, firstly, the consistency of the driver during the driving trips. When the lines are close, then this shows that the drivers are similar to each other. The second observation is the comparison between each driver's trips and the all drivers line, shows how an individual drove compared to the rest of the group.

From the cumulative plots of the Vauxhall Corsa drivers, it can be concluded that drivers 7 and 11 had similar vehicle speed preferences, whilst driver 14 tended to use a higher speed. Among the Nissan Note drivers, the variation was greater than for the Vauxhall drivers; for instance, driver 12 was always below the speed limit, while driver 13 was above it for a significant amount of the time. Perhaps the most important finding from this study can be seen in the final graph, where the speed cumulative frequencies for all the drivers are combined into one plot. This shows that the drivers who were more cautious on the left and those less so on the right.

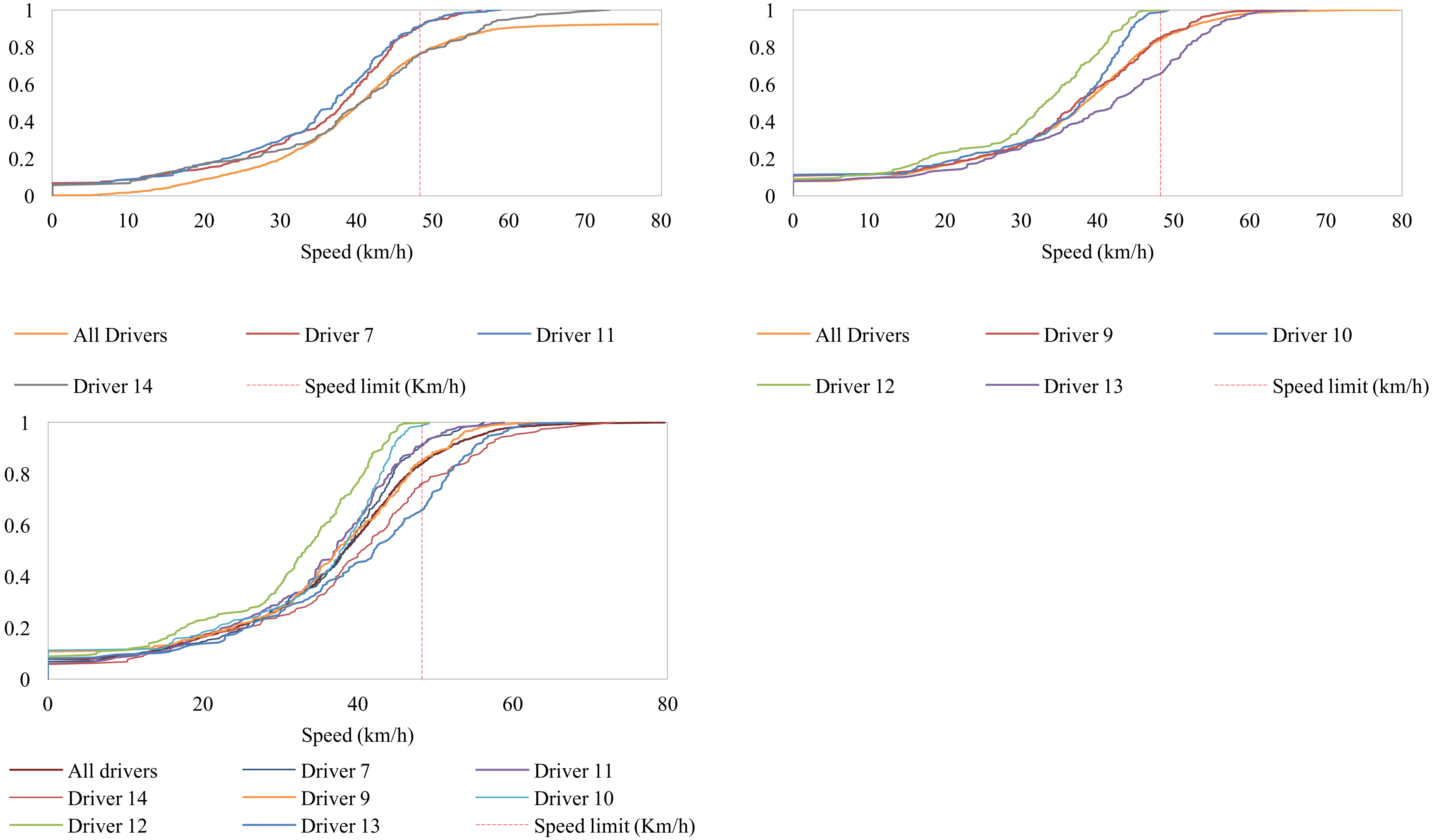


Figure 73. Speed cumulative frequency graph, all Vauxhall Corsa drivers, (top left), All Nissan Note drivers, (top right), and All Class B drivers (bottom left): Speed limits are marked with dashed red lines

The second study in this part compares the drivers' vehicle speed based on the number of times they travelled within specific speed ranges. The following speed range criteria were chosen:

1. Between 0 to 5 mph (time at rest);
2. Between 5 to 24 mph;
3. Between 25 to 30 mph;
4. Over the speed limit.

While from previous charts the consistency of drivers' speed during the driving event can be examined, the following graph aims to provide an overall view of each driver's tendency to drive in the four mentioned zones. Visualisation of the data in this manner provides an instant view of every drivers' behaviour. For example, drivers 10 and 12 never went over the speed limit.

The results are plotted in the percentage stacked bar chart format to show which ranges the drivers mainly drove in. The data for drivers 8 and 15 are also included in this study. As has been mentioned before, the speed limit is 30 mph[92] for the entire driving route and it can be seen that drivers 10 and 12 rarely, if ever, went over the speed limit. Drivers 7, 11 and 12 spent around 50% or more of their time between 5 mph and 24 mph, while the more aggressive drivers, such as drivers 13, 14 and were at at least 25 mph for over half of their time. In fact, these drivers spent 20% or more of the whole event over the speed limit.

[92] For this graph, drivers speed is presented in mph, because the speed limited is also given in this unit.

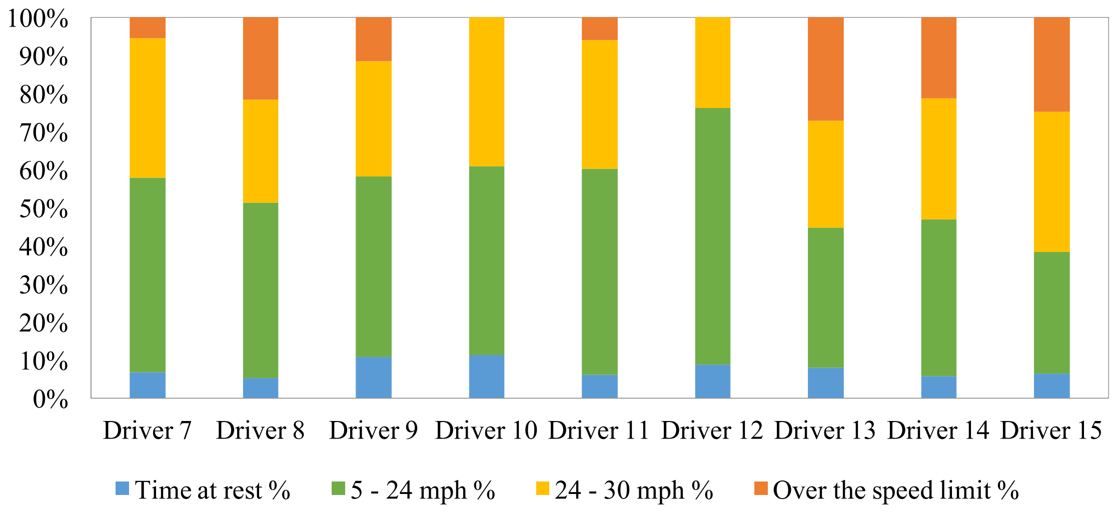


Figure 74. Stacked bar chart of the drivers' chosen range of vehicle speeds

The final part of this section involves a comparison of the drivers' vehicle speeds in the box-and-whisker plot (box plot) format. The box diagram carries information that is not identifiable from the previously discussed methods. To plot drivers speed range box plot, the median as well as the first and third quartiles of their vehicle speed samples were calculated. The computed values are presented in the table below. See a full list of the computed parameters in Appendix F.

Table 49. Drivers' speed box plot parameter values

	Driver ID								
Parameters	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
Min	0	0	0	0	0	0	0	0	0
Q_1	28.4	28.2	28.5	27.42	26.92	22.27	29.9	31.2	32.9
Median	38.3	39.3	37.3	37.8	37.2	33.2	42.1	40.5	43.8
Q_3	43.7	48.3	45.4	42.27	42.82	39.57	50.8	48	50
Max	56.3	79.6	63.1	49.3	58.9	49.2	67.6	73.2	61.8

From the box plots, the range of vehicle speed for each driver, their maximum speed, and the interquartile range, which indicates the speed range each driver most often drove within, can be ascertained. The box part of the graph shows the speed range between Q_1 and Q_3, while lines within each box show the median value. The boxes with their whiskers show the full spectrum of the drivers speed from minimum to maximum and the dashed red line represents the speed limit (30 mph or 48 km/h). As the plot demonstrates, drivers 7, 9, 11, 10 and 12 appear to have been calmer compared to drivers 8, 11, 13, 14 and 15. However, being calm does not necessary mean a driver was driving within the speed limit, for example, except for drivers 10 and 12, the other drivers were over the speed limit at least once. Among the class B drivers, driver 11 and driver 12 had the lowest Q_2 values, while driver 13 and 14 has the highest. The outcomes of the studies in this part have shown the significance of vehicle speed when comparing driving performance and identifying differences in driving behaviours. In order to come to any conclusions about classifying driver behaviours, it is necessary to assess the nature of each driver's sample speed distribution. Specifically, the data need to be tested to see whether they are normally distributed or not. For, such analysis will determine the appropriate statistical method to be used for their modelling.

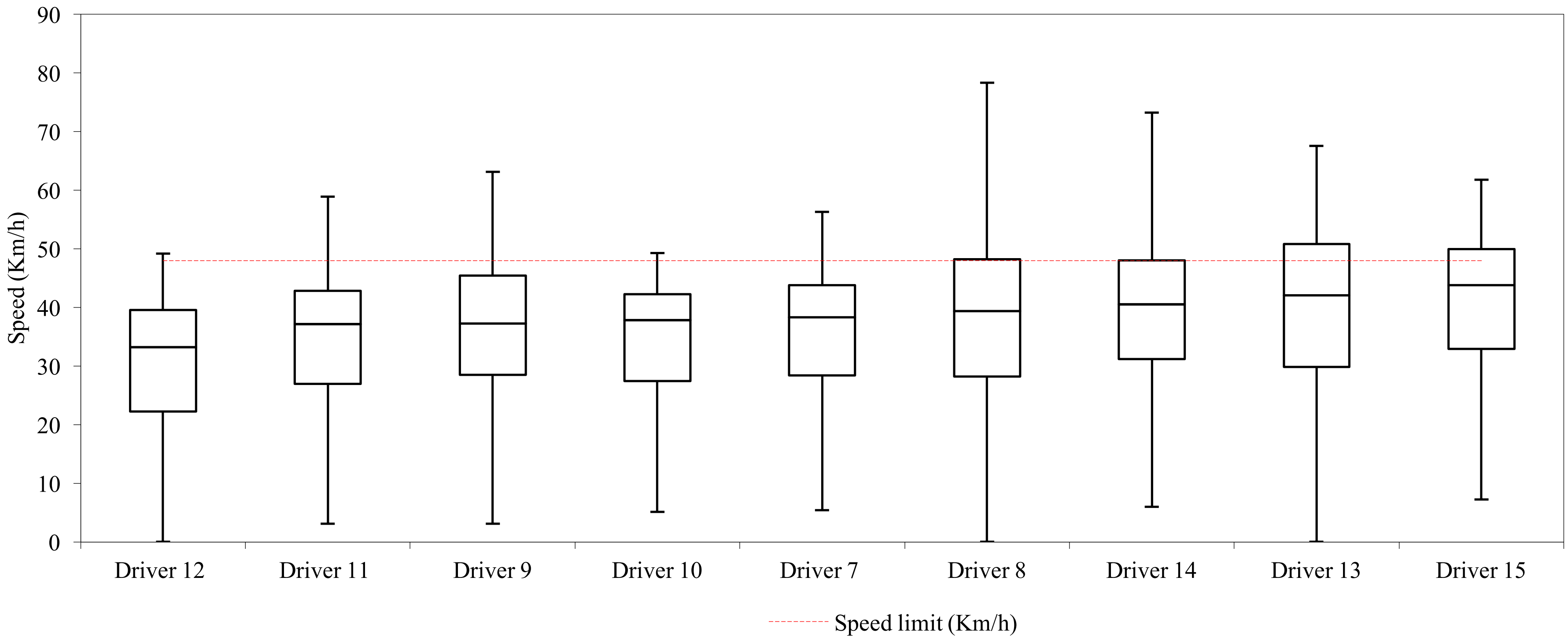


Figure 75. Box-and-whisker plots of the drivers' vehicle speed ranges

- Drivers are arranged according their speed distribution median, from the lowest value (driver 12) to the highest (driver 15)

4.2.1.5.2 Results of the descriptive analysis of the drivers speed distribution

As has been shown in the above studies, driver vehicle speed has the potential to represent driving behaviour differences and preferred driving speed range. Prior to drawing any conclusions about vehicle speed profiles, it is important to establish their type of distribution. As has been discussed in the previous chapter (3.4.3.5.2), it is very unlikely that the driving speed for a single driver is normally distributed, since drivers speed are commonly around the speed limit, and tilting to either side depending on choice of speed.

To this end, the sample mean ($\bar{x}$), standard deviation (σ), skewness (S) and excess kurtosis (K) values were calculated for all the drivers' vehicle speeds. The skewness and excess kurtosis values were obtained to determine whether the drivers' speed data were normally distributed. This is the case for any given dataset, if the values of skewness and excess kurtosis are between -2 and +2. Subsequently, the Jarque-Bera normality test was applied to the speed profiles to check these distributions and the results are presented in the tables below. As aforementioned, as the drivers would generally be expected to travel close to the speed limit, it is highly unlikely that their values would be normally distributed. However, it is important to confirm that this is so as it determines which inferential analysis for classifying the drivers' in terms of their vehicle speed is appropriate. Table 50 summarises the results of the mean, standard deviation, skewness and excess kurtosis values for each driver. The second table (Table 51) shows the outcome of the Jarque-Bera normality test outcome, where the p-value, skewness outcome, and Jarque-Bera normality test outcome are tabulated for comparison. The results show that even with the Jarque-Bera normality test, which is suitable for this type of distribution, only drivers 8 and 14 passed this test.

Table 50. Descriptive values of the drivers' speed datasets (the mean value is in km/h)

Drivers ID	Mean	Standard Deviation	Skewness value	Excess kurtosis
7	36.86	10.54	-0.83	0.41
8	39.97	15.51	0.03	-0.04
9	38.54	10.53	-0.45	-0.06
10	37.34	7.52	-1.05	0.47
11	35.76	10.98	-0.68	0.11
12	32.51	9.77	-0.89	0.11
13	41.51	12.40	-0.60	-0.20
14	40.31	13.76	-0.36	-0.07
15	42.66	11.34	-0.76	0.18
All Drivers	38.19	11.89	-0.32	0.30

Table 51. Results of the Jarque-Bera normality test outcome

Drivers ID	Z Score	P-Value	Skewness outcome	Jarque-Bera normality test
7	28.18	0.0%	True	False
8	0.06	97.2%	True	True
9	7.68	2.1%	True	False
10	38.10	0.0%	True	False
11	15.80	0.0%	True	False
12	23.17	0.0%	True	False
13	10.54	0.5%	True	False
14	3.88	14.4%	True	True
15	16.34	0.0%	True	False
All Drivers	38.69	0.0%	True	False

1. The normality test for all the drivers' speed distribution and the Q-Q plot (probability plot);
2. Individual histograms for all nine drivers so as to compare drivers' speed distributions. (see Appendix E).

The final stage in this subsection is to present and interpret these results. The histogram plots have been used to demonstrate drivers' vehicle speed distributions, with the continuous lines representing what they would look like if they were normally distributed (see Appendix E). Despite the skewness and excess kurtosis values being in the normal distribution range for all drivers and taken all together the drivers' distribution being normal, based on the individual Jarque-Bera normality tests it can be concluded that except for driver 8, driver 14 and driver 15, the drivers' vehicle speeds are not distributed in this way (see Appendix E). These two drivers exhibit this distribution because they spent so much time over the speed limit. The rest of the drivers' distributions show negative skewness with larger tails for the lower vehicle speeds (see Appendix E). Figure 76 shows the overall speed distribution of all the drivers' driving speed data (average 38.20 km/h). With the JB test, the sample showed that it is not normally distributed and hence more trials would be needed for this situation to be reached. This is consistent with the conclusion made by other scholars regarding the normally distributed driving speed on stretches of roads (Berry and Belmont, 1951; Center for Transportation Research and Education, 2001; Souleyrette, Stout and Carriquiry, 2009). Figure 77 (Q-Q plot) demonstrates how closely the sample is distributed as normally distributed data.

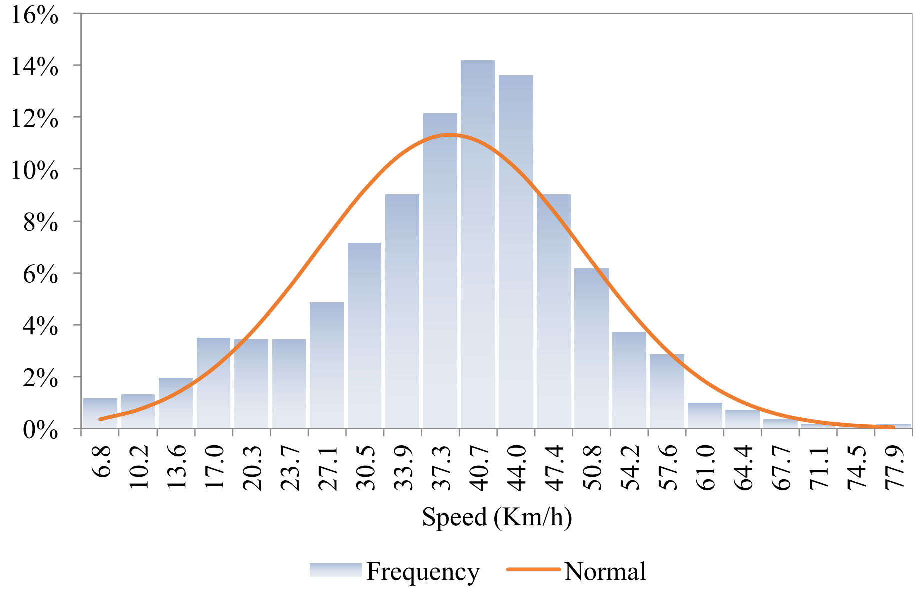


Figure 76. The speed distribution of all the drivers

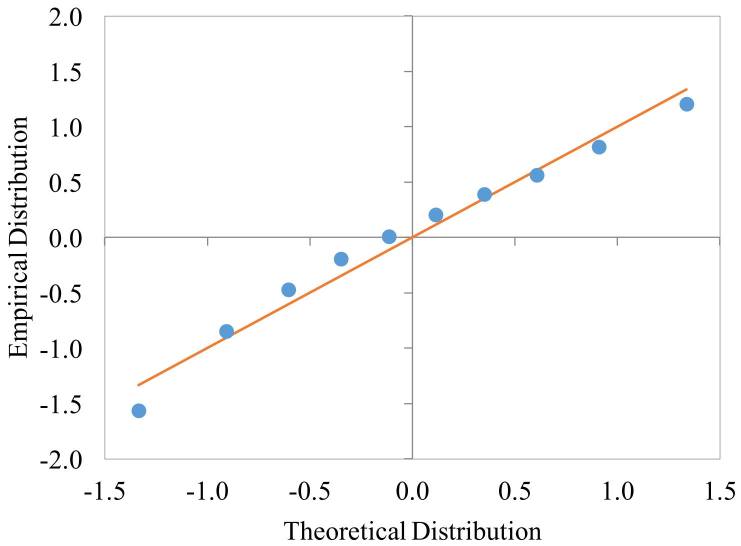


Figure 77. The Q-Q plot of all the drivers' speed distribution

4.2.2 Results of the drivers' performance classification study

This subsection presents the results of the classification methods used to group drivers with similar driving data together. It also involves ranking them based on their fuel usage, according to a scoring method commonly used for usage-based vehicle insurance and by fleet management service providers. This part contains three subsections, the first, provides the results of the statistical power and the effective sample size analysis from using the Wilcoxon-Mann-Whitney (WMW) method. The outcomes of this analysis are presented in the form of a matrix of effective sample sizes. Next, the matrix results are visualised by network analysis and the drivers are grouped together based on their vehicle speed similarities. The second part of this subsection contains the results of the study categorising the drivers based on their fuel usage. This includes, first, comparing and reporting drivers' fuel consumption by using two common metrics widely used to report the fuel economy of cars and drivers. Secondly, the results of the Vehicle Specific Power – Fuel Consumption metric are presented and discussed. The final part of the classification study provides the classification results of the drivers according to the fleet management scoring method.

4.2.2.1 Classifying the drivers based on the similarity of their speed distribution

This subsection presents the results of the drivers' classification based on their speed distribution, which involved calculating the effective sample size between every pairing of the drivers. As explained in the previous chapter, the concept of effective sample size was used to represent the difference between two drivers' driving behaviour. Specifically, if their vehicle speed had a similar sample distribution (i.e. sample mean and sample standard deviation), then the effective sample size value would be a very large number. This would mean that these two drivers would have to undertake certain number of trial laps before their driving behaviour could be distinguished. The table below summarises the drivers' average speed and standard deviation.

Table 52. Drivers speed mean and standard deviation values (the mean value is in km/h)

	Drivers ID								
Parameters	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
Mean	36.86	39.97	38.54	37.34	35.76	32.51	41.51	40.31	42.66
Standard Deviation	10.54	15.51	10.53	7.52	10.98	9.77	12.4	13.76	11.34

4.2.2.1.1 The matrix results of the effective sample size analysis

As the matrix shows, driver 7, had very similar speed distribution characteristics to drivers 9 and 11, for the reasons explained above. With the same logic, the effective sample size between driver 7 and 8 is comparably small and this indicates that these two drivers were very different in their driving styles. This is true between driver 7 and drivers 12, 13, 14, and 15. The best way to show these relationships is by using network analysis, as with such matrix data it is possible to use it in order to visualise these relationships in a meaningful way. In the following subsection, they are displayed using a node and edge type of network.

Table 53. The effective sample size matrix of the drivers' speed profiles

Driver ID	7	8	9	10	11	12	13	14	15
7	-	262	12214	2366	22785	305	243	356	161
8		-	257	144	131	42	18385	7558	889
9			-	1027	3016	175	239	381	151
10				-	3578	993	138	179	104
11					-	284	129	180	92
12						-	43	52	37
13							-	3322	1606

Driver ID	7	8	9	10	11	12	13	14	15
14								-	595
15									-

4.2.2.1.2 The results of the network analysis of drivers' speed distribution similarity

The network analysis was conducted to demonstrate the associations of drivers according to their vehicle speeds. The first visualisation was based on distance of the nodes, whereby driver pairs with high effective sample size values (similar speed distribution) are drawn close to each other, as illustrated in the three-dimensional network diagram below. In addition to the distance of the nodes, the thickness shows the degree of association between the drivers' speed profiles. This visualisation can be interpreted as follows: first, numbers are drivers ID and second the thicker the arrow between any two drivers means a higher level of similarity between the two.

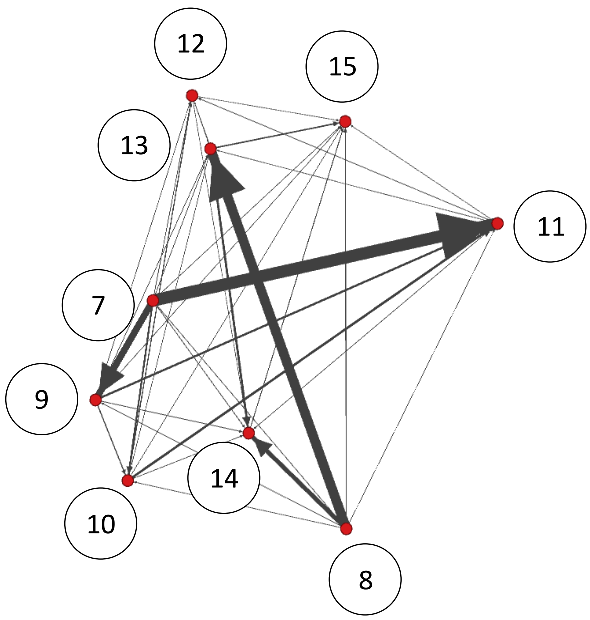


Figure 78. 3D driver association based on the similarity of their effective sample size values

When using Gephi software to demonstrate these relationships, on screen, it is evident what the distances are, because it is possible to rotate the network like a three-dimensional object. However, in a two-dimensional snapshot this is not the case, for instance, whilst drivers 7 and 11 appear to be distant, in fact, they were very close. Since the range of sample sizes values is from 37 and 22,785, the graph is not entirely capable of demonstrating the association between the drivers and so the effective sample size values were standardised with a range of 0 for the lowest value and 100 for the highest. The standardised values of the effective sample size values are presented the table below.

Table 54. The standardised, effective sample size matrix the drivers' speed profiles

Driver ID	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
7	-	0.98	53.53	10.24	100.0	1.18	0.91	1.40	0.55
8		-	0.97	0.47	0.41	0.02	80.66	33.06	3.75
9			-	4.35	13.10	0.61	0.89	1.51	0.50
10				-	15.57	4.20	0.44	0.62	0.29
11					-	1.09	0.40	0.63	0.24
12						-	0.03	0.07	0.00
13							-	14.44	6.90
14								-	2.45
15									-

After the standardisation process, the spherical arrangement format was applied to visualise the association between the drivers. In this case, as illustrated below, the thicknesses of the links (edges) presents the strength of the relationships between them (nodes) and the distances are no longer of relevance. The improved diagram was created using the NodeXL network analysis and visualisation software. As can be observed in the following graph, driver 7 exhibits strong similarity, in terms of speed profile, to drivers 9, 10, and 11. Driver 8 is very similar to drivers

13 and 14. Drivers 9 and 11 share similar traits of driving. Drivers 10 and 11 have strong association, while driver 10 has slightly less association with driver 12 than driver 11. Driver 11 is not similar at all to drivers 13, 14, and 15, as is also the case for driver 12. Finally, driver 13 shows a strong similarity to drivers 14 and 15 (see figure below).

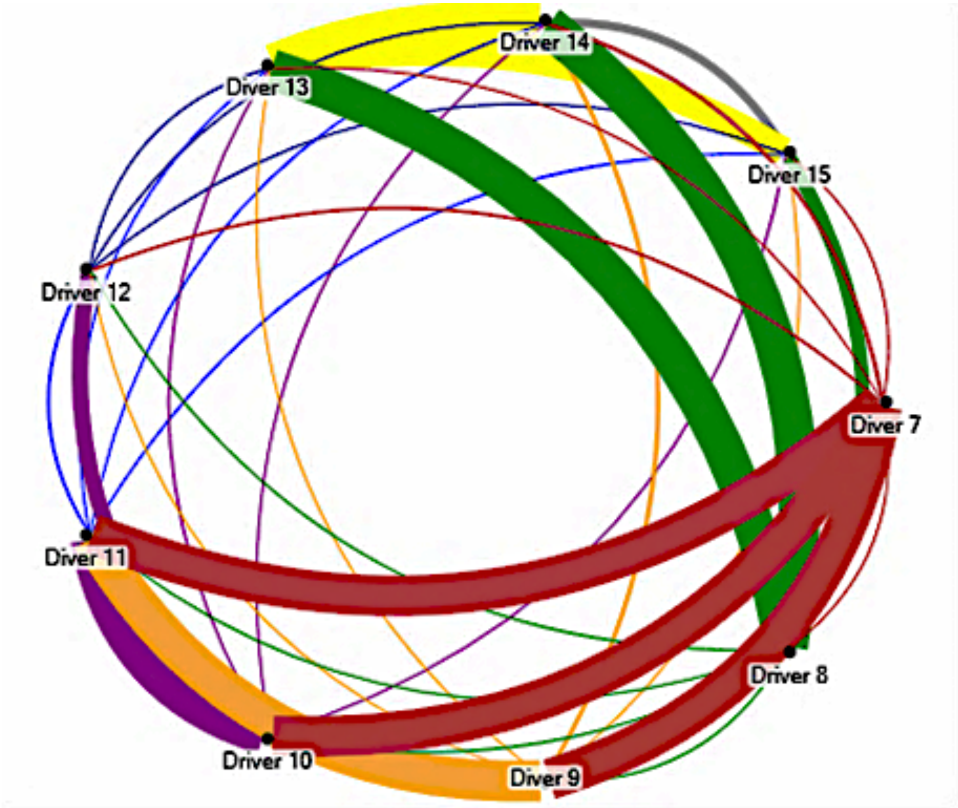


Figure 79. Network analysis of drivers' similarity of speed profiles based on the standardised effective sample size values

4.2.2.2 Ranking drivers based on their fuel economy and fuel consumption

The second approach towards classifying drivers was to rank them based on the amount of their consumed fuel at the driving event. In the first part, the results of three methods of reporting their fuel economy are presented. In the second, the findings from calculating fuel consumption efficiency and the adapted method for calculating drivers' fuel consumption based on their specific vehicle power and specific fuel consumption are given. The three fuel economy metrics used were:

- The amount of consumed fuel per fixed unit of distance ($l/100km$);
- The amount of consumed fuel per distance per unit of fuel (km/l).

4.2.2.2.1 Results of the fuel economy and fuel consumption rating

The following graphs show the results of two metrics for all the class B and class C cars. Vehicles 8 and 15 are included so as to provide a comparison between the two vehicle groups and these single vehicles. The following charts exhibit the participants' overall fuel consumption according to the two mentioned metrics. For both metrics, drivers' total fuel consumption reported from the OBD dongles in litres has been used. In the first metric, the measurement presents fuel consumption, where the lower the value is the less fuel has been used. The second chart shows the most fuel-efficient drivers among the participants. It can be interpreted as that the higher the value is the more efficient a driver performed.

1. The amount of consumed fuel per fixed unit of distance ($l/100km$)

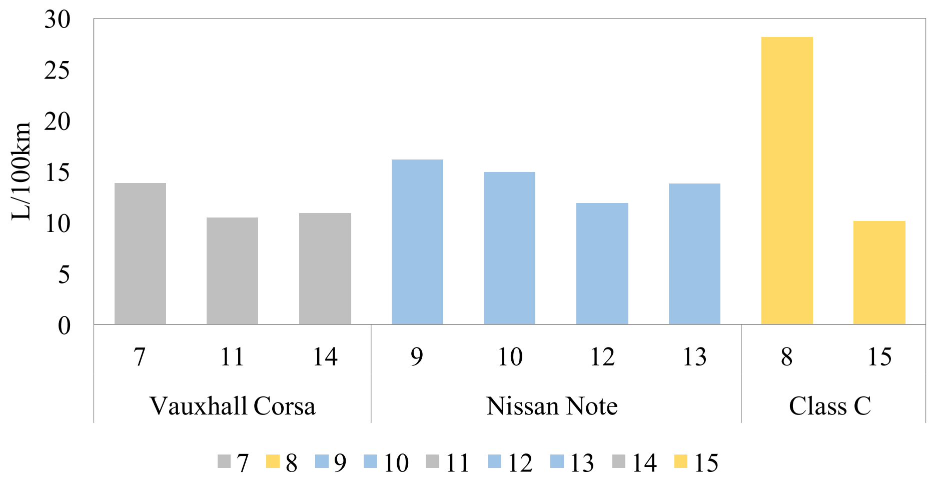


Figure 80. Comparison of drivers' fuel usage in l/100km

2. The amount of consumed fuel distance per unit of fuel (km/l)

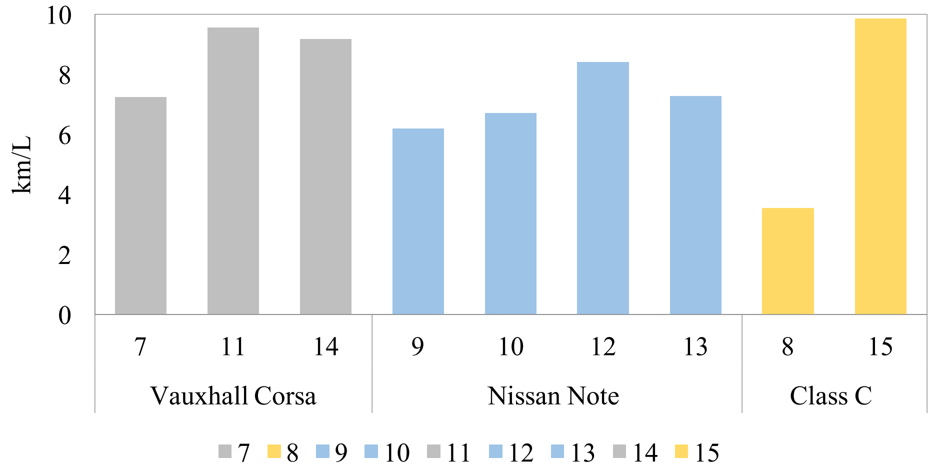


Figure 81. Drivers fuel consumption in km/l

The above metrics, despite being in different units are in agreement with each other in terms of showing the amount of consumed fuel and who were the most fuel efficient drivers among the participants. However, both fail to take into account the difference that might occur if the drivers drove in different driving and road conditions. Whilst both metrics are valid in showing what they aimed to communicated, the problematic part is when a comparison has to be made between two drivers where one drove on hilly roads and the other drove on a flat route. Moreover, these metrics do not take into account parameters such as road gradient or the aerodynamic characteristics of the vehicle and hence, are falling short of comparing drivers on the same footing .

4.2.2.2.2 The results of the Vehicle Specific Power – Fuel Consumption metric

The Vehicle Specific Power – Fuel Consumption metric (VSP – FC) allows for comparison of the drivers on equal grounds since vehicle engine capacity and its characteristics are excluded from the computing model. Instead, parameters including the road gradient, rolling resistance and the vehicle drag coefficient, are used to calculate the rate of fuel consumption. Whilst the comprehensive VSP values were available throughout the driving rounds, this was not the case for instantaneous fuel consumption. Hence, the metric was calculated for all Vauxhall Corsa and Nissan Note drivers by dividing the total amount fuel consumed in (g/s) by the modulus values for VSP for each driver ($VSP - FC = {}^{B_i}/_{|Total\ VSP|}$).

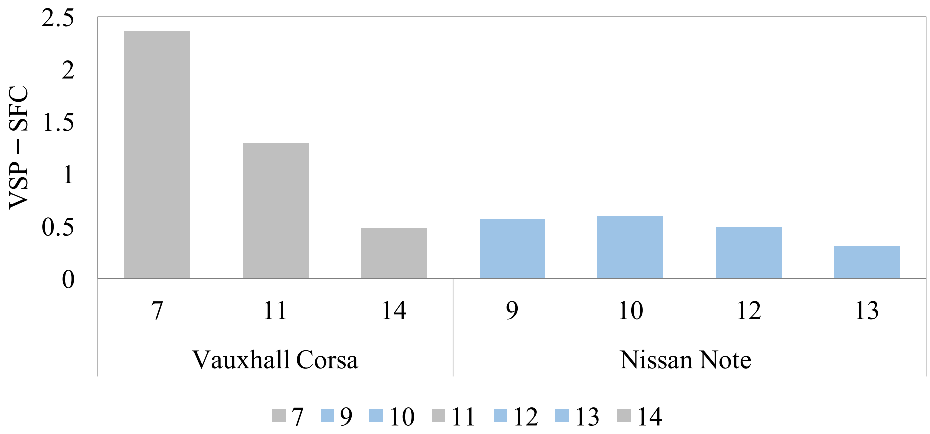


Figure 82. The fuel efficiency of the drivers based on the VSP – FC metric[93]

In the graph above, for each driver the total grams of fuel is obtained at the end of five laps and this total is divided by total time spent driving (s) to calculated the B value. Then, the total modulus value of VSP is computed for each driver and the VSP – FC values are calculated. The VSP-FC value incorporates individual driving behaviour in terms of driver speed profiles and hence, fuel efficiency can be compared regardless of the road gradient and characteristics of the vehicle. For instance, calm drivers, such as drivers 7, 11, 10, scored higher than aggressive ones, who had excessive vehicle and engine speeds, such as driver 14 and driver 13. Table 55 below shows the exact values of the drivers' efficiency ratios.

[93] the Vehicle Specific Power – Fuel Consumption metric $\frac{g_{fuel}/_s}{W/Kg_{vehicle\ mass}}$

Table 55. Drivers VSP – FC values

	ID	VSP – FC
Vauxhall Corsa	7	2.36
	11	1.29
	14	0.48
Nissan Note	9	0.56
	10	0.60
	12	0.49
	13	0.30

4.2.2.3 Results of the drivers' eco performance score according to the fleet management scoring system

The standard scoring practice among fleet management service providers and the usage-based vehicle insurers is to rank drivers based on a set of time interval threshold rules. In order to score the drivers eco-driving performance, first, the total number of times they exhibited sudden acceleration, sudden deceleration and exceeding engine speed (above 4500 rpm) was counted. Then, the eco scores were calculated based on the total number of occurrences of these parameters, with a coefficient being assigned to each of them. The outcomes are provided in the table below. The drivers scores were then arranged in the form of an energy efficiency diagram. (For the methods employed, please see subsection 3.4.4.3). The scores for drivers 8 and 15, the class C drivers, are only included in the table below and not on the score visualisation graphs.

Table 56. Total number of occurrences of dangerous behaviours and eco-driving score

		Number of sudden accelerations	Number of sudden decelerations	Number of incidents of high engine speed	Eco driving score	Alphabetic score
Driver ID	7	0	1	0	47	A
	8	14	3	1	0	D
	9	4	1	1	29	C
	10	1	1	0	43	B
	11	2	1	0	39	B
	12	6	0	0	26	C
	13	6	1	0	23	C
	14	14	5	0	0	D
	15	5	4	0	18	D

From the table above, it can be seen that of the Vauxhall Corsa drivers, driver 7 has the highest score at 47, driver 11 comes second and driver 14 is the worst driver, with zero score. Regarding the Nissan Note drivers, driver 10 is the best, with a score of 43, whilst driver 9 and driver 12 come second and third, with 29 and 26 scores, respectively. Driver 13 is slightly the worst driver amongst the Nissan Note drivers, with a score of 23. What is salient about the above results, is that they show that by the chosen driver scoring method aggressive drivers can clearly be identified (driver 14 and 13), since they scored lower than other drivers. In addition, the calmer drivers with higher scores can also be singled out. However, the method is not always reliable, for example, driver 12 scored quite badly and yet in the earlier studies this driver was considered a calm driver, which raises an inconsistency and hence, brings into question the reliability of this type of threshold rules scoring model.

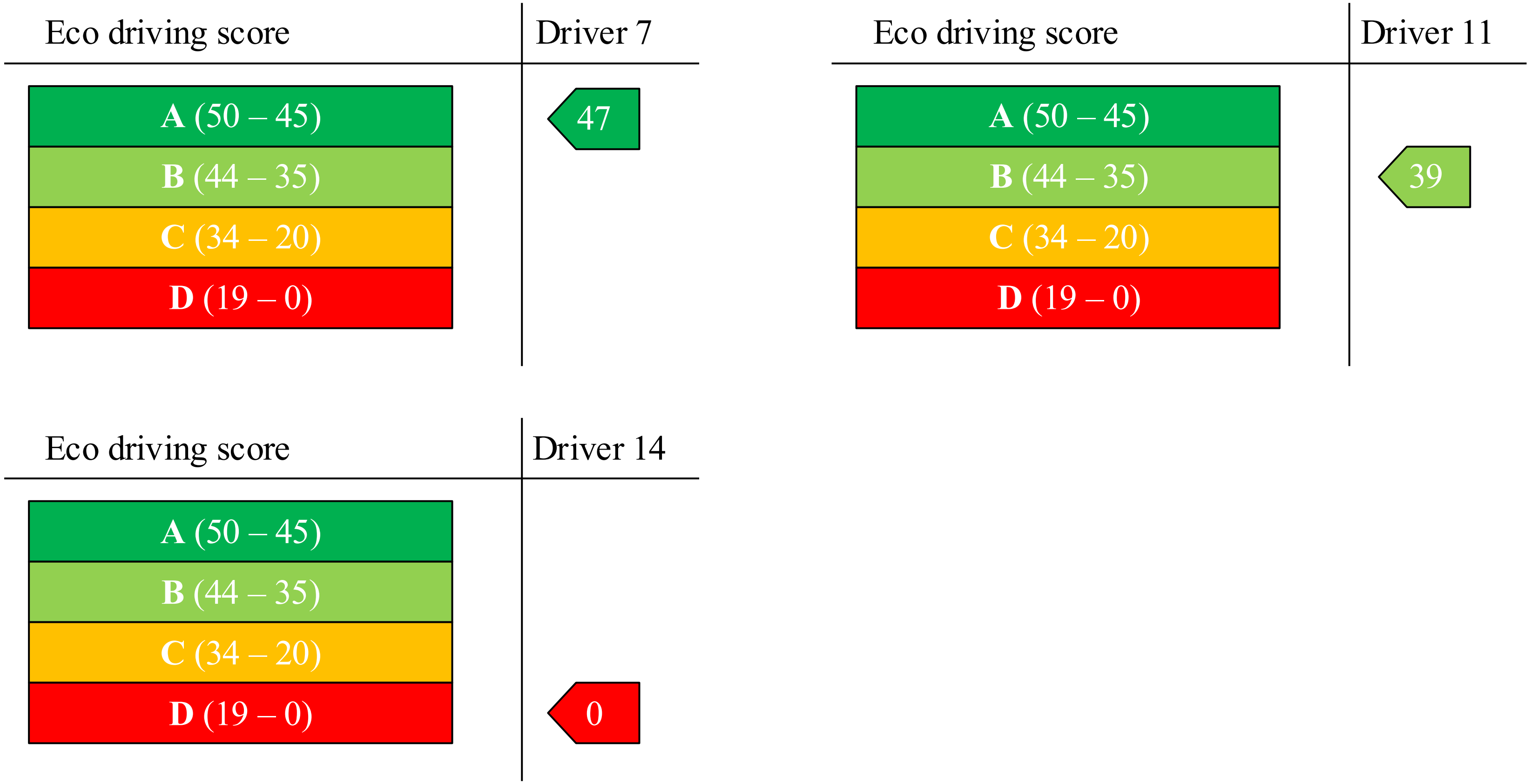


Figure 83. The Vauxhall Corsa drivers' average eco-driving score based on the fleet management method of profiling driver performance

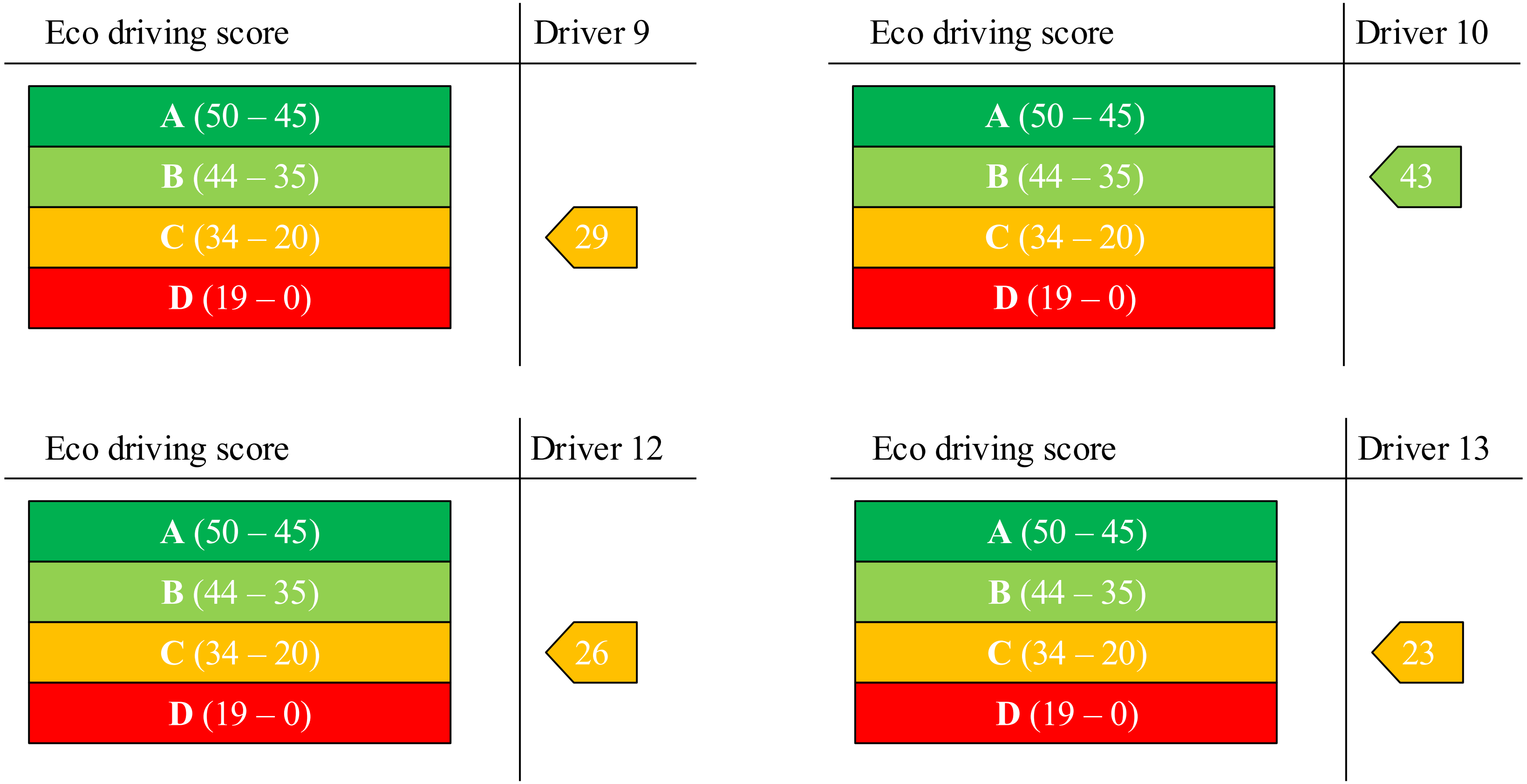


Figure 84. The Nissan Note drivers' average eco-driving score based on the fleet management method of profiling driver's performance

4.2.3 Results of developing a fuel consumption forecasting model and modelling drivers' differences by virtual simulation

In this subsection, the results of the fuel consumption forecasting model and the outcomes of IPG Carmaker virtual simulations of drivers' real driving data are presented. The first study involved using the data collected from vehicles to establish the factors greatly affecting fuel usage and subsequently, constructing a regression model based on those factors to predict the fuel consumption of the tested vehicles. For this study, the IBM SPSS package has been used to build the fuel consumption regression model and outputs. The second study involved using the IPG Carmaker driving simulation software to model and simulate two simulation scenarios. The scenarios were designed and tested to examine the feasibility of using this software package to model variation in driving behaviours.

4.2.3.1 Results of the study on the factors affecting fuel consumption and the construction of a fuel consumption forecasting model

The regression model was constructed in two steps. Firstly, the correlation between all the available parameters was examined, and those highly correlated (at the 1% significance level) parameters were eliminated from the model. Out of all recorded parameters collected with OBD dongles the following 6 parameters were found to be highly correlated with the amount of fuel used by the drivers. The 6 parameters are: average engine coolant temperature, average vehicle speed, average distance, number of sudden accelerations, number of high engine speeds, number of sudden accelerations, number of high engine speeds, and the number of high engine temperature events. The following table shows the parameters were nominated as having a significant effect on fuel usage. An ANOVA (analysis of variance or analysis of means) test was conducted to understand the level of differences amongst all the parameters.

Table 57. Parameters that highly affect fuel usage based on the correlation analysis

Driving and vehicle engine parameters
Average engine coolant temperature (ºC)
Average vehicle speed (km/h)
Average distance (Km)
Number of sudden accelerations
Number of high engine speeds (rpm)
Number of high temperatures

The summarised calculation for the initial regression model (model 1) is tabulated as follows. Model 1 is constructed based on the retained uncorrelated parameters.

Table 58. Model 1 ANOVA analysis values

Model		Sum of Squares	df	Mean Square	F	Sig.[94]
1	Regression	110933.97	6	18488.995	63.192	0.00
	Residual	11410.75	39	292.583		
	Total	122344.72	45			

Table 59. Model 1 summary

Model	R[95]	R Square	Adjusted R Square	Std. Error of the Estimate
1	0.95	0.907	0.892	17.1051

Table 60. Model 1, regression model coefficients

Model		Unstandardised Coefficients		Standardised Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	451.35	41.305		10.927	.000
	Average vehicle speed	-1.143	0.562	-0.109	-2.036	0.049
	Number of sudden accelerations	-0.017	0.630	-0.002	-0.028	0.978
	Number of high temperatures	-0.406	0.618	-0.037	-0.657	0.515
	Average of total distance recorded	-43.186	4.363	-0.568	-9.899	0.00

94 Predictors, i.e. independent variables.
95 Dependent variable: drivers' real fuel consumption.

	Average engine coolant temperature	0.036	0.321	0.006	0.113	0.91
	Number of high engine speeds	-87.345	6.888	-0.722	-12.68	0.00

For the first model (model 1), the R square value was 90.7%, which means that 90% of the variation of the dependent variable can be explained by the independent variables. The aggregate model F value was 63.19 and it was significant at the 5% level, whilst the P value was P <0.0005 for the model. However, the individual coefficients of the independent variables were not all significant and hence, those that were not were eliminated from the model.

Table 61. Parameters removed from model 1

Driving and vehicle engine parameters
Average engine coolant temperature (C)
Number of sudden accelerations
Number of high temperatures

The final linear multiple regression model was developed based on the rest of the parameters, which thus had the following independent variables.

Table 62. Final driving and engine parameters included to predict fuel consumption

Driving and vehicle engine parameters
Average vehicle speed (Km/h)
Average distance recorded (km)
Number of high engine speeds (rpm)

The summarised calculation of the final linear multiple regression model (final model) is tabulated as follows.

Table 63. Final linear multiple regression model: ANOVA analysis values

Model		Sum of Squares	df	Mean Square	F	Sig.
Final	Regression	110760.584	3	36920.195	133.860	0.000
	Residual	11584.135	42	275.813		
	Total	122344.719	45			

Table 64. Final linear multiple regression model summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
Final	0.951	0.905	0.899	16.6076

Table 65. Final linear multiple regression model coefficients

Model			Unstandardised Coefficients		Standardised Coefficients	T	Sig.
			B	Std. Error	Beta		
Final	(Constant)	β_0	461.40	34.715		13.29	0.000
	Average vehicle speed	β_1	-1.12	0.516	-0.107	-2.178	0.035
	Average distance recorded	β_2	-44.14	3.747	-0.581	-11.78	0.000
	Number of high engine speeds	β_3	-87.90	5.772	-0.727	-15.22	0.000

For the final model, the R-square value stayed approximately the same at 90.5% and the aggregate model was significant, with an F value of 133.86 and a P <0.0005. Two parameters were highly significant, namely, average distance and total number of high engine speed occurrences (<1%), whilst average driving speed was so at the 5% level. By looking at standardised coefficient values, it can be seen that all had an adverse impact on the real consumption value. The number of high engine speeds (rpm) had the greatest effect, the average distance (km) was second, and the average vehicle speed (km/h) had the least effect on consumption out of the three parameters.

Table 66. Final linear multiple regression model residual statistics values

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	100.388	282.184	145.50	49.6120	46
Residual	-39.1445	27.0452	0.000	16.0445	46
Std. Predicted Value	-0.909	2.755	0.000	1.000	46
Std. Residual	-2.357	1.628	0.000	0.966	46

The predictive fuel consumption model could now be constructed based on these parameters: average vehicle speed, average distance, and total number of high engine speeds. The linear regression model was constructed based on the followed equation.

Equation 31. The fuel consumption multivariable linear regression model

$$Y = \beta_0 + \beta_1 \times \text{Average vehicle speed} + \beta_2 \times \text{Average distance} + \beta_3 \times \text{Number of high engine speeds}$$

$$Y = 461.40 - 1.12 \times \text{Average vehicle speed} - 44.14 \times \text{Average distance} - 87.80 \times \text{Total number of high engine speeds}$$

First, the model error was tested, and the histogram of the standardised residual of the regression model was plotted. The main objective of plotting the frequency – regression standardised residual chart is to understand whether or not the model residuals are well-behaved or not. That is, to know that the model assumptions are reasonable and the model performs appropriately. The histogram below is a way of visualising whether the residuals are normally distributed or not, which can be interpreted as part of the variation unexplained by the model. As can be observed, the graph the distribution is not normal, which is because the sample size is small and hence, methods such as P-P plot or PRESS are recommended (Tsay, 2010; *NIST/SEMATECH e-Handbook of Statistical Methods*, 2012; Tabachnick and Fidell, 2012).

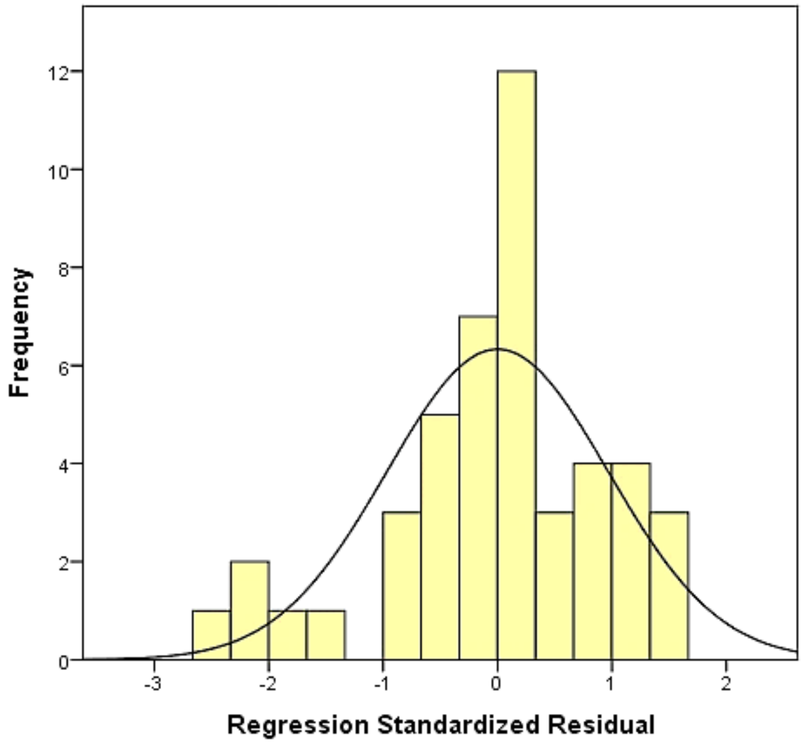


Figure 85. The histogram of the regression model

Rather than computing the model error by PRESS (Predicted Residual Error Sum of Squares) the P-P plot has been used to explain the predictability of the model. The P-P plot of the error was plotted to check how closely the model (expected cumulative probability, y-axis) agreed with the observed data (cumulative probability, x-axis), and it can be seen that the deviation of its outcomes from the real fuel consumption line is subtly different. Hence, the results of the regression show that model is able to predict fuel consumption by parameters other than through direct reading.

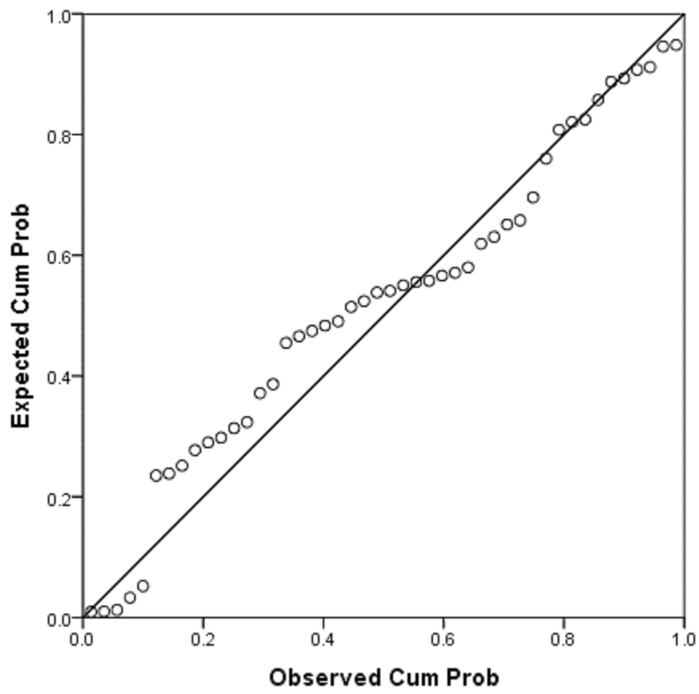


Figure 86. The P-P plot of the regression model

4.2.3.2 Modelling variations of driving styles using virtual driving simulator software

In this subsection, the outcomes from modelling and comparing drivers' driving performances using the IPG Carmaker virtual driving simulation software package are presented. The following two simulation scenarios were designed based on the Vauxhall Corsa drivers' road speeds and vehicle characteristics.

4.2.3.2.1 Scenario I: Results of the study modelling and comparing Vauxhall Corsa drivers' eco-driving performances

The road speeds of the first, third and final round of driving of the Vauxhall Corsa drivers were selected and assigned to the simulation model as the inputs for driving manoeuvres. In total, four simulations were conducted for the first scenario. The Brake Specific Fuel Consumption (BSFC) contour maps were plotted to display, firstly, the capability of the Carmaker software package to deliver distinguishable results from only variations in vehicle speeds. Secondly, the aim was to compare whether the drivers' eco-driving scores were an accurate representation of their performance and finally, this was to confirm that vehicle speeds are a useful metric for distinguishing driving behaviour differences.

Table 67. List of the simulation specifications for Scenario I

ID	Route condition	Vehicle Model	Manoeuvre input to the model
1 to 3	With elevation	Vauxhall Corsa	Corsa drivers' fifth trial road speeds
4	With elevation	Vauxhall Corsa	Ideal driver, 30 mph speed limit max

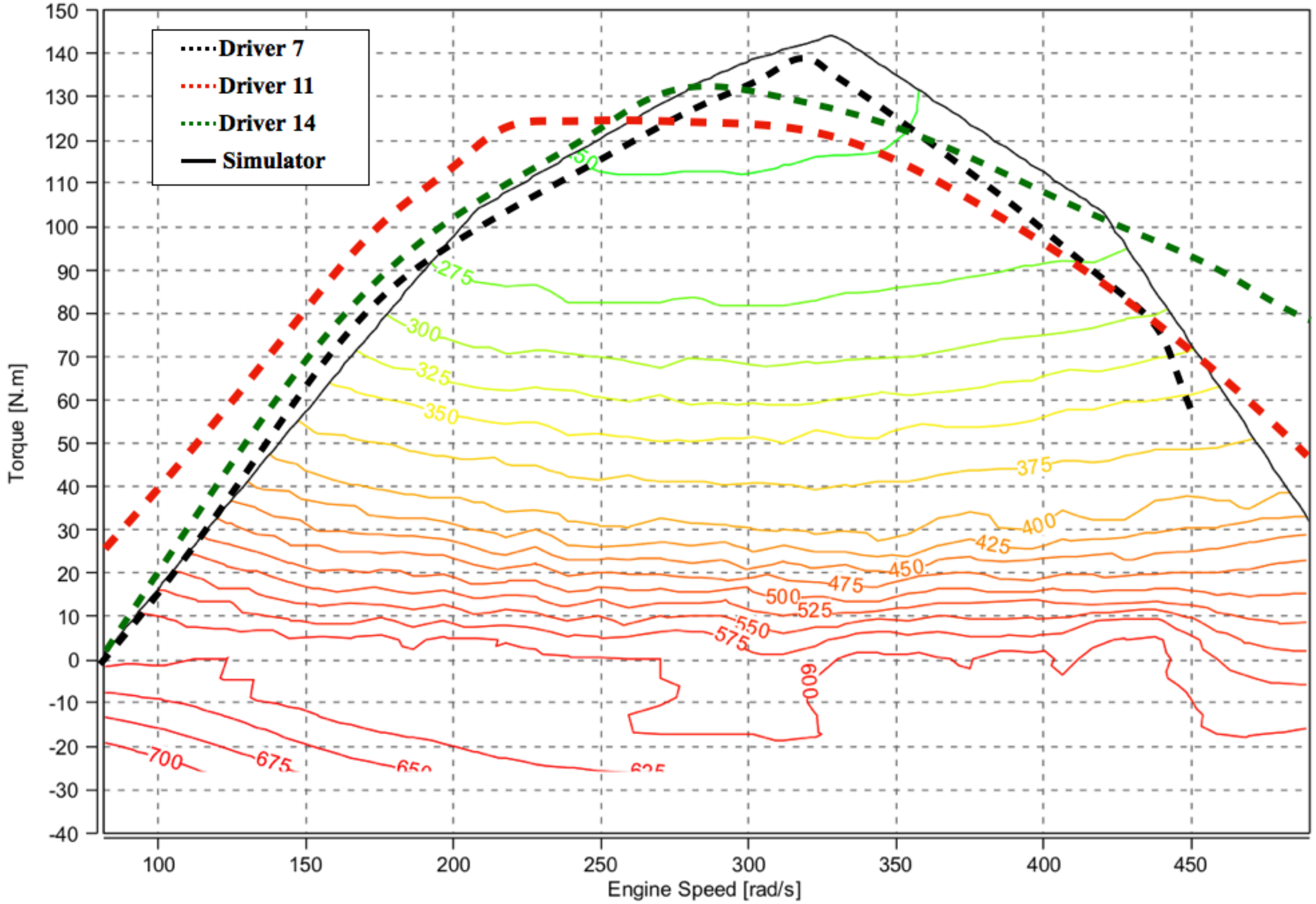


Figure 87. The Brake Specific Fuel Consumption (BSFC) contour maps of three Vauxhall Corsa drivers, Driver 7 (top left), Driver 11 (top right), Driver 14 (bottom left) and a simulator (bottom right)

- Based on drivers' actual speed, with the simulator assigned to keep the speed constant at 30 mph. (Results were plotted using AVL concerto)

By over lapping drivers' contour graphs on the simulator, this sows how well each driver drove in comparison with the ideal driver presented as simulator driver. In the chart above, the black line show the simulator driver, with the dashed lines from driver 7, driver 11 and driver 14 being, black red and green, respectively. The two objectives of conducting this analysis were satisfied as the simulation confirmed that the simulation model is, in fact, capable of delivering distinguishable outcomes by only using drivers' vehicle speeds. Secondly, it can be seen that based on the classification results, driver 7 scored the highest, driver 11 came second and driver 14 was the worst driver. The simulation results, moreover, show that driver 7 spent a significant amount of time in the efficient region of the vehicle engine performance map, while driver 14 drove inefficiently for much of the time. The final consideration is in relation to the consumption map of the simulator, which was assumed to be the perfect performance of the vehicle calculated by simulation using the IPGDriver module. As has been explained in Chapter 3, the results of the engine efficiency values of the drivers and simulator were calculated and these can be seen in the table below. The results show that drivers 7 and 11 with, 16.82% and 16.81%, respectively, were very close in terms engine efficacy and had high eco-driving scores. It is important to highlight that engine's efficiency values in the table below are entirely computed from output data of the simulator model and the eco-driving score is based on drivers' driving data. Interestingly, drivers who were shown as being efficient in the computer model were score highly through the eco-driving score calculated in subsection 4.2.2.3.

Table 68. The results of drivers' efficiency values based on average brake specific fuel consumption

Test ID	The engine 's efficiency	Eco-driving score
Driver 7	16.82%	47
Driver 11	16.81%	39
Driver 14	17.69%	0
Simulator	18.18%	-

Hence, the efficiency values are consistent with the fleet scoring method presented in subsection 4.2.2.3. It can be concluded that the fleet scoring method is robust in identifying outliers (Driver 14), and accurate when it comes to ranking the most efficient drivers, as can be seen from the simulator software outcomes. Finally, the findings of this study also confirm the final objective of this test scenario, whereby the modelling of different driving performances by only using drivers' speed has been demonstrated as being achievable by using Carmaker and hence, this method is valid for modelling a range of aspects regarding driver behaviour.

4.2.3.2.2 Scenario II: Results of the study modelling and comparing the effect of road slope on drivers' fuel usage

In this study, the Vauxhall Corsa drivers' fifth lap driving data was used as the manoeuvres input for the simulation model. The model was then simulated under two different types of road conditions, first, without any elevation and the second, it was run with a digitalised road that matched the elevation profile of the actual route. The drivers' instantaneous and absolute fuel consumption were recorded from these simulations. This study was aimed at checking the capability of the IPG's Carmaker software of simulating roads without slope.

Table 69. List of simulation specifications for Scenario II

	Route condition	Vehicle type	Manoeuvre input for the model
1	Without elevation	Vauxhall Corsa VX	Driver 7 final driving speed data
2	Without elevation	Vauxhall Corsa VX	Driver 11 final lap driving speed data
3	Without elevation	Vauxhall Corsa VX	Driver 14 final lap driving speed data
4	Without elevation	Vauxhall Corsa VX	Simulator, with 30 mph speed limit
5	With elevation	Vauxhall Corsa VX	Driver 7 final lap driving speed data
6	With elevation	Vauxhall Corsa VX	Driver 11 final lap driving speed data
7	With elevation	Vauxhall Corsa VX	Driver 14 final lap driving speed data
8	With elevation	Vauxhall Corsa VX	Simulator, with a 30 mph speed limit

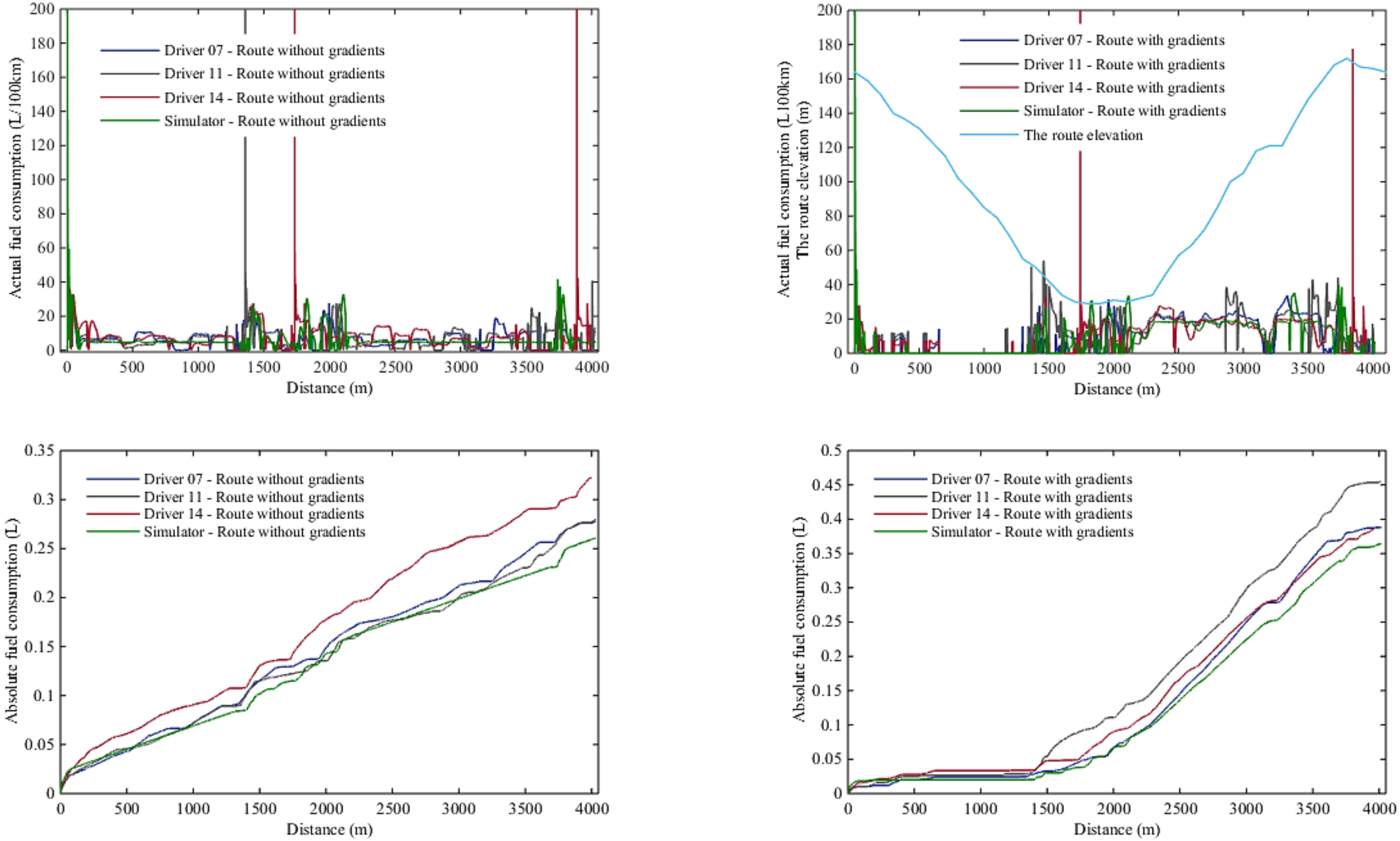


Figure 88. Drivers' actual instantaneous and absolute fuel consumption results for simulation on a road without elevation (left) and with elevation (right)

Since every parameter was kept the same for all the simulations, the effect of road elevation and drivers' vehicle speeds on fuel consumption could be investigated. As it can be seen from the right-hand side graphs, uphill driving increases fuel consumption dramatically. The simulation results show that the ideal driving behaviour created by the simulator, under both test conditions (road with and without elevation), uses less total fuel than the real drivers did, according to their vehicle speed profiles.

In order to find the impact of road slope on fuel consumption, the drivers and simulator's fuel usage for both test conditions (without elevation and with elevation), i.e. those sections of the routes downhill (distance from start to 1,800 metres) and uphill (distance from 2,200 metres to the end)) were compared. As expected, drivers used less fuel during downhill driving (when the elevation is included in the model), compared to when the simulation was running on a flat road. The important observation is the effect that uphill driving has on fuel usage. For every simulation (including drivers and the simulator), more than 40% increase in fuel usage is recorded, according to Carmaker (Fig 45). In the case of flat road simulations, it can be observed that despite the fact that the digitalised road does not have any elevation to it, the results show that drivers used more fuel on the hypothetical downhill section of road than the uphill one. To explain this outcome, it should be noted that regarding the first section (downhill part) on the actual road and on the digitalised road this is very close to a straight route, hence drivers travelled at a higher speed than had there been bends in the road. Since their actual speed was used in the software, Carmaker generated the consumption correspondence to their speed (based on the powertrain information provided to the software). Regarding the simulation results, the following table shows the increase in fuel usage as the route setting changes from being a flat road to the actual real-world elevation profile.

Table 70. The effect of uphill driving on fuel usage based on Carmaker simulation results

ID	Percentage of fuel usage that is increased
Driver 7	50%
Driver 11	41%
Driver 14	67%
Simulator	49%

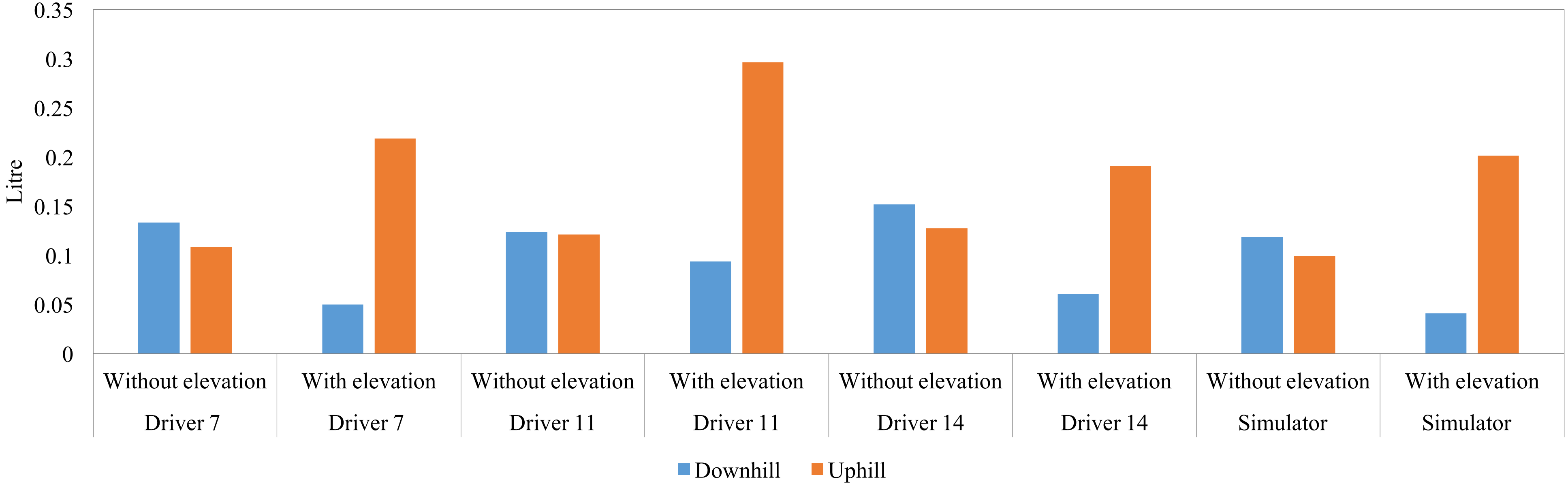


Figure 89. The Carmaker downhill and uphill fuel consumption simulation results based on the real-world speed data of the Vauxhall Corsa drivers and the simulator driver

This study confirms three important points: firstly, based on the virtual simulation, the road gradient can affect drivers' fuel efficiency dramatically. Secondly, in the case of the simulator driver, when the software was set to maintain the speed limit, it was not representative of real-world driving or the most efficient driving results, which confirms the need to have better representations of real-world driving in virtual driving simulator software. The final point is that Carmaker confirms that using drivers' actual speed in the model can provide valuable insight into driver differences even when speed is the only parameter that varies.

4.2.3.3 Summary of the modelling methods

In this subsection, the results of two approaches to modelling driving data have been presented and discussed. The linear regression model in the first part, presented a metric for predicting drivers' fuel consumption based on their vehicle speed, distance travelled and the total number of high engine speed incidents. The model can be used to predict fuel consumption with parameters other than direct measurement and this can then be used to compare across drivers or to predict the fleet fuel consumption amount.

The second approach to modelling driving data has been to use the real world driving data to simulate virtual driving scenarios. The two presented scenarios (see 4.2.3.2.1 and 4.2.3.2.2) confirm the effectiveness and accuracy of the Carmaker driving simulation software, whereby it has shown that the simulation model has the ability to distinguish drivers' differences only using their vehicle speed profiles. Regarding the first study, drivers' efficiency was compared by using their actual speed profiles. The results show that in the event where every setting of the simulation remained constant, the efficient drivers could be determined by using just the variation in driving speed between drivers. The second scenario involved incorporating road elevation into the simulation, which made it possible to uncover the effect that gradient has on fuel consumption. The results of the simulations confirm that with a 16% increase in the road gradient, on average, a 41% increase in fuel consumption can be observed. This conclusion is drawn based on the computer model that was constructed to replicate the exact conditions of driving (road conditions, vehicle specifications) and the computerised driver constructed based on drivers' real-world speed using the manoeuvre settings in the IPGDriver module.

4.3 Result of the studies relating to drivers' driving safety

In this section, the findings of the studies conducted to examine the drivers' perception of safe driving are presented. Similar to the studies on driver performance in section 4.2, this investigation was focused on drivers' driving abilities in terms of their attitude to road safety based on their driving data. This included identifying unsafe behaviours and habits and classifying drivers in terms of whether they were driving safely or not. Since the safety of drivers is not determined by the make of vehicle, the driving data for all of the participant drivers was included in the following studies. To recall, the driving information was from those drivers who participated in the driving event on 9th February 2014.

4.3.1 Results of identifying drivers' safety perception by using vehicle engine data

This subsection presents the results of four studies, each of which was aimed at determining aspects of the drivers' perception towards road safety and risky driving behaviour. The following is the list of studies covered in this subsection:

1. Study on drivers' vehicle speeds prior to approaching to various road settings;
2. Study on the total number of incidents and locations of hard acceleration/deceleration by the drivers;
3. Study in identifying drivers with coasting downhill habits;
4. Identifying collision-prone road segments on the testing route based on historical collision data.

4.3.1.1 Results identifying drivers with a tendency to drive over the speed limit in urban road settings

The interest here was in the identification of the level of unsafe driving behaviour of all the drivers at locations where caution is required. The total number of driving trials was 46, and so the outcomes are reported in terms the percentage of the number of times that a particular behaviour was observed. Specifically, drivers approaching speed was examined when entering the main road, approaching the roundabout and finally, prior to the pedestrian crossing on Bathwick Hill. Table 71 below, shows these locations as distances from the starting point and the criteria, according which the drivers' vehicle speeds were examined.

Table 71. Locations, where drivers' safe driving was compared and each site's testing criteria

Location	Distance	Traffic setting	Algorithm criteria
End of North Road	1.4 km	Entering the main road	$0 \le \text{Speed} \le 5\text{mph}$
Bathwick Hill	2.1 km	The roundabout	$\text{Speed} \le 20$ mph
			$\text{Speed} > 30$ mph
39 Bathwick Hill	2.4 km	The pedestrians crossing	$\text{Speed} \le 20$ mph
			$\text{Speed} > 30$ mph

Table 72. The percentage of times the drivers approached the focal locations with a safe driving speed

Distance	Traffic setting	Algorithm criteria	Percentage
1.4 km	Entering the main road	$0 \le \text{Speed} \le 5\text{mph}$	22%
2.1 km	The roundabout	$\text{Speed} \le 20$ mph	22%
		$\text{Speed} > 30$ mph	10%
2.4 km	The pedestrians crossing	$\text{Speed} \le 20$ mph	13%
		$\text{Speed} > 30$ mph	28%

As presented in the table above, the percentage of times drivers made a full stop at the end of North Road, prior to entering the main road was 22%, which indicates that their perception about stopping at this location was poor. Similar analysis regarding the drivers reaching the

Bathwick Hill roundabout and the pedestrian crossing show that speeds of less than 20 mph occurred only 22% and 13% of the time, respectively. What is more alarming is that 10% of the time the roundabout was approached at more than 30 mph and in relation to the pedestrian crossing, this figure was much higher, standing at 28%. Overall, the reported results indicate that most of the participants had a poor attitude to safety, as judged by their speeds when approaching certain road settings.

4.3.1.2 Study of the drivers' hard acceleration and deceleration behaviours

In this subsection, the results in respect of the drivers' harsh acceleration and deceleration behaviour, in terms of the number of occurrences each lap at certain locations, are presented. This was addressed in two ways, firstly, the number of times and on which driving laps the drivers used the accelerator and brake pedals sharply. Secondly, this was achieved by making a comparison between the locations of these occasions and those with actual vehicle collisions in the past.

4.3.1.2.1 Results of the study on repetition of harsh acceleration and deceleration behaviour

The incidents of using the accelerator and brake pedals harshly can be indicators of the drivers' perceptions towards road safety. The total number of times such behaviours happened per lap was recorded and then put in graphical form, as shown in the graph below. As can be observed, there is an overall ascending trend in the total number of times the drivers performed sudden acceleration/deceleration. To visualise the overall trend of these two alarms, the logarithm lines were fitted to the samples.

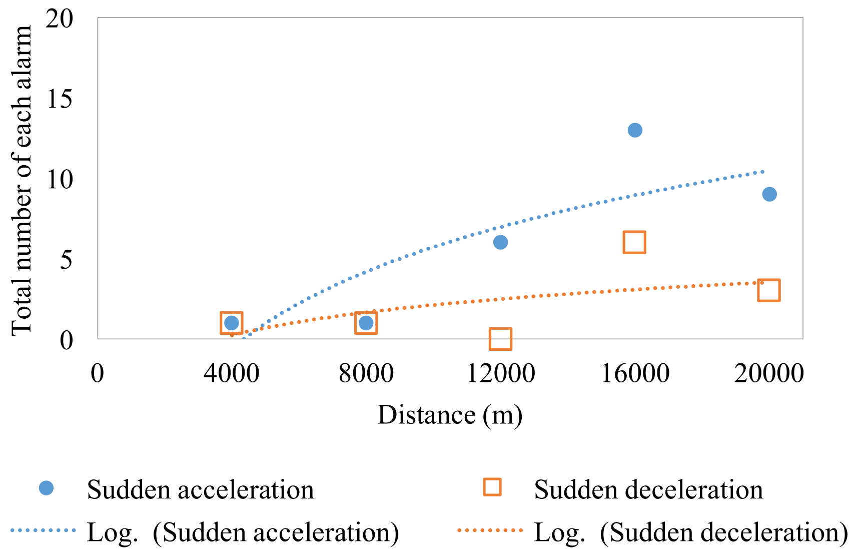


Figure 90. The overall trend of total number of harsh acceleration and deceleration occurrences per lap (distance travelled)

Since the total number of both sudden accelerations and decelerations was not significant, the following analysis is conducted by combining all the events from all the participants. As table below presents on the first lap there had been only one sudden acceleration and one sudden deceleration. The exact reason behind these trends is difficult to explain, but it could be that the drivers' fatigue or being bored of driving the same route five times contributed to their more aggressive usage of the acceleration pedal and engaging in harsh braking towards the end of the driving task. In the table below, it can be seen that the total number of sudden accelerations or decelerations started low for the first two laps rising to 19 for the fourth lap and falling back to 12 for the fifth.

Table 73. The drivers total number of sudden accelerating and decelerating behaviours[96]

	Lap number | Distance	1 | 4 km	2 | 8 km	3 | 12 km	4 | 16 km	5 | 20 km
Alarms	Sudden acceleration	1	1	6	13	9
	Sudden deceleration	1	1	0	6	3

[96] Table is organised based on driving lap and distance driven.

4.3.1.2.2 Results of the spatial interpolation of locations with a high number of dangerous behaviours and collision-prone history (pre-collision study)

It was deemed appropriate to investigate the importance of these two behaviours, namely, sudden acceleration/deceleration, in more depth. This study involved investigating the correlation between the locations that drivers made sudden acceleration/deceleration and the total number of recorded accidents at those locations. To this end, the driving route was divided into 10 segments, each being 400 metres in length. The total number of collisions that happened between 2005 and 2013 for each segment was extracted from the Department for Transport published database (Graves *et al.*, 2014). The total number of sharp accelerations and decelerations for each segment were also allocated to each section. The following two maps were created based on, first, the total number of times collisions happened for each segment of the road (Figure 91) and second, and the total number of times the drivers made harsh acceleration/deceleration in the same segments (Figure 92). Road segments are to scale, but the double arrow guides are not.

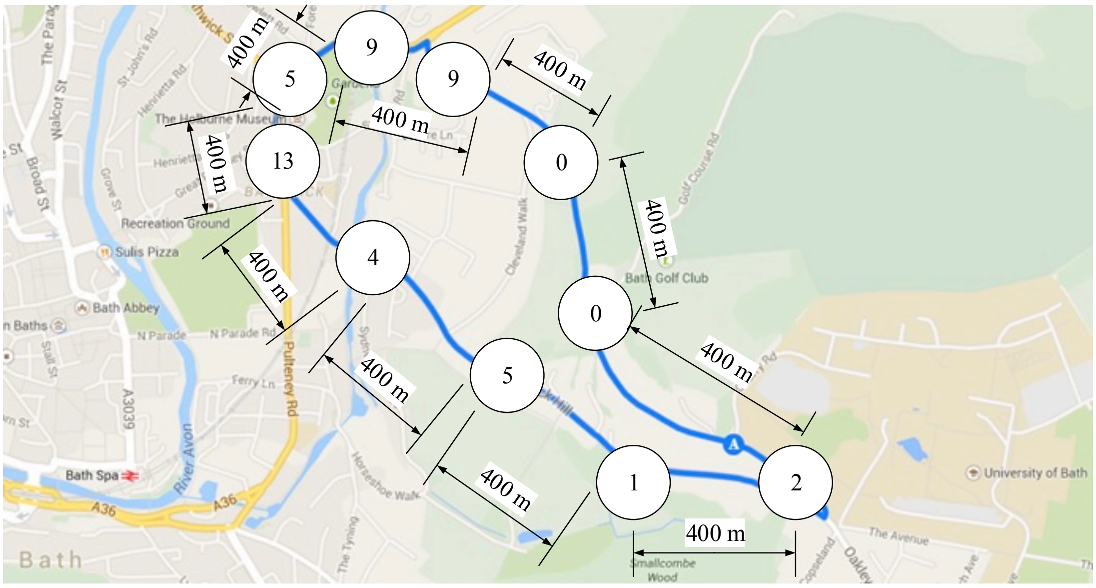


Figure 91. Total number of recorded collisions in each segments between 2005 and 2013

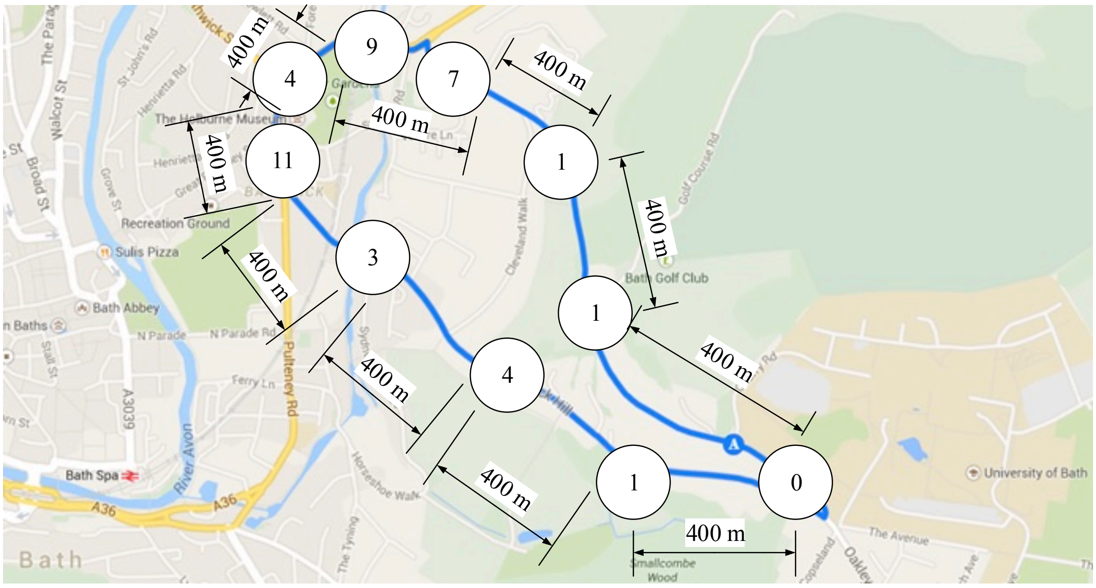


Figure 92. Total number of drivers' harsh acceleration or deceleration incidents based on their location

Finally, the level of association between these two datasets, (places with previous collision histories and locations where drivers made harsh acceleration/deceleration) is investigated. The measure of correlation (correlation coefficient) was calculated using three methods, the Pearson, Spearman and Kendall methods. The results are presented below. As it has been emphasised in chapter 3, in this section the aim is to explore potential and meaningful correlation between the two aforementioned datasets. Hence, for the exploratory data analysis (EDA) the decision was taken to compare all the appropriate correlation analysis methods.

Table 74. Results of correlation tests under the Pearson, Spearman, and Kendall correlation methods

Method	Correlation target	Standard error	Target	P-Value	T-Stat	Critical value
Pearson	97.0%	8.5%	75.0%	1.6%	2.58	2.31
Spearman	91.3%	33.3%	75.0%	31.2%	0.49	1.96
Kendall	81.9%	26.9%	75.0%	39.8%	0.26	1.96

According to the Pearson method, a 75% target figure achieved a correlation of 97%, thus showing a fundamental association between the locations where drivers made harsh accelerations/decelerations and previous collision history. The two other methods (Spearman and Kendall) also reported close links between the two parameters, but not as strong as with the Pearson method. As these data sets are completely independent, the result that the simultaneous occurrence of the sudden changes in speed and collisions occurred 75% of the time most likely establishes the importance of the latter in identifying risky behaviour. This indicates that from only 46 trials and using the collision data for recorded accidents over an eight-year period an association between them can be found, which can lead to risky behaviour being identified. This outcome not only confirms that the drivers' sudden acceleration/deceleration is a major factor regarding their safety level, for it also provides a new method of finding collision-prone locations in terms of their prioritisation for road improvements.

4.3.1.3 Results of identifying drivers with a coasting downhill habit

As has been explained in previous subsections, coasting downhill can be dangerous for drivers. Identifying those with a tendency to engage in such a behaviour can provide a new metric for classifying drivers' safety levels. Based on the route specific threshold rule developed for the participating vehicles, those drivers who exhibited hypothetical coasting downhill behaviour were detected. The results are provided in the following table.

Table 75. Drivers' total number of hypothetical coasting downhill incidents based on the algorithm

	ID	Coasting downhill incidents
Vauxhall Corsa	7	3
	11	0
	14	1
Nissan Note	9	1
	10	0
	12	0
	13	1

(see subsection 3.4.7.3)

The outcomes suggest that driver 7 coasted downhill the highest number of times and to ratify whether this behaviour actually took place, the speed – distance and engine speed – distance plot from a previous study (study 4.2.1) is presented again below (Figure 93). The actual coasting downhill occasions can be spotted by comparing the drivers' speeds in conjunction with their vehicle speed and engine speed, whereby if a driver was coasting downhill, there are identifiable gaps between these two lines. These occasions are highlighted on the graph with the blue dashed lines. This would appear to be an effective method for identifying drivers with a coasting downhill habit and later, this method is used to improve the assessment of drivers' safe driving scores.

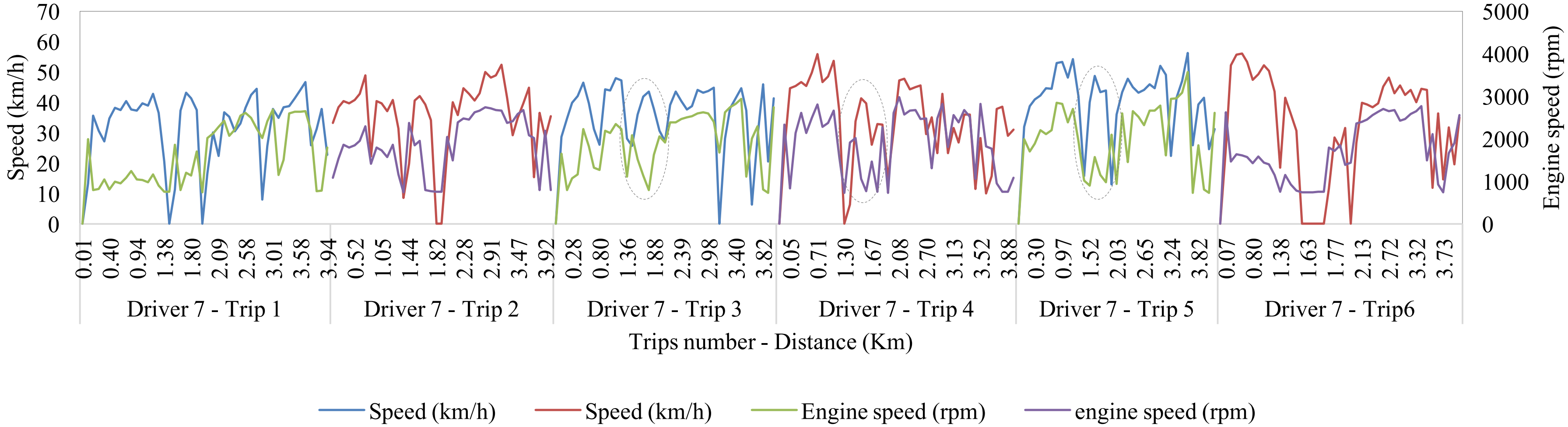


Figure 93. Driver 7, speed - distance and engine speed - distance graph: the coasting downhill incidents can be spotted in the graph

- Dashed lines are occasions where the driver coasted downhill.

4.3.1.4 Result of identifying collision-prone road segments on the testing route

In order to determine collision-prone segments on the driving route (see 2.9 and 3.4.6.4), firstly, descriptive analysis was conducted on the historical data, which confirmed that the sample was, in fact, under-dispersed (sample-variance > sample mean). Moreover, as two of the segments (1, 2) have had no collisions reported, (zero observation), this, along with the under-dispersion would indicate that the results are biased. Hence, the identification of hotspots using the above method would not be without error. However, the aim here is not to predict where accidents will occur, but rather, to add parameters to the fleet driving score approach that will lead to its improvement and hence, these errors are not of great importance. As with the study in 4.3.1.2.2, the roads of the driving route were segmented into 10 sections of 400 metres in length and the total number of collisions in each between 2005 and 2013 was assigned to each section. Since the value of C_V is less than 1.0, it is possible to investigate the collision-prone locations by using the statistical confidence intervals method (see 2.9 and 3.4.6.4) The results of the descriptive analysis and the statistical confidence intervals method are provided in Table 76 below.

Table 76. Descriptive analysis of the historical collision data and results of the crash-prone analysis

Parameters	Values	Segment	Number of actual collisions in the past	Zone status
Variance	17.16	1	0	-
Average	4.80	2	0	-
Standard deviation	4.37	3	9	Collision-prone
Skew	0.67	4	9	Collision-prone
Median	4.50	5	5	-
Min	0.00	6	13	Collision-prone
Max	13.00	7	4	-

Parameters	**Values**	**Segment**	**Number of actual collisions in the past**	**Zone status**
Q 1	1.25	8	5	-
Q 3	8.00	9	1	-
C_V	0.91	10	2	-

The figure below shows the collision-prone zones on the route. The numbers represent the segment numbers, and the highlighted ones (segments 3, 4, and 6) are collision-prone zones, according to the statistical confidence intervals approach. The segments are marked with dimension guides that are not to scale. It is standard practice that road safety engineers, when identifying previously unknown hot spots, eliminate those they already know about when performing tests on suspected hazardous locations, which includes deleting junctions and roundabouts from the analysis. Segment number 6 covered the roundabout, i.e. a specific road feature and hence, was eliminated from the collision-prone zone in the subsequent modelling, whilst 3 and 4 just covered stretches of road, being thus retained.

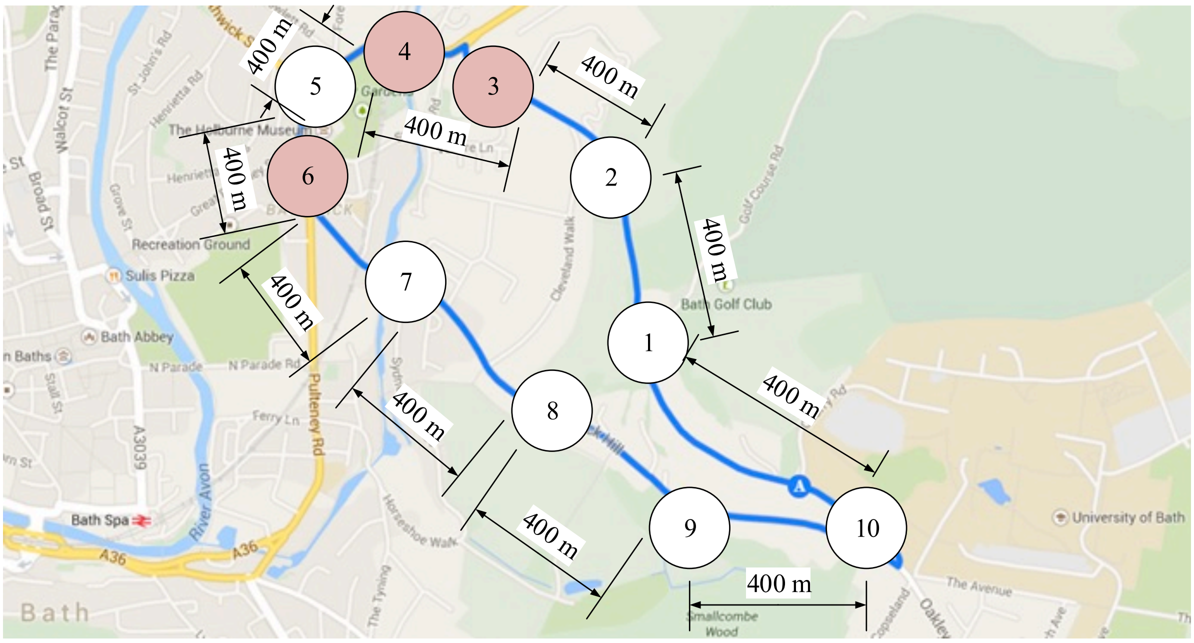


Figure 94. The results from identifying the collision-prone locations of the driving route: numbers represent road segments

4.3.2 Results of classifying and scoring drivers' safety differences from vehicle primary data

In this subsection, drivers' safety scores based on a standard scoring method are presented. This method is employed by usage-based insurers and fleet management service providers to score drivers in terms of safety. The second part of this subsection provides the results of a modified scoring model that includes drivers' coasting downhill behaviour (4.3.1.3) and the locations where they made sudden acceleration or deceleration (4.3.1.4). It is contended that these changes make the standard scoring model one that is more sophisticated, in particular, because it is location sensitive.

4.3.2.1 Results of drivers' safety score according to the fleet management scoring system

In order to score the safety aspects of drivers, first, the total number of times the drivers exhibited sudden acceleration/deceleration or exceeded the engine speed (above 4,500 rpm) were counted. Using these data, the safety scores were calculated, and the results are presented in the table below.

Table 77. Total number of occurrences of dangerous behaviours, drivers' safety ratings, and their alphabetic classification

		Number of sudden accelerations	Number of sudden decelerations	Number of high engine speeds	Drivers' safety score under the fleet management system	Alphabetic score
Driver ID	7	0	1	0	46	A
	8	14	3	1	0	D
	9	4	1	1	28	C
	10	1	1	0	43	B
	11	2	1	0	38	B
	12	6	0	0	25	C

		Number of sudden accelerations	**Number of sudden decelerations**	**Number of high engine speeds**	**Drivers' safety score under the fleet management system**	**Alphabet ic score**
	13	6	1	0	22	C
	14	14	5	0	0	D
	15	5	4	0	17	D

According to this method, driver 7 was the safest driver, while driver 14 was the most unsafe. The only difference between this approach and the ranking model commonly used to score drivers' eco performance is that the coefficients assigned to each of the alarms are different. Given the outcomes of some of the previous studies in this chapter that have proven there are driving conditions and situations that this scoring method fails to take into account, it is contended that it is inadequate for fulfilling this task effectively.

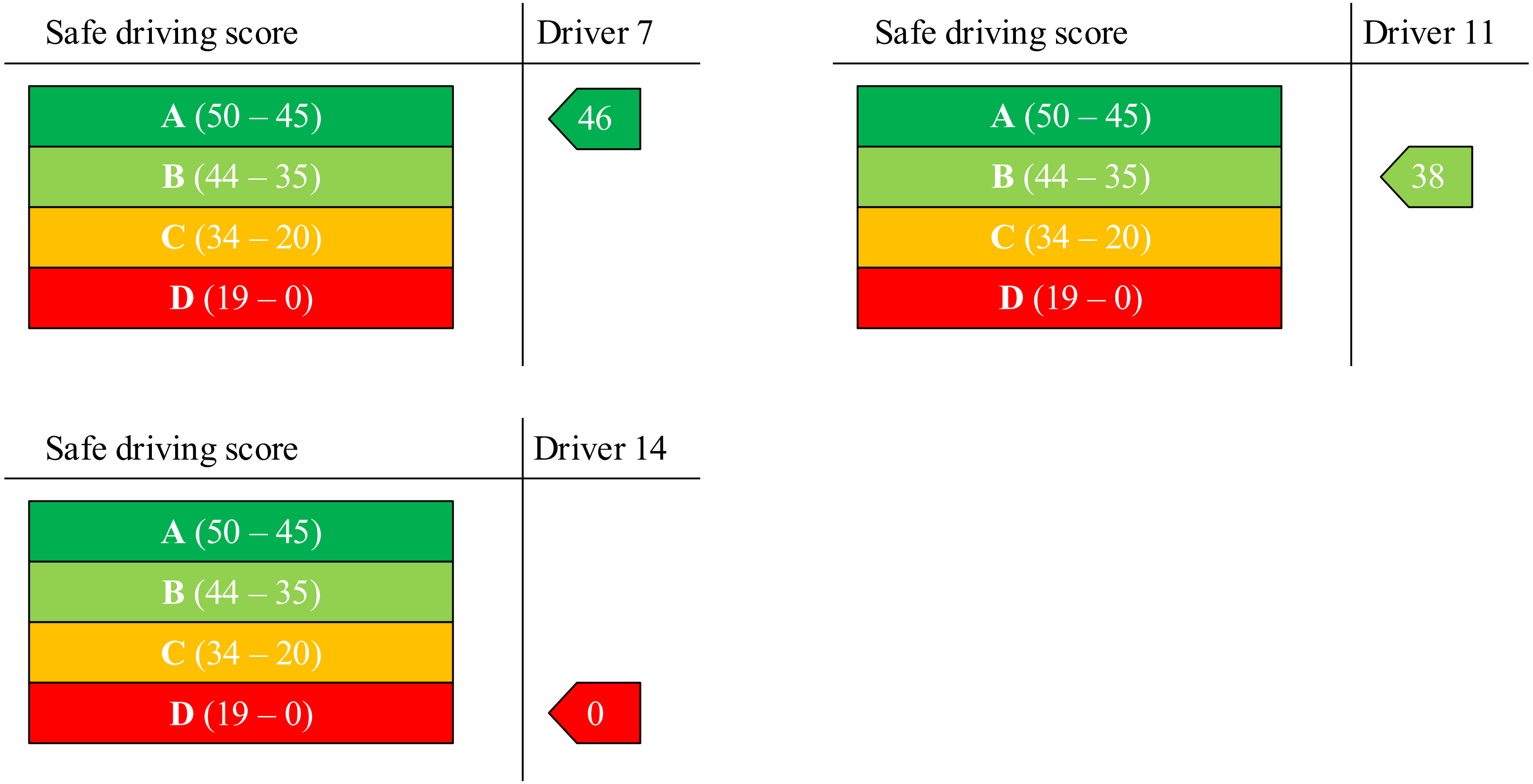


Figure 95. Vauxhall Corsa drivers' safe driving score based on the fleet management method of profiling their safety

Figure 96. Nissan Note drivers' safe driving score based on the fleet management method of profiling their safety

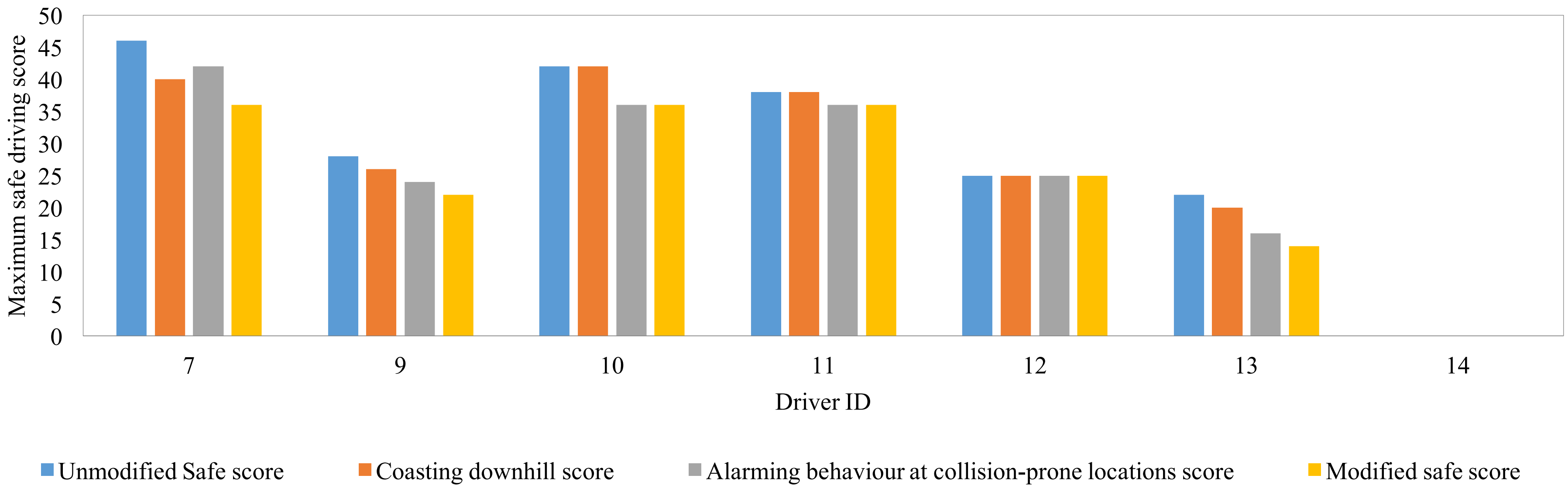
4.3.2.2 Results of modification of the fleet management drivers' safety scoring method

This subsection presents the results of two modifications aimed at improving the capacity of the fleet management scoring system so as to reflect real driving behaviour. As has been discussed previously, there are two aspects of driving behaviour, amongst others, that the existing scoring method does not address. The first matter is with regards to including the number of occasions when drivers were coasting downhill and based on the results of the study in 4.3.1.3, the negative effect of this was incorporated into the scoring system. The second modification to the existing scoring method draws on the fact that there is a high correlation between the locations of previous accidents and those where drivers made sudden accelerations/decelerations. Segments 3 and 4, as explained above in subsection 4.3.1.4, were taken as being stretches of road that represented hotspots in terms of collisions and hence, such behaviour by drivers at these locations was included in the scoring process. That is, whilst the total number of sudden acceleration and deceleration occasions of drivers was taken into account for all segments for all the Vauxhall Corsa and Nissan Note drivers, this was more severely penalised for hot spots. The drivers' scores were updated by building in the negative effect that these two behaviours had on their driving scores, individually, as explained in subsection 3.4.7.2. The final action was to combine the impact of both of these behaviours to arrive at a new safe driving score. In sum, the proposed model presents a driver safe driving score that includes the negative impacts of coasting downhill and making sudden accelerations or decelerations at collision-prone locations.

Table 78. Drivers safe driving score based on the standard method and the new safe driving score based on the modifications made to the scoring method

ID	Unmodified Safe score	Number of coasting downhill incidents	Number of sudden accelerations or decelerations at collision-prone locations	Coasting downhill score	Alarming behaviour at collision-prone locations score	Modified safe score
7	46	2	2	40	42	36
9	28	2	2	26	24	22
10	42	1	3	42	36	36
11	38	0	1	38	36	36
12	25	0	0	25	25	25
13	22	1	3	20	16	14
14	0	4	0	0	0	0

Figure 97. Drivers' safe driving score improvement from unmodified to modified safe driving scores



Chapter 5

Discussion

“Pick a flower on Earth and you move the fartherst star.”

– Paul Dirac

5.1 Introduction

This section is dedicated to providing further explanation of the analysis conducted in this study. The aim here is link the work that has been undertaken with the five questions posed in Chapter 1.

1. What driving parameters (e.g. vehicle speed, acceleration, and engine speed, road geometry) can identify drivers’ driving differences?

2. What is an effective way to classify, rank, and group drivers based on driver differences and similarities?

3. Is it feasible to model variations of driving behaviour based on collected real driving data?

4. What driving parameters (e.g. vehicle speed, acceleration, and engine speed, road

geometry) can identify drivers' attitudes to driving safely?

5. What is an effective way to classify, rank, and group drivers' safe driving based on dangerous driving habits?

5.2 The conducted analysis regarding the drivers' driving behaviour

This first set of analyses was aimed addressing the first research question as to whether distinct differences in drivers' driving performance can be identified through their collected driving data. As has been demonstrated, several of the vehicle parameters as well as the drivers' approach to driving can be used to gain insights into their behaviour. These include vehicle speed, engine speed and acceleration/deceleration, with the latter being an indicator of whether people drive calmly or aggressively.

As has been demonstrated in the first part of this subsection, for similar driving conditions (same route, same vehicle) drivers' driving data can substantially vary (see study 4.2.1.5.1 and 4.2.1.5.2 outputs). However, as they become more familiar with the route and its traffic settings (traffic lights and roundabout locations), drivers with similar driving behaviour and temperament begin to follow similar driving patterns. Specifically, they tend to have similar vehicle speed as well as similar acceleration and deceleration patterns. From the comparison of the drivers' vehicle speed data against their engine speed data, it has been demonstrated that at low speeds some drivers shifted gears more efficiently than others and were able to control the car in a more efficient manner.

Considering driving parameters in conjunction with locations has provided significant further insight into the drivers' driving behaviours. By including the route elevation, it has been possible to discover the reason as to why some drivers used higher engine speed than others, which turned out to be because when they started going uphill, they put their foot down harder on the accelerator than other drivers. If this were hypothesised, it could not be verified without this topographical data. In the geo-analysis of vehicle speed and engine speed subsection (4.2.1.4.1), the route was split into downhill, urban, and uphill driving. The outcomes made it possible to consider the possibility of the coasting downhill habits of some of the drivers (see further study in 4.3.1.3). This analysis also identified those drivers who drove rather recklessly in the urban area, particularly in terms of their speed and sharp acceleration/deceleration. In the final subsection (4.2.1.4.3) of this identification subsection, the outcomes regarding the drivers' differences in terms of their vehicle specific power values were analysed. Including the fixed parameters, such as the aerodynamic characteristics of car and road friction value as well as vehicle dynamics variables, such as speed and acceleration, made it possible to compare all the class B drivers' (Vauxhall Corsa and Nissan Note cars) driving performances, according to their vehicle specific power values. It is clear that some drivers used higher vehicle specific power than others, which indicated that they were more aggressive in their driving in terms of their choice of speed and engaging in sharp acceleration/deceleration.

The last work in this section involved determining the distribution of the vehicle speeds taken across the five laps of the event at 10-second intervals (see 4.2.1.5). This was necessary to complete the next phase of the research, i.e. the classification of drivers with similar behaviour. First, the range and consistency of speeds were analysed using box plots and cumulative frequency graphs to see whether there were clear differences between the drivers and it emerged that there were. Next, the descriptive statistics of the vehicle speeds were calculated and subsequently, subjected to distribution testing to determine whether they were normally distributed. It was concluded that the drivers' driving data were not normally distributed even with a JB normality test and in order to compare any two drivers' speed data a method had to be selected that did not require a normal distribution assumption, for example, a U-test.

5.3 The conducted analysis on classifying drivers' driving behaviour

In a later set of analyses, the outcomes of three methods of classifying, ranking, and scoring drivers were presented. These studies were carried out to address the second research question, which pertains to whether it is possible to classify drivers according to their driving data. In the first part, the drivers' degree of similarity in terms of vehicle speed profiles was visualised using network diagrams. Effective sample size analysis revealed that drivers 7, 9, 10, and 11 had similar vehicle speed profiles, whilst the other drivers (drivers 8, 13, 14, and 15) could also be considered as driving mainly in the same manner and hence, they too were grouped together (see 4.2.2.1.2).

The second part of this subsection provided the results of ranking the drivers based on two common metrics for reporting drivers and vehicles' fuel efficiency. As discussed, these metrics are not effective when it comes to representing real driver behaviour, as they do not take into account parameters such as road gradient, vehicle speed, or acceleration/deceleration. As a result, a new metric was introduced to compare drivers' efficacy, that of their vehicle specific power value in regard to their fuel usage. In the final part of this subsection, the scores for the drivers based on the total number of sudden accelerations, decelerations and incidents of high engine speed were reported. The threshold rule scoring method allowed numerical values to be assigned to each driver, but this falls short of representing real driving behaviour.

The most significant findings in this subsection are, first, it emerged that drivers who exhibited similar vehicle speed profiles in terms of their effective sample sizes, also did so under the fleet management scoring system. For example, drivers 9, 10 and 11 had very similar speed profiles, and they all scored highly as well. Secondly, except for driver 14, the drivers' VSP – FC values were consistent with their fleet management ratings. This is important, for it means that the VSP – FC values can be used to classify driver fuel efficiency and score their driving behaviour. The evidence of this conclusion is summarised in the table below.

Table 79. Comparison between drivers' average eco score and their VSP – FC values

	ID	Average score	VSP – FC
Vauxhall Corsa	7	47	2.36
	11	39	1.29
	14	0	0.48
Nissan Note	9	29	0.56
	10	43	0.60
	12	26	0.49
	13	23	0.30

5.4 The conducted analysis on the identification of the drivers' safety attitudes

This subsection's studies have all been aimed at addressing the question of identifying risky drivers. Drivers' driving data has been used to distinguish their different apparent attitudes about safety. Their vehicle speeds in conjunction with location were used to examine their approaching speeds at three points on the road, where the road setting required extra caution. Drivers were identified as exhibiting risky behaviours based on the recommendations in the UK Highway Code. The second part focused on drivers' sudden acceleration and deceleration behaviour, investigating repetition of their occurrence and location. As has been shown above, the total number of sudden accelerations and decelerations increased as drivers completed more driving rounds. The geo-study elicited that there was a 75% correlation between such behaviour and the number of recorded accidents at the same location. In general, the findings of these two studies have established the fact that drivers' sudden acceleration and deceleration behaviour has a major impact on their safety.

The final study in this subsection pertained to identifying those drivers who engaged in downhill coasting. As has been pointed out previously, this is a risky behaviour that puts the driver, car passengers and other road users in danger, as the foremost can lose control of the vehicle when operating it in this way. The total number of coasting downhill incidents for both the Vauxhall Corsa and Nissan Note drivers was identified. In the next section, the locations where drivers performed sudden acceleration or deceleration and total number of times they coasted downhill behaviour were used to modify the standard driver safety score metric.

5.5 The conducted analysis on classifying drivers in terms of their attitudes to safety

In this section, the research question in relation to an improved method for measuring drivers' safe driving is addressed. Whilst assigning a numerical value to the safety aspects of an individual's driving is not a true representation of their driving ability, nor can it predict their future involvement in road accidents, it can help insurers and fleet managers to have better understanding of the attitudes to driving of the drivers they are providing services to. As has been contended here, the scoring achieved by the standard model do not clearly identify unsafe drivers. In contrast, the two modification that are suggested for improving this method do allow for drivers to be categorised according their level of carelessness when driving. For instance, driver 7 scored 46 under the standard method, but because he coasted downhill twice and made a sudden acceleration/deceleration twice, his score fell by 10 points to 36. To conclude, by including drivers' total number of times that they coasted downhill and the total number of times they undertook sudden acceleration or deceleration at collision-prone locations, their safe driving score measurement can be improved significantly and thus, become a more realistic representation of their attitude regarding their own and others' safety.

5.6 Project achievement in the context of naturalistic driving behaviour

To this end, it is important to revisit the V-model and explain its salience in providing a framework to conduct naturalistic driving behaviour. The first part of the framework provided a structure to design an experiment. The table below summarise how each part of the framework helped to address the design of the experimental part of this research.

Table 80. Summary of the design of the study

Listing research questions	• Understanding and characterising the impact of real world driving on fuel economy • Establishing the contributing factors of drivers' driving behaviour to road safety • Investing potential analytics approaches to rank drivers and the effect of drivers' driving behaviour differences
Defining monitoring devices	• OBD II and an GPS sensor with a sim card
Defining the test procedure and data collection approach	• 4 km long route • Cars that were built after 2001 (to be compatible with OBD II device) • Drivers from similar demographic
Address what data has to be collected, anonymised and stored	• No personal data has been collected or stored • Drivers' identity was anonymised from the start and driver ID in terms of number rather than name was linked to a car and monitoring device

The second and the third stage helped in conducting a series of analysis to provide potential answers to the research questions. In Chapter 3, list of the analytical approaches was introduced and their relevance to addressing the research questions was explained. In chapter 4, the results of each analysis were presented and explained. The following table summarises the methods, the findings and the reasoning behind these analyses.

Table 81. Summary of the methods, the findings, and the analyses

		List of studies	Analysis approach and the reasoning
Eco Driving Studies	**Identifying driver differences**	Study of vehicle speed – distance and engine speed – distance analysis Study of the relationship between drivers' road speed and engine speed Study identifying drivers' aggression by comparing their speed vs. acceleration Geo-analysis of vehicle parameters in conjunction with the route elevation profile Descriptive analysis of vehicles' speed profiles	Exploratory data analysis of driving data by comparing drivers' vehicle speed, engine speed, acceleration, deceleration, and gear changing strategies The underling relationship between the drivers' aggressiveness with the scatter distribution of their speed and acceleration Spatial analysis of driving parameters with the route elevation to investigate drivers' differences including their differences when facing uphill and downhill sections of the road as well as the use of excessive vehicle power to overcome a slope Establishing the idea that individual speeds are rarely normally distributed, while the speed distribution on a given segment of a road is normally distributed around the actual average speed of the segment, if enough driving data has been collected
	Classifying and scoring drivers	Evaluating the similarity of drivers' speed profiles Classifying drivers based on their fuel usage and proposing the Vehicle Specific Power – Fuel Consumption metric (VSP-FC) Eco driving classification according to the fleet management scoring system	Introducing the concept of the similarity of two individuals' driving performance through the similarity of their speed distribution. That is, two drivers' driving speeds can be compared according to the value of statistical effective sample size, with those having a high score, having similar driving behaviour Proposing the concept of VSP-FC metric as a worthy metric to compare drivers driving and Eco driving performance since the VSP (vehicle specific power) formula has information about the key characteristics of the road, namely, gradient and the car drag coefficient, which makes a comparison of fuel usage closer to the real-world consumption.

Eco Driving Studies	**Modelling and simulating different driving behaviour**	Study of the factors affecting fuel consumption and constructing a fuel consumption forecasting model Modelling different driving styles and road condition effects on fuel usage and car emissions by using IPG Carmaker	Using a linear regression model to examine the most influential factors on total fuel usage. Feasibility study of using real-world driving parameters, such as driving speed, to build a visual computer model
Safe Driving Studies	**Identifying driver differences**	Study of drivers' safe speed perception when approaching road settings Study of drivers' hard acceleration and deceleration habits Study for identifying drivers with coasting downhill habits Identifying collision-prone zones	Investigating drivers' driving differences by comparing their driving speed while approaching a roundabout and pedestrian crossing Examining driving approach towards making harsh accelerating and deceleration, which led to investigating the correlation between their locations with those of historical crashes Developing a metric for identifying the dangerous behaviour of coasting downhill, which subsequently was added to the scoring equation
	Classifying and scoring drivers	Safe driving classification according to the fleet management scoring system Two modifications to the fleet management scoring system	Investigating a linear method for aggregating driving events, which have been evaluated as being dangerous, including driving above the speed limit, making harsh acceleration and/or deceleration and excessive usage of engine power above certain thresholds. The modification to the initial equation was based on adding the coasting downhill habits and making acceleration and deceleration on locations with historical crash records

Chapter 6

Conclusion

"Science is a way of thinking much more than it is a body of knowledge."

– Carl Sagan

6.1 Introduction

As has been discussed, studying differences between drivers' driving behaviour is an important topic. By understanding the impact of drivers on vehicle fuel consumption and emissions, it is possible to estimate the amount of the transport sector's greenhouse gas contribution and then implement plans to reduce GHGs, thus improving air quality. Moreover, such studies can help identify drivers who do not operate vehicles in a safe manner, which increases the probability of them being involved in road accidents. Having reviewed the existing published studies on understanding the impacts of driving behaviour, it can be concluded that it is important to study drivers driving behaviours in conjunction with fuel-efficient driving and safe driving. Understanding drivers' differences can be completed in three stages: first, by identifying abnormal and outlier driving habits (in this research in relation to high engine power, high VSP values as well as sudden acceleration/deceleration), and secondly, by classifying drivers with similar driving styles, that is, by scoring them for ranking purposes or finding patterns of similar behaviour that allows for them to be grouped together. Finally, based on these outcomes it becomes possible to model driving behaviour differences. Modelling variations of driving style in a virtual environment allows for the conducting of many tests at very little cost. For this project, the three stages mentioned for both eco-driving and safe driving were followed. This chapter concludes by considering the findings for each stage.

6.2 Summary of the studies and findings

Identifying drivers' driving differences using naturalistic driving parameters

- Drivers' driving patterns can be compared using the speed – distance and engine speed – distance plots. At this level of observation, drivers with significant differences are identifiable (i.e.. those regularly exceeding the speed limit or driving with a much higher engine speed than other drivers) (see 4.2.1.1).

- Observing gear shifting strategy is a robust indicator for identifying drivers with inefficient gear shifting habits. There are two types of drivers in this regard: first, there are those with the habit of remaining in a low gear for longer than is necessary (causing high engine power at a low speed). Secondly, there are those that tend to go to a higher gear at a low speed (causing inefficient engine performance) (see 4.2.1.2).

- A comparison was made between drivers' acceleration behaviour by using speed vs. acceleration scatter plots. It was observed that there was a distinct separation between drivers who accelerated and decelerated sharply and those who did so smoothly. That is, the observation of the scatter plots of acceleration vs. speed showed that aggressive drivers had a wide distribution of data points, while those of calm drivers were concentrated near to each other (see 4.2.1.3).

- Spatial analysis of the vehicle parameters provided valuable insights into driver differences when approaching different parts of the route. A driver's vehicle speed, engine speed, acceleration and VSP values were plotted against distance travelled and elevation. The outcomes indicated that while a driver's vehicle speed vs. route elevation did not provide new information, his or her engine speed and VSP values vs. route gradient profile provided valuable insights into differences among drivers' driving approaches. For example, a driver who was identified as someone with a habit of exceeding the speed limit and making sharp accelerations and decelerations frequently had the tendency to use more VSP than was required to overcome an uphill road gradient (see 4.2.1.4).

- As has been demonstrated in previous studies, a driver's vehicle speed is very much indicative of driving behaviour. Accordingly, further studies were conducted on driver vehicle speed profiles. Firstly, it was shown that, as was expected, the vehicle speeds were not normally distributed. Secondly, drivers speed consistency was illustrated in cumulative speed graphs as these are excellent indicators of drivers' driving behaviours. In this study, drivers' driving differences and outliers in terms of vehicle speed were identified using cumulative graphs, stack bar charts and box plots (see 4.2.1.5).

Classifying drivers' driving differences using naturalistic driving parameters

Three methods of classification were introduced for consideration.

- The first involved comparing speed distribution profiles. For this particular study, the matrix of similarity was calculated based on the vehicle speeds for all pair arrangements of the drivers. The results were visualised by using a node and link network diagram and as such, provided a simple method for classifying similar drivers (see 4.2.2.1).

- The second method of classifying drivers was based on their fuel usage. As only their total fuel consumption data were available, three fuel economy metrics (/100km, km/l, and kg/h) were compared based on these values. However, it emerged none of these metrics was capable of providing a meaningful comparison between the focal drivers and as a result, the VSP – Specific Fuel ratio was introduced. This metric contains information about driver speed, acceleration and road gradient and hence, is more meaningful when comparing driver behaviour using total fuel consumption (see 2.2.2).

- The third driving classification was based on the method currently used by fleet management service providers and motor insurers. The threshold rule based method scores drivers according to the total number of times they use a high engine speed or exhibit sharp acceleration/deceleration. The cross comparison of this score with the VSP – SF ratio, showed that the aforementioned ratio has the credentials to be employed by the relevant industries as a method to score drivers' fuel efficiency as well as driving behaviour (see 4.2.2.3).

Developing a fuel consumption forecasting model and modelling drivers' differences by virtual simulation

- A linear regression model was developed based on the available driving parameters, with the final model being developed based on the average vehicle speed (km/h), average distance travelled (km), and number of times using a high engine speed (rpm). The model provides an estimation of the fuel consumption of drivers when the aforementioned parameter measurements are used as inputs (see 4.2.3.1).

- The modelling of three driving classes was performed using the IPG Carmaker virtual driving simulation software under two scenarios. The aim was to establish the feasibility of using real-world driving data to represent variations in drivers' behaviour in simulations using this software and a driver's vehicle speed was used as the only input to the simulator. The first simulation scenario showed that drivers with a good driving score from using the fleet management method (4.2.2.3) were not necessarily the most fuel efficient. Hence, this has elicited that more effective ways are required to score and compare drivers' eco performances (see 4.2.3.2.1).

- The second scenario pertained to investigating the effect of road slope on fuel usage by simulating two versions of the same route (with gradient and without gradient). It was observed that inputting a road gradient of 16% increased. on average, the fuel usage by up to 41% for drivers in the simulation environment (see 4.2.3.2.2).

Identifying drivers' safety attitude using naturalistic driving parameters

- Drivers' driving speed was compared at three locations: (1) entering a main road from a minor one, (2) approaching a roundabout, and (3) coming towards a pedestrian crossing. It was observed that only 22% of the time did the drivers stop at the junction. Moreover, they were over the speed limit at the roundabout and the pedestrian crossing for 22% and 28% of the time, respectively (see 4.3.1.1)

- The total number of sharp accelerations and decelerations made by the drivers were calibrated for each lap and the occurrences of both increased per lap as the route progressed (see 4.3.1.2.1). For the second part of this study, locations at which drivers made sharp accelerations/decelerations were compared with those of previous recorded collisions. Pearson correlation analysis indicated that locations where drivers made harsh acceleration/braking were highly correlated (75%) with segments that had records of prior accidents (see 4.3.1.2.2).

- Drivers' habit of coasting downhill was selected as one of the unsafe behaviours that has been given little attention by other researchers. Accordingly, in the study in 4.3.1.3, those drivers who exhibited this sort of behaviour were identified. The results show that 4 out of 7 drivers coasted downhill at least once (see 4.3.1.3).

- The final part of this subsection investigated the collision-prone locations on the route on which the driving test was conducted. The route hotspots were identified using the statistical confidence intervals method and two out of 10 segments emerged as being so (see 4.3.1.4).

Classifying drivers' safety differences using naturalistic driving parameters

- Drivers' safe driving was scored using the method that is widely used by fleet management providers and usage-based motor insurance providers. As the results show (see 4.3.2.1), the scoring method's ability to distinguish drivers' differences is very limited and therefore, two modifications were made to the model to make it more realistic. That is, a driver's total number of times coasting downhill and total number of occasions making sharp accelerations/decelerations at collision-prone locations were added to the scoring model. These modifications introduced greater accuracy to measuring drivers' actual behaviour (see 4.3.2.2).

6.3 Limitations

This project, like every field study, by its nature, had a number of limitations, which manifested themselves at three levels. Firstly, there are avoidable limitations, which are possible to address in future studies. In this study, the OBD dongle was set to record five driving parameters (in order to have sufficient internet credit to complete driving laps) every 10 seconds, which generated a substantial amount of data. It is possible to use an advanced dongle to record data second by second for 18 parameters, but this would have brought the data processing to an unmanageable scale for a project of this size. Pedal position and instantaneous fuel consumption were two parameters that may be should have been included for comparing and classify drivers' driving behaviour differences, and it is important for future studies to consider them. The second level of limitation is the semi-avoidable constraints. These are limitations that are partly avoidable by detailed consideration at the design of the experiment stage of the project. For instance, this could be a driver's demographics and vehicles types used. In this study, special care was taken to have a homogeneous demographic of drivers with similar vehicle types. The final level of limitation comprises those that are unavoidable. As is self-explanatory, parameters such as traffic flow conditions, weather (temperature, humidity), and to some extent the psychological status of a driver are parameters that are not controllable and hence, the impacts of them on the study (if any) are unavoidable and unquantifiable.

6.4 Contributions and their applications

This project was aimed at contributing effective methods for identifying and classifying drivers in terms of efficient eco-driving and safe driving. As has been outlined in the previous chapters, identifying abnormal driving styles and classifying drivers with similar driving behaviour attributes is of high importance, in particular, because it can benefit fleet management service providers and usage-based insurance companies. Of all the studies that were conducted for this project, the following findings were selected as the major contributions to this field of study.

1 **Identifying a metric for ranking drivers' fuel consumption**: The Vehicle Specific Power (VSP) – Specific Fuel (SF) ratio has been proposed as a method to rank drivers. Since the VSP value is calculated based on drivers' vehicle speed, acceleration, road gradient, and car drag coefficient, it provides common ground for comparing driver performances. The ratio that is calculated by dividing the fuel consumption ratio (g/s) to the VSP value is a very robust metric for comparing and ranking drivers based on their fuel usage and driving behaviour (driving and vehicle conditions). The suggested method is useful for telematics providers, and fleet management service providers since the VSP – SF ratio affords an opportunity for them to compare fuel efficient drivers with simple parameters that are available to them (vehicle speed, amount of fuel usage and location information).

2 **Classifying drivers based on their speed similarity**: Power analysis (effective sample size analysis) method has been used in this project in a new way in that it has been shown that drivers' speed profiles can be compared so as to demonstrate their similarity or difference. By calculating the effective sample size required based on every pair of

drivers, it has been shown that drivers can be classified purely based on their speed similarity. The method is a very useful tool for industries that want to classify driver similarity without the usage of complex pattern classification methods.

3 **Scoring drivers**: Based on the method that is frequently used by industries, such as motor insurance providers and fleet management service providers, two additional adjustments were added to the model in order to make the method realistic and accurate. The two modifications that were added were the following: firstly, the drivers received negative points when they coasted downhill; and secondly, if they made harsh accelerations or decelerations at locations that were identified as route hotspots, they would receive twice the number of negative points, as when they had performed in this manner in non-hotspot locations. The direct application of these two modifications can benefit those industries that need to rank drivers, such as those developing usage-based vehicle insurance policies.

4 **Establishing the level of association between the locations where drivers made harsh accelerations/decelerations at the event and those with previous collision records.** First, places where drivers exhibited either harsh acceleration or sudden braking were identified, according to 10 road segments. Next, the historical records of collisions were extracted from published records between 2005 and 2013 for these segments. Finally, the Pearson correlation measures of these two datasets for each segment were calculated. It emerged that 75% of the time that drivers made harsh acceleration and deceleration occurred in the two segments that had the highest number of recorded accidents.

5 **Using Carmaker virtual driving simulator to compare drivers' driving**: Based on the two scenarios designed in this research, it has been shown that IPG Carmaker is capable of generating distinguishable driving results when a driver's vehicle speed is the only parameter that is changed. According to the findings of first study, it is possible to estimate driver efficiency based on driving data generated by Carmaker using only drivers' real-world speed as the software input value. Regarding the second study using this software, it has been shown that a driver's fuel usage is highly influenced by the nature of the road he / she is driving on. In the case of this study, the simulation based on real-world conditions has shown that a 16% increase in road gradient will result in around a 41% increase in fuel consumption

6.5 Recommendations

This work is aligned with other ongoing naturalistic driving behaviour studies in the EU, for instance, the UDRIVE and the EcoDriver. It has been established by this work and other scholars that field studies are a reliable and valuable method of studying driver driving behaviour differences and hence, they provide a better understanding of the impact of real-world driving conditions on fuel consumption and driver safety. It is recommended that further projects be conducted with the aims and objectives of naturalistic driving behaviour studies and frameworks, as has been proposed in the PROLOGUE documentation (Van Schangen *et al.*, 2011).

Developing a driving behaviour database

It is logical to consider constructing a universal driving behaviour database, which can then be shared by academics and industry. Such a platform would provide the opportunity to understand drivers' driving behaviours on a worldwide basis. In this thesis, the fundamentals of creating such a database have been briefly discussed. However, a collaborative effort is required to create such an open access platform. Consequently, it is logical to argue that there should be a database compiled of different driver behaviours and accordingly, efforts are now being made to this end in both the UK and other countries in Europe, with the aim being to collect such data for 5,000 drivers.

The VSP – FC metric

It has been demonstrated that the VSP-FC metric has the capacity to be used as a method to rank drivers' driving behaviour as well as their eco-driving performances. However, this conclusion has been drawn based on this study only. It is recommended that further projects evaluate and validate the effectiveness of this metric, by comparing the VSP-FC metric values of the driving data of drivers who do not drive on the same routes.

Driving behaviour scoring

In this thesis, various methods and metrics for comparing and ranking drivers' driving performances have been utilised. It has also been demonstrated that some of the existing models for scoring drivers are not capable of doing so fairly or realistically. Moreover, if the aim of the scoring is in relation to usage-based insurance, it is important to ensure the homogeneity of the risk model in that sector, if it is to work effectively. In this work, driver similarity in terms of driving speed has been proved as being an insightful metric. Moreover, the outcomes from two modifications to the existing fleet scoring model, namely, coasting behaviour and speeding/braking at accident hotspots, have indicated that the calculations from this model could more closely reflect real-world driving through further improvement of the methods used in this thesis. Regarding future research on this topic, consideration of advanced pattern classifications and predictive analytics is recommended, so as to develop a more advanced approach to classifying and scoring driving behaviours.

Using vehicle simulation software: IPG Automotive Carmaker software

Using the Carmaker software provided an opportunity to model and compare drivers' driving performance in a virtual environment. The package is capable of modelling and simulating complex driving conditions in great detail. It is contended that using the package is a very cost effective method for testing a vehicle under various driving conditions and in identifying differences in drivers' driving behaviours. It is proposed here that future research would benefit from using this software, such as by conducting virtual driving tests in order to develop realistic driver models. Moreover, since the software is open source, it is possible to integrate the package with in-house driver driving models and existing hardware in loop chassis dynamometer testing.

Appendix A: Eco Safe Driving Challenge Final Report

This report was published in December 2014.

Background

The idea of gathering driving enthusiasts and giving them an opportunity to drive is not particularly remarkable. However, providing insights about their performance as well as advising them how to be safer or how to be eco-friendly drivers, we believe would attract a wider audience than simply ardent car lovers. That is, from motorsport enthusiasts to parents worrying about their novice driver offspring among many others, knowing about car performance as well as level of driving skills is valuable information. Moreover, given the rise in fuel costs in recent times, and increasingly expensive car insurance, there is economic incentive to improve car performance, more importantly, greater fuel efficiency, will help towards achieving the ultimate goal of pollution free air, something we have a duty to try to pass on to future generations. Including motorists' carbon footprints in every report produced by the Department of Transport and Department for Environment highlights awareness that individual road users can have a marked impact on the environment, by reducing their carbon emissions. The Eco Safe Driving Challenge (**ESDC**) event has provided understandable information to the general public and driving enthusiasts about how their driving behaviour affects fuel usage, vehicle performance and road safety. From the other side, the research team in the department of Mechanical Engineering at the University Bath wanted to collect reliable data on the effects of driver behaviour on fuel usage in a real world setting.

Aims

The ESDC aim is to deliver the core message: "your driving habits, will determine your motoring future". More specifically, it is a given fact that improving driving behaviour can lower fuel costs, reduce annual vehicle insurance and lead to better road safety. Moreover, deaths and injuries from driving accidents are not only tragedies for those close to the victims, but also take millions of pounds out of the future economy from lost potential earnings and taxation as well as unnecessarily wasting precious NHS funds. The ESDC has tasked itself with raising awareness among students, driving enthusiasts and the general public about their driving habits and the ways in which they could improve their fuel consumption, improve car performance and drive with more care. To this end, the team has the intention of using affordable, user-friendly and modern technologies to inform drivers about their driving as well as to give advice on how they can improve it to their own benefit.

Project milestones

The project started on February 2013 based on a research proposal to investigate driving behaviours, which was supported by the University of Bath's research-led innovation team. Since its inception, it has steadily made progress in achieving its objectives and gathering a wide audience of those interested in its outcomes, including students, driving aficionados and the general public. It has been a challenging project, not least because everything had to be started from scratch and consequently, it evolved during the process. Since the official launch of the ESDC, a community has been steadily built around its cause, with its message having reached out to drivers across the South West region. 250 miles has been safely driven by driver participants during three events and hopefully they now drive more safely than before. Most significantly, an 8% reduction in fuel usage was recorded during the last event. In February this year, we conducted our largest driving event, with two different car categories, in the University of Bath West Car Park. We used similar cars for our volunteer drivers and the EE mobile network provider kindly supported us with their service that enabled us to monitor them live while driving. More than 45 circuits were driven at the event, which resulted in our biggest reliable data set being gathered to date. Since then we have been analysing this data and have been able to identify particular driving behaviours that cause motorists to use more fuel and occasionally act unsafely. We are aiming to publish our very interesting findings in both academic journals and for the general public by the end of this summer.

Public engagement

The team has been engaged with the public both through social media as well as the aforementioned driving events. Several families came and competed as teams in these events, with strong friendship bonds being forged amongst all the participants. Through the social media channels, we were able convince potential event participants that their engagement would make a tremendous contribution to this academic research about driving behaviour and its consequent effects on society as a whole. We replayed for them their driving activity after each round through the provided online platform, with the aim of encouraging them to take charge their driving outcome by modifying their behaviour in subsequent trips. Regarding which, they showed a keen interest to drive differently when they saw their results, for these acted as a positive incentive to improve between rounds in terms of driving more safely and in a more eco-friendly way.

Knowledge exchange

By providing information to drivers and participants about the outcomes of their driving and their contribution to the study, they themselves not only learnt about the effect of the former on fuel consumption and the environment, but also, started sharing their experience and understanding about the topic. They told us about good/bad driving experiences, previous incidents/accidents and their points of view about the types of driving challenges we had set them. Consequently, a substantial amount of knowledge was exchanged by both parties throughout the events. In particular, the more experienced drivers passed on driving tips that provided useful awareness and insight for our research.

We had an excellent opportunity to discuss and share our ideas about the research with the UK’s leading transport research team: Transport Research Laboratory or TRL. Moreover, their presentation at our awards ceremony about their cutting-edge approach to transport challenges further motivated us in our research endeavours. In particular, their constructive feedback and encouragement helped us to expand our industrial network, thereby increasing the potential for exploiting collaborative opportunities in transport studies that will enhance the University of Bath’s reputation in this field.

Appendix B: Driver’s eligibility criteria

- All drivers must be aged over 18 years old on the day of a driving event (Sunday 9th February). In case you want to have your co-driver with you, your co-driver must be over 18 and be able to give you a running commentary about a route and road conditions ahead.
- In respect of your driving experiences, you must have a full UK driving licence with no points or endorsements.
- All UK road-legal petrol or diesel vehicles built after 2001 are able to participate, proof of MOT and road tax may be required.

Appendix C: Weather conditions report on 09/02/2014

The 9th February 2014 was the main driving event date. As part of the project, the weather conditions were recorded and are provided in the table below.

		Actual	Average	Record
Temperature	Mean Temperature	4 °C	-	-
	Max Temperature	6 °C	7 °C	13 °C (2008)
	Min Temperature	2 °C	2 °C	-4 °C (2004)
Degree Days	Heating Degree Days	26	-	-
Moisture	Dew Point	2 °C	-	-
	Average Humidity	86	-	-
	Maximum Humidity	93	-	-
	Minimum Humidity	75	-	-
Precipitation		0.0 m	-	-
Sea Level Pressure		983.54hPa (29.04)	-	-
Wind	Wind Speed	30 km/h	-	-
	Max Wind Speed	50 km/h	-	-
	Max Gust Speed	70 km/h	-	-
Visibility		8.7 km	-	-
Condition		Rain	-	-

Source: Weather Underground at https://www.wunderground.com/

Appendix D: Example of acceleration vs. speed

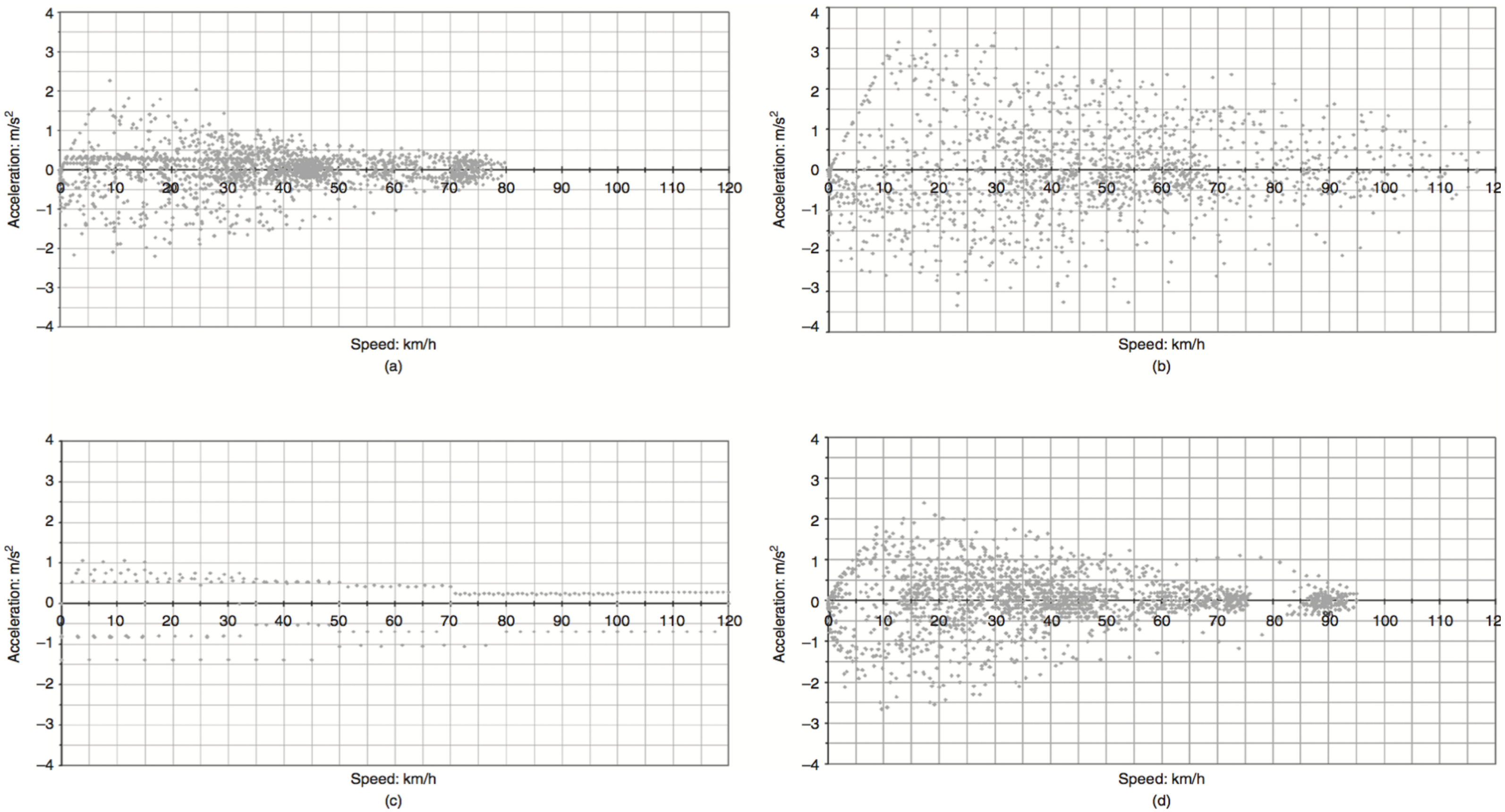
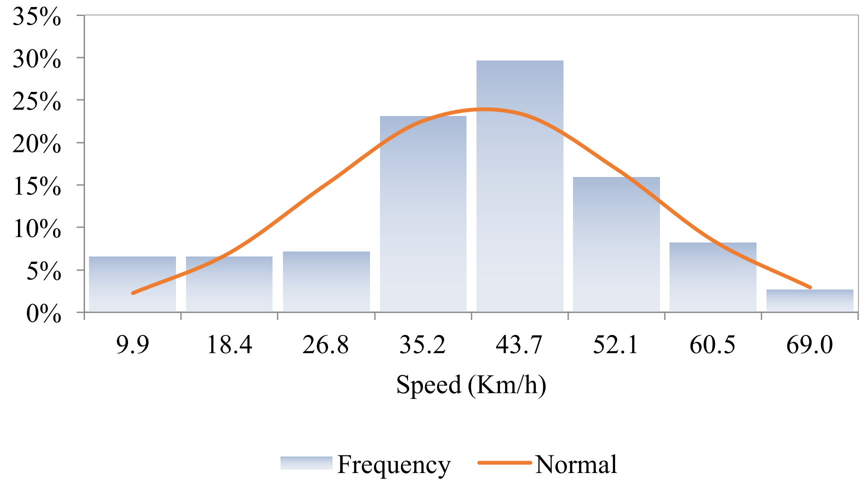
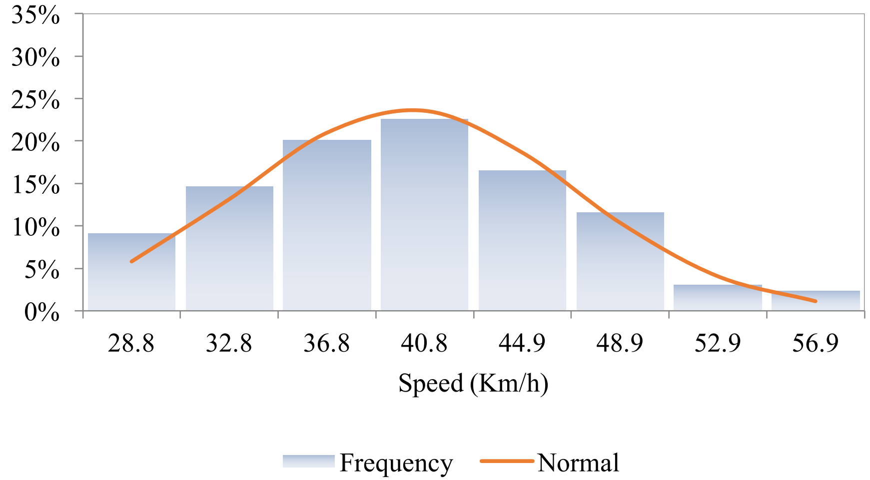


Figure 98. (a) Passive driver, (b) Aggressive driver, (c) ECE combined driving cycle, and (d) Spring Laboratory combined driving cycle (Felstead, McDonald and Fowkes, 2009)

Appendix E: Individual histograms for all nine drivers overall speed distribution





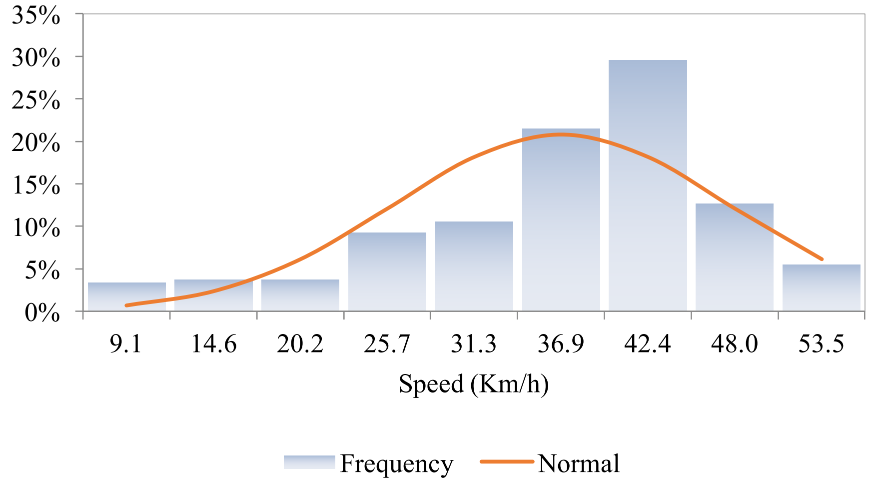


Figure 99. The Vauxhall Corsa drivers' speed profiles distribution histograms: Driver 7 (top left), Driver 11 (top right), Driver 14 (bottom left)

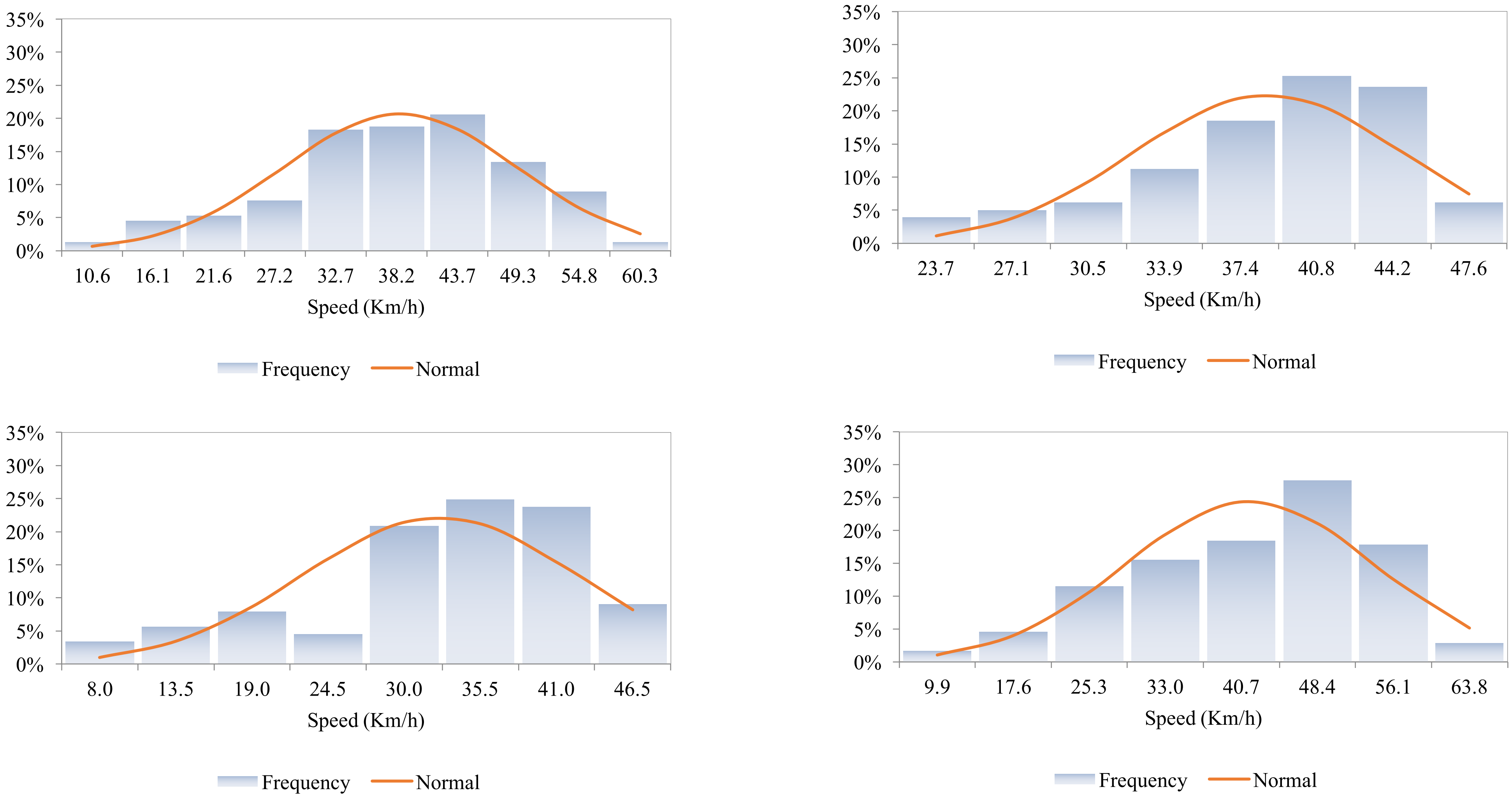
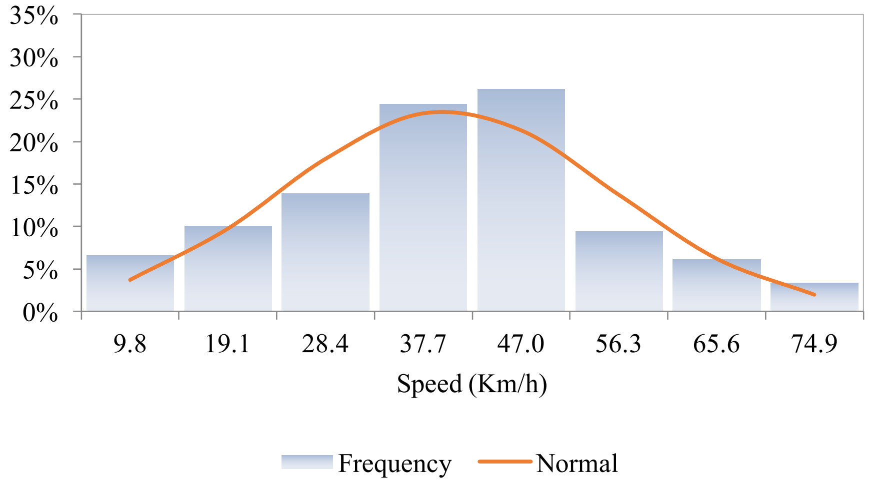


Figure 100. The Nissan Note drivers' speed profiles distribution histograms: Driver 9 (top left), Driver 10 (top right), Driver 12 (bottom left), and Driver 13 (bottom right)



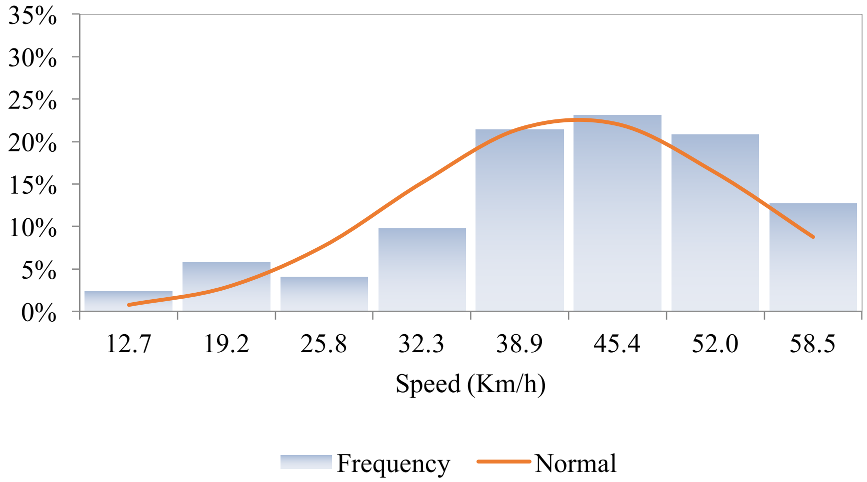


Figure 101. The Class C drivers' speed profiles distribution histograms: Fiat 500L - Driver 8 (left), Nissan Juke - Driver 15 (right)

Appendix F: Drivers' full speed parameters of the-box-and-whisker plot

	Driver ID								
Parameters	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
Min	0	0	0	0	0	0	0	0	0
Q_1	28.4	28.2	28.5	27.42	26.92	22.27	29.9	31.2	32.9
Median	38.3	39.3	37.3	37.8	37.2	33.2	42.1	40.5	43.8
Q_3	43.7	48.3	45.4	42.27	42.82	39.57	50.8	48	50
Max	56.3	79.6	63.1	49.3	58.9	49.2	67.6	73.2	61.8
IQR[97]	15.3	20.0	16.9	14.85	15.9	17.3	20.9	16.8	17.1
Lower Outliers	17	0	27	25	13	0	0	12	12
Q2-Q1	9.8	11.1	8.8	10.37	10.27	10.92	12.2	9.3	10.9
Q3-Q2	5.4	8.92	8.1	4.47	5.625	6.37	8.7	7.5	6.2
Q_3+1.5*IQR	66.8	78.38	70.75	64.55	66.67	65.52	82.15	73.2	75.65
Q_1-1.5*IQR	5.4	-1.91	3.15	5.15	3.075	-3.67	-1.45	6	7.25
Upper Whisker	56.3	78.38	63.1	49.3	58.9	49.2	67.6	73.2	61.8
Lower Whisker	5.4	0	3.15	5.15	3.075	0	0	6	7.25
W_{upper}-Q_3	12.5	30.11	17.7	7.025	16.07	9.625	16.8	25.2	11.8
Q_1-W_{lower}	23	28.2	25.35	22.27	23.85	22.27	29.9	25.2	25.65

[97] Interquartile range IQR= $Q_3 - Q_1$.

Appendix J: Calculating the value of the total consumption for downhill and uphill driving

The following lines of codes in the table below summarise the logic and syntax used to calculate the absolute consumption values. IPG post processing procedure in matlab.

Matlab code	**Discerption**
cd('C:\CM_Projects\initial_test\src_cm4sl');	Define and change the current folder to the IPG folder
cmenv	Initialising Carmaker in Matlab
a=cmread('C:\CM_Projects\initial_test\SimOutput******');	Importing driver 7's without elevation simulation result and call it (a)
Dista=a.Vhcl_sRoad.data; Consmpa=a.PT_Engine_Consump_Abs.data;	Defining the distance and consumption values of (a)
[i1a,j1a]=find(Dist<=1800); [i2a,j2a]=find(Dist<=2200);	Selecting the sections of the route that involve downhill and uphill driving
C_DHa=Consmpa(j1a(end))-Consmpa(1) C_UHa=Consmpa(end)-Consmpa(j2a(end)	Calculating the value of the total consumption for downhill and uphill driving

References

[1] A. T. Kearney, Reifel, J., Hales, M., Xu, G. and Lala, S. (2010) *Telematics: The Game Changer*.

[2] Aarts, L. and van Schagen, I. (2006) 'Driving Speed and the Risk of Road Crashes: A review', *Accident; analysis and prevention*, 38(2), pp. 215–24. doi: 10.1016/j.aap.2005.07.004.

[3] Aarts, L., Wegman, F. and SWOV Institute for Road Safety Research (2008) 'Sustainable Safety: The Safe System Approach of The Netherlands', *Behavioural Research in Road Safety 2008 Seminar*, (September), pp. 1–15.

[4] ACEA (2014) *The Automobile Industry Pocket Guide*, *European Automobile Manufacturers Association*. Available at: https://www.acea.be/uploads/publications/POCKET_GUIDE_2014-1.pdf.

[5] Ajzen, I. (1985) 'From intentions to actions: A theory of planned behavior', in Kuhl, J. and Beckman, J. (eds) *Action-control: From cognition to behavior*. Heidelberg: Springer, pp. 11–39.

[6] Akçelk, R., Smit, R. and Besley, M. (2012) 'Calibrating Fuel Consumption and Emission Models for Modern Vehicles', in *IPENZ Transportation Group Conference*. Available at: http://www.sidrasolutions.com/cms_data/contents/sidra/folders/resources/articles/articles/~contents/8kr2vybbfm8vpnls/akcelik_fuel-emissionmodels-ipenz2012-.pdf.

[7] Al-shihabi, T. and Mourant, R. R. (2001) 'Toward More Realistic Driving Behavior Models for Autonomous Vehicles in Driving Simulators', (June), pp. 1–20.

[8] Alessandrini, A., Filippi, F., Orecchini, F. and Ortenzi, F. (2006) 'A New Method for Collecting Vehicle Behaviour in Daily Use for Energy and Environmental Analysis', *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, 220(11), pp. 1527–1537. doi: 10.1243/09544070JAUTO165.

[9] Alimoradian, B. (1983) 'Predicting Fuel Consumption and Emissions-Transferring Chassis Dynamometer Results to Real Driving Conditions', *SAE paper 830435.*, 830435. doi: 10.4271/830435.

[10] Altman, I., Wohlwill, J. F. and Everett, P. B. (1981) *Transportation and Behavior*. Edited by I. Altman, J. F. Wohlwill, and P. B. Everett. Boston, MA: Springer US. doi: 10.1007/978-1-4684-4001-0.

[11] Amata, H., Miyajima, C., Nishino, T., Kitaoka, N. and Takeda, K. (2009) 'Prediction Model of Driving Behavior Based on Traffic Conditions and Driver Types', in *International IEEE Conference on Intelligent Transportation Systems*, pp. 747–752. Available at: http://ir.nul.nagoya-u.ac.jp/jspui/bitstream/2237/13899/1/Amata_ITSC2009-10.pdf (Accessed: 12 April 2016).

[12] Amditis, A., Bailly, B., Baumann, M., Bekiaris, E., Bellet, T., Bengler, K., Cacciabue, P. C., Carsten, O., Chao, K. C., Cody, D., Engström, J., Fuller, R., Georgeon, O., Gordon, T., Hollnagel, E., Inagaki, T., Janssen, W., Jürgensohn, T., Krems, J. F., Macchi, L., Marzani, S., Meyenobe, P., Montanari, R., Nilsson, L., Panou, M., Papakostopoulos, V., Parker, D., Peters, B., Polychronopoulos, A., Re, C., Saad, F., Summala, H., Tango, F., Vaa, T., van der Horst, R. and Weir, D. H. (2007) *Modelling Driver Behaviour in Automotive Environments*. Edited by P. C. Cacciabue. London: Springer London. doi: 10.1007/978-1-84628-618-6.

[13] American Automobile Association (AAA) (2009) *Aggressive Driving* : Research Update. Available at: AAAFoundation.org.

[14] An, F., Barth, M., Norbeck, J. and Ross, M. (1997) 'Development of Comprehensive Modal Emissions Model Operating Under Hot-Stabilized Conditions', *Transportation research record*, 1587, pp. 52–62. Available at: http://www.cert.ucr.edu/cmem/docs/TRR1997-Hot-Stabilized.pdf (Accessed: 23 April 2015).

[15] Ando, R., Nishihori, Y. and Ochi, D. (2010) 'Development of a System to Promote Eco-Driving and Safe-Driving', in Balandin, S., Dunaytsev, R., and Koucheryavy, Y. (eds) *Smart Spaces and Next Generation Wired/Wireless Networking*. Berlin, Heidelberg: Springer Berlin Heidelberg (Lecture Notes in Computer Science), pp. 207–218. doi: 10.1007/978-3-642-14891-0_19.

[16] Ardekani, S., Hauer, E. and Jamei, B. (2001) 'Traffic Impact Models'. Available at: http://www.fhwa.dot.gov/publications/research/operations/tft/chap7.pdf (Accessed: 23 April 2015).

[17] Arthur, W., Barrett, G. V. and Alexander, R. A. (1992) 'Prediction of Vehicular Accident Involvement: a Meta-Analysis', *Journal of Safety Research*, 23, pp. 73–80.

[18] Ashbaugh, L. and Lawson, D. (1991) 'A Comparison of Emissions from Mobile Sources Using Random Roadside Surveys Conducted in 1985, 1987, and 1989', *The Air & Waste Management Association*, p. 91–180.58. Available at: https://hero.epa.gov/hero/index.cfm/reference/details/reference_id/46441 (Accessed: 13 April 2015).

[19] Ayyildiz, K., Cavallaro, F., Nocera, S. and Willenbrock, R. (2017) 'Reducing fuel consumption and carbon emissions through eco-drive training', *Transportation Research Part F: Traffic Psychology and Behaviour*, 46. doi: 10.1016/j.trf.2017.01.006.

[20] Azzopardi, M. and Cortis, D. (2013) 'Implementing Automotive Telematics for Fleet Insurance', *Journal of Technology Management & Innovation*, pp. 59–67. Available at: http://www.jotmi.org/index.php/GT/article/view/1308 (Accessed: 31 August 2015).

[21] Backer-Grøndahl, A., Lotan, T. and Schagen, I. van (2011) *Summary and integration of a series of Naturalistic Driving field trials | www.swov.nl*. Available at: https://www.swov.nl/en/publication/summary-and-integration-series-naturalistic-driving-field-trials.

[22] Bandivadekar, A. P. (2008) 'Evaluating the impact of advanced vehicle and fuel technologies in U.S. light duty vehicle fleet'. Massachusetts Institute of Technology. Available at: https://dspace.mit.edu/handle/1721.1/43856 (Accessed: 25 July 2017).

[23] Barić, D., Zovak, G. and Periša, M. (2013) 'Effects of eco-drive education on the reduction of fuel consumption and CO2 emissions', *PROMET-Traffic&Transportation*, 25(3), pp. 265–272. doi: 10.7307/ptt.v25i3.1260.

[24] Barlow, T. j, Latham, S., McCrae, I. S. and Boulter, P. G. (2009) *A reference book of driving cycles for use in the measurement of road vehicle emissions*. 3rd edn. TRL Limited. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/4247/ppr-354.pdf (Accessed: 20 January 2016).

[25] Bartlett, J. E., Kotrlik, J. W. and Higgins, C. C. (2001) 'Organizational research: Determining appropriate sample size in survey research', *Information Technology, Learning, and Performance Journal*, 19(1), pp. 43–50. doi: 10.1109/LPT.2009.2020494.

[26] van Basshuysen, R. and Schäfer, F. (2008) 'Fuel Consumption', in *Internal Combustion Engine Handbook - Basics, Components, Systems, and Perspectives*. SAE International, pp. 737–751. Available at: http://app.knovel.com/hotlink/toc/id:kpICEHBCS1/internal-combustion-engine-3/internal-combustion-engine-3.

[27] Bates, L. J., Davey, J., Watson, B., King, M. J. and Armstrong, K. (2014) 'Factors Contributing to Crashes among Young Drivers.', *Sultan Qaboos University medical journal*. Sultan Qaboos University, 14(3), pp. e297-305. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25097763 (Accessed: 1 August 2017).

[28] Bera, A. K. and Jarque, C. M. (1982) 'Model specification tests. A simultaneous approach', *Journal of Econometrics*, 20(1), pp. 59–82. doi: 10.1016/0304-4076(82)90103-8.

[29] Berry, D. S. and Belmont, D. M. (1951) *Distribution of vehicle speeds and travel times*. Institute of Transportation and Traffic Engineering. Available at: http://www.worldcat.org/title/distribution-of-vehicle-speeds-and-travel-times/oclc/9560630 (Accessed: 2 September 2017).

[30] Berry, I. M. (2010) *The Effects of Driving Style and Vehicle Performance on the Real-world Fuel Consumption of U.S. Light-duty Vehicles*. Available at: http://web.mit.edu/sloan-auto-lab/research/beforeh2/files/IreneBerry_Thesis_February2010.pdf.

[31] Bokare, prashant shridhar and Maurya, akhilesh kumar (2013) 'Study of Effect of Speed, Acceleration and Deceleration of Small Petrol Car on its Tail Pipe Emission', *international journal of traffic and transport engineering*, 3(4), pp. 465–478. Available at: http://www.ijtte.com/uploads/2013-12-30/5ebd908d-451f-a5dfIJTTE_Vol 3(4)_9.pdf (Accessed: 7 October 2015).

[32] Bolderdijk, J. W. and Steg, L. (2011) 'Pay-as-You-Drive Vehicle Insurance as a Tool to Reduce Crash Risk: Results so Far and Further Potential', *Roundtable on Insurance Costs and Accident Risks*, pp. 1–27.

[33] Bonnington, P. (2009) *Vehicle Duty Cycle Analysis*.

[34] Boriboonsomsin, K., Barth, M. and Vu, A. (2011) 'Evaluation of Driving Behavior and Attitude towards Eco- Driving : A Southern California Limited Case Study', *Transportation Research Record:*

Journal of the Transportation Research Board, (January). Available at: ftp://ftp.hsrc.unc.edu/pub/TRB2011/data/papers/11-0151.pdf.

[35] Boyce, T. E. and Geller, E. S. (2001) 'A technology to measure multiple driving behaviors without self-report or participant reactivity.', *Journal of applied behavior analysis*, 34(1), pp. 39–55. doi: 10.1901/jaba.2001.34-39.

[36] Boyce, T. E. and Geller, E. S. (2002) 'An instrumented vehicle assessment of problem behavior and driving style: Do younger males really take more risks?', *Accident Analysis and Prevention*, 34(1), pp. 51–64. doi: 10.1016/S0001-4575(00)00102-0.

[37] Brown, I. and Groeger, J. (1990) 'Editorial: A Way with Errors', *Ergonomics*, 33, pp. 1183–1184.

[38] Bruneteau, F., Hallauer, T., Noël, M. and Tusa, S. (2013) *Usage-Based Insurance global Study*. Available at: http://www.ptolemus.com/content/uploads/2013/10/UBI-Global-Study-Free-Abstract1.pdf (Accessed: 1 January 2014).

[39] BS ISO 4138:2012 (2012) *British Standard BS ISO 4138: 2012 Passenger cars. Steady-state circular driving behaviour. Open-loop test methods*. Geneva. Available at: https://bsol.bsigroup.com/Search/Search?searchKey=BS+ISO+4138%3A2012 (Accessed: 27 October 2014).

[40] BS ISO 7975 (2006) *British Standard BS ISO 7975: 2006 Passenger cars. Braking in a turn. Open-loop test method.* Available at: https://bsol-bsigroup-com.ezproxy1.bath.ac.uk/Search/Search?searchKey=ISO+7975%3A2006&OriginPage=Homepage%2Bsearch&autoSuggestion=false (Accessed: 28 February 2016).

[41] BS ISO 9816 (2006) *British Standard BS ISO 9816: 2006 Passenger cars. Power-off reaction of a vehicle in a turn. Open-loop test method.* Available at: https://bsol-bsigroup-com.ezproxy1.bath.ac.uk/Search/Search?searchKey=ISO+9816%3A2006&OriginPage=Homepage%2Bsearch&autoSuggestion=false (Accessed: 28 February 2016).

[42] Burke, R., Brace, C. J. and Moffa, J. (2009) 'Increasing accuracy and repeatability of fuel consumption measurement in chassis dynamometer testing', *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, 223(9), pp. 1163–1177. doi: 10.1243/09544070JAUTO1084.

[43] Byon, Y., Shalaby, A. and Abdulhai, B. (2006) 'Travel Time Collection and Traffic Monitoring Via GPS Technologies', *2006 IEEE Intelligent Transportation Systems Conference*. IEEE, pp. 677–682. doi: 10.1109/ITSC.2006.1706820.

[44] Cacciabue, P. C., Re, C. and Macchi, L. (2007) 'Simple simulation of driver performance for prediction and design analysis', *Modelling Driver Behaviour in Automotive Environments: Critical Issues in Driver Interactions with Intelligent Transport Systems*, (1948), pp. 344–375. doi: 10.1007/978-1-84628-618-6_19.

[45] Carsten, O. (2007) 'From driver models to modelling the driver: What do we really need to know about the driver?', *Modelling Driver Behaviour in Automotive Environments: Critical Issues in Driver Interactions with Intelligent Transport Systems*, pp. 105–120. doi: 10.1007/978-1-84628-618-6_6.

[46] Casadei, anrico, Broda, R. and Ricardo Inc. (2008) *Impact of Vehicle Weight Reduction on Fuel Economy for Various Vehicle Architectures*. Available at: http://www.drivealuminum.org/research-resources/PDF/Research/2008/2008-Ricardo-Study.pdf (Accessed: 6 April 2015).

[47] CASTEL Company (2014) *CASTEL OBD Telematics presentation*.

[48] Castignani, G., Derrmann, T., Frank, R. and Engel, T. (2015) 'Driver behavior profiling using smartphones: A low-cost platform for driver monitoring', *IEEE Intelligent Transportation Systems Magazine*, 7(1). doi: 10.1109/MITS.2014.2328673.

[49] Center for Transportation Research and Education (2001) *Speed Limits and Road Safety*. Available at: http://www.ctre.iastate.edu/pubs/itcd/speedlimits.pdf.

[50] Chen, Y.-L. (2007) 'Driver personality characteristics related to self-reported accident involvement and mobile phone use while driving', *Safety Science*, 45(8), pp. 823–831. doi: 10.1016/j.ssci.2006.06.004.

[51] Cheng, W. and Washington, S. P. (2005) 'Experimental evaluation of hotspot identification methods.', *Accident; analysis and prevention*, 37(5), pp. 870–81. doi: 10.1016/j.aap.2005.04.015.

[52] Clarkson, H. and Hicks, R. G. (1982) *Highway engineering*. Available at:

http://trove.nla.gov.au/work/15379560?selectedversion=NBD2064562 (Accessed: 15 April 2016).
[53] Conger, J. J., Gaskill, H. S., Gladd, D. D., Hassell, L., Rainey, R. V. and Sawrey, D. W. . (1959) 'Psychological and Psychophysical Factors in Motor Vehicle Accidents', *American Medical Association*, 169, pp. 1581–1587.
[54] Daniel, R., Brooks, T. and Pates, D. (2009) 'Analysis of US and EU drive styles to improve understanding of market usage and the effects on OBD monitor IUMPR', *SAE World Congress & Exhibition*, pp. 2009-01–0236. doi: 10.4271/2009-01-0236.
[55] Darby, P., Murray, W. and Raeside, R. (2009) 'Applying online fleet driver assessment to help identify, target and reduce occupational road safety risks', *Safety Science*, 47(3), pp. 436–442. doi: 10.1016/j.ssci.2008.05.004.
[56] Delphi (2013) 'Worldwide Emissions Standards', pp. 39–51. Available at: https://delphi.com/pdf/emissions/Delphi-Passenger-Car-Light-Duty-Truck-Emissions-Brochure-2012-2013.pdf.
[57] Department for Energy and Climate Change (2015) *UK Greenhouse Gas Emissions, 2013 Final Figures*. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/407432/20150203_2013_Final_Emissions_statistics.pdf (Accessed: 17 March 2016).
[58] Department for Transport (2012) *Guidance on Road Classification and the Primary Route Network*. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/315783/road-classification-guidance.pdf.
[59] Department for Transport (2014) *National Travel Survey: England*. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/342160/nts2013-01.pdf (Accessed: 7 September 2015).
[60] Department for Transport (2015) *Road Safety Data - Datasets*, *Department for Transport*. Available at: http://data.gov.uk/dataset/road-accidents-safety-data/ (Accessed: 17 February 2015).
[61] Detroit Diesel Corporation - Demand Detroit (2015) *How to Get Great Fuel Efficiency*. Available at: https://www.demanddetroit.com/performance/efficientdriving.aspx (Accessed: 11 April 2016).
[62] Ben Dhaou, I. (2011) 'Fuel estimation model for ECO-driving and ECO-routing', in *IEEE Intelligent Vehicles Symposium, Proceedings*, pp. 37–42. doi: 10.1109/IVS.2011.5940399.
[63] DIN 1940 (1976) *Standard DIN 1940:1976 Reciprocating internal combustion engines; terms, formulae, units*. Available at: http://www.beuth.de/en/standard/din-1940/714377 (Accessed: 1 January 2014).
[64] DIN 70020-7 (2013) *Standard DIN 70020-7:2013 Road vehicles - Automotive engineering - Part 7: Engine mass of combustion engines in passenger cars, which are only propelled by combustion engines*. Available at: http://www.beuth.de/en/standard/din-70020-7/170094456 (Accessed: 27 October 2015).
[65] Dingus, T. A., Klauer, S. G., Neale, V. L., Petersen, A., Lee, S. E., Sudweeks, J. D., Perez, M. A., Hankey, J., Ramsey, D. J., Gupta, S., Bucher, C., Doerzaph, Z. R., Jermeland, J. and Knipling, R. R. (2006) *The 100-Car Naturalistic Driving Study, Phase II - Results of the 100-Car Field Experiment*. Available at: http://trid.trb.org/view.aspx?id=783477 (Accessed: 27 August 2015).
[66] Driver and Vehicle Standards Agency (2015) *The official highway code*.
[67] Duarte, G. O., Gonçalves, G. A., Baptista, P. C. and Farias, T. L. (2015) 'Establishing bonds between vehicle certification data and real-world vehicle fuel consumption – A Vehicle Specific Power approach', *Energy Conversion and Management*, 92, pp. 251–265. doi: 10.1016/j.enconman.2014.12.042.
[68] Dula, C. and Geller, e S. (2003) 'Risky, Aggressive, or Emotional Driving: Addressing the Need for Consistent Communication in Research', *Journal of Safety Research*, 34, pp. 559 – 566. Available at: http://www.etsu.edu/apl/Documents/APL Materials/Dula%26Geller(2004).pdf.
[69] Duri, S., Gruteser, M., Lin, X., Moskowitz Paul, Perez, R., Singh, M. and Tang, J.-M. (2002) 'Framework for Security and Privacy in Automotive Telematics'.
[70] Eenink, R., Barnard, Y., Baumann, M., Augros, X. and Utesch, F. (2014) 'UDRIVE: the European

naturalistic driving study', *Transport Research Arena*, (2). Available at: http://eprints.whiterose.ac.uk/93078/1/Paper - UDRIVE the European naturalistic driving study %283%29.pdf (Accessed: 17 September 2017).

[71] Eenink, R. and European Commission (2016) *European Commission: Projects and Results* : Periodic Report Summary 2 - UDRIVE (eUropean naturalistic Driving and Riding for Infrastructure & Vehicle safety and Environment). Available at: http://cordis.europa.eu/result/rcn/182466_en.html.

[72] Eidson, J. (2005) 'IEEE-1588 Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems -A Tutorial'. Available at: https://www.nist.gov/sites/default/files/documents/el/isd/IEEE/tutorial-basic.pdf.

[73] Elliott, M. A., Armitage, C. J. and Baughan, C. J. (2003) 'Drivers' compliance with speed limits: An application of the theory of planned behavior.', *Journal of Applied Psychology*, 88(5), pp. 964–972. doi: 10.1037/0021-9010.88.5.964.

[74] Elliott, M. A., Armitage, C. J. and Baughan, C. J. (2007) 'Using the theory of planned behaviour to predict observed driving behaviour', *British Journal of Social Psychology*. Blackwell Publishing Ltd, 46(1), pp. 69–90. doi: 10.1348/014466605X90801.

[75] Elliott, M. A. and Thomson, J. A. (2010) 'The social cognitive determinants of offending drivers' speeding behaviour', *Accident Analysis & Prevention*, 42(6), pp. 1595–1605. doi: 10.1016/j.aap.2010.03.018.

[76] Ellis, P. D. (2010) *The essential guide to effect sizes* : statistical power, meta-analysis, and the interpretation of research results. Cambridge University Press. Available at: https://books.google.be/books/about/The_Essential_Guide_to_Effect_Sizes.html?id=5obZnfK5pbsC&redir_esc=y.

[77] Elvik, R., Christensen, P., Amundsen, A. and The Institute of Transport Economics (TOI) (2004) *Speed and road accidents An evaluation of the Power Model*. Available at: http://www.trg.dk/elvik/740-2004.pdf (Accessed: 13 October 2015).

[78] EOS Gallup Europe (2003) 'Aggressive driving behaviour is a global issue'.

[79] European Commission (2014) *EU transport in figures*. Available at: http://ec.europa.eu/transport/facts-fundings/statistics/doc/2014/pocketbook2014.pdf (Accessed: 17 March 2015).

[80] EuroRAP (2015) *How much do road crashes cost where you live?*

[81] Evans, L. (1979) 'Driver Behavior (behaviour) Effects on Fuel Consumption in Urban Driving', *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 21(4), pp. 389–398. doi: 10.1177/001872087902100401.

[82] Fairchild, R. G., Brake, J. F., Thorpe, N., Birrell, S. a., Young, M. S., Felstead, T. J. and Fowkes, M. (2009) 'Using on-board driver feedback systems to encourage safe, ecological and efficient driving: the Foot-LITE Project', in *AISB 2009 Convention*, pp. 4–7. Available at: http://eprints.soton.ac.uk/73799/.

[83] Faul, F. and Erdfelder, E. (2014) *G*Power 3.1 manual*. Available at: http://www.gpower.hhu.de/fileadmin/redaktion/Fakultaeten/Mathematisch-Naturwissenschaftliche_Fakultaet/Psychologie/AAP/gpower/GPowerManual.pdf (Accessed: 22 March 2016).

[84] Faul, F., Erdfelder, E., Buchner, A. and Lang, A.-G. (2009) 'Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses', *Behavior research methods*, 41(4), pp. 1149–1160. Available at: http://www.gpower.hhu.de/fileadmin/redaktion/Fakultaeten/Mathematisch-Naturwissenschaftliche_Fakultaet/Psychologie/AAP/gpower/GPower31-BRM-Paper.pdf (Accessed: 28 May 2016).

[85] Felstead, T., McDonald, M. and Fowkes, M. (2009) 'Driving style extremes and potential vehicle emission effects', *Proceedings of the ICE - Transport*, 162(3), pp. 141–148. doi: 10.1680/tran.2009.162.3.141.

[86] FESTA Project Consortium (2014) *FESTA Handbook (Field Operational Tests Networking and Methodology Promotion)*. Available at: http://fot-net.eu/wp-content/uploads/sites/7/2014/06/FESTA_HBK_rev5_ref3-2.pdf.

[87] Fildes, b. n., Rumbold, G. and Leening, A. (1991) *Speed Behaviour and Drivers' Attitude to Speeding*. victoria. Available at: http://www.monash.edu/miri/research/reports/muarc016.pdf (Accessed: 13

October 2015).
[88] Fontaras, G., Franco, V., Dilara, P., Martini, G. and Manfredi, U. (2014) 'Development and review of Euro 5 passenger car emission factors based on experimental results over various driving cycles', *Science of the Total Environment*, 468–469, pp. 1034–1042. doi: 10.1016/j.scitotenv.2013.09.043.
[89] Fontaras, G., Zacharof, N.-G. and Ciuffo, B. (2017) 'Fuel consumption and CO 2 emissions from passenger cars in Europe Laboratory versus real-world emissions', *Progress in Energy and Combustion Science*. Elsevier Ltd, 60, pp. 97–131. doi: 10.1016/j.pecs.2016.12.004.
[90] Forbes, N. L. (2009) *BEHAVIOURAL ADAPTATION TO IN-VEHICLE NAVIGATION SYSTEMS*. University of Nottingham.
[91] Frey, H. C. C., Rouphail, N. M. M., Unal, A. and Colyar, J. D. D. (2001) *Emissions Reduction Through Better Traffic Management: An Empirical Evaluation Based Upon On-Road Measurements, FHWY/NC/2002-001, Prepared by NC State University for NC Department of Transportation, Raleigh*. Available at: https://trid.trb.org/view.aspx?id=661327 (Accessed: 23 April 2015).
[92] Frey, H. C., Zhang, K. and Rouphail, N. M. (2008) 'Fuel Use and Emissions Comparisons for Alternative Routes, Time of Day, Road Grade, and Vehicles Based on In-Use Measurements', *Environmental Science & Technology*. American Chemical Society, 42(7), pp. 2483–2489. doi: 10.1021/es702493v.
[93] Friedman, S. and Canaan, M. (2014) *Overcoming speed bumps on the road to telematics* .
[94] Fuller, R. (1984) 'A Conceptualization of Driver Behavior as Threat Avoidance', *Ergonomics*, 27, pp. 1139–1155.
[95] Fuller, R. (2005) 'Towards a General Theory of Driver Behaviour', *Accident; analysis and prevention*, 37(3), pp. 461–72. doi: 10.1016/j.aap.2004.11.003.
[96] Fuller, R., Bates, H., Gormley, M., Hannigan, B., Stradling, S., Broughton, P., Kinnear, N. and O'Dolan, C. (2006) 'Inappropriate High Speed: Who Does it and Why?', in *Behavioural Research in Road Safety 2006 Seminar*, pp. 70–84.
[97] Fuller, R., Hannigan, B., Bates, H., Gormley, M., Stradling, S., Broughton, P., Kinnear, N. and O'Dolan, Ac. (2007) 'Understanding Inappropriate High Speed: Qualitative Results from the HUSSAR Project', in *Behavioural Research in Road Safety 2007 Seventeenth Seminar*, pp. 225–235.
[98] Fulton, L. and Eads, G. (2004) *IEA/SMP Transport Model*. Available at: http://oldwww.wbcsd.org/plugins/DocSearch/details.asp?type=DocDet&ObjectId=MTE0Njc.
[99] Gibson, J. J. and Crooks, L. E. (1938) 'A Theoretical Field-Analysis of Automobile-Driving', *The American Journal of Psychology*, 51(3), pp. 453–471.
[100] Goering, C. E. and Cho, H. (1988) 'Engine model for mapping BSFC contours', *Mathematical and Computer Modelling*. Pergamon, 11, pp. 514–518. doi: 10.1016/0895-7177(88)90546-8.
[101] Goldenbeld, C., Levelt, P. B. M. and Heidstra, J. (2000) 'Psychological perspectives on changing driver attitude and behaviour', *Recherche - Transports - Sécurité*, 67, pp. 65–81. doi: 10.1016/S0761-8980(00)90108-0.
[102] Gonder, J., Earleywine, M. and Sparks, W. (2012) 'Analyzing Vehicle Fuel Saving Opportunities through Intelligent Driver Feedback', *Sae*, (April), pp. 24–26. doi: 10.4271/2012-01-0494.
[103] Graves, F., Lloyd, D., Wilson, D., Mais, D., Deda, W. and Bhagat, A. (2014) *Department for Transport Reported Road Casualties Great Britain: 2013*. Available at: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/359311/rrcgb-2013.pdf (Accessed: 1 September 2015).
[104] Greening, P. (1992) *On-Board Diagnostics for Control of Vehicle Emissions*.
[105] Halim, Z., Kalsoom, R., Bashir, S. and Abbas, G. (2016) 'Artificial intelligence techniques for driving safety and vehicle crash prediction', *Artificial Intelligence Review*. Springer Netherlands, 46(3), pp. 351–387. doi: 10.1007/s10462-016-9467-9.
[106] Hamilton, J. D. (1994) *Time Series Analysis*. Princeton University Press. Available at: https://books.google.com/books?id=B8_1UBmqVUoC&pgis=1 (Accessed: 7 September 2015).
[107] Han, Y., Yao, Y. and Liu, H. (2014) 'Design and Implementation of Driving Behavior Analysis System', *JOURNAL OF NETWORKS*. IEEE, 9(8), pp. 2114–2120. doi: 10.1109/DBTA.2010.5658945.
[108] Handel, P., Skog, I., Wahlstrom, J., Bonawiede, F., Welch, R., Ohlsson, J. and Ohlsson, M. (2014)

'Insurance Telematics: Opportunities and Challenges with the Smartphone Solution', *IEEE Intelligent Transportation Systems Magazine*. IEEE, 6(4), pp. 57–70. doi: 10.1109/MITS.2014.2343262.

[109] Hanowski, R. J., Perez, M. A. and Dingus, T. A. (2005) 'Driver distraction in long-haul truck drivers', *Transportation Research Part F: Traffic Psychology and Behaviour*, 8(6), pp. 441–458. doi: 10.1016/j.trf.2005.08.001.

[110] Hasan, N. N., Arif, A., Pervez, U., Hassam, M. and Husnain, S. S. U. (2011) 'Micro-controller Based On-Board Diagnostic (OBD) System for Non-OBD Vehicles', in *2011 UkSim 13th International Conference on Computer Modelling and Simulation*. IEEE, pp. 540–544. doi: 10.1109/UKSIM.2011.109.

[111] Hashim, I. H. (2011) 'Analysis of speed characteristics for rural two-lane roads: A field study from Minoufiya Governorate, Egypt', *Ain Shams Engineering Journal*. Faculty of Engineering, Ain Shams University, 2(1), pp. 43–52. doi: 10.1016/j.asej.2011.05.005.

[112] Haworth, N. and Symmons (2001) *Drving to reduce fuel consumption and improve road safety*. Available at: http://acrs.org.au/files/arsrpe/RS010036.pdf (Accessed: 26 August 2017).

[113] Ho, S. H., Wong, Y. D. and Chang, V. W. C. (2015) 'What can eco-driving do for sustainable road transport? Perspectives from a city (Singapore) eco-driving programme', *Sustainable Cities and Society*, 14(1), pp. 82–88. doi: 10.1016/j.scs.2014.08.002.

[114] Hong, J.-H., Margines, B. and Dey, A. K. (2014) 'A smartphone-based sensing platform to model aggressive driving behaviors', in *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14*. doi: 10.1145/2556288.2557321.

[115] Hong, J., Zhang, X., Chen, J., Wei, Z., Cao, J. and Ren, Y. (2007) 'Analysis of Time- and Space-domain Sampling for Probe Vehicle-based Traffic Information System', *2007 IEEE Intelligent Transportation Systems Conference*. IEEE, pp. 836–841. doi: 10.1109/ITSC.2007.4357698.

[116] Horberry, T., Anderson, J., Regan, M. A., Triggs, T. J. and Brown, J. (2006) 'Driver distraction: The effects of concurrent in-vehicle tasks, road environment complexity and age on driving performance', *Accident Analysis and Prevention*, 38(1), pp. 185–191. doi: 10.1016/j.aap.2005.09.007.

[117] Hual, T., Durbin, T., Younglove, T., Scora, G., Barth, M. and Norbeck, J. (2005) 'Vehicle Specific Power Approach to Estimating On-Road NH3 Emissions from Light-Duty Vehicles', *Environmental Science & Technology Journal*, 39(24), pp. 9595–9600. Available at: http://cichlid.cert.ucr.edu/research/pubs/6.pdf (Accessed: 23 April 2015).

[118] Husnjak, S., Peraković, D., Forenhacher, I. and Mumdziev, M. (2015) 'Telematics System in Usage Based Motor Insurance', *Procedia Engineering*, 100, pp. 816 – 825.

[119] Ichihara, T., Kumano, S., Yamaguchi, D., Sato, Y. and Suda, Y. (2009) 'Driver assistance system for eco-driving', in *16th Intelligent Transport Systems World Congress*, p. 8.

[120] Inata, K., Raksincharoensak, P. and Nagai, M. (2008) 'Driver Behavior Modeling Based on Database of Personal Mobility Driving in Urban Area', *2008 International Conference on Control, Automation and Systems*. IEEE, pp. 2902–2907. doi: 10.1109/ICCAS.2008.4694252.

[121] Intel (2015) *Designing Next-Generation Telematics Solutions*.

[122] International Energy Agency (2012) 'CO2 Emissions from fuel Combustion', *IEA Statistics*, p. 138. doi: 10.1787/co2-table-2011-1-en.

[123] Ippisch, T. (2010) *Telematics Data in Motor Insurance: Creating Value by Understanding the Impact of Accidents on Vehicle Use*.

[124] Iversen, H. (2004) 'Risk-taking Attitudes and Risky Driving Behaviour', *Transportation Research Part F: Traffic Psychology and Behaviour*, 7(3), pp. 135–150. doi: 10.1016/j.trf.2003.11.003.

[125] Jamson, S. L., Hibberd, D. L. and Jamson, A. H. (2015) 'Drivers' ability to learn eco-driving skills; effects on fuel efficient and safe driving behaviour', *Transportation Research Part C: Emerging Technologies*, 58(PD), pp. 657 – 668. doi: 10.1016/j.trc.2015.02.004.

[126] Jaramillo, D. M. G. and Narvaez, C. V. I. (2014) 'Vehicle Online Monitoring Sytem Based on Fuzzy Classifier', in *VEHICULAR 2014, The Third International Conference on Advances in Vehicular Systems, Technologies and Applications*, pp. 33–38. Available at: http://www.thinkmind.org/index.php?view=article&articleid=vehicular_2014_2_30_30039 (Accessed: 11 November 2014).

[127] Jiménez-Palacios, J. L. (1999) *Understanding and Quantifying Motor Vehicle Emissions with Vehicle Specific Power and TILDAS Remote Sensing*. Massachusetts Institute of Technology. Available at: http://cires1.colorado.edu/jimenez/Papers/Jimenez_PhD_Thesis.pdf (Accessed: 28 July 2015).

[128] Jiménez, J. L., Mcclintock, P., Nelson, D. D. and Zahniser, M. S. (1999) 'Vehicle Specific Power : A Useful Parameter for Remote Sensing and Emission Studies Effect of Driving Conditions on Emissions • Driving conditions may strongly influence emissions', *9th CRC On-Road Vehicle Emissions Workshop*, p. 26.

[129] Johansson, H., Färnlund, J. and Engström, C. (1999) *Impact of Eco Driving on Emissions and Fuel Consumption*.

[130] Johnston, I. R. and Perry, D. R. (1980) *Driver Behavior Research Needs and Priorities*. Victoria, Australia.

[131] Joumard, R., André, M., Laurikko, J., Anh, T. L. E., Geivanidis, S., Samaras, Z., Oláh, Z., Devaux, P., André, J., Cornelis, E., Lacour, S., Prati, M. V., Vermeulen, R. and Zallinger, M. (2006) *Accuracy of Exhaust Emissions Measurements on Vehicle Bench*.

[132] Joumard, R., André, M., Vidon, R., Tassel, P. and Pruvost, C. (2000) 'Influence of driving cycles on unit emissions from passenger cars', *Atmospheric Environment*, 34(27), pp. 4621–4628. doi: 10.1016/S1352-2310(00)00118-7.

[133] Jubraj, P. and Moneta, A. (2014) *Insurance telematics: A game-changing opportunity for the industry*.

[134] Karapiperis, D., Birnbaum, B., Brandenburg, A., Castagna, S., Greenberg, A., Harbage, R. and Obersteadt, A. (2015) *Usage-Based Insurance and Vehicle Telematics: Insurance Market and Regulatory Implications*. Available at: http://www.naic.org/documents/cipr_study_150324_usage_based_insurance_and_vehicle_telematics_study_series.pdf (Accessed: 20 May 2004).

[135] Kawashima, A., Kobayashi, K. and Watanabe, K. (2001) 'Implementation of Human-Like Driving for Autonomous Vehicle', in. doi: 10.4271/2001-01-0805.

[136] Kelly, N. A. and Groblicki, P. J. (1993) 'Real-World Emissions from a Modern Production Vehicle Driven in Los Angeles', *The Air & Waste Management Association*. Taylor & Francis Group, 43(10), pp. 1351–1357. doi: 10.1080/1073161X.1993.10467209.

[137] Kidd, E. A. and Laughery, K. R. (1964) *A Computer Model of Driving Behavior: The Highway Intersection Situation*. Buffalo, NY.

[138] Kiencke, U., Majjad, R. and Kramer, S. (1999) 'Modeling and performance analysis of a hybrid driver model', *Control Engineering Practice*, 7(8), pp. 985–991. doi: 10.1016/S0967-0661(99)00053-2.

[139] Klauer, S. G. (2005) *Assessing The Effects of Driving Inattention on Relative Crash Risk*.

[140] Klauer, S. G., Dingus, T. A., Neale, T. A., Sudweeks, J. D. and Ramsey, D. J. (2006) *The Impact of Driver Inattention On Near-Crash / Crash Risk* : An Analysis Using the 100-Car Naturalistic Driving Study Data.

[141] Kloeden, C., McLean, A., Moore, V. and G, P. (1997) *FEDERAL OFFICE OF ROAD SAFETY DOCUMENT RETRIEVAL INFORMATION Title and Subtitle Travelling Speed and the Risk of Crash Involvement Volume 1: Findings*. Available at: https://infrastructure.gov.au/roads/safety/publications/1997/pdf/Speed_Risk_1.pdf.

[142] Koh, D. W. and Kang, H. B. (2015) 'Smartphone-based modeling and detection of aggressiveness reactions in senior drivers', in *IEEE Intelligent Vehicles Symposium, Proceedings*. doi: 10.1109/IVS.2015.7225655.

[143] Kono, T., Fushiki, T., Asada, K. and Nakano, K. (2008) 'Fuel Consumption Analysis and Prediction Model for "Eco" Route Search', in *In 15th World Congress on Intelligent Transport Systems and ITS America's*, pp. 1–9.

[144] Kruse, R. E. and Huls, T. A. (1973) 'Development of the Federal Urban Driving Schedule', in. doi: 10.4271/730553.

[145] Kuhn, M. and Johnson, K. (2016) *Applied Predictive Modeling*. Springer. doi: 10.1007/978-1-4614-6849-3 ISBN.

[146] Laurell, H. (1985) *The Impact of Driving Style on Fuel Consumption (in Swedish: Körsättets betydelse för bränsleförbrukningen)*.

[147] LeBlanc, D. J., Sivak, M. and Bogard, S. (2010) *Using Naturalistic Driving Data to Assess Variations in Fuel Efficiency among Individual Drivers*.

[148] Lee, J. D., Moeckli, J., Brown, T., Roberts, S. and Schwarz, C. (2013) *A comparative approach to identify unsafe older drivers.* doi: 10.1177/0018720812462029.

[149] Lenth, R. V (2001) 'Some Practical Guidelines for Effective Sample Size Determination', *The American Statistician*, 55(3), pp. 187–193. doi: 10.1198/000313001317098149.

[150] Li, H., Andrews, G. E., Khan, A. A., Savvidis, D., Daham, B., Bell, M., Tate, J. and Ropkins, K. (2007) 'Analysis of Driving Parameters and Emissions for Real World Urban Driving Cycles using an on-board Measurement Method for a EURO 2 SI car', *Water*, pp. 862–872. doi: 10.4271/2007-01-2066.

[151] Li, W., Chuanjiu, W., Xiaorong, S. and Yuezu, F. A. N. (2005) 'Probe vehicle sampling for real-time traffic data collection', *Proceedings. 2005 IEEE Intelligent Transportation Systems, 2005.* IEEE, (1), pp. 886–888. doi: 10.1109/ITSC.2005.1520167.

[152] Loo, B. P. Y. and Anderson, T. K. (2015) *Spatial Analysis Methods of Road Traffic Collisions*. CRC Press. Available at: https://books.google.com/books?id=ysB5CgAAQBAJ&pgis=1 (Accessed: 13 March 2016).

[153] Lord, D., Mannering, F. and Pankow, C. (2010) 'The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives', *Forthcoming in Transportation Research, Part A*, 44(5), pp. 291–305. doi: 10.1016/j.tra.2010.02.001.

[154] Luke, R. and Heyns, G. J. (2014) 'Reducing risky driver behaviour through the implementation of a driver risk management system', *Journal of Transport and Supply Chain Management*, 8(1), p. 10 pages. doi: 10.4102/jtscm.v8i1.146.

[155] Ma, H., Xie, H., Chen, S., Yan, Y. and Huang, D. (2014) *Effects of Driver Acceleration Behavior on Fuel Consumption of City Buses*. doi: 10.4271/2014-01-0389.

[156] Madireddy, M., De Coensel, B., Can, A., Degraeuwe, B., Beusen, B., De Vlieger, I. and Botteldooren, D. (2011) 'Assessment of the impact of speed limit reduction and traffic signal coordination on vehicle emissions using an integrated approach', *Transportation Research Part D: Transport and Environment*. Elsevier Ltd, 16(7), pp. 504–508. doi: 10.1016/j.trd.2011.06.001.

[157] Malek, S., Brace, C. and Liu, S. (2012) 'Effect of driving behaviour on fuel consumption', in *International Conference on Driver Behaviour and Training, 5th, 2011, Paris, France*. Available at: http://trid.trb.org/view.aspx?id=1236933 (Accessed: 16 January 2015).

[158] Mann, H. B. and Whitney, D. R. (1947) 'On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other', *The Annals of Mathematical Statistics*. Institute of Mathematical Statistics, 18(1), pp. 50–60. doi: 10.1214/aoms/1177730491.

[159] Martens, M., Kaptein, N., Claessens, F. and Tinok van Hattum, S. (1998) 'Road Design, Cognitive Road Classification and Driving Behaviour', 1.

[160] Mathew, T. V. (2014) 'Fuel Consumption and Emission Studies', in *Transportation Systems Engineering*. Available at: https://www.civil.iitb.ac.in/tvm/1111_nptel/583_FuelEmi/fancy.pdf (Accessed: 15 April 2016).

[161] May, O., Bosteels, D. and Favre, C. (2014) 'An Assessment of Emissions from Light-Duty Vehicles using PEMS and Chassis Dynamometer Testing', *SAE Int. J. Engines*, 7(3), pp. 1326–1335. Available at: http://digitallibrary.sae.org.ezproxy1.bath.ac.uk/content/2014-01-1581.

[162] McCarthy, P. (2001) 'Effect of speed limits on speed distributions and highway safety: a survey of recent literature', *Transport Reviews*. Taylor & Francis Group, 21(1), pp. 31–50. doi: 10.1080/014416400750059275.

[163] McEvoy, S. P., Stevenson, M. R. and Woodward, M. (2006) 'The impact of driver distraction on road safety: results from a representative survey in two Australian states.', *Injury prevention* : journal of the International Society for Child and Adolescent Injury Prevention, 12(4), pp. 242–7. doi: 10.1136/ip.2006.012336.

[164] McGordon, A., Poxon, J. E. W., Cheng, C., Jones, R. P. and Jennings, P. A. (2011) 'Development of a driver model to study the effects of real-world driver behaviour on the fuel consumption', *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, 225(11), pp. 1518–1530. doi: 10.1177/0954407011409116.

[165] McLaughlin, S., Hankey, J. and Dingus, T. A. (2009) 'Driver Measurement: Methods and Applications', in *Engineering Psychology and Cognitive Ergonomics*. Springer-Verlag Berlin Heidelberg, pp. 404–413.

[166] McNally, M. G., Marca, J. E., Rindt, C. R. and Koos, A. M. (2003) *TRACER: In-vehicle, GPS-based, Wireless Technology for Traffic Surveillance and Management California PATH Research Report*.

[167] McRuer, D. T., Allen, R., Weir, D. H. and Klein, R. (1977) 'New Results in Driver Steering Control Models', *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 19, pp. 381–397.

[168] McTavish, N. and BP Castrol Technology Centre (2008) 'Enhancing Fuel Economy in Transport Applications', in *Technical and maintenance conference*.

[169] Mejuto, P. (2016) *How does the ecoDriver assistance affect the driver?* Available at: https://goo.gl/qNLqqN.

[170] Mensing, F., Bideaux, E., Trigui, R., Ribet, J. and Jeanneret, B. (2014) 'Eco-driving: An economic or ecologic driving style?', *Transportation Research Part C: Emerging Technologies*, 38, pp. 110–121. doi: 10.1016/j.trc.2013.10.013.

[171] Meseguer, J. E., Calafate, C. T., Cano, J. C. and Manzoni, P. (2013) 'DrivingStyles: A smartphone application to assess driver behavior', in *Proceedings - International Symposium on Computers and Communications*. doi: 10.1109/ISCC.2013.6755001.

[172] Michon, J. A. (1985) 'A Critical View of Driver Behavior Models: What Do We Know, What Should We Do?', in *Human Behaviour and traffic safety*, pp. 485–520. Available at: http://www.jamichon.nl/jam_writings/1985_criticial_view.pdf (Accessed: 20 May 2014).

[173] Van Mierlo, J., Maggetto, G., Van de Burgwal, E. and Gense, R. (2004) 'Driving Style and Traffic Measures—Influence on Vehicle Emissions and Fuel consumption', *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, 218(1), pp. 43–50. doi: 10.1243/095440704322829155.

[174] Mierlo, J. Van, Maggetto, G., Burgwal, E. Van De and Gense, R. (2004) 'Driving style and traffic measures, influence on vehicle emissions and fuel consumption', 218, pp. 43–50.

[175] Miyajima, C., Nishiwaki, Y., Ozawa, K., Wakita, T., Itou, K., Takeda, K. and Itakura, F. (2007) 'Driver Modeling Based on Driving Behavior and Its Evaluation in Driver Identification', *Proceedings of the IEEE*, 95(2), pp. 427–437. doi: 10.1109/JPROC.2006.888405.

[176] De Mol, J., Vlassenroot, S., Zwerts, E., Allaert, G. and Witlox, F. (2009) 'The evolution of car power, weight and top speed during the last twenty years in Belgium: a consideration for future policies', *3rd Transport Research Day*. VUB Press, pp. 607–620. Available at: https://biblio.ugent.be/publication/669979/file/6818612.pdf (Accessed: 16 March 2016).

[177] Molen, H. H. and Botticher, A. M. T. (1988) 'A Hierarchical Risk Model for Traffic Participants', *Ergonomics*, 31, pp. 537–555.

[178] Montazeri-Gh, M. and Naghizadeh, M. (2003) 'Development of Car Driver cycle for Simulation of Emissions and Economy', in *15th European Simulation Symposium*, pp. 1–6.

[179] Montella, A. (2010) 'A comparative analysis of hotspot identification methods.', *Accident; analysis and prevention*, 42(2), pp. 571–81. doi: 10.1016/j.aap.2009.09.025.

[180] Mujanovic, A. (2011) *Influence on driving style and fuel consumption with powertrain calibration*. Available at: http://publications.lib.chalmers.se/records/fulltext/153110.pdf (Accessed: 10 December 2015).

[181] Näätänen, R. and Summala, H. (1976) 'Road User Behavior and Traffic Accidents'. Amsterdam: North Holland Publishing Company.

[182] Nam, E. K. (2004) *Advanced Technology Vehicle Modeling in PERE*.

[183] Nam, E. K., Gierczak, C. A., Butler, J. W. and Ford Scientific Research Laboratory (2003) *A Comparison Of Real-World and Modeled Emissions Under Conditions of Variable Driver Aggressiveness*, *TRB Annual Meeting*. Available at: http://www.ltrc.lsu.edu/TRB_82/TRB2003-002236.pdf (Accessed: 10 February 2015).

[184] Naranjo, J. E., Gonzalez, C., Garcia, R. and de Pedro, T. (2008) 'Lane-Change Fuzzy Control in Autonomous Vehicles for the Overtaking Maneuver', *IEEE Transactions on Intelligent Transportation*

Systems, 9(3), pp. 438–450. doi: 10.1109/TITS.2008.922880.
[185] Neale, V. L., Dingus, T. A., Klauer, S. G., Sudweeks, J. and Goodman, M. (2005) 'An overview of the 100-car naturalistic driving study and findings', *Proceedings of the 19th International Technical Conference on Enhanced Safety of Vehicles*, (05–0400), pp. 1–10.
[186] NHTSA (2011) *Countermeasures that work: A highway safety countermeasure guide for state highway safety offices*. Washington, DC.
[187] *NIST/SEMATECH e-Handbook of Statistical Methods* (2012). Available at: http://www.itl.nist.gov/div898/handbook/pri/section2/pri24.htm.
[188] Noland, R. B., Ochieng, W. Y., Quddus, M. A., North, R. J., Polak, J. W. and Centre for Transport Studies Dept. of Civil and Environmental Engineering Imperial College London (2004) 'The Vehicle Emissions and Performance Monitoring System: Analysis of Tailpipe Emissions and Vehicle Performance', *Transportation Planning and Technology*, 27(6), pp. 431–447. Available at: https://dspace.lboro.ac.uk/dspace-jspui/bitstream/2134/4846/1/Noland.pdf (Accessed: 23 April 2015).
[189] Norris, F. H., Matthews, B. A. and Riad, J. K. (2000) 'Characterological, situational, and behavioral risk factors for motor vehicle accidents: a prospective examination', *Accident Analysis & Prevention*, 32(4), pp. 505–515. doi: 10.1016/S0001-4575(99)00068-8.
[190] O'Neill, B. (1977) 'A Decision-theory Model of Danger Compensation', *Accident Analysis and Prevention*, 9, pp. 157–165.
[191] Olivier, J. G. J., Muntean, M. and Peters, J. A. H. W. (2015) 'Trends in global CO2 emissions: 2015 report', *PBL Netherlands Environmental Assessment Agency & European Commission's Joint Research Centre (JRC)*, pp. 1–78.
[192] Olsen, E. C. B. and Wierwille, W. W. (2001) 'A Unique Approach for Data Analysis of Naturalistic Driver Behavior', *SAE International*. doi: 10.4271/2001-01-2518.
[193] Onoda, T. and Gueret, T. (2007) *Fuel Efficient Road Vehicle None-Engine Components*.
[194] Osborne, L., Brummond, J., Hart, R., Zarean, M. and Conger, S. (2005) *Clarus: Concept of Operations*.
[195] Owsley, C., Ball, K., Sloane, M. E., Roenker, D. L. and Bruni, J. R. (1991) 'Visual/Cognitive Correlates of Vehicle Accidents in Older Drivers', *Psychology and Aging*, 6, pp. 403–415.
[196] Paefgen, J., Staake, T. and Thiesse, F. (2013) 'Evaluation and aggregation of pay-as-you-drive insurance rate factors: A classification analysis approach', *Decision Support Systems*, 56, pp. 192–201. doi: 10.1016/j.dss.2013.06.001.
[197] Parry, M. H. (1968) *Aggression on the Road: a pilot study of behaviour in the driving situation*. Tavistock Publications.
[198] Pelkmans, L. and Debal, P. (2006) 'Comparison of On-road Emissions with Emissions Measured on Chassis Dynamometer Test Cycles', *Transportation Research Part D: Transport and Environment*, 11(4), pp. 233–241. doi: 10.1016/j.trd.2006.04.001.
[199] Pentland, A. and Liu, A. (1999) 'Modeling and Prediction of Human Behavior', *Neural computation*, 11(1), pp. 229–42. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9950731.
[200] Potter, C. J. and Savage, A. C. (1982) *The evaluation of the Warren Spring Laboratory vehicle exhaust gas proportional*.
[201] Quimby, A. R. and Department for Transport (2005) 'UK and European drivers on speed and speeding issues: some results from SARTRE 3 survey', in *Behavioural Research in Road Safety 2005* : Fifteenth Seminar, pp. 49–67.
[202] Quimby, a, Maycock, G., Palmer, C. and Buttress, S. (1999) 'The factors that influence a drivers choice of speed'. Available at: http://swov.summon.serialssolutions.com/2.0.0/link/0/eLvHCXMwY2AwNtIz0EUrE1KBFRvoZC5D0CKMxEQTU6Mkk7RkozRga9ks2TQZtP3Z1cnMK9IkLMg8BKmAdxNiYErNE2WQdHMNcfbQLS7PL4uHjmvEA3sKlsaGYgwswH5yqgSDgqkFsD-RbGqabGqeYpJmkAhsOwLbAhZm5omgU-KMDQDpKiRk%5Cnwww.summon.com%5Cn.
[203] Rakha, H., Du, J., Park, S., Guo, F., Doerzaph, Z., Viita, D., Golembiewski, G., Katz, B., Kehoe, N. and Rigdon, H. (2011) *Feasibility of Using In-Vehicle Video Data to Explore How to Modify Driver Behavior That Causes Nonrecurring Congestion*.

[204] Ranney, T. A. (1994) 'Models of Driving Behavior: a Review of their Evolution', *Accident Analysis & Prevention*, 26(6), pp. 733–50. Available at: http://www.ncbi.nlm.nih.gov/pubmed/7857489.

[205] Rapone, M., Ragione, L. Della, Prati, M. V and Violetti, N. (2000) 'Driving Behavior and Emission Results for a Small Size Gasoline Car in Urban Operation', *SAE Technical paper, International Fall Fuels and Lubricants Meeting and Exposition*, (724).

[206] Reason, J. T., Manstead, A., Stradling, S., Baxter, J. and Campbell, K. (1990) 'Errors and Violations on the Roads: A Real Distinction?', *Ergonomics*, 33, pp. 1315–1332.

[207] Redsell, M., Lucas, G. G. and Ashford, N. J. (1988) 'Comparison of On-road Fuel Consumption For Diesel and Petrol Cars', *Transport and Road Research Laboratory (TRRL)*. Available at: http://trid.trb.org/view.aspx?id=295251 (Accessed: 24 January 2016).

[208] Redsell, M., Lucas, G. G. and Ashford, N. J. (1993) 'Factors Affecting Car Fuel Consumption', *Journal of Automobile Engineering*, 207(1), pp. 1–23. doi: 10.1243/PIME.

[209] Reed, N. (2007) 'Assessing Fuel-efficient Driver Behaviour Through Tachograph Information', in *Behavioural Research in Road Safety 2007 Seventeenth Seminar*, pp. 191–20.

[210] Regan, M. A., Hallett, C. and Gordon, C. P. (2011) 'Driver distraction and driver inattention: definition, relationship and taxonomy.', *Accident; analysis and prevention*, 43(5), pp. 1771–81. doi: 10.1016/j.aap.2011.04.008.

[211] Regan, M. a, Williamson, A., Grzebieta, R. and Tao, L. (2012) 'Naturalistic Driving Studies: Literature Review and Planning for the Australian Naturalistic Driving Study', *Australasian College of Road Safety National Conference*, pp. 1–13. Available at: http://acrs.org.au/wp-content/uploads/57_Regan-PR.pdf.

[212] Robert Bosch GmbH (2011) *Automotive Handbook*. 8th edn. Edited by K. Dietsche. Robert Bosch GmbH.

[213] Roberts, A., Brooks, R. and Shipway, P. (2014) 'Internal combustion engine cold-start efficiency: A review of the problem, causes and potential solutions', *Energy Conversion and Management*, 82, pp. 327–350. doi: 10.1016/j.enconman.2014.03.002.

[214] Rolim, C., Baptista, P., Duarte, G., Farias, T. and Shiftan, Y. (2014) 'Quantification of the impacts of eco-driving training and real-time feedback on Urban buses driver's behaviour', in *Transportation Research Procedia*, pp. 70–79. doi: 10.1016/j.trpro.2014.10.092.

[215] Ron, U. (2011) *GreenRoad Optimizes Fleet Driver Performance, Safety Practices and Fuel Consumption with New Features - Greenroad*, *GreenRoad press release* . Available at: https://greenroad.com/greenroad-optimizes-fleet-driver-performance-safety-practices-and-fuel-consumption-with-new-features/.

[216] Sa, J. S., Chung, N. H. and Sunwoo, M. H. (2003) 'Experimental Analysis of Driving Patterns and Fuel Economy for Passenger Cars in Seoul', *International Journal of Automotive Technology*, 4(2), pp. 101–108.

[217] Sadigh, D., Campbell, K. D., Alessandro, A., Puggelli, A., Li, W., Shia, V., Bajcsy, R., Sangiovanni-Vincentelli, A. L., Shankar, S., Sanjit, S., Seshia, A., Sadigh, D., Driggs-Campbell, K., Puggelli, A., Li, W., Shia, V., Bajcsy, R., Sangiovanni-Vincentelli, A. L., Sastry, S. S. and Seshia, S. A. (2013) 'Data-Driven Probabilistic Modeling and Verification of Human Driver Behavior'. Available at: http://www.eecs.berkeley.edu/Pubs/TechRpts/2013/EECS-2013-197.html (Accessed: 29 November 2016).

[218] Saffarzadeh, M. and Arjroody, A. R. (2003) 'An Optimum Fuel Consumption Model for the Transport of Cargoes by Road Vehicles', in *International Association for Energy Economics*. Available at: http://www.iaee.org/documents/Prague/p03saffarzadeh.pdf (Accessed: 8 August 2014).

[219] Saiprasert, C. and Pattara-Atikom, W. (2013) 'Smartphone enabled dangerous driving report system', in *Proceedings of the Annual Hawaii International Conference on System Sciences*. doi: 10.1109/HICSS.2013.484.

[220] Salter, R. J. (1974) 'The distribution of vehicular speeds in a highway traffic stream', in *Highway Traffic Analysis and Design*. London: Macmillan Education UK, pp. 135–144. doi: 10.1007/978-1-349-02093-5_14.

[221] Van Schangen, I., Welsh, R., Backer-Grondahl, A., Hoedemaeker, M., Lotan, T., Morris, A., Sagberg,

F. and Winkebauer, M. (2011) *Towards a Large-scale European Naturalistic Driving Study* : Final Report of PROLOGUE.

[222] Schipper, L. (2011) 'Automobile use, fuel economy and CO2 emissions in industrialized countries: Encouraging trends through 2008?', *Transport Policy*, 18(2), pp. 358–372. doi: 10.1016/j.tranpol.2010.10.011.

[223] Sentoff, K. M., Aultman-Hall, L. and Holmén, B. A. (2015) 'Implications of driving style and road grade for accurate vehicle activity data and emissions estimates', *Transportation Research Part D: Transport and Environment*, 35, pp. 175–188. doi: 10.1016/j.trd.2014.11.021.

[224] Shinar, D. (1998) 'Aggressive Driving: The Contribution of the Drivers and the Situation', *Transportation Research Part F Traffic Psychology and Behaviour*, pp. 137–160. Available at: https://www.researchgate.net/profile/David_Shinar/publication/222351338_Aggressive_Driving_the_Contribution_of_the_Drivers_and_the_Situation/links/0f317534f9e8cd4921000000.pdf (Accessed: 12 April 2016).

[225] Sivak, M. and Schoettle, B. (2011) 'Eco-Driving: Strategic, tactical, and operational decisions of the driver that improve vehicle fuel economy', *University of michigan*, 3(August), pp. 483–496.

[226] Skog, I. and Handel, P. (2014) *Indirect Instantaneous Car-Fuel Consumption Measurements*, *IEEE Transactions on Instrumentation and Measurement*. doi: 10.1109/TIM.2014.2315739.

[227] Sodré, J. R. and Soares, S. M. C. (2003) 'Comparison of Engine Power Correction Factors for Varying Atmospheric Conditions', *ABCM*, XXV(3), pp. 279–285. Available at: http://www.scielo.br/pdf/jbsms/v25n3/a10v25n3 (Accessed: 10 February 2015).

[228] Souleyrette, R., Stout, T. and Carriquiry, A. (2009) *Evaluation of Iowa's 70 mph Speed Limit — 2.5 Year Update*. Available at: http://www.ctre.iastate.edu/reports/70mph_speed.pdf.

[229] Stradling, S., Gormley, M., Fuller, R., Broughton, P., Kinnear, N., O'Dolan, C. and Hannigan, B. (2007) 'A typology of speeding drivers: extent of, and motives for, exceeding the speed limit', in *Behavioural Research in Road Safety 2007 Seventeenth Seminar*, pp. 236–252.

[230] Strecher, V. J., Shope, J., Bauermeister, J. A., Chang, C., Newport-Berra, M., Boonin, A., Giroux, A., Guay, E., Candee, E., Ewing, L., West, R., McMahon, K., Forsyth, E., O'reilly, D., Fuller, R., Corbett, C., Read, L. and Pidgeon, N. (2006) 'Predictors of Safe Driving Behaviour: Towards an Integrative Model', in *Behavioural Research in Road Safety 2006 Seminar*, pp. 39–48.

[231] Summala, H. (1985) 'Modeling driver behavior: A pessimistic prediction?', in Evans, L. and Schwing, R. C. (eds) *Human behavior and traffic safety*. New York: Plenum Press.

[232] Summala, H. (1988) 'Risk Control is Not Risk Adjustment: The Zero-risk Theory of Driver Behavior and its Implications', *Ergonomics*, 31, pp. 491–506.

[233] Tabachnick, B. G. and Fidell, L. S. (2012) *Using Multivariate Statistics*. Pearson Education. Available at: http://books.google.com/books?id=ucj1ygAACAAJ&pgis=1 (Accessed: 15 July 2014).

[234] Takano, W., Matsushita, A., Iwao, K. and Nakamura, Y. (2008) 'Recognition of human driving behaviors based on stochastic symbolization of time series signal', in *IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, pp. 167–172. doi: 10.1109/IROS.2008.4650671.

[235] Tasca, L. (2000) *A Review of the Literature on Aggressive Driving Research*. Available at: http://www.stopandgo.org/research/aggressive/tasca.pdf.

[236] Taylor, D. H. (1964) 'Drivers' Galvanic Skin Response and the Risk of Accident', *Ergonomics*, 7, pp. 439–451.

[237] Taylor, M. C., Lynam, D. A. and Baruya, A. (2000) 'The effects of drivers' speed on the frequency of road accidents'. Available at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.476.7564&rep=rep1&type=pdf (Accessed: 26 August 2017).

[238] Thadewald, T. and Büning, H. (2007) 'Jarque–Bera Test and its Competitors for Testing Normality – A Power Comparison', *Journal of Applied Statistics*, 34(1), pp. 87–105. doi: 10.1080/02664760600994539.

[239] Thatcham Research (2013) *The Group Rating System How the System Works*.

[240] The Association of British Insurers (ABI) (2015) *Motor Insurance (Consumer) Market Reports*. Available at: file:///C:/Users/Sahand/Downloads/Motor_Insurance_(Consumer)_70945.pdf (Accessed:

1 July 2015).
[241] The Goodyear Tire & Rubber Company (2012) *Factors Affecting Truck Fuel Economy*.
[242] The International Council on Clean Transportation (2013) *Reducing CO2 and Fuel Consumption from New Cars: Assessing the Near-term Technology Potential in the EU*. Available at: http://www.theicct.org/sites/default/files/publications/Briefing Technology Potential Long EN v3.pdf (Accessed: 6 April 2015).
[243] The International Council on Clean Transportation (2015) *CO2 emissions from new passenger cars in the EU: Car manufacturers' performance in 2014*. Available at: http://www.theicct.org/sites/default/files/publications/ICCTbriefing_EU-CO2_201507.pdf (Accessed: 5 April 2016).
[244] The International Council on Clean Transportation Europe (2015) *European Vehicle Market Statistics*. Available at: http://www.theicct.org/sites/default/files/publications/ICCT_EU-pocketbook_2015.pdf (Accessed: 2 February 2016).
[245] Tillmann, W. A. and Hobbs, G. E. (1949) 'The Accident-prone Automobile Driver; a Study of the Psychiatric and Social Background', *The American journal of psychiatry*. American Psychiatric Publishing, 106(5), pp. 321–31. doi: 10.1176/ajp.106.5.321.
[246] Toledo, G., Shiftan, Y. and Hakkert, S. (2007) 'Framework for Analysis and Modeling of Driving Behavior Icorporating in vehicle data recorders'.
[247] Toledo, T. (2003) *Integrated Driving Behavior Modeling*, *Massachusetts Institute of Technology*.
[248] Toledo, T. and Lotan, T. (2006) 'In-Vehicle Data Recorder for Evaluation of Driving Behavior and Safety', *Journal of the Transportation Research Board*, 1953, pp. 112–119. doi: http://dx.doi.org/10.3141/1953-13.
[249] Tong, D., Merry, C. J. and Coifman, B. (2005) *Traffic Information Deriving Using GPS Probe Vehicle Data Integrated with GIS*.
[250] Tong, S., Lloyd, L., Durrell, L., Mcrae-Mckee, K. and Lloyd, L. (2015) *Provision of telematics research*.
[251] Troncoso, C., Danezis, G., Kosta, E., Balasch, J. and Preneel, B. (2011) 'PriPAYD: Privacy-Friendly Pay-As-You-Drive Insurance', *IEEE Transactions on Dependable and Secure Computing*. IEEE, 8(5), pp. 742–755. doi: 10.1109/TDSC.2010.71.
[252] True, H. (2003) *The dynamics of vehicles on roads and on tracks*. Swets & Zeitlinger.
[253] Tsay, R. S. (2010) *Analysis of Financial Time Series*. John Wiley & Sons. Available at: https://books.google.be/books?id=OKUGARAXKMwC&dq=skewness&source=gbs_book_similarbooks (Accessed: 13 January 2016).
[254] Tsay, R. S. (2013) *Multivariate time series analysis* : with R and financial applications. Available at: https://books.google.be/books?id=A4QVAgAAQBAJ&source=gbs_similarbooks.
[255] Tselentis, D. I., Yannis, G. and Vlahogianni, E. I. (2016) 'Innovative insurance schemes: pay as/how you drive', *Transportation Research Procedia*, 14, pp. 362–371. doi: 10.1016/j.trpro.2016.05.088.
[256] Tutuianu, M., Marotta, A., Steven, H., Ericsson, E., Haniu, T., Ichikawa, N. and Ishii, H. (2013) *Development of a World-wide Worldwide harmonized Light duty driving Test Cycle (WLTC)*. Available at: https://www.unece.org/fileadmin/DAM/trans/doc/2014/wp29grpe/GRPE-68-03e.pdf.
[257] Unal, A. (2002) *On-board Measurement and Analysis of On-Road Vehicle Emission*, *North Carolina State University*. Available at: http://repository.lib.ncsu.edu/ir/bitstream/1840.16/4901/1/etd.pdf (Accessed: 23 April 2015).
[258] US Environmental Protection Agency (1993) *Federal Test Procedure Review Project: Preliminary Technical Report*. Available at: http://www.epa.gov/oms/regs/ld-hwy/ftp-rev/ftp-tech.pdf (Accessed: 23 April 2015).
[259] De Vlieger, I. (1997) 'On-board Emission and Fuel Consumption Measurement Campaign on Petrol-driven Passenger Cars', *Atmospheric Environment*, 31(22), pp. 3753–3761. doi: 10.1016/S1352-2310(97)00212-4.
[260] De Vlieger, I., De Keukeleere, D. and Kretzschmar, J. . (2000) 'Environmental effects of driving behaviour and congestion related to passenger cars', *Atmospheric Environment*, 34(27), pp. 4649–4655. doi: 10.1016/S1352-2310(00)00217-X.

[261] Van der Voort, M., Dougherty, M. S. and van Maarseveen, M. (2001) 'A Prototype Fuel-Efficiency Support Tool', *Transportation Research Part C: Emerging Technologies*, 9(4), pp. 279–296. doi: 10.1016/S0968-090X(00)00038-3.

[262] Wåhlberg, A. E. af (2003) 'Some Methodological Deficiencies in Studies on Traffic Accident Predictors', *Accident; analysis and prevention*, 35(4), pp. 473–86. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12729811 (Accessed: 4 July 2014).

[263] Wåhlberg, A. E. af (2006) 'Short-term Effects of Training in Economical Driving: Passenger Comfort and Driver Acceleration Behavior', *International Journal of Industrial Ergonomics*, 36(2), pp. 151–163. Available at: http://www.sciencedirect.com/science/article/pii/S0169814105001496 (Accessed: 8 January 2014).

[264] Wåhlberg, A. E. af (2007) 'Long-term Effects of Training in Economical Driving: Fuel Consumption, Accidents, Driver Acceleration Behavior and Technical Feedback', *International Journal of Industrial Ergonomics*, 37(4), pp. 333–343. doi: 10.1016/j.ergon.2006.12.003.

[265] Wåhlberg, A. E. af (2008) 'Driver Celeration Behaviour and Accidents – an analysis', *Theoretical Issues in Ergonomics Science*. Taylor & Francis Group, 9(5), pp. 383–403. doi: 10.1080/14639220701596722.

[266] Wåhlberg, A. E. af (2009) *Driver Behaviour and Accident Research Methodology: Unresolved Problems*. Ashgate. Available at: https://www.routledge.com/products/9780754670766.

[267] Wåhlberg, A. E. af and de Winter, J. C. F. (2012) 'Commentaries and Responses to "The Driver Behaviour Questionnaire as a predictor of accidents: A meta-analysis"', *Journal of Safety Research*, 43(1), pp. 83–99. doi: 10.1016/j.jsr.2011.06.011.

[268] Wahlstrom, J., Skog, I. and Handel, P. (2015) 'Detection of Dangerous Cornering in GNSS-Data-Driven Insurance Telematics', *IEEE Transactions on Intelligent Transportation Systems*, 16(6), pp. 3073–3083. doi: 10.1109/TITS.2015.2431293.

[269] Wang, C., Quddus, M. A. and Ison, S. G. (2013) 'The effect of traffic and road characteristics on road safety: A review and future research direction', *Safety Science*, pp. 264–275. doi: 10.1016/j.ssci.2013.02.012.

[270] Waters, M. H. L. and Laker, I. B. (1980) 'Research on Fuel Conservation for Cars', *Transport and Road Research Laboratory (TRRL)*, pp. 1–34. Available at: http://www.trl.co.uk/online_store/reports_publications/trl_reports/cat_traffic_and_the_environment/report_research_on_fuel_conservation_for_cars.htm (Accessed: 26 February 2014).

[271] Weiss, J. and Smollik, J. (2012) 'Beginner's Roadmap to Working with Driving Behavior Data', *Casualty Actuarial Society E-Forum (CAS Forum)*, pp. 1–35. Available at: http://www.casact.org/pubs/forum/12wforumpt2/Weiss-Smollik.pdf.

[272] Whitlock, F. A. (1971) *Death on the road: A study in Social Violence*. London: Tavistock Publications.

[273] Wilde, G. J. S. (1982) 'The Theory of Risk Homeostasis: Implications for Safety and Health', *Risk Analysis*, 2, pp. 209–225.

[274] Wolf, J. D. and Barrett, M. F. (1978) *Driver-vehicle Effectiveness Model: Final report*. Washington, DC.

[275] World Health Organisation (2013) *WHO | Global Status Report on Road Safety 2013*. World Health Organization. Available at: http://www.who.int/violence_injury_prevention/road_safety_status/2013/en/ (Accessed: 7 September 2015).

[276] Wyatt, D. W., Li, H. and Tate, J. (2013) 'Examining the Influence of Road Grade on Vehicle Specific Power (VSP) and Carbon Dioxide (CO2) Emission over a Real-World Driving Cycle', *SAE Int. J. Engines*. doi: 10.4271/2013-01-1518.

[277] Wyatt, D. W., Li, H. and Tate, J. E. (2014) 'The Impact of Road Grade on Carbon Dioxide (CO2) Emission of a Passenger Vehicle in Real-world Driving', *Transportation Research Part D: Transport and Environment*, 32, pp. 160–170. doi: 10.1016/j.trd.2014.07.015.

[278] Yamada, T., Kayano, K. and Funabiki, M. (1993) 'The effectiveness of Pd for converting hydrocarbons in TWC catalysts', *SAE Technical Papers*, (412). doi: 10.4271/930253.

[279] Yong He and Lei Feng (no date) 'Diesel fuel injection system faults diagnosis based on fuzzy injection

pressure pattern recognition', in *Fifth World Congress on Intelligent Control and Automation (IEEE Cat. No.04EX788)*. IEEE, pp. 1654–1657. doi: 10.1109/WCICA.2004.1340935.

[280] Young, M. S., Birrell, S. A. and Stanton, N. A. (2009) 'Design for Smart Driving: A Tale of Two Interfaces', in *Engineering Psychology and Cognitive Ergonomics*, pp. 477 – 485.

[281] Young, M. S., Birrell, S. A. and Stanton, N. A. (2011) 'Safe driving in a green world: A review of driver performance benchmarks and technologies to support "smart" driving', *Applied Ergonomics*, 42(4), pp. 533–539. doi: 10.1016/j.apergo.2010.08.012.

[282] Yun, D. S., Lee, J. W., Lee, S. K. and Kwon, O. C. (2012) 'Development of the eco-driving and safe-driving components using vehicle information', in *International Conference on ICT Convergence*, pp. 561–562. doi: 10.1109/ICTC.2012.6387198.

[283] Zarkadoula, M., Zoidis, G. and Tritopoulou, E. (2007) 'Training urban bus drivers to promote smart driving: A note on a Greek eco-driving pilot program', *Transportation Research Part D: Transport and Environment*, 12(6), pp. 449–451. doi: 10.1016/j.trd.2007.05.002.

[284] Zhai, H. (2007) *Regional On-Road Mobile Source Emissions Characterization for Conventional and Alternative Vehicle Technologies.* North Carolina State Univeristy. Available at: http://repository.lib.ncsu.edu/ir/bitstream/1840.16/5150/1/etd.pdf (Accessed: 1 January 2014).

[285] Zhang, G., Yau, K. K. W., Zhang, X. and Li, Y. (2016) 'Traffic accidents involving fatigue driving and their extent of casualties', *Accident Analysis and Prevention*, 87, pp. 34–42. doi: 10.1016/j.aap.2015.10.033.

[286] Zhao, H., Zhou, H., Chen, C. and Chen, J. (2013) 'Join driving: A smart phone-based driving behavior evaluation system', in *GLOBECOM - IEEE Global Telecommunications Conference*. doi: 10.1109/GLOCOM.2013.6831046.

[287] Zhao, N., Mehler, B., Reimer, B., D'Ambrosio, L. A., Mehler, A. and Coughlin, J. F. (2012) 'An investigation of the relationship between the driving behavior questionnaire and objective measures of highway driving behavior', *Transportation Research Part F: Traffic Psychology and Behaviour*, 15(6), pp. 676–685. doi: 10.1016/j.trf.2012.08.001.