Predicting Sleepiness from Driving Behaviour

Pablo Puente Guillen

Submitted in accordance with the requirements for the degree of Doctor of Philosophy.

The University of Leeds

School of Computing
Institute for Transport Studies
School of Psychology

November 2016

Acknowledgements

In the present section, an analysis is done to understand the effect different variables had on the completion of the present PhD study. The variables were classified as “academic” and “non-academic”. The “academic” class contained the variables “supervisors” and “colleagues”. It was found that the variable “supervisors” had the biggest effect in the completion of the PhD. The variable “supervisors” were divided in four sub-variables named as following: Professor Anthony Cohn, Professor Oliver Carsten, Dr. Richard Wilkie and Dr. Faisal Mushtaq. Each sub-variable contained a number of attributes: allocation of time, motivation given and idea generation. It was found that the attributes of each sub-variable of “supervisors” were highly correlated with the researcher’s achievements. It was also found that there was a significant effect between the variable “researcher’s incredible gratitude” and the “supervisors” variable. The results led to the conclusion that the completion of the PhD would not have been achieved without the above-mentioned sub-variables.

The second variable contained in the “academic” class was “colleagues”. An experiment was designed by University of Leeds et al. (2012) where multiple participants, hereafter called “students”, were allocated in a confined space for a long period of time. They found that after a certain period of time, which they call “adaptation” period, the participants started interacting with each other, leading to collaboration on each other’s research. It is worth mentioning that several outliers were found. Participants with id tag “Zeynep Uludag” and “David Aguilar Lleyda” presented a higher than “normal collaboration” in the task “completion of PhD experiments” and participants with id tag “Aryana Tavanai”, “Eris Chinellato”, “Panagiotis Spyridakos” and “Daryl Hibberd” presented high collaboration in the task “providing valuable information and Matlab codes”. It was concluded that the collaboration by the “students” has led to a successful completion of the PhD. It is also worth mentioning that a subjective value of “forever grateful” was achieved.

The variable “colleagues” also contained other important sub-variables that affected the completion of the PhD. The sub-variables “Roy Ruddle” and “Natasha Merat” were related to the task “incredible guidance”. The sub-variables “Michael Daly” and “Anthony Horrobin” had a great effect in the task “designing and developing the experiment”, one of the main tasks in the present PhD study. Using the Karolinska Thankfulness Scale, a value of 10 (“eternally grateful”) was given to each sub-variable.

The “non-academic” class contained very important variables: “family” (parents and brother), “spouse” and “cousins/uncles/aunts/friends”. In accordance to the results found by many other researchers in literature, there is a statistically reliable relationship between “family” and “strength and courage to achieve my goals in life”. The Pearson’s correlation statistic showed a result of 1, reflecting complete dependency between these two measures. In similar way, the variable “spouse” had one of the biggest effects in the overall achievement of the PhD, although from past research, it has been found that the variable “spouse” has been positively related to every achievement of previous endeavours. The so-called “without my spouse I would not be where I am now” coefficient has been and will continue having an effect in future research. Finally, the variable “cousins/uncles/aunts/friends” had a big effect in the task “unwind and relax”, which has led to a value of 10 out of 10 in the Happiness Scale. For further information regarding the sub-variable “cousins/uncles/aunts/friends” please refer to Facebook (Facebook Inc., Cambridge MA, 2014).

The present PhD study was funded by CONACYT. All hail, CONACYT!

A mi abuelita y a nuestros angelitos que nos cuidan todos los dias. Este y cualquier otro logro es gracias a su ayuda.

Abstract

This research investigates the use of objective EEG analysis to determine multiple levels of sleepiness in drivers. In the literature, current methods propose a binary (awake or sleep) or ternary (awake, drowsy or sleep) classification of sleepiness. Having few classification of sleepiness increases the risk of the driver reaching dangerous levels of sleepiness before a safety system can prevent it. Also, these methods are based on subjective analysis of physiological variables, which leads to lack of reproducibility and loss of data, when a lack of consensus is reached amongst the EEG experts. Therefore, the doctoral challenge was to determine whether multiple levels of sleepiness could be defined with high accuracy, using an objective analysis of EEG, a reliable indicator of sleepiness. The study identified awake, post-awake, pre-sleep and sleep as the multiple levels of sleepiness through the objective analysis of EEG. The research used Neural Networks, a type of Machine Learning algorithm, to determine the accuracy of the proposed multiple levels of sleepiness. The Neural Networks were trained using driving and physiological behaviour. The EEG data and the driving and physiological variables were obtained through a series of experiments aimed to induce sleepiness, conducted in the driving simulator at the University of Leeds. As the Neural Network obtained high accuracy when differentiating between awake and sleep and between post-awake and pre-sleep, it led to the conclusion that the proposed objective classification based on objective EEG analysis was suitable. However, this study did not reach the highest levels of accuracy when the 4 levels of sleepiness are combined, nevertheless the solutions proposed by the researcher to be carried in future work can contribute towards increasing the accuracy of the proposed method.

Table of Contents

Abbreviations

The following key would be beneficial for decoding the abbreviations used in this thesis.

CAR	Center for Automotive Research
ICCT	International Council for Clean Transportation
NHTSA	National Highway Traffic Safety Administration
WHO	World Health Organisation
IRTU	International Road Transport Union
EEG	Electroencephalogram
FFT	Fast Fourier Transform
REM	Rapid eye movement
fMRI	Functional magnetic resonance image
fNIR	Functional near-infrared spectroscopy
MLA	Machine learning algorithm
PERCLOS	Percentage of eye closure
ECG	Electro-cardiography
RBF	Radial Basis Function Network
ESS	Epsworth Sleepiness Scale
SDLP	Standard deviation of lane position
SDSpeed	Standard deviation of speed
SDSteering	Standard deviation of steering wheel angle
HFS	High frequency steering
TTLC	Time to lane crossing
OOL	Out of lane
EGI	Electrical Geodesics Inc.
ICA	Independent component analysis
BESA	Brain Electrical Source Analysis
CGF	Centre of gravity frequency
KDS	Karolinska Drowsiness Scale
NASA	The National Aeronautics and Space Administration
DTREG	Nonlinear Regression and Curve Fitting
UoLDS	University of Leeds Driving Simulator

SMOTE	Synthetic Minority Oversampling Technique
RMSE	Root mean square error

Chapter 1: Introduction

1. Introduction

1.1 Background and research focus

This thesis describes a programme of research that investigates a solution to determine and predict different levels of sleepiness in drivers using brain wave activity as well as driving and physiological behaviour. The impetus for carrying out this research was an opportunity to develop a solution that could reduce accidents due to falling asleep while driving. Today, there are more than one billion cars circulating around the world and it is expected that this will double during the next decade (Center for Automotive Research, 2011; Dargay et al., 2007; International Council for Clean Transportation, 2014; WHO, 2009; Sperling & Gordon, 2008; Sousanis, 2011). The prevalence of automotive vehicles means that ensuring the safety of drivers, passengers and pedestrian is a top priority for researchers and car manufacturers (Royal, 2003; Philip & Akerstedt, 2006; Maycock, 1997; Horne & Reyner, 1999; NHTSA, 1999). Unfortunately, there are around 50 million road accidents every year and approximately 1.2 million of them result in fatalities (WHO, 2009). From all the road accidents, approximately 85% are accounted to human errors (International Road Transport Union, 2007).

One of the most common and dangerous causes of road accident due to human-error is sleeping while driving (Royal, 2003; Philip & Akerstedt, 2006; Connor et al., 2001; NHTSA, 2015). It has been estimated that, across the world, around 20% of road crashes are related to sleepiness (MacLean et al., 2003). In the U.S.A. alone, around 1,550 crashes due to sleepiness resulted in a fatality per year (NHTSA, 1999). Based on the frequency and magnitude of sleep-related driving errors there has been a push towards exploring ways in which one might be able to predict high levels of sleepiness in the driver to avoid road accidents while driving.

In order to tackle this problem, the first steps taken by a number of researchers and car companies has been to identify the most suitable variables to predict the level of sleepiness in the driver. For example, recently Lexus (2012) presented a device that would determine the level of sleepiness of the driver by analysing their blinking behaviour- as research indicates that longer and more frequent blinks are related to higher levels of sleepiness (Yeo et al., 2009). On the other hand, Bosch (2012) developed a system that used lateral movement of the car as a determinant of the level of sleepiness.

Although physiological and driving behaviour are highly correlated with sleepiness, electrical activity produced by the brain has been found to be more reliable to determine sleepiness. One of the methods to measure electrical activity produced by the brain is using an electroencephalogram (EEG; electrodes in contact to the scalp of the human). Brain wave activity suffers fewer changes due to changes in the driving environment (Artaud et al., 1994; Lal et al., 2003; Jap et al., 2009). EEG provides a highly reliable biological marker of different levels of physiological arousal (Jap et al., 2009; Lal & Craig, 2001a,b; Stern & Engel, 2005; Yeo et al., 2009; Zhao et al., 2012; Akerstedt & Gillberg, 1990; Campagne, Pebayle & Muzet, 2004; Jap et al., 2009; Kecklund & Akerstedt, 1993; Zhao et al., 2012; Lowden et al., 2009). To separate this EEG activity data into different levels of sleepiness, expert clinicians subjectively classify the data using visual indicators in the brain wave activity data. Unfortunately, when clinicians attempt to classify the data into different levels of sleepiness, many discrepancies arise due to the subjectivity of the visual analysis of the EEG signal (Yeo et al., 2009; Vuckovic et al., 2002; Knoblauch et al., 2003; Andrillon et al., 2011; Benbadis, 2006; Carskadon and Dement, 2011; Gennaro and Ferrara, 2003; Devuyst et al., 2010; Cantero et al., 2002; Parekh et al., 2015; Teplan, 2002). This relatively poor inter-rater reliability decreases the amount of data that can be used to inform sleep-related decisions. By developing an objective analysis of this brain wave activity data, the classification process can be reproduced in any data set without the need to consult clinicians. Such an objective analysis would also reduce the amount of data lost due to a lack of agreement from the clinicians. Therefore, one of the primary goals of this PhD project is to produce an objective analysis of brain wave activity to identify sleepiness.

Once the variables being used to determine the levels of sleepiness have been decided upon, the next step is to determine how to define and predict different levels of sleepiness. Although many researchers and car companies have developed systems to predict and prevent sleepiness, these systems only react to a binary or ternary number of levels of sleepiness (Lal et al., 2003; Sayed & Eskandarian, 2001; Yeo et al., 2009; Shuyan & Gangtie, 2009; Yang et al., 2010; Patel et al., 2011;

NapZapper, 2016; Bosch, 2012; Lexus, 2012), i.e. the sleepiness level changes from "awake" to "sleep" in a binary system and from "awake" to "drowsy" to "sleep" in a ternary system. During the "sleep" level, the driver's capabilities reduce and the probability of a road accident is high (Royal, 2003; Philip & Akerstedt, 2006; Connor et al., 2001; Klauer et al., 2006; Lamond & Dawson, 1999). This means that when a system that can only determine binary states of sleepiness predicts a "sleep" state, the driver is already in a high risk of being involved in a road accident.

A system with ternary states of sleepiness (awake, drowsy and sleep) allows for smoother and safer transitions between the levels of sleepiness. Once the system has determined the instantaneous level of sleepiness of the driver, the system can take actions to warn the driver (if the level of sleepiness is low) or aid the driver through the automation of certain driving tasks (if the level of sleepiness is high). If the system detects a medium level of sleepiness ("drowsy" state), a low level of automation action, e.g. an alarm suggesting the driver to take a break or coffee, would suffice to alert the driver of the potential danger if he/she continues to drive while his/her level of sleepiness increases. On the other hand, if the system determines a high level of sleepiness, a high level of automation action would be needed from the system, e.g. the car would take partial or complete control of the driving tasks, and this has a high risk of being involved in a road accident. Therefore, reducing or completely removing control of the driving tasks from the driver would reduce the possibility of the driver performing an incorrect driving action. Unfortunately, research found that jumping from a high level of automation action to a low level of automation action has serious consequences on the driver, leading to serious accidents (Merat et al., 2014; Endsley, 1995; Carsten et al., 2012). By determining a higher number of levels of sleepiness, it is possible to have a smoother and safer transition between the levels of automation of the actions required for each level of sleepiness. Therefore, the present PhD study proposed a multiple classification of the levels of sleepiness.

In the present PhD, the researcher explored different algorithms used to determine and predict different levels of sleepiness using physiological and driving behaviour. An understanding of the brain wave activity data as well as the driving and physiological behaviour allowed the adoption of a Machine Learning Algorithm that could predict and determine the multiple levels of sleepiness aimed to obtain at the end of the present PhD study.

1.2 Research intent

This PhD project presents a novel approach to classifying EEG data analysis. By adopting an objective analysis of the data, it is possible to avoid the problems that arise when analysing the data according to subjective visual interpretation. It also allows for reproducibility and repeatability of the analysis method when the amount of data or the data sets is high.

This PhD also explores the possibility of detecting multiple levels of sleepiness using physiological and driving behavioural data. Using Machine Learning algorithms, analysis was conducted to determine the possibility of creating a system that would predict different levels of sleepiness of a driver and therefore allow different level of action with different levels of automation. It is hypothesised that such system could reduce and prevent accidents related to sleepiness while driving by detecting premature signs of sleepiness and react preventively to avoid any incident.

1.3 Aims and objectives

The intention and purpose of the research was to determine the role that driving and physiological behaviour plays when detecting sleepiness in drivers. The research also seeks to provide a more robust and objective quantification and classification of the brain wave activity data, i.e. the basis for the classification of the sleepiness' states. By using Machine Learning algorithms, it is possible to determine the accuracy of the different levels of sleepiness classified through brain wave activity as well as the role that each physiological and driving behavioural variable plays when detecting the mentioned levels of sleepiness.

The research in the first year focused on determining the different variables used in research and industry to classify the different levels of sleepiness. An analysis was done to understand the effect of changes in sleepiness on different driving and physiological variables. An analysis was also done on the different Machine Learning algorithms used in research to predict sleepiness using physiological and driving data.

This allowed a better understanding of the design of experiments, in order to obtain sleeping data from participants while driving. Following this first year, the study conducted a number of experiments focused on obtaining physiological and driving data from drivers while their sleepiness increased. These experiments were conducted using a driving simulator to reduce the risks of danger of driving while being in a high state of sleepiness. Finally, the data obtained from the experiments conducted was used to predict the levels of sleepiness through a Machine Learning algorithm. The broad aims are described below followed by the objectives devised to achieve them:

Aim 1: Identifying the literature related to physiological and driving behaviour of people driving under a high state of sleepiness.

- Distinguish how physiological and driving variables are modulated as sleepiness increases
- Determine the "ground truth" used in literature to classify sleepiness into different clusters

Aim 2: Provide an objective analysis of the EEG data

- Determine the challenges and advantages of using visual analysis (subjective) of the brain wave activity data
- Define the variables and factors used as identifiers by clinicians when visually analysing brain wave activity data
- Develop a reproducible program able to obtain high accuracy when objectively analysing the brain wave activity data of different participants

Aim 3: Adopt a Machine Learning algorithm to detect multiple levels of sleepiness using the drivers' data

- Determine the advantages and disadvantages of different Machine Learning algorithms
- Define the parameters of the algorithm that would achieve the highest accuracy when detecting different levels of sleepiness

- Analyse the role of physiological and driving data towards the detection of the different levels of sleepiness

1.4 The thesis structure

Chapter 1 Introduction: the study's focus, scope, aims, objectives and significance

Chapter 2 Literature review: Sleepiness while driving: correlation of the study's central argument and theoretical foundation with existing literature

Chapter 3 Literature review: Classification of sleepiness in drivers: analysis of the different algorithms that have been used in previous research focusing on sleeping while driving

Chapter 4 Identifying markers of sleepiness in secondary data: Data analysis using blinking behaviour as the factor that classifies the levels of sleepiness and driving behaviour as the predictor data. The data used in this chapter was obtained from past research conducted by another researcher in the driving simulator.

Chapter 5 Inducing high levels of sleepiness in drivers: Narrative of the experiment design that was conducted in a static driving simulator during this PhD study to obtain sleep data from participants and the subsequent results.

Chapter 6 Identifying markers of sleepiness using EEG: A Machine Learning algorithm is trained using the driving and physiological data obtained during the experiment conducted presented in Chapter 5. The levels of sleepiness are determined by classifying the brain wave activity into different clusters.

Chapter 7 Discussion and conclusion: Relates the finding from the data analysis to results obtained by other researchers and presents a statement of the contribution the research makes to new knowledge.

Chapter 2:
Literature review: Sleeping while driving

2. Literature review: Sleeping while driving

The following chapter starts by discussing the complexity of the driving task and the consequences of human errors (non-performance errors) while driving. One of the most common human errors while driving is sleeping while driving (Internatioanl Road Transport Union, 2007). The chapter further develops the statistics regarding crashes of people falling asleep on the wheel and the different approaches taken to predict and prevent this type of human error. Sleepiness is the major contributor of this type of error. The different stages of sleepiness, as well as its consequences in the physiology of people, are further explained in this chapter. At the end of the chapter, a summary of different types of automation approaches that can lead to a reduction on accidents due to sleepy drivers is presented.

2.1 Driving: a complex task

Driving is an everyday task that allows millions of people around the world to transport themselves and/or goods from one place to another with relative ease. There has been a substantial rise in driving over the past three decades (Grove, 2015; George & Kershaw, 2016; Leibling, 2008). In the United Kingdom alone, the number of vehicles on the road has increased dramatically every year since 1950 and there are now 35.6 million vehicles licensed in total (Grove, 2015; George & Kershaw, 2016). According to research, in 2020 there will be over 37 million vehicles on the road in the United Kingdom (Leibling, 2008). Whilst this appears to be a routine task in many people's lives, the act of driving is layered with complexity.

A simple model explaining the different levels of complexity present in driving was described in Plankermann (2013). The model is a combination of Michon's (1985) task hierarchical model and Rasmussen (1983) task performance. The model presents three hierarchical levels (strategic, manoeuvring and control) and each one is associated to a task performance.

(i) The first level is the strategic level (knowledge task). In driving, this first level is associated with the knowledge the driver has of the road, i.e. mental map of the directions to a destination, knowledge of streets, etc.

(ii) The second level is the manoeuvring level (rule task). In driving, this is represented by the knowledge the driver has regarding the driving rules, i.e. distance the driver should maintain with the car in front, speed needed to take a gentle curve, etc.

(iii) The third level is the control level (skill task). This level is how well the driver performs the desired action planned in the manoeuvring level. A failure in any of these levels (human error accident) could lead to a collision that could endanger the people inside and outside the car (National Highway Traffic Safety Administration [NHTSA], 2015).

According to the International Road Transport Union, the number of road freight accidents due to human error is approximately 85% (Figure 2-1) (International Road Transport Union, 2007). In 2009, more than 2,000 people died in the UK due to road traffic accidents (Box, 2011). Although it only represented 0.5% of all deaths in that year, for young people (between 15 and 19 years old) it represented 25% of all deaths (Box, 2011). Crashes due to human error can be categorised in recognition error (strategic level), decision error (manoeuvring level), performance error (control level) and non-performance error (National Highway Traffic Safety Administration [NHTSA], 2015). Non-performance factors (sleep, alcoholic or drug intoxication, etc.), although not that common, can affect the three task hierarchy levels outlined above and thus cause serious accidents (European Agency for Safety and Health at Work [EU-OSHA], 2010). One of the most dangerous and common non-performance errors is sleeping while driving (Royal, 2003; Philip & Akerstedt, 2006; Connor et al., 2001, National Highway Traffic Safety Administration [NHTSA], 2015).

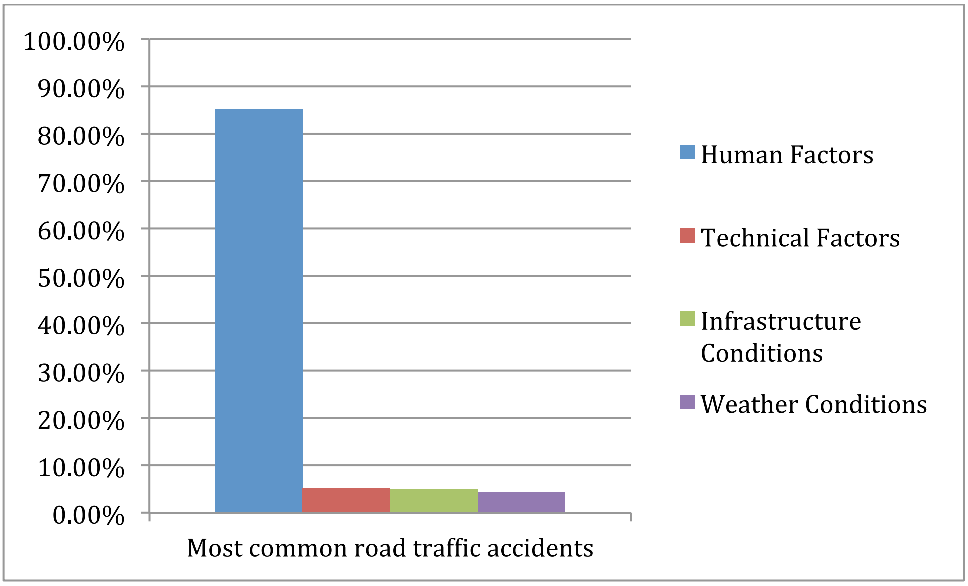


Figure 2-1 Most common road traffic accidents factors according to the European Truck Accident Causation (Adapted from: IRTU, 2007)

2.2 Falling asleep while driving

Sleepiness is one of the major causes of road traffic accidents (Royal, 2003; Philip & Akerstedt, 2006; Connor et al., 2001). MacLean et al. (2003) estimated that, around the world, approximately 20% crashes are related to sleepiness, where off-road crashes, i.e. driving out of the road, are the most common (George, 2005). In the United Kingdom 15-20% of the accidents are related to sleepiness (Maycock, 1997; Horne & Reyner, 1999) and in the U.S.A. it is estimated that there are 56,000 sleep-related crashes each year of which 1,550 resulted in a fatality (NHTSA, 1999).

The risk of having an accident while being sleepy is four to six times higher than when the driver is alert (Klauer et al., 2006). Researchers have demonstrated that a driver who has sleep deprivation has the same driving skills as a person with an illegal high concentration of alcohol in their system (0.1 g/l) (Lamond & Dawson, 1999). In order to understand how we might build interventions to address this important topic of research, it is necessary first to examine the physiology of sleep and distinguish the difference between alertness and sleepiness as well as the different "sleep" and "awake" stages before engaging in methods to predict and prevent people from falling asleep while driving.

2.3 The physiology of sleep

According to Sadock and Sadock (2000, p. 281), "sleep is a state of decreased awareness of environmental stimuli that is distinguished from states such as coma or hibernation by its relatively rapid reversibility". Another important characteristic of sleep is that as sleepiness increases, the individual is capable to recognize higher levels of sleepiness (Sadock & Sadock, 2000; Carskadon & Dement, 2011). In addition, there is a correlation between the feeling of "sleepy" and visible behavioural changes as well as changes in the physiology of the individual (Sadock & Sadock, 2000; Carskadon & Dement, 2011, Boyle et al., 2008; Brookhuis & de Waard, 2010; Filtness, Reyner & Horne, 2012).

The changes in behaviour due to sleepiness have been widely studied in literature (Sadock & Sadock, 2000; Carskadon & Dement, 2011, Boyle et al., 2008; Brookhuis & de Waard, 2010; Filtness, Reyner & Horne, 2012; Jap, Lal & Fisher, 2011). In many studies, it has been found that as sleepiness increases, the blinking behaviour of an individual modifies by an increase in the frequency of blinks, a longer duration of blink and shorter inter-blinking time (Yang et al., 2010; Bergasa et al., 2006; Wierwille et al., 1994; Dinges et al., 1998). Many other researchers have reported that as sleepiness increases head nodding as well as yawning increases (Hartley et al., 2000; Haworth & Vulcan, 1991). Although these are clear visible indicators of sleep, there are other physiological changes, which are not as overtly obvious, but are tightly correlated with sleepiness.

As sleepiness increases, the human body experiences physiological changes (Chua et al., 2012; Elsenbruch et al., 1999; Patel et al., 2011). There is a long history of researchers using physiological and psychophysiological measures to detect these changes. Chua et al. (2012) and Elsenbruch et al. (1999) found examples of changes in physiological and psychophysiological variables as sleepiness increases. Using electrocardiogram (ECG), Chua et al. (2012) found that there is a strong positive correlation between increase in sleepiness (determined through a psychomotor vigilance task) and heart rate ($r = 0.68$), while Elsenbruch et al. (1999) found that heart rate changes in different levels of sleepiness (determined by recording brain wave activity). Chua et al. (2012) also found that eyes closure behaviour ($r = 0.77$), blinking pattern (r = -0.51) and subjective sleepiness ratings (r = 0.56) have a strong positive correlation with changes in sleepiness.

Another major change due to increase in sleepiness happens at a neural level (Eoh et al., 2005a,b; Lal et al., 2003; Gillberg et al., 1996; Artaud, 1994; Vuckovic et al., 2002; Yeo et al., 2009; Filtness et al., 2012; Zhao et al., 2012; Lowden et al., 2009; Jap et al., 2009). The brain sends signals to the body by firing small electrical activity waves between its cells, called neurons (Cohen, 2014). These electrical activity waves (often referred to as "brain wave activity") contain information that allows the brain to act in response to the outside world (Cohen, 2014). The electrical activity waves can be classified in fast wave activity or slow wave activity (Cohen, 2014) depending on their frequency.

This type of electrical waveform (e.g. AC current in appliances in the house) can be represented as a sine wave composed of a magnitude and a frequency (frequency = 1/period T), as presented in Figure 2-2. The magnitude refers to the instant value in any given time of the electrical waveform (in the case of brain activity is measured in micro Volts) (Johnson, 2013). The frequency (inverse of the period) represents the speed of the wavelength (Johnson, 2013).

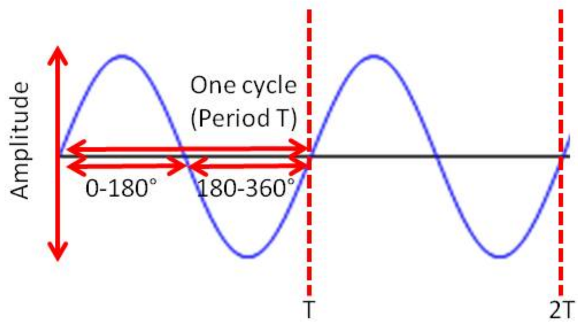


Figure 2-2 Sinusoidal waveform and its components. The amplitude is the value measured from peak to peak in the vertical axis. The period is the time it takes to complete a cycle (one positive and one negative peak). The frequency is then the amount of cycles in one second and is measures as the inverse of the period.

When combining different sine waves with different frequencies and different amplitudes, it is possible to create a complex electrical waveform as shown in Figure 2-3. This is the case with the brain wave activity. Brain wave activity is composed of many single electrical sine waveforms as presented in Figure 2-4. It is possible to separate each component of the brain wave activity with a method called Fast Fourier Transform (FFT) as presented in Figure 2-5, which will be explained in more detail in chapter 5.

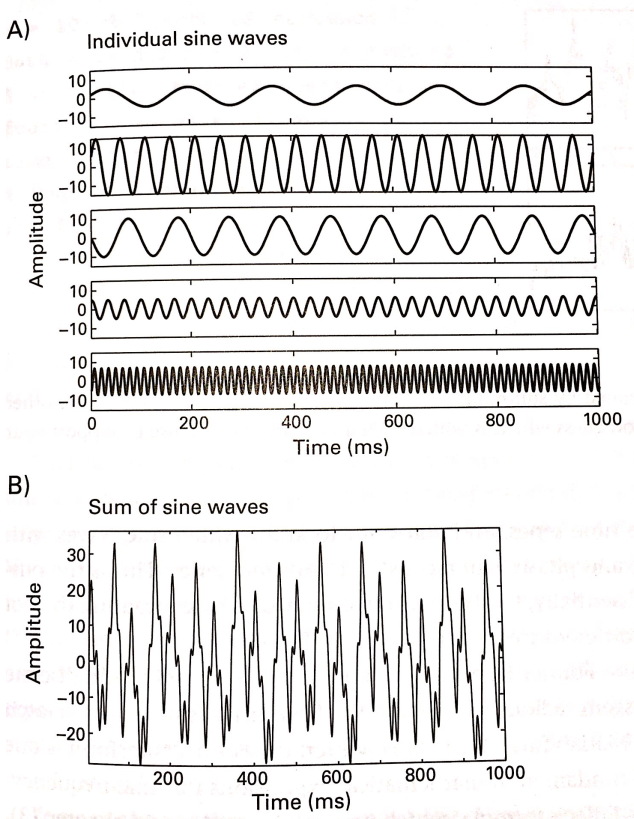


Figure 2-3 Visual representation of different sine waves a) Presents many different sine waves, each with different magnitude and frequency. b) All the sine waves from a) can be sum resulting in a new complex signal (sum of sine waves) as presented in b) (Source: Cohen, 2014) Reprinted from "Analysing Neural Time Series Data: Theory and Practice" by Mike X. Cohen. Copyright © 2014 by Mike X. Cohen. Used by permission of The MIT Press, Cambridge, MA, USA.

When the brain wave activity signal is separated in its different components (using FFT), it is seen that the behaviour of the brain wave activity changes depending if an individual is awake or asleep (Cohen, 2014; Vuckovic et al., 2002).

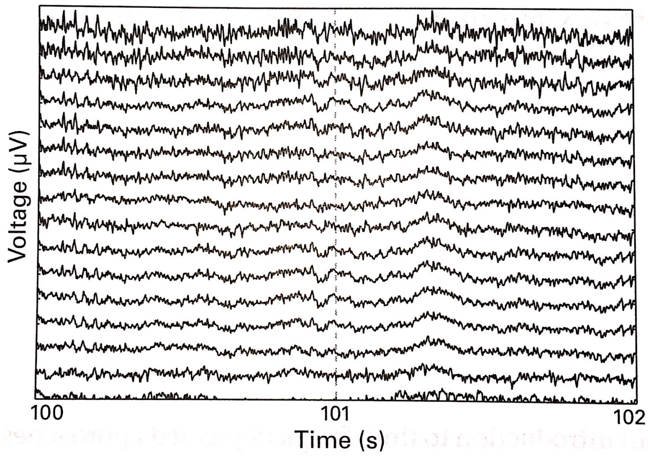


Figure 2-4 Example of brain wave activity recorded from a human participant. Each line represents an electrode, which measures the brain wave activity (in microvolts), in different locations in the scalp of an individual (Source: Cohen, 2014) Reprinted from "Analyzing Neural Time Series Data: Theory and Practice" by Mike X. Cohen. Copyright © 2014 by Mike X. Cohen. Used by permission of The MIT Press, Cambridge, MA, USA.

When an individual is awake, signals with frequencies around 13 to 20 Hertz (Hz) have a high amplitude compared to signals with frequencies around 4 to 13 Hz (Filtness et al., 2012; Zhao et al., 2012; Lowden et al., 2009; Jap et al., 2009, Vuckovic et al., 2002). When an individual's sleepiness is increasing, signals with frequencies around 13 to 20 Hz decrease in magnitude and signals with frequencies around 4 to 13 Hz have an increase in magnitude, specifically 8 to 13 Hz are the frequencies with the highest increase. When an individual is asleep, it has been found that frequencies around 30 Hz and above have an increase in amplitude (Vuckovic et al., 2002). For analysis purposes, many researchers have grouped and labelled the frequency ranges of the brain wave activity. There are four main frequency bands; the frequency interval of 4 to 8 Hz is called theta band, from 8 to 13 Hz is called alpha band, from 13 to 30 Hz is called beta band and frequencies above 30 Hz are categorise in the gamma band (Filtness et al., 2012; Zhao et al., 2012; Lowden et al., 2009; Jap et al., 2009, Vuckovic et al., 2002).

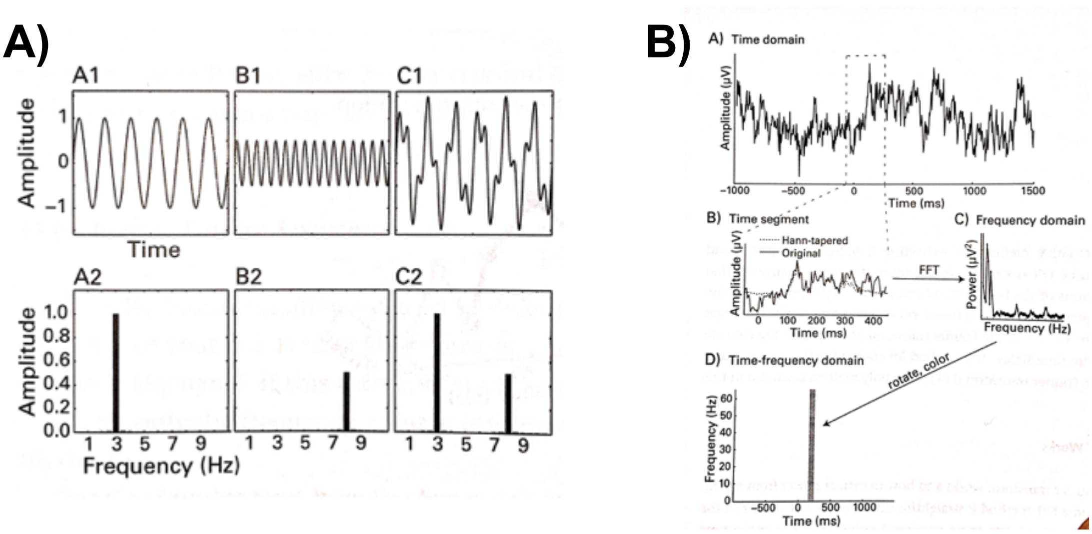


Figure 2-5 Effect of FFT in different sine waves a) Sine wave A1 (with amplitude 1 and frequency of 3 Hertz) and sine wave B1 (with amplitude 0.5 and frequency of 8 Hz) are combined to create signal C1. When FFT is performed in each of these sine waves, the result is decomposition of the frequencies belonging in each signal (as the sine waves A1 and B1 are just composed of a single sine wave with unique frequency the plot A2, B2 just present activity in one frequency. On the other hand, C2 presents activity in two frequency, as C1 is composed of two sine waves with different frequencies). b) An EEG (brain wave activity) signal (Figure B-A) is a composition of many sine waves with different frequencies; therefore, FFT analysis can be performed. In this case, a segment of the EEG data (Figure B-B) was analysed and the result can be plotted in the frequency domain (Figure B-C) or in the time-frequency domain (Figure B-D) (Source: Cohen, 2014) Reprinted from "Analysing Neural Time Series Data: Theory and Practice" by Mike X. Cohen. Copyright © 2014 by Mike X. Cohen. Used by permission of The MIT Press, Cambridge, MA, USA.

Using brain wave recording techniques to determine changes in different frequency bands, researchers have been able to determine and differentiate the awake and sleepy states in an individual, as presented in Figure 2-6. Within the sleep state, it has been possible to determine four different sleep levels (stage N1 to N3 and REM), as shown in Figure 2-6. Unfortunately for the awake state, is not easily classified.

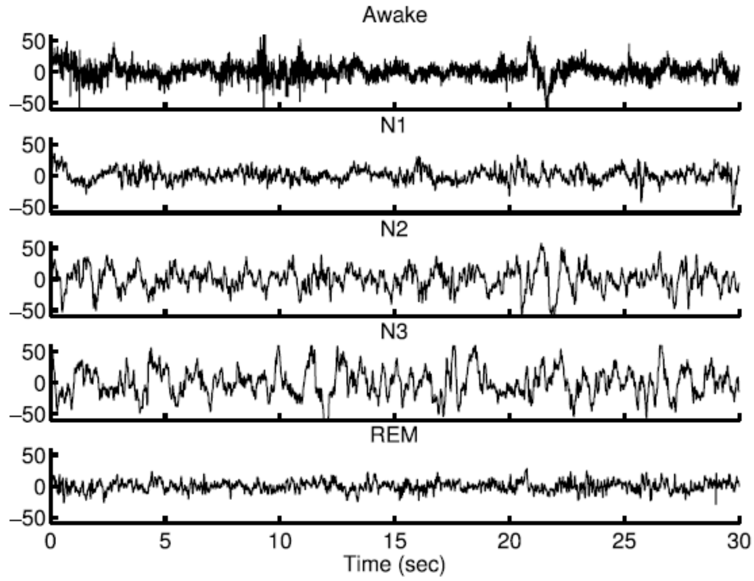


Figure 2-6 Brain wave activity during awake and sleep stages. Sleep stages are categorised in 4 stages (N1, N2, N3 and REM) whilst awake is more difficult to classified in many stages (Source: Shi et al, 2017) Reprinted from "A comparison study on stages of sleep: Quantifying multiscale complexity using higher moments on coarse-graining" by Shi et al. Copyright © 2017 by Shi et al. Used by permission of Elsevier, Amsterdam, NL.

Once the effects of sleep in an individual and the different stages in awake and sleep have been presented, it is important to define the difference between the terms that used to describe awake and sleepy state in a driving scenario. In the following section, the different terms related to awake and sleep are explained in order to have a common understanding of the concepts that are referred to in later chapters of this thesis.

2.3.1 Alertness, drowsiness and sleepiness

In many research papers the terms drowsiness, sleepiness, and fatigue are used interchangeably to refer to a state where a person has the inability to stay awake (Filtness et al., 2012; Reyner et al., 2012; Wu & Chen, 2008; Yang et al., 2010; Zhao et al., 2012; Lowden et al., 2009; Jap et al., 2009, Vuckovic et al., 2002). In order to determine the effects of a driver falling asleep and waking up while driving, it is necessary to have a clear definition of sleepiness and its difference with other 'awake' and 'sleepy' stages such as alert, fatigue and drowsiness. For the purpose of this research, a common definition for these terms was determined using existing definitions in literature.

2.3.1.1 Alertness (alert wakefulness)

This is the state of being awake and is commonly characterized by eye blink duration of 0.3 to 0.4 seconds and inter-eye blink of 6 to 8 seconds (Yeo et al., 2009; Hart, 1992; Doughty, 2002). During this state, beta activity in the brain is dominant (Hart, 1992).

2.3.1.2 Drowsiness (quiet wakefulness)

The stage prior falling asleep is called drowsiness (Johns, 2000). This state can be recognized by difficulty in maintaining visual focus attention, limitation of the higher cognitive functions, and limitations in the visual perception (Lamond & Dawson, 1999; Thomas et al., 1998). This level is also characterized by the appearance of microsleeps (Broughton & Hasan, 1995; Tanaka, Hayashi, & Hori, 1996) as well as increase in the alpha and theta activity in the brain and a decrease in the beta activity in the brain (Jap et al., 2009; Lal & Craig, 2001a,b; Stern & Engel, 2005; Yeo et al., 2009; Zhao et al., 2012; Akerstedt & Gillberg, 1990; Campagne, Pebayle & Muzet, 2004; Jap et al., 2009; Kecklund & Akerstedt, 1993; Lowden et al., 2009). The concept of microsleeps and alpha, theta and beta activity in the brain is explained further in section 2.5.

Drowsiness can also be distinguished by some physiological factors: eye blink duration of more than 0.5 seconds (Yeo et al., 2009), head nodding (Hartley et al., 2000; Haworth & Vulcan, 1991; Kaplan et al., 2007; Lal & Craig, 2002), increase in head movement (Berg & Landstrom, 2006) and yawning (Gu et al., 2002; Kaplan et al., 2007).

During drowsiness state, research found differences in driving performance such as an increase in lateral position variability as well as higher steering movements (Liu et al., 2009; Arnedt et al., 2001; De Valck & Cluydts, 2001; Ingre et al., 2006). Thiffault and Bergeron (2003) found that when drivers are in a drowsiness state, they make larger steering wheel movements (6-10 degrees) and fewer small steering wheel movements (1 – 5 degrees).

2.3.1.3 Fatigue

There is no agreed consensus in the literature on this term. For Brown (1994), fatigue is a state where a person would experience an unwillingness to continue a particular task. On the other hand, Bartlett (1953) stated that fatigue is a process that can be determined by changes in the performance of an activity in time. This term has been associated to sleepiness, as both terms are related to a reduction in attention and cognitive abilities.

2.3.1.4 Sleepiness

According to Johns (1998) and Curcio, Casagrande & Bertini (2001) sleepiness refers to the transition for someone to go from the stage of being awake to being in a drowsy or sleep stage at a particular time. Throughout this thesis, sleepiness is used as the measurement of sleep in the driver. Sleepiness can be affected by different factors:

1. Arousal level of the task being performed (Curcio et al., 2001). Monotonous[1] roads are one of the main examples of low arousal tasks that affect the increase of sleepiness (Zhao & Rong, 2013). It has been found that driving on monotonous roads can lead to larger steering movements (overcorrections) compared to non-monotonous road (Thiffault & Bergeron, 2003). May & Baldwin (2009) defined this concept as task related sleepiness.
2. The length of time a person has been awake (Johns, 2000). This is mainly related to sleep deprivation and it is defined as a sleep related factor (May & Baldwin, 2009). After sleep deprivation, drivers have greater lane position variability, drive closer to the lane, have a higher standard deviation of their speed and have more unwanted lane crossings (Lenne et al., 1998; Philip et al., 2005). It has also been found that reaction time performance decreases with sleep deprivation (Graw et al., 2004; Philip et al., 2005).
3. The circadian rhythm (Borbély, Achermann, Trachsel, & Tobler, 1989). The circadian rhythm refers to a 24-hour biological cycle of every being that determines the patterns of sleeping and feeding (Vitaterna,

[1] "According to MacBain (1970) a situation is said to be monotonous when then stimuli remain unchanged or change in a predictable manner, resulting in sensory stimulation that is constant or highly repetitive" (Thiffault & Bergeron, 2003, p.383)

Takahashi, & Turek, 2001). This means that during a 24-hour period, there are moments when human beings are more prone to fall asleep, usually being maximal at 3 a.m. to 4 a.m. (Borbély et al., 1989; Monk, 2005; Dinges, 1989; Tune, 1969). Pack et al. (1995) found that sleep crashes occurred more often between 2am and 6am and between 2pm and 4pm, this last one as a consequence to an increase in sleepiness due to a circadian rhythm period called 'Afternoon dip' (Monk, 2005).

2.4 Variables used to predict sleepiness

One important step to prevent accidents of people falling asleep while driving is to detect sleepiness before an accident occurs. This issue arises because there is not yet a consensus in literature regarding if drivers are unable to determine or predict that they will fall asleep with enough accuracy. According to Di Stasi et al. (2012), drivers are not aware and/or deny impairment of their driving skills due to fatigue. It is also very difficult to assess sleepiness in driving condition as drivers try to resist sleep while struggling to maintain alertness (Yeo et al., 2007; Eoh et al.; Monk 2005; Thiffault & Bergeron, 2003; Liang et al., 2005; Moller et al., 2006). On the other hand, Williamson et al. (2014) found that drivers were well aware of changes in the level of sleepiness. Further investigation in this topic needs to be done to reach a consensus regarding the subjective assessment of drivers' own levels of sleepiness.

2.4.1 PERCLOS

One of the most common indicators used to predict sleepiness is PERCLOS (PERcentage of eye CLOSure), which uses the percentage of time the pupil is 80% covered by the eyelid within a 1 to 3 minutes (Lal & Craig, 2001a,b; Hayami et al., 2002; May & Baldwin, 2009, Wierwille et al., 1996, Dinges & Grace, 1998). PERCLOS is considered one of the most accurate ways to predict sleepiness (Bergasa et al., 2006; Dinges et al., 1998; Mallis, 1999) and it is presently used commercially (Lexus-Europe, 2012; LumeWay, 2014).

The problem with PERCLOS is that environmental factors such as changes in the lighting in the road, headlight from other cars and air temperature might change the behaviour of the eye blinks, as well creating problems for the device to correctly determine the eye closure (Horne & Reyner, 1996).

2.4.2 Driving behaviour

Another way to determine sleepiness is by analysing the driving behaviour (Wakita et al., 2006; Takei & Furukawa, 2005; McCall et al., 2005). As stated before, an increase in standard deviation and steering wheel movement is related to an increase in sleepiness (Arnedt et al., 2001; Liu et al., 2009; De Valck & Cluydts, 2001; Ingre et al., 2006). This method to determine sleepiness is also used commercially by Bosch and Daimler who have created a device that depending on the lateral deviation of the car, the sleepiness of the driver can be determined (Bosch, 2012).

Unfortunately, driving behaviour changes from driver to driver, therefore it is difficult to assess changes in sleepiness in relation to driving behaviour (Liu et al., 2009). As such, driving behaviour has often been used in combination with other physiological measures to determine sleepiness in a more accurate way (Risser et al., 2000; Lal & Craig, 2002). As discussed previously, sleepiness affects the physiology and behaviour of an individual, so it is possible to use those physiological and behavioural changes as an indicator of sleepiness.

2.4.3 Physiological variables

The analysis of physiological variables such as head nodding, heart rate and body movement have been found to be closely related with sleepiness (Hartley et al., 2000; Haworth & Vulcan, 1991; Zilberg et al., 2007, 2009; Apparies et al., 1998; Li et al., 2004; Hartley et al., 1994). The relation between some of these measures with sleepiness is so strong that it has led to the development of devices to predict sleepiness. For example, NapZapper (2008-2016) created a commercial device using head nodding as a method to determine sleepiness.

The downside of using physiological variables is that the frequency and length of appearances of these physiological variables changes between individuals (Lal & Craig, 2002; van den Berg & Landstrom, 2006). Similar to the case of using driving behaviour to predict sleepiness, the individuality of driver's physiological behaviour makes it difficult to predict sleepiness with high accuracy.

2.4.4 EEG as the ground truth

Electroencephalogram (EEG) is one of the most reliable and precise indicators of sleepiness (Artaud et al., 1994; Lal et al., 2003; Jap et al., 2009). As such, it often serves as the gold standard, or a "ground truth" measure of sleepiness. It has been found that most of the drivers' EEG measurements have a common behaviour, i.e. EEG is less affected by individuality of the participants (Jap et al., 2009; Lal & Craig, 2001a,b; Stern & Engel, 2005; Yeo et al., 2009; Zhao et al., 2012; Akerstedt & Gillberg, 1990; Campagne, Pebayle & Muzet, 2004; Jap et al., 2009; Kecklund & Akerstedt, 1993; Lowden et al., 2009). EEG has not been used commercially due to the difficulty to set up an EEG system in a real car (space is limited and the noise recorded in real driving is very high; Lal et al., 2003).

2.5 Brain wave and EEG as predictor of sleepiness

As mentioned before, brain wave activity has been used to determine awake and sleepy states. Brain wave activity can be recorded with different non-intrusive methods depending on the purpose and design of the research. The most common ways to record brain wave activity is using electroencephalogram (EEG), functional magnetic resonance image (fMRI) and Functional near-infrared spectroscopy (fNIR). In the following section, advantages and disadvantages of these techniques are discussed.

fMRI is a technique that uses a standard magnetic resonance scanner to create images of the brain (Lindquist & Wager, 2014). fMRI has a great spatial resolution, i.e. it is possible to determine the location in the brain where the electrical activity occurred (Lindquist & Wager, 2014; Cohen, 2014). Unfortunately, the temporal resolution of the fMRI is poor compared to other techniques, i.e. there is latency between the time the electrical activity occurs and the time is recorded by the fMRI (Lindquist & Wager, 2014). Finally, it is difficult to use fMRI in a driving experiment due to the size of most fMRI and the constraint position of the participants while using fMRI. In terms of the environment, there is also the constraint of objects around the fMRI, e.g. metallic objects, which could harm the subjects and create interference with the fMRI (European Commission, 2013).

EEG is a technique of positioning a net with electrodes (the number of electrodes can differ from 32 to more than 200) on the head of an individual (Cohen, 2014), as shown in Figure 2-7. The electrical activity in the brain travels across the neurons until it reaches the scalp where the EEG records it. Compared to fMRI, EEG has very good temporal resolution but poor spatial resolution.

Figure 2-7 EEG net cap with 129 electrodes (used to record brain wave activity) held together with a transparent rubber net. Electrodes are soaked in potassium chloride electrolyte to allow better conductivity. The impedance, which measures the conductivity between the scalp and the electrode, was kept under 100 ohms.

fNIR is a technique that uses infrared sensors to detect changes in the concentration of oxygenated and deoxygenated haemoglobin in the blood, which is related to cerebral activity (León-Carrión & León-Domínguez, 2012). Although this technique is easier than positioning multiple electrodes, as with the EEG, the temporal resolution of the fNIR is worse than the temporal resolution of the EEG. EEG has the best characteristics to record brain wave activity in situations where temporal resolution, is very important, such as driving.

Brain wave activity, analysed through EEG, has been found to be a strong predictor of sleepiness while driving (Lal & Craig, 2002; Eoh, Chung & Kim, 2005;

Jap et al., 2009). Specifically, the brain wave activity related to sleeping while driving consists of four main frequency bands. Delta $\boldsymbol{\delta}$ (0-4 Hz) and theta $\boldsymbol{\theta}$ (4-8 Hz) are known as slow waves activity, and alpha $\boldsymbol{\alpha}$ (8-13 Hz) and beta $\boldsymbol{\beta}$ (13-20 Hz) are known as fast wave activities (Lal & Craig, 2002; Eoh, Chung & Kim, 2005; Jap et al., 2009; Fisch, 2000; Hallvig et al., 2013; Filtness et al., 2012). Although these frequency ranges are commonly used, the frequency interval for each band may differ between different researchers.

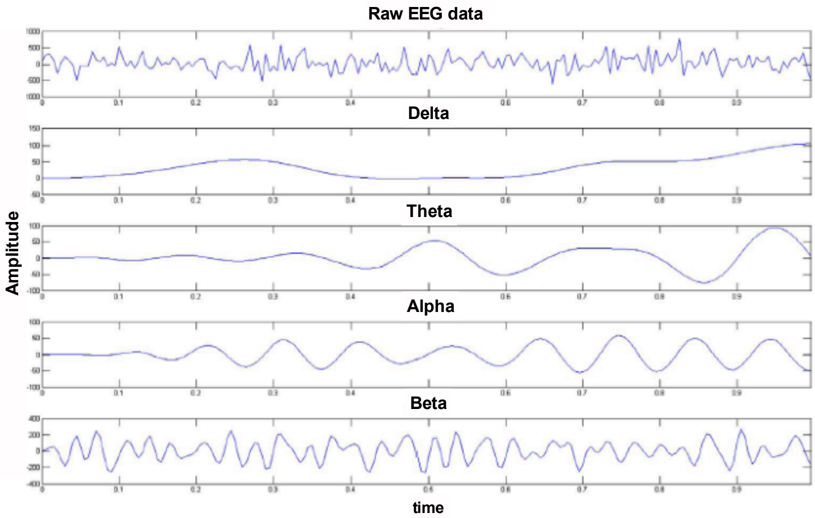


Figure 2-8 Depending on the activity, EEG data (brain wave activity) can be classified in different frequency bands. After performing an FFT in raw EEG data, frequencies from 0-4 Hz belong to the delta band, 4-8 Hz to the theta band, 8-13 to the alpha band and 13-30 to the beta band. Different levels of sleepiness can be determined by analysing the power of each frequency band. (Source: Mohamed et al., 2017) Reprinted from "Towards automated quality assessment measure for EEG signals" by Mohamed et al. Copyright © 2017 by Mohamed et al. Used by permission of Elsevier.

A decrease in beta activity, especially in the temporal and frontal region of the scalp, is related to the increase of sleepiness while driving (Jap et al., 2009; Lal & Craig, 2001a,b; Stern & Engel, 2005; Yeo et al., 2009; Zhao et al., 2012). In addition, an increase in alpha (especially in the occipital region; Yeo et al., 2009) and theta activity (particularly in the frontal, temporal and occipital regions; Yeo et al., 2009) is closely related to an increase in sleepiness of the driver (Akerstedt & Gillberg, 1990; Campagne, Pebayle & Muzet, 2004; Jap et al., 2009; Kecklund & Akerstedt, 1993; Zhao et al., 2012; Lowden et al., 2009) although Eoh, Chung & Kim (2005) found that as sleepiness increases, alpha activity decreases. It was also found that sleep deprivation is associated with an increase activity in alpha and theta activity (Horne & Reyner, 1996; Otmani et al., 2005; Anund et al., 2008). Most of the highest changes in EEG have been found in the occipital and posterior region of the scalp (Cantero et al., 2002).

Another important feature of analysis of sleepiness with EEG is the appearance of microsleeps. Microsleeps are a potential indicator of a person falling asleep as they appear at the end of the drowsy state, i.e. the onset of sleep (Boyle, Tippin, Paul, & Rizzo, 2008). Although the characteristics of microsleep stage are loss of attention, blank stares (Thorpy & Yager, 1991) and long blinking duration (~3 seconds; Boyle et al., 2008), the most accurate way to identify the microsleeps is through analysis of brain waves activity (high bursts in the theta and alpha frequency bands) using EEG (Harrison & Horne, 1996; Moller, Kayumov, Bulmash, Nhan, & Shapiro, 2006). In simulated driving, the appearance of microsleep stage has been correlated with poor performance while driving (Moller et al., 2006; Paul, Boyle, Tippin, & Rizzo, 2005; Risser, Ware, & Freeman, 2000), i.e. difficulty to detect and act to critical situations due to attention lapses happening during microsleeps (Dinges & Kribbs, 1991).

Another commonly used measurement to determine sleepiness through EEG is brain wave activity ratios (Campagne, Pebayle & Muzet, 2004; Otmani et al., 2005; Eoh, Chung & Kim, 2005; Jap et al., 2009). The most commonly used ratios are $\frac{\boldsymbol{\theta}+\boldsymbol{\alpha}}{\boldsymbol{\beta}}$ and $\frac{\boldsymbol{\beta}}{\boldsymbol{\alpha}}$, as an increase in the first ratio and a decrease in the second ratio have been found to be related to the increase in sleepiness while driving. These ratios and the aforementioned frequency bands have been also correlated with physiological and driving behaviour variables as explained next.

Campagne, Pebayle & Muzet (2004) found that while driving in a simulated highway, alpha activity has a positive correlation to the number of times the driver exceeding the lane markings ($r = .407$). There is also a significant difference in the levels of beta frequency and the ratio $\frac{\boldsymbol{\theta}+\boldsymbol{\alpha}}{\boldsymbol{\beta}}$ before an accident and during the accident onset (Eoh, Chung & Kim, 2005). Eoh, Chung and Kim (2005) also found that during curved sections of the road the activity of beta and $\frac{\beta}{\alpha}$ are larger than in straight sections of the road opposed to alpha and $\frac{\theta+\alpha}{\beta}$ (higher activity during straight sections of the roads than during curved sections of the road) (Eoh, Chung & Kim, 2005).

As stated in section 2.1.1, young drivers are the group most at risk to crash due to falling asleep while driving. Due to this reason, it is important to determine whether previous research has found differences in EEG while driving as well as driving behavior between young drivers and older drivers.

2.6 Differences between young and old drivers

Although sleepiness affects every person, the age group most at risk to crash due to falling asleep while driving are young drivers, especially those under the age of 30 (Pack et al., 1995; Akerstedt & Kecklund, 2001; Johns, 2000; Horne & Reyner, 1995). Most of the crashes related to sleeping while driving in young drivers happen when they are driving alone and during the night on long monotonous roads (Horne & Reyner, 1999; Sagberg, 1999; McCartt et al., 2000; Akerstedt et al., 2005; Connor et al., 2001).

Campagne, Pebayle & Muzet (2004) studied the differences in three age groups (20-30 years old, 40-50 years old and 60-70 years old) in sleeping while driving. Although they found that theta, alpha and $\frac{\theta+\alpha}{\beta}$ increases on time irrespective of the age group, only alpha band showed a positive correlation with drivers running off the lane for all age groups (r =.407). For theta band, only old drivers showed a strong positive correlation with running out of lane (r =.882) and the ratio $\frac{\theta+\alpha}{\beta}$ showed no correlation between any age group (Campagne, Pebayle & Muzet, 2004). They also found that young drivers (20-30 years old) run off the road over time segments more frequently than other age groups and that lane speed variation over time is higher in older drivers than in the other groups.

Lowden et al. (2009) found that young drivers are sleepier while driving at night than older drivers. There is a significant difference in the alpha range between young and old drivers during the night drive as well as a higher perceived subjective sleepiness in young drivers compared to older drivers (Lowden et al., 2009). On the other hand, older drivers showed an increment in beta activity, which is associated with wakefulness, i.e. older drivers were less sleepy. These findings are in line with the finding by Philip et al. (2004), who found that at night younger drivers could not maintain a normal driving behaviour (slower reaction times) compared to the older drivers. Filtness et al. (2012) also found a significant difference in the 4-11 Hz frequency band (alpha and theta bands) region of EEG being higher in young drivers compared to older drivers in an afternoon drive.

One of the hypothesis for younger people having a higher increase in sleepiness than older people, is thought to be due to the fact that in old people has been found high cortisol levels, which promotes vigilance and impairs sleep (Steiger, 2002; Chang & Opp, 2001; Gronfier et al., 1999). It was also found that, in terms of risk assessment, young drivers take more risk while driving than old drivers (Rafaely et al., 2006; Ferguson, 2003).

2.7 Types of automation to prevent sleeping while driving

Once sleepiness has been predicted, the following step is to determine the action to be taken by the system to prevent a collision due to falling asleep while driving. The action taken by the system depends on the level of automation of the car, as the system could present a simple warning to the driver or it could take complete control of the driving tasks. The prediction of sleepiness and action to be taken are part of the pre-collision state of a driver avoidance system. If no action is taken, the collision and post-collision state might endanger the driver's life. In the following section, it will be presented the current proposed solutions to predict and prevent (pre-collision) accidents due to falling asleep while driving.

2.7.1 Manual driving

The principle of human-centred automation states that the human should always have the final decision while performing a task (Woods, 1989; Billings, 1997). Present commercial devices in cars that try to reduce accidents related to sleeping while driving work under the principle of human-centred automation. An example is Bosch's drowsiness detection system which predicts drowsiness by analysing the driving behaviour and presents an alarm to the driver, in which case the driver is in full responsibility of deciding to stop driving or not (Sgambati, 2012). Lexus (2012) uses blinking behaviour and the gaze direction of the driver to determine if the driver is paying attention to the road. If the driver is not gazing towards the road, due to inattentiveness or sleepiness, and the car detects a threat in front, the car takes control by reducing the speed autonomously until the driver attempts to avoid the collision, either by reducing speed or steering the car (Lexus, 2012). The problems arising here with these types of devices are small number of levels of sleepiness defined and the large and fast jumps between high and low level of automation.

According to Sheridan (1992a,b) and Parasuraman et al. (2000), there are 10 levels of automation, from the lowest level being the human has complete control without any help from the computer to the highest level being the computer has complete control without any help from the human. This is just one example of the many pieces of work that attempt to determine different levels of automation (Parasuraman et al., 2000). According to SAE (On-Road Automated Vehicle Standards Committee, 2014), there are five levels of automation (No automation, Driver Assistance, Partial Automation, Conditional Automation, High Automation and Full Automation). This is one of the most commonly used hierarchies in the driving domain. Either of the proposed hierarchies can be used to explain the problem of high jumps in automation. In a case where the human is in control, the level of automation is low (Sheridan, 1992a,b; Parasuraman et al., 2000). As sleepiness increases, the level of automation remains the same until the point when the driver falls asleep completely or loses attention of the road completely. When the drivers falls asleep completely the car can trigger an alarm or a warning, i.e. a low level of automation, or it can take complete control, i.e. a high level of automation (Sheridan, 1992a,b; Parasuraman et al., 2000). In the first case (an alarm is trigger, i.e. low level of automation), the problem arising is the lack of support from the system to the driver in the driving tasks during dangerous levels of sleepiness. As stated before, during high levels of sleepiness the driver has reduced capabilities concerning his/her driving skills. Therefore, defining a small number of levels of sleepiness leads to the situation where driver is still in charge of the driving tasks during dangerous levels of sleepiness. In the second case (when the car takes complete control), the driver shows an attempt to avoid the collision, either by braking or a steering manoeuvre. In this case, the automation will go from high level back to low level, i.e. the driver is in control, and the system will not know if the action taken by the driver is a correct one or not (Sheridan, 1992a,b; Parasuraman et al., 2000). Parasuraman et al. (1991) found these fast changes between high and low level of automation are disruptive to performance. Even if the driver did not showed an attempt to avoid the collision, there is another problem arising in regards to knowing and assessing when to give back control taking into account the condition of the driver. (Merat et al., 2014; Endsley, 1995; Carsten et al., 2012).

If the system returns control to a driver with high level of sleepiness, the consequence might be a collision as presented in this section (Merat et al, 2014; Endsley, 1995; Carsten et al., 2012). It has been found that only naps and caffeine reduces fatigue effectively (Reyner & Horne, 1997, 2000, 2002; Philip et al., 2005). When drivers nap or were given coffee, the drivers' frequency of incorrect lane departures reduced (De Valck & Cluydts, 2001; Horne & Reyner, 1996). This means that any system dedicated to predict and prevent should take in consideration the state of the driver before giving back control. As a solution to this case, many companies are working towards fully autonomous cars (Markoff, 2010; Volvo-Car-Group, 2014), i.e. removing the driver from the driving scenario.

2.7.2 Autonomous cars

Another solution to people falling asleep while driving is autonomous cars, where the car is always in complete control of the driving actions. Many companies (Google, Volvo, Tesla, etc.) are in the race to develop and bring to the market the first commercial autonomous car (Markoff, 2010; Volvo-Car-Group, 2014; McHugh, 2015). Unfortunately, several problems arise due to autonomous systems. One of the biggest problems is the "out of the loop" state (Woods, 1989; Wickens, 1994; Endsley & Kiris, 1995; Sarter and Woods, 1995; Parasuraman & Riley, 1997; Sarter et al., 1997; Young & Stanton, 2002; Inagaki & Stahre, 2004). When the driver is not in charge of any driving action, the level of involvement in the driving task from the driver reduces, as he/she is "out of the loop" of the task (Woods, 1989; Wickens, 1994; Endsley & Kiris, 1995; Sarter and Woods, 1995; Parasuraman & Riley, 1997;

Sarter et al., 1997; Young & Stanton, 2002; Inagaki & Stahre, 2004; Carsten et al., 2012, Kaber & Endsley, 2004). If the driver has to take control again over the car, the performance of the driver will be poor due to a decrease in driving experience and slow reaction due to an over-trust on the autonomous system (Young & Stanton, 2002).

Many researchers have studied the problems with automation and the transition between autonomous and manual driving (Merat et al., 2014; Endsley, 1995; Carsten et al., 2012). Merat et al. (2014) found that eye movement fixations and lateral control were highly variable when changing between autonomous and manual driving. Carsten et al. (2012) found that as automation increases, drivers are more likely to divert attention from the road to secondary non-driving related tasks and when presented with the need to resume control their performance deteriorated (Merat et al., 2012). Finally, Endsley (1995) found that automation in one area of the car affects the driver in other areas of the driving task. All these studies suggest that automation can negatively affect the behaviour of the drivers.

A certain challenge with autonomous cars is already happening in another transport field that relies in autonomous systems, i.e. the aviation industry. Autonomous systems in airplanes have presented many challenges that have led to crashes: pilots have to rely on their memory to determine the actions of the automated system; there is complacency and over-trust in the system; lack of understanding on the functionality of the automated decision making process; and a decline in the skills of the pilots due to being out of the loop (Billings, 1997; Wiener, 1998). These problems could be some of the problems that autonomous cars could also encounter.

2.7.3 Adaptive automation

Adaptive automation, unlike manual driving or completely autonomous cars, can change dynamically over time the tasks being held by the human and the tasks being held by the car depending on changes in the performance of the human driver (Rouse, 1994; Inagaki, 2003). Although adaptive automation has been studied as a solution to many automation problems in the aviation industry, adaptive automation has not been developed deeply as a solution in the automotive industry (Inagaki, 2003, 2006, 2009). Goodrich et al. (2006) and Flemisch et al. (2003) explained adaptive automation with the H metaphor. The H metaphor (the H refers to an analogy from horse riding) is presented as a hypothetical situation when an individual riding a bicycle in the forest and when the individual is riding a horse in the forest.

In the case of an individual riding a bicycle, the person is in control of the direction, and speed of the bicycle. If the person wants to see the map for directions, the person will have to stop biking, look at the map and then continue the journey. When the individual is riding a horse, the person on top of the horse is in control of the direction, and speed of the horse. In contrast with riding a bicycle, if the person on top of the horse decides to see the map for direction, he/she would be able to leave the reins, check the map and then hold the reins back again for control of the horse. During the time the reins are loose, the horse is in control of its movement, speed and direction. The person has complete trust that the horse can manage its way across the forest without bumping into a tree or falling into a hole. This means that if the rider is not in a state to control the horse (due to inattention, physical inability, etc.), the horse will be able to control itself in a safe manner. In the same way, if the person on top of the horse decides to take an action that could put the horse in danger, the horse will instinctively try to avoid or refuse to do the action. This outlines the fundamental rationale of the H metaphor (H-mode; Goodrich et al., 2006; Flemisch et al., 2003). In an experiment, conducted by NASA, to examine the public approval of adaptive automation (H-mode), participants stated that they would use the system as they felt confortable with it (Goodrich et al., 2006).

One case being studied with adaptive automation in the automotive field is lane departure (Inagaki, 2007). The allocation of tasks in this system depends on the attention of the driver of possible cars in an adjacent lane when trying to enter that lane. If the driver tries to change lanes and fails to see an upcoming car in that lane or misjudges the velocity and position of another car coming in that lane, the adaptive automation system of the car will block the steering wheel (car in control), preventing the driver from changing lanes and this way avoids the possibility of a crash. If the driver assessed the change of lane correctly and there is no probability of a crash, the driver is in control and will be able to change lane freely. In Inagaki (2007), the adaptive automation system varies the level of control the car should take depending on the attention and the assessment the driver has of the situation while changing lanes. Mulder et al. (2012) also tested the advantages of adaptive automation through haptic control and found that adaptive automation increased the performance of the driver and reduced the out of the loop problem. By detecting multiple levels of sleepiness, an adaptive automation system can be developed to ensure the safety of the drivers in cases of falling asleep while driving.

2.8 Conclusion

This chapter has highlighted the hidden complexity of a task such as driving, demonstrating that a high level of attention and a high level of skills required to complete the task safely and successfully. Any situation that reduces the attention and performance skills of the driver, e.g. sleepiness, can increase the probability of an accident. It has been found that sleeping while driving is a serious cause of danger in young drivers. Many companies have developed solutions to predict and prevent accidents due to people falling asleep at the wheel. Unfortunately, most solutions predict sleep when it is already too risky for the driver to be driving. The current methods use blinking and driving behaviour as the estimator for sleep. Unfortunately, these measures lack sensitivity. EEG has been found to be a more accurate predictor of increase in sleepiness.

The other problem with these solutions is that the action taken is human centred, i.e. the human is always in control. When predicting sleepiness, an audio or a visual warning will let the driver know he/she is sleepy and it is the driver who has to take a decision accordingly. Allowing the driver to decide when he/she is sleepy might not be the best solution. This chapter stated that it is not well understood if drivers tend to estimate incorrectly their sleepiness. This could lead to the driver taking an incorrect action under the false assumption that he/she is not that sleepy. In addition, the solutions discussed in this chapter face the problem of high jumps in the levels of automation, which leads to accidents.

The present PhD planned to use brain wave activity to determine different levels of sleepiness between alert and asleep (in contrast with solutions that only have a binary or tertiary classification of the levels of sleepiness). Machine Learning algorithms (MLAs) are a potential tool that can be used to predict the proposed multiple levels of sleepiness, using EEG as ground truth, and driving and physiological variables as features. However, before using MLA to determine the different levels of sleepiness, it is important to analyse the advantages of using MLA's to predict the data. The following section describes the theory and uses of different types of MLAs, as well as the MLA's used by researchers to determine different levels of sleepiness.

Chapter 3:

Literature survey: Classification of sleepiness in drivers

3. Literature survey: Classification of sleepiness in drivers

3.1 Introduction

As presented in the previous chapter, high levels of sleepiness are one of the major causes of accidents amongst young drivers, leading to large number of fatalities (Royal, 2003; Philip & Akerstedt, 2006; Connor et al., 2001; NHTSA, 1998). As such, predicting high levels of sleepiness is a critical step of action needed to prevent this type of accident.

Sleepiness has been correlated with a number of different driving and physiological variables (Sadock & Sadock, 20002005; Carskadon & Dement, 2011, Boyle et al., 2008; Brookhuis & de Waard, 2010; Filtness, Reyner & Horne, 2012; Jap, Lal & Fisher, 2011). This leads to the possibility of developing systems that can determine different levels of sleepiness using the information of the person's physiological and driving behaviour. With the development of faster and better Machine Learning Algorithms (MLAs) it is now possible to have a better prediction of the levels of sleepiness without relying in determine specific threshold values. Machine Learning is an Artificial Intelligence technique that allows computers to acquire knowledge without explicitly being told. The approach has gained recognition across many different fields, from the gaming industry to aerospace exploration missions.

In the following chapter, a detailed insight into a method of manual classification of sleepiness is presented. The chapter then continues to describe Machine Learning and analyses the advantages of using MLAs to predict different levels of sleepiness. An explanation of the evolution of Machine Learning and the different types of algorithms is presented. The chapter concludes by presenting different MLAs that could be used to predict sleepiness in drivers.

3.2 Manual prediction of sleepiness

As presented in previous chapters, it has been found that in high levels of sleepiness, drivers tend to deviate more in their lane position, their speed changes and they present physiological indicators of "sleep" state, e.g. yawning and head nodding (Hartley et al., 2000; Haworth & Vulcan, 1991; Kaplan et al., 2007; Lal & Craig, 2005). This would suggest the possibility of defining threshold values using driving and/or physiological variables to index sleepiness. For example, Lal et al. (2003) developed a system that determines specific EEG threshold values for different levels of sleepiness and this has one of the highest accuracies presented in literature. In this system, the threshold values are determined using the mean and standard deviation of the EEG baseline of the drivers, i.e. the EEG data of the drivers during a period of time considered to be an "alert" state, and different coefficients for each specific levels of sleepiness determined by the researchers (Lal et al., 2003). The different levels of sleepiness that they defined where "awake", "transitional", "transitional-posttransitional" and "posttransitional". This meant that the researchers defined four sets of coefficients, one for each level of sleepiness.

Although high accuracy was achieved in the system presented by Lal et al. (2003), certain problems were found with this approach. The first problem is that the system requires determining the mean and standard deviation of all the drivers before assessing the levels of sleepiness, i.e. the system does not work without the baseline information of the all the drivers involved. This means that if a new driver is to be assessed by the system, it is necessary to add the baseline information of this new driver for the system to update the mean and standard deviation of the EEG baseline of the group of drivers. Machine Learning Algorithms learn from a number of examples and develops a system that does not have to be updated every time a new driver needs to be assessed. Another issue found with the manual methods to determine different levels of sleepiness is individuality of the drivers' behaviour (Campagne et al., 2004; Lowden et al., 2009; Filtness et al., 2012). It has been found that EEG and sleepiness behaviour changes depending on characteristics like age and body-mass index (Campagne et al., 2004; Lowden et al., 2009; Filtness et al., 2012). This means that if the group of drivers involved in the manual system have different EEG behaviour due to age, this will increase the variability of the EEG baseline, therefore increasing the standard deviation value. As the manual system depends on the calculation of the standard deviation of the EEG baseline, having a big standard deviation will increase the number of false positives in the prediction system. Finally, manual systems as the one developed by Lal et al. (2003) use intrusive systems (every drivers EEG data needs to be recorded) that in the real world are not easily implemented in a car (Lal et al., 2003).

Contrary to manual methods, i.e. methods where the threshold is defined by the user, to predict sleepiness, MLAs have the advantage to learn from a set of examples. Thus, when assessing new drivers it is not necessary to update the system with the information of these new drivers. The problem of individuality is also assessed by MLAs, as the MLAs can learn from big datasets to determine patterns that will reduce the variability of the drivers. Finally, MLAs determine the levels of sleepiness by assessing different features (Alpaydin, 2010; Harrington, 2012). This means that it can learn from a combination of driving and physiological features, which are non-intrusive and easy to implement in a real car scenario (Lal et al., 2003; Sayed & Eskandarian, 2001; Yeo et al., 2009; Shuyan & Gangtie, 2009; Yang et al., 2010; Patel et al., 2011). The following sections explain in more detail the definition of MLAs and the advantages of using this technique to determine different levels of sleepiness.

3.3 Machine Learning Algorithms

3.3.1 Definition of Machine Learning Algorithms

3.3.1.1 Machine and algorithms

In order to understand the term Machine Learning Algorithms (MLA), the term will be decomposed into the three words that comprise it. The first word - “Machine” - refers to a computer system capable of acting according to a specific set of instructions developed by a user, i.e. a program (Evans, 2011). The second word that will be defined is algorithm- which can be understood as a set of instructions defined by the user, i.e. a program, which transforms a specific input into a desired output every time the program runs (Alpaydin, 2010). These types of algorithms are called deterministic algorithms (Marion, 2008; Alpaydin, 2010). This algorithm is then converted into a programming language that a computer can understand. With these two terms defined, it is possible to conclude that the term “Machine Algorithm” refers to a computer which outputs an answer according to a specific input following the set of instructions developed by the user. However, what happens when the set of instructions are not clear or not known by the user? “Learning” is the solution to this question and this is the primary reason why MLAs have garnered such interest in many different fields.

3.3.1.2 Learning

As discussed before, there are many situations when it is difficult or even impossible for the user to set up an explicit set of instructions for the computer to follow, as presented in the following section.

One example is where a user is not able to explain the process except by example data. There are many situations where the user will not be able to explain a process or his/her experience of a process (Alpaydin, 2010; Harrington, 2012). A common example of this type of situation is the recognition of spoken communication. Humans are capable of understanding spoken speech, largely irrespective of the differences in age, gender or accent. Describing this process in an explicit set of instructions is difficult, as it is still not completely known how this process works in the human brain. In this case, the algorithm cannot be defined by the user as a set of explicit instructions. Instead, the computer is presented with a large set of data of different people speaking. By having a large set of data of different people speaking, the computer is trained to find patterns and adapt to different accents, age and speech forms to be able to “understand” spoken speech.

Another situation arises when the volume of data are too large for the user to be able to analyse effectively and define a deterministic algorithm. Due to the increase in digital storage space capacity over time and the increase of the digital communication speed, there is a vast amount of data available to a great majority of people (Murphy, 2012). Every minute more than 100 hours of video are uploaded to Youtube, there are more than 40 billion websites on the internet, laboratories have access to the genome information of thousands of people and supermarkets such as Walmart see more than 1 million transactions per hour which leads to a database of more than $\mathbf{2.5x10^{15}}$ bytes of information. With this amount of big data, it is not possible to analyse specific cases manually (Murphy, 2012; Marsland, 2015; Alpaydin, 2010). In addition, due to the amount of data, there might be hidden patterns not possible to detect by the user developing the algorithm (Marsland, 2015; Alpaydin, 2010). In this case, the data are fed to the computer, which is trained to find patterns in the data and come up with conclusions by itself.

A third situation arises when users not know certain characteristics of the environment where the machine will be performing the task. There are certain environments where humans cannot explore due to physiological limitations, e.g. the bottom of the sea, beyond certain point of the universe, the core of the earth (Marsland, 2015; Murphy, 2012; Alpaydin, 2010). Due to the lack of experience of the humans in these environments, it is impossible to predict and exactly describe the conditions to a machine and how to adapt to those conditions (Marsland, 2015; Murphy, 2012; Alpaydin, 2010; Castaño et al., 2003). An example is the Mars Exploration Rover mission (National Aeronautics and Space Administration, 2015; Castaño et al., 2003; McGovern & Wagstaff, 2011). The environment of Mars is largely unknown to the humans so the machine has to adapt to an unknown environment without an explicit set of instructions developed by the user (Castaño et al., 2003; McGovern & Wagstaff, 2011).

The final situation where it is not possible for the user to set up an explicit set of instructions for the computer to follow is when the environment is volatile-changing over time. There are everyday situations where the environment changes constantly, e.g. driving (Marsland, 2015; Murphy, 2012; Alpaydin, 2010). If a machine is intended for driving purposes, it will have to adapt to the constantly changing driving scenarios (Tamatsu & Nitanda, 2014). A system that can learn will tackle many of the problems presented in the cases where a deterministic algorithm is not convenient (Harrington, 2012; Bell, 2015; Marsland, 2015; Murphy, 2012; Alpaydin, 2010).

In MLA terms, 'learning' refers to the ability of a system to find patterns in the data (Harrington, 2012; Bell, 2015; Marsland, 2015; Murphy, 2012; Alpaydin, 2010). Taken together, the term MLA is perhaps best defined by Arthur Samuel, a pioneer in self-learning computer programmes, as a "Field of study that gives computers the ability to learn without being explicitly programmed" (Bell, 2015, p. 2). Now that MLAs have been defined, the next section focuses in the different types of MLAs.

3.3.2 Evolution of Machine Learning Algorithms

MLAs have evolved in the recent years and are widely used in many different research fields (Harrington, 2012; Bell, 2015; Marsland, 2015; Murphy, 2012; Alpaydin, 2010). The development of new algorithms has risen since the past decade, as seen in Figure 3-1, which has allowed MLAs to be present in many aspects of our daily life. Every day activities such as using social media, having an email account which can classify important and unimportant emails and receiving personalised advertisements from online shopping services, are few of the many examples where people have an interaction with MLA. MLAs have also been used in more specific fields such as in the development of a machine capable of playing a game like chess (Alpaydin, 2010; Block et al., 2008; Campbell, Hoane Jr. & Hsu, 2002), the Mars Exploration Rover robot currently exploring Mars surface (Castaño et al., 2003; McGovern & Wagstaff, 2011) or intelligent cars which are able to predict and react according to different situations (Davis, 2014).

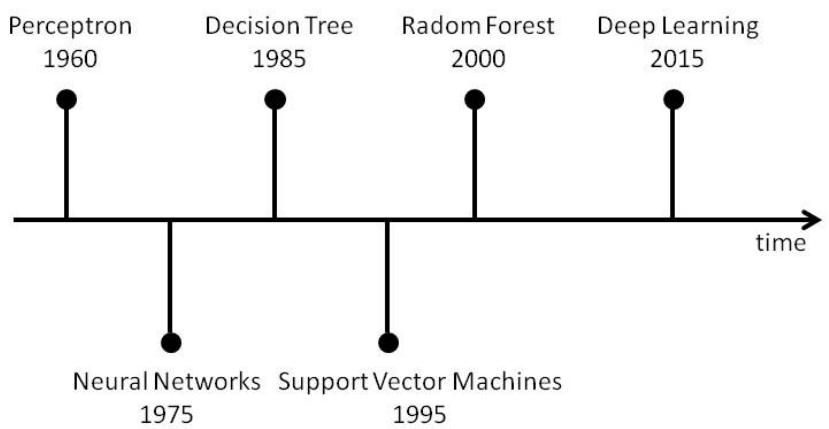


Figure 3-1 Timeline of the development of different MLAs that are still used nowadays (Adapted from: Bi, 2014)

In the context of driving, MLA has been used by many researchers to predict sleepiness while driving and react accordingly to ensure the safety of the driver (Yeo et al., 2009; Shuyan & Gangtie, 2009; Yang et al., 2010; Patel et al., 2011). This means that MLAs enable machines to adapt themselves to many different situations and act to solve a specific problem (Harrington, 2012; Bell, 2015; Marsland, 2015;

Murphy, 2012; Alpaydin, 2010). The following section presents different types of MLAs, their advantages and disadvantages, before presenting an analysis of the MLAs used to predict sleepiness in driving.

3.3.3 Types of Machine Learning Algorithms

MLA can be classified in three different categories: supervised learning, unsupervised learning and reinforcement learning (Harrington, 2012; Bell, 2015; Marsland, 2015; Murphy, 2012; Alpaydin, 2010).

3.3.3.1 Supervised learning

In supervised learning, examples containing information of the desired output are presented to the MLA. For the MLAs presented in this chapter, every dataset that is input into a MLA is called a feature set. A feature set is the set of variables in the dataset, e.g. size of a house in a specific area, size of an engine in certain car models, the lane deviation of the driver. The feature vector, i.e. an instantiation of the feature set, is the data the MLA will use to learn and train itself. In supervised learning, the dataset also contains the expected value associated to the feature set. These values are called target set, e.g. price of a house according to specification of the house, acceleration power of a car according to specifications of the engine, level of sleepiness of a driver according to the driving behaviour. Most of the times, the feature set is not specified as a continuous value but as a class. The feature sets are defined into classes (discrete sets or categories) according to their properties or attributes. The target vector, an instantiation of the target set, contains the class that the MLA is trying to predict using the feature vectors. The MLA will try to learn by reducing the error between the predicted value by the MLA and the expected value. This can be explained in the following example.

Imagine that a person decides to collect the data of X houses around a specific area. The person obtains the size of the houses around a specific area, the number of rooms and the prices of each house. This means that the person has a dataset composed by a features set characteristics of the house and target set the price of the house. That person is interested in using a MLA that given the size of any house and the number of rooms, the MLA can predict the price of the house. For each set of feature vectors in the feature set, i.e. characteristics of the house, the MLA will predict a price, which will be compared with the real price of the house. In each 'learning' iteration process, the MLA will update its parameters to reduce the error, i.e. the difference between the predicted value and the real value of the house.

3.3.3.2 Unsupervised learning

The second types of MLA are unsupervised learning algorithms. In comparison to the supervised learning algorithms, the unsupervised learning algorithms do not have a target set, i.e. the user does not know the expected output. These types of algorithms are used when the user does not have the knowledge of the real value for the features obtained. In this case, the MLA will try to find patterns in the features input by the user and create its own target class. These types of MLA are used by companies like Amazon (Linden, Smith & York, 2003). Researchers at Amazon are not sure how people can be categorised for targeted marketing (by novels genre, preferred sports, preferred videogames, etc.). Instead they allow the MLA to find clusters depending on the items people buy through the website. After finding different clusters, e.g. people who like crime novels, if a person is considered to be part of that cluster, Amazon's MLA will recommend books that belong to the crime novel cluster.

3.3.3.3 Reinforcement learning

The final type is reinforcement learning. These type of algorithms are not concerned with a specific output from a specific set of features, instead they are concerned with a correct sequence of actions that might reach a desired goal (Alpaydin, 2010). One example for these types of algorithms can be found in machines dedicated to playing chess (Alpaydin, 2010; Block, et al., 2008; Campbell, Hoane Jr. & Hsu, 2002; Hsu, 1999). The output of a specific play (action) might not be important meanwhile the policy, i.e. sequence of correct actions, lead to the desired output, i.e. winning the game (Harrington, 2012; Bell, 2015; Marsland, 2015; Murphy, 2012; Alpaydin, 2010; Block, et al., 2008; Campbell, Hoane Jr. & Hsu, 2002).

This means that the type of MLA that should be chosen depends on the goal, the features set and the target set. In the field of driving and sleeping, many researchers have used MLAs that can predict sleepiness while driving (Yeo et al., 2009; Shuyan & Gangtie, 2009; Yang et al., 2010; Patel et al., 2011). To train the MLA, the researchers input the behaviour of people when they are driving while being awake and their behaviour when they are driving whilst sleepy. The MLA will then obtain the behaviour of a new driver and predict if the driver is in an awake or sleepy state. This is a supervised learning approach, i.e. learning by example. In the following section, the different types of supervised learning algorithms used to predict sleepiness in driving presented in literature will be discussed.

3.4 Machine Learning Algorithms to predict sleep in driving

As discussed in chapter 2, sleeping while driving is one of the biggest causes of accidents on the road in young people. Being able to predict increase in sleepiness in the driver is the first step to prevent accidents due to sleepy drivers. Although many researchers have developed MLAs that can predict with high accuracy different levels of sleepiness, the systems only work with a binary classification of sleepiness (Yeo et al., 2009; Shuyan & Gangtie, 2009; Yang et al., 2010; Patel et al., 2011). As presented in chapter 2, a binary or ternary classification of sleepiness might lead to detection of high levels of sleepiness when the driver is already at high risk of having an accident (Klauer et al., 2006; Lamond & Dawson, 1999). Also, this might lead to high jumps of automation in the actions taken by the system, e.g. if the system gives a warning at low level of sleepiness and takes partial or complete control of the driving task at high levels of sleepiness. High jumps in the level of automation of a system have been found to lead to accidents (Merat et al., 2014; Endsley, 1995; Carsten et al., 2012).

Although only binary and ternary classifications of sleepiness have been achieved thus far, it is important to analyse the variables and MLAs used by other researchers to solve the problem of predicting sleepiness in drivers. Patel et al. (2011) used a supervised learning Artificial Neural Network algorithm that achieved 90% accuracy in sleep prediction. Yeo et al. (2009) and Shuyan & Gangtie (2009) developed supervised learning Support Vector Machines algorithms that achieved around 99% and 85% accuracy, respectively, in predicting sleepiness in drivers. Yang et al. (2010) developed a supervised learning dynamic Bayesian Network algorithm, which effectively predicted the sleepy states of drivers. Although all results are presented in accuracy percentages, there are different types of evaluation measures that can be used to determine the suitability of the MLA. The following section discusses different types of evaluation measures for MLAs. Once the evaluation measures are described, the types of supervised learning MLAs used to predict sleepiness in drivers is discussed.

3.4.1 Evaluation measures for Machine Learning Algorithms

As all the results presented above are measured by the accuracy of the MLA, it is important to determine how the accuracy is calculated and other evaluation measures that can be used in MLAs. As discussed before, the target vectors contain the expected results to be predicted by the MLA (Lavesson, 2006; Costa, 2007). The values of the target vectors are discrete classes. In the case of predicting different levels of sleepiness in drivers, the classes of the target set are “awake” or “sleep”. If the aim of the system is to predict a “sleep” level of sleepiness in the driver, the class “sleep” is a positive outcome of the MLA. Therefore, the class “awake” is considered the negative outcome of the MLA. Whenever the MLA predicts a positive outcome (“sleep”) and the expected outcome is positive as well, then this is known as a true positive. A positive outcome predicted when a negative (“awake”) outcome is expected, is a false positive. Vice versa, when a negative prediction is given and the expected outcome is negative then is called true negative; and a negative outcome when a positive out come is expected, is called false negative. Accuracy depends on both, the number of true positives and the number of true negative, as presented in equation (1).

$$\boldsymbol{A} = \frac{\boldsymbol{True\ Positive + True\ Negative}}{\boldsymbol{True\ Positive + True\ Negative + False\ Positive + False\ Negative}} \qquad (1)$$

Although accuracy is a measure often used, it is worth mentioning different measures to determine the suitability of the MLA. Error rate is the measure of incorrect predictions by the MLA and it is calculated as Error rate = 1-Accuracy. The error rate and the accuracy can be used in multiclass problems, i.e. when the target set has more than two types of classes. The following measures are used only in binary classification problems. The first measure is called recall or true positive rate. This can be used when the algorithm needs to have a high number of correct true positive predictions. The recall can be calculated using the following equation (2)

$$\boldsymbol{R} = \frac{\boldsymbol{True\ Positive}}{\boldsymbol{True\ Positive+False\ Negative}} \tag{2}$$

The second measure that can be used in binary classification problems is the specificity. Contrary to the recall measure, specificity is focused in determining how good the algorithm is predicting true negatives. To calculate the specificity of an algorithm, equation (3) can be used. Finally, precision is a measure used in binary classification to determine the probability that a positive prediction is right. Equation (4) is used when precision is needed to determine the suitability of the MLA. For the problem of detecting sleepiness in drivers, it is important to have a high level of correct negative and positive prediction in sleepiness, i.e. it is important to avoid false positive so the user do not get annoyed by the system and avoid false negative which could put the driver in danger. As such, for the problem of detecting sleepiness, accuracy appears to be a suitable measure to assess the efficacy of the MLA. With one of the aims of the PhD study being to determine multiple levels of sleepiness, the evaluation measures for binary classification problems are not appropriate.

$$\boldsymbol{S} = \frac{\boldsymbol{True\ Negative}}{\boldsymbol{False\ Positive+True\ Negative}} \tag{3}$$

$$\boldsymbol{P} = \frac{\boldsymbol{True\ Positive}}{\boldsymbol{True\ Positive+False\ Positive}} \tag{4}$$

3.4.2 Artificial Neural Networks

Artificial Neural Networks (ANN) is one of the first MLAs that were developed, as shown in Figure 3-1. The basic functionality of the ANN was inspired in the architecture and the functionality of the brain (at a very basic level) (Hagan et al., 2014; Cohen, 2014; Bell, 2015; Marsland, 2015). As the "learning" process occurs within the brain, researchers thought that by mimicking the way the brain works it might be possible to create a MLA that could learn as well as the humans. Unfortunately, the processes happening within the brain are not completely known.

Interestingly, in the beginning, it was thought that ANN modelled exactly the functionality of a human brain. Later, it was found that it resembles more, in a very basic way, the functionality of an animal brain.

The brain is composed of more than $\mathbf{10^{11}}$ cells (Hagan et al., 2014). These cells are called neurons and are the processing units of the brain (Hagan et al., 2014; Cohen, 2014; Marsland, 2015). Each neuron is interconnected with other neighbouring neurons (around $\mathbf{10^{4}}$ connections per element) (Hagan et al., 2014). Each neuron is divided in three sections: the cell body, the axon and the dendrites. Each neuron is interconnected with other neurons' axons through its dendrites. The connections between the neurons are called synapses.

A communication between two neurons happens due to changes in the transmitter chemicals within the brain. Consequently, the electrical potential inside a neuron changes. If the change of the electrical potential surpasses a certain threshold, this specific neuron will 'fire' an electrical pulse, which will travel down its axon to the other neurons connected to that axon. This electrical pulse, and all the other electrical pulses received by a neuron through its dendrites are then summed in the cell body. If the summation of the pulses surpasses the threshold of that specific neuron, a new pulse is sent from that neuron, through its axon, to all the neurons connected to that axon.

There are many neural structures and connections defined at birth. In addition, it has been found that throughout neural development, it is possible to modify "the strength of synaptic connections between neurons" (Marsland, 2015, p. 40) and create new connections (Hagan et al., 2014; Cohen, 2014; Marsland, 2015). This is called plasticity and it is believed to be the main process of learning (Marsland, 2015). Figure 3-2 shows the mathematical model of a neuron produced by McCulloch and Pitts (Marsland, 2015) and combines all the concepts that have been presented in this section. The $\boldsymbol{X_n}$ represent the electrical pulse sent by other neurons through their axons to the dendrites of a specific neuron. The cell body then sums all the electrical pulses and compares it to a pre-determined threshold. If the sum is higher than the threshold, the neuron will 'fire', otherwise it will not. The $\boldsymbol{W_n}$ are called the weights and it represents how strong is the connection between the two neurons. The plasticity term is represented here. The weights will change depending on the learning process, i.e. the synaptic connection between two neurons will change. At the end of every "learning" iteration, the ANN adapts these weights and the threshold in order to 'learn' from a dataset.

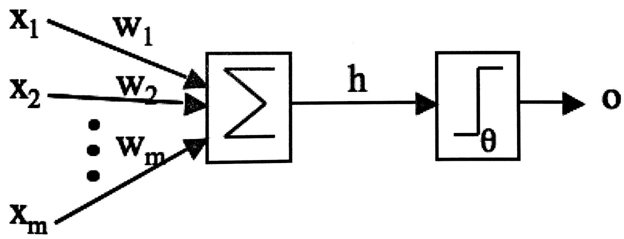


Figure 3-2 Mathematical model of a neuron produced by McCulloch and Pitts in 1943 (Source: Marsland, 2015) Reprinted from "Machine Learning: An Algorithmic Perspective" by Stephen Marsland. Copyright © 2015 by Stephen Marsland. Used by permission of CRC Press, Florida, USA.

The mathematical model presented in Figure 3-2 can be adapted to create a MLA, i.e. the ANN. Figure 3-3 shows a single-input neuron algorithm, i.e. one neuron connected to another neuron. In an ANN, every input neuron and every output from a neuron is called a node. The input node p is multiplied by the weight w (strength of the synaptic connection). The input b is called the bias node. This bias node will allow an algorithm to adapt even if the neuron input is a zero value. The input node and the bias node are sum up and compare to a threshold function f. The output a is defined by the equation $\boldsymbol{a} = \boldsymbol{f}(\boldsymbol{w} * \boldsymbol{p} + \boldsymbol{b})$. The threshold function, also called transfer function, determines the output of the neuron. The most common transfer functions are hard limit, linear and sigmoid.

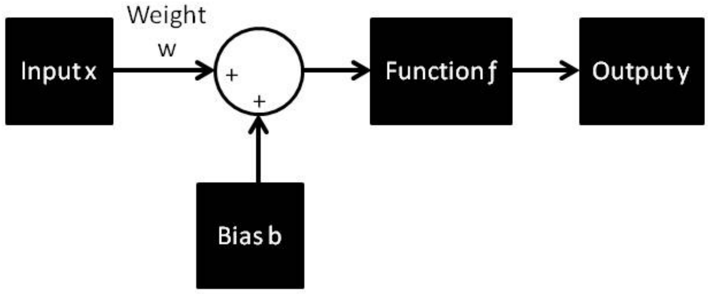


Figure 3-3 Single-input neuron ANN design (Adapted from Hagan et al., 2014)

Once the input p and the bias b are summed and analysed by the transfer function, in a supervised learning approach, output a is compared to the expected (real) value. The difference is then be used to update the value of weight w, as shown in equation (5).

$$\boldsymbol{w_n \leftarrow w_n - \eta(y_n - a_n) * p_n} \tag{5}$$

In equation (5), $\boldsymbol{w_n}$ is the weight that will be update, i.e. learning process; $\boldsymbol{y_n}$ is the expected (real) value and $\boldsymbol{a_n}$ is the output of the neuron; $\boldsymbol{p_n}$ is the input value; $\boldsymbol{\eta}$ is a parameter called the learning rate. The learning rate parameter is a predefined value that will determine the rate of change of the weight value. If the learning rate is too high, the learning rate is faster but it might overshoot the best weight value. If the learning rate is too low, the learning rate is slow but it is more accurate in reaching the best weight value.

An ANN can have multiple input nodes (many neurons connected to one neuron) as shown in Figure 3-4. More complicated ANN can be designed with multiple inputs and output nodes within many layers (the layer in between the input nodes layer and the output nodes layer are called hidden layers). In some cases, connecting the output of one neuron back as input can generate feedback. If no feedback is present, the layer is called feed-forward. Independently of the number of input, output and hidden layers, the process is the same as explained for the single input ANN (Figure 3-3).

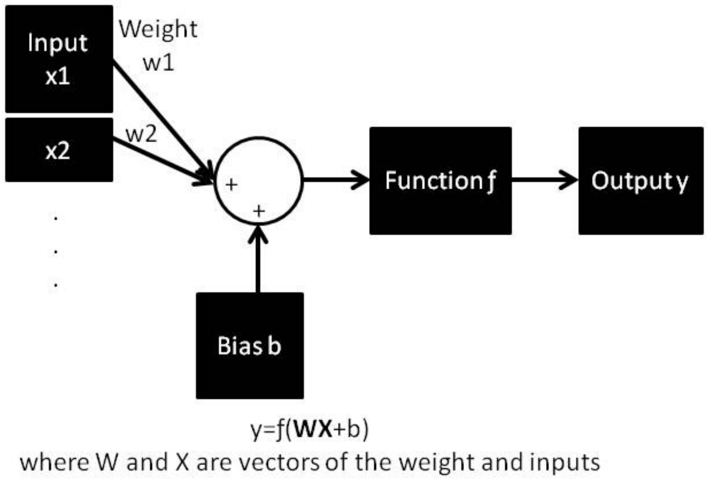


Figure 3-4 Multiple input neuron ANN design (Adapted from: Hagan et al., 2014)

3.4.2.1 ANN to predict sleep while driving

ANN has been used to predict the sate of sleep in drivers (Patel et al., 2011; Vuckovic et al., 2002). For example, Patel et al. (2011) used the heart rate variability of 12 truck drivers (mean age 47, SD=11) to determine different levels of sleepiness. Heart rate variability was used as it has been related to a decrease in mental workload and increase in driver fatigue, as presented in chapter 2. A single layer feed-forward ANN was used. The transfer function was a bipolar logistic function (Patel et al., 2011; Hagan et al., 2014). To train the ANN, heart rate variability of awake truck drivers and heart rate variability of sleepy truck drivers was input into the MLA. The heart rate variability was presented as a 30x30 pixel spectral image of the power spectral density analysis obtained through Fast Fourier Transformation of the electrocardiography data obtained from each participant, as seen in Figure 3-5. The ANN had 900 input nodes (30x30 pixels) and two output nodes (awake and sleep nodes). The activation function was a bipolar logistic function and a learning constant of 0.5 gave the best results. The NN was trained using five data sets (5 participants) and was tested using five data sets. The dataset was obtained from Lal and Craig (2002) and was classified into alert and sleep by subjective visual analysis (using a video image of the participants). The algorithm gave 90% accuracy when predicting sleepiness in drivers.

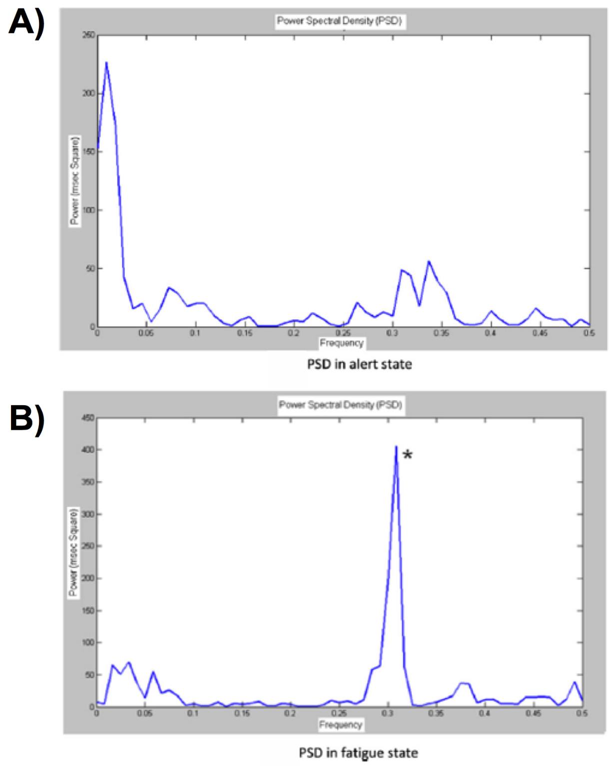


Figure 3-5 Power spectra density of the heart rate variability of “awake” drivers (a) and “sleepy” drivers (b) during a driving simulator study. (Source: Patel et al., 2011). Reprinted from “Applying neural network analysis on heart rate variability data to assess driver fatigue” by M. Patel et al. Copyright © 2011 by M. Patel et al. Used by permission of Elsevier.

3.4.3 Support Vector Machines

Support Vector Machines (SVM) are one of the most novel MLAs that has been designed. Developed in 1992 by Vladimir Vapnik, they have presently become one of the most used MLA due to their advantages (Marsland, 2015). SVM has been found to have great capabilities in problems where the data are not linearly separable (Marsland, 2015; Murphy, 2012; Bell, 2015; Harrington, 2012), as shown in Figure 3-6. The possibility of automatically transforming the dataset into a higher dimension dataset is one of the biggest advantages of SVM.

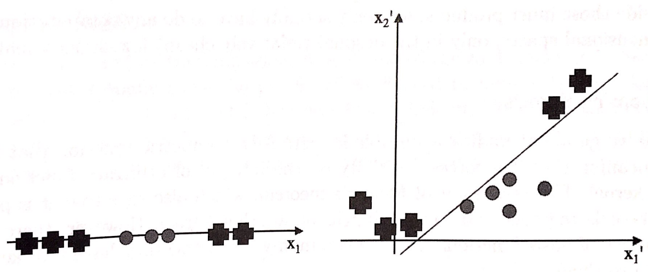


Figure 3-6 Benefits of SVM. On the left, the data (crosses and dots) are not linearly separable. The data can be transformed into a higher dimension (plot on the right) where the data become linearly separable. This is one of the advantages of using SVM (SVM can use functions to transform the data into higher dimensions where the data might be more easy to separate) (Source: Marsland, 2015). Reprinted from "Machine Learning: An Algorithmic Perspective" by Stephen Marsland. Copyright © 2015 by Stephen Marsland. Used by permission of CRC Press, Florida, USA.

The main purpose in classification problems is to find a linear function that can separate the data into different groups. For a dataset, there are multiple classification lines, as shown in Figure 3-7. The three classification lines presented in Figure 3-7 solve the problem of separating the data in different classes. The problem with the left and the right plot is that the line passes on top or close to some data point. There is a probability that it might classify some of the data points in the wrong class. By allowing the line to be far away from the data points, the probability of misclassifying any data point reduces as presented in the middle plot in Figure 3-7. This is the principle of a SVM- to find the optimal line and space (called margin) between the different classes.

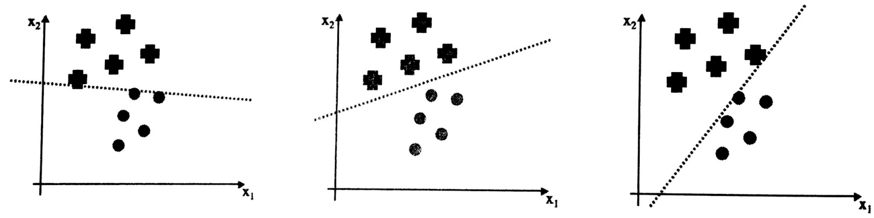


Figure 3-7 Dataset classified by three different classification lines. Although the left one and the right one are correctly separating the data, due to the linear classifier's proximity to the data points, it is probable that when classifying new points, the classification error might be high when these new data points are close to the data points closer to the line. The graph in the middle has a linear classifier which is far

away from the data points, diminishing the possibility of misclassifying new data points near to the data points closer to the line (Source: Marsland, 2015). Reprinted from “Machine Learning: An Algorithmic Perspective” by Stephen Marsland. Copyright © 2015 by Stephen Marsland. Used by permission of CRC Press, Florida, USA.

A SVM will maximize the space of the margin by maximizing the space of the data points from the classifying line. The distance is measured as the perpendicular distance from each point to the classifying line. In reality, the SVM will only need to maximize the distance of the closest points for each class from the classifying line. These points are called the support vectors, as they are the ones that will determine where the line is, as shown in Figure 3-8. The classifying line can be also a hyperplane if the dataset is in a higher dimension.

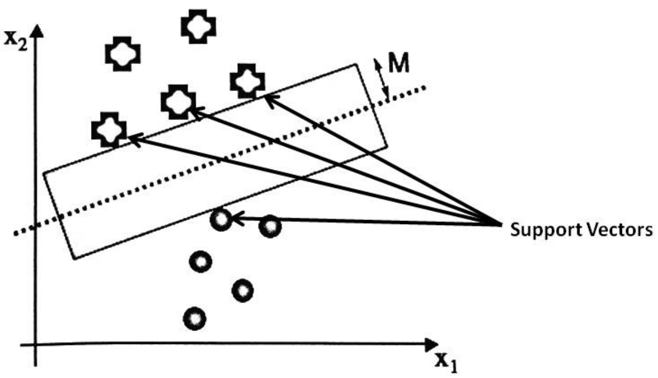


Figure 3-8 The data points used to determine the maximum space between the linear classifier are called Support Vectors. Support vectors of each class are used to determine the optimal margin. (Source: Marsland, 2015). Reprinted from “Machine Learning: An Algorithmic Perspective” by Stephen Marsland. Copyright © 2015 by Stephen Marsland. Used by permission of CRC Press, Florida, USA.

Before the SVM can find the optimal linear classifier, the data has to be linearly separable. A SVM makes use of kernel functions to be able to transform the data into a higher dimension where the data can be linearly separable, as shown in Figure 3-6. In Figure 3-6, it is possible to see that if the data are squared and plotted in y, the dataset will be now linearly separable. Unfortunately, sometimes it is not possible to have an insight into which kernel function will transform the data into a linearly separable dataset. In addition, if the data are in a N-dimension it is not possible to be able to plot and visualize the data. Therefore, it is necessary to use a kernel function that can be generalised to many problems. Although in principle any function can be used as a kernel function, the most used ones are: polynomial kernels, sigmoid functions, and radial basis function expansion. There is not yet a rule regarding when is more convenient to use a specific kernel and many times a test is done with many kernels to determine the best one for the specific dataset. The kernel functions are mathematically implemented in a variety of readily available software packages.

3.4.3.1 SVM to predict sleep while driving

SVMs have been used to classify sleep and awake states in drivers (Yeo et al., 2009; Shuyan & Gangtie, 2009). Yeo et al. (2009) used SVM to classify drivers in three different classes: alert and drowsy. EEG data were used to determine the different levels of sleepiness. The EEG data were divided in 10 seconds epoch and then manually classified by two raters. Epochs were classified into alert if there was were eye blink artifcats lasting 0.3 to 0.4 seconds, inter-blink intervals lasting 6 to 8 seconds and EEG activity in the beta frequency. The epochs with eye closures lasting longer than 0.5 seconds, EEG showing alpha activity in the occipital region (more than 50% of the epoch) and with appearances of alpha dropout events were classified as drowsy. Epochs were discarded due to lack of consensus between the raters, especially in cases where the alpha dropout was not very prominent. During the experiment, twenty young (10 males and 10 females) students between 20 to 25 years old were recruited to take part in a driving simulator task. They had to drive for one hour on a monotonous highway. Their blinking behaviour (through EOG) and brain activity was recorded using a 17 channels EEG (Figure 3-9 shows an example of the EEG recording). As described before, the EEG and the EOG were used to determine the different levels of sleepiness. The brain activity data (split into 10 seconds epochs) was then transformed using Fast Fourier Transformation to obtain four features per epoch, which were used to train the SVM. The four features obtained per epoch were dominant frequency (frequency with the highest power), average power of dominant peak (average power of the full width half maximum band using the dominant peak), centre of gravity (using the formula $\boldsymbol{GF} = \frac{\sum_i \boldsymbol{P(f_i)xf_i}}{\sum_i \boldsymbol{P(f_i)}}$, where $\boldsymbol{f_i}$ is frequency and $\boldsymbol{P(f_i)}$ is the estimated power density), and frequency variability (using the formula $\boldsymbol{FV} = \frac{\sum_i \boldsymbol{P(f_i)x f_i^2} - (\sum_i \boldsymbol{P(f_i)x f_i})^2 / \sum_i \boldsymbol{P(f_i)}}{\sum_i \boldsymbol{P(f_i)}}$). Each feature was done for the four different frequency bands (delta, theta, alpha and beta). This meant that each epoch was a 272x1 vector (4 features x 17 EEG channels x 4 frequency bands). The SVM was trained with 239 epochs of alert and 702 epochs of drowsiness and it was tested with 239 epochs of alert and 702 epochs of drowsiness. The SVM obtained 99.3% accuracy when detecting the different stages (awake and sleep) and 90% accuracy when detecting transition stages, e.g. from alert to drowsy and from drowsy to sleepy.

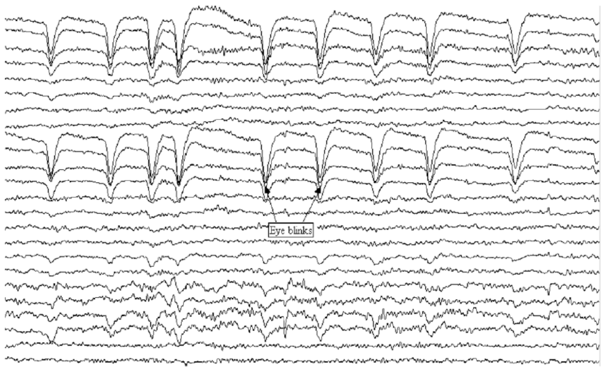


Figure 3-9 Eye blink and brain activity sample used to determine the different stages of sleepiness and train the SVM. (Source: Yeo et al., 2009). Reprinted from “Can SVM be used for automatic EEG detection of drowsiness during car driving?” by Mervyn V.M. Yeo et al. Copyright © 2009 by Mervyn V.M. Yeo et al. Used by permission of Elsevier.

Shuyan & Gangtie (2009) also used a multiclass SVM to predict increase of sleepiness in drivers. Thirty-seven sleep-deprived participants took part in a driving simulator study lasting 45 minutes. Subjective sleepiness (using 9 scale Karolinska Sleepiness Scale; KSS), EEG and eye movement (EOG) behaviour was recorded while the participants were in the driving task. The EEG and subjective data were used to classify the data into three different categories: alert, sleepy and very sleepy. The data were separated into 20 seconds epochs. Each EEG epoch was separated in 2 seconds bins (10 x 2 second bins). Each bin was visually analysed to determine if there were signs of high levels of sleepiness (slow eye movements, alpha activity and/or theta activity). If there were signs of high levels of sleepiness, the bin would be assigned a value of 10; otherwise, the bin would be assigned a value of 0. At the end, each epoch had a value between 0 and 100 (the sum of the values of each of the 10 x 2 seconds bins) and this value was called Karolisnka Drowsiness Scale (KDS). The KDS and the KSS were used to classify the epochs into alert, sleepy and very sleepy. An epoch was classified as alert if the epoch was from the first 5 minutes of driving, the KSS value during those 5 minutes was less or equal to 7 and the KDS value of the epoch was less than 10; the epoch was classified as sleepy if the epoch was part of the middle of the driving time, the KSS value during those 5 minutes is more or equal to 7 and the KDS values of the epoch was more than 15 and less than 25; the very sleepy epochs were the ones that were part of the 5 minutes before an accident happened (the car going out of the road), a KSS value of 8 or more and a KDS value of the epoch with a value more than 25.

The SVM was trained with 11 features (per participant) obtained from the eye movement behaviour. The features were the following: blink duration, blink duration 50-50 (from the half rise point to the half fall point of the blinking), amplitude of the blink (measured in microVolts), the average speed of the closure of the eye, the peak value of the speed of the closure of the eye, the average speed of the opening of the eye, the peak value of the speed of the opening of the eye, the delay of eyelid opening from previous blink, time from 80% of eyelid opening at rise to 20% of eyelid closure at fall, length of time for complete closure of the eye and length of time for complete opening of the eye. The features obtained from the eye behaviour are presented in Figure 3-10. Same as the research conducted by Yeo et al. (2009), the dataset of Shuyan & Gangtie has a high dimension. The epochs of five participants were used for training of the SVM. The same data were used for validation of the SVM. The SVM in this research obtained an accuracy of around 85% when identifying the different stages the driver was in (Shuyan & Gangtie, 2009).

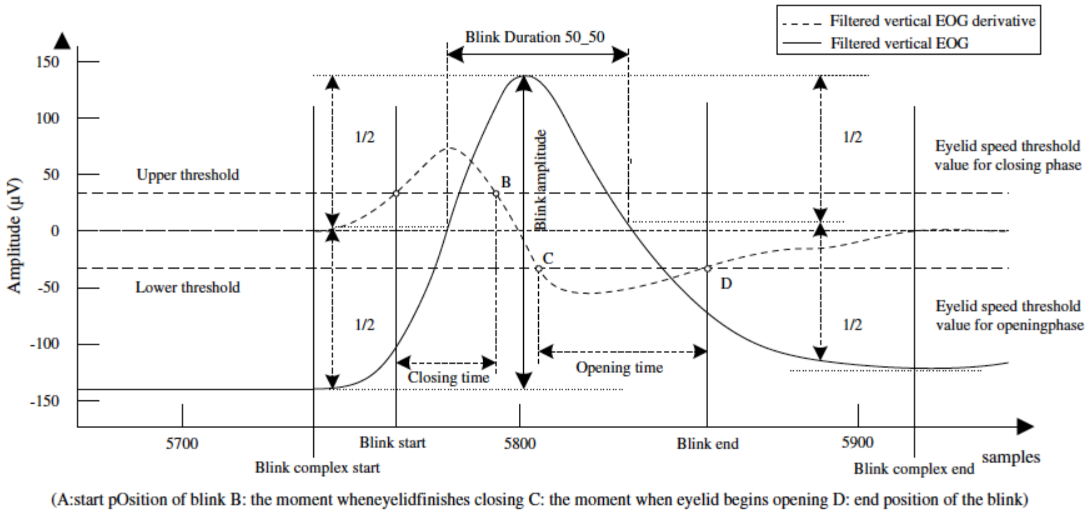


Figure 3-10 Features extracted from a single blink of a participant to train the SVM. (Source: Shuyan & Gangtie, 2009). Reprinted from "Driver drowsiness detection with eyelid related parameters by Support Vector Machine" by Shuyan Hu and Gangtie Zheng. Copyright © 2009 by Shuyan Hu and Gangtie Zheng. Used by permission of Elsevier.

3.4.4 Dynamic Bayesian Networks

Dynamic Bayesian Networks (DBN) are a type of MLA that predict outcomes depending on the probabilities of certain events happening (Marsland, 2015; Murphy, 2012; Bell, 2015; Harrington, 2012; Korb & Nicholson, 2004). These events are related to each other and the outcome of one event affects the probability of the events related to it. To determine the outcome of an uncertain event, Bayesian Networks combine three different fields: graph theory, probability theory and Bayes' theorem. Although being widely used in other fields, DBN has not been used by many researchers to predict sleepiness in drivers. One of the research focus in predicting different levels of sleepiness in drivers using DBN is presented in the next section.

3.4.4.1 DBN to predict sleepiness while driving

Yang, Lin & Bhattacharya (2010) developed a DBN to detect two sleepiness states (awake and fatigue) in the drivers. The features that were input into the DBN were a combination of casual/contextual and physiological features. The complete set of features used by Yang, Lin & Bhattacharya (2010) are presented in Table 3.1.

Table 3-1 Variables used by Yang, Lin & Bhattacharya (2010)

Variable name	Definition
Circadian rhythm	Casual/contextual feature that determines probability of increase of sleepiness due to the time of the day. The states that could take this feature are low and high.
Work environment	A casual/contextual feature with two parent nodes: Temperature and Noise. Temperature and noise referred to the environment where the participant was driving. Temperature could take two states (high and normal) and noise could take two states (high or normal). According to these nodes, work environment could take two states (bad or good).
Sleep Quality	A casual/contextual feature with two parent nodes: Sleep environment and Sleep time. Sleep environment referred to the sleep place of the driver and could take two states (poor or normal). Sleep time referred to the amount of time slept by the driver and could take two states (sufficient or deprived). Depending on the state of its' parents node, sleep quality could take two states (bad or good).
Eye movements	A physiological feature that determines the percentage of time the eyelid was closed under a certain threshold. This node could take one of three states (large, medium or small).
Electrocardiograph	A physiological feature that determines the changes in the ratio between low frequency and high frequency in the heart rate variability of the driver. The node could take one of three states (decrease, no-change and increase). This variable took advantage of the temporal feature of DBN, as its state depend on the comparison to the previous state of the same node.
Electroencephalogram	A physiological feature that determines the changes in the alpha frequency band of the EEG recording after Fast Fourier Transform has been

	applied. The node could take one of three states (decrease, no-change or increase). In the same way as electrocardiograph, this variable took advantage of the temporal feature of DBN, as its state depend on the comparison to the previous state of the same node.

Figure 3-11 shows the DBN developed by Yang, Lin & Bhattacharya (2010). The conditional probabilities were calculated from conclusion obtained from previous research. Although being an interesting attempt to detect different levels of sleepiness due to the amount and type of features selected, no results have been published yet for the present DBN, so no comparison can be drawn between DBN and NN or SVM.

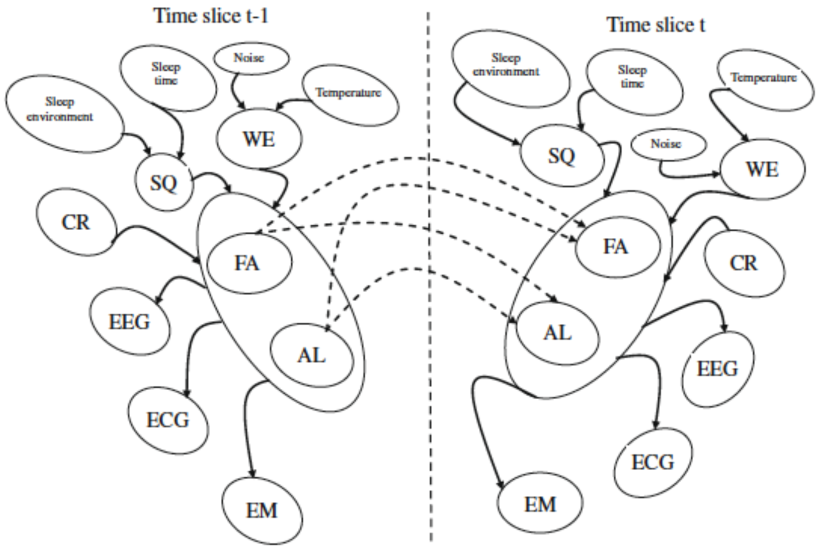


Figure 3-11 DBN developed to detect sleepiness states in the driver. The left BN is the state of the participant at time t-1. FA and AL is the probability of the participant of being in a fatigue stage (FA) or in an alert stage (AL). This probability is then passed to the next BN in time t, where the probability in t-1 plus the other factors and variables will provide a new probability for the participant to be in a fatigue stage (FA) or in an alert stage (AL) (Source: Yang, Lin & Bhattacharya, 2010). Reprinted from "A driver fatigue recognition model based on information fusion and dynamic Bayesian network" by Guosheng Yang et al.. Copyright © 2010 by Guosheng Yang et al.. Used by permission of Elsevier.

3.5 Conclusion

In the present chapter, a detailed definition of different methods to determine levels of sleepiness was presented, i.e. manual classification and classification using Machine Learning Algorithms. Afterwards, the term Machine Learning Algorithms was explained by defining three different concepts: machine, learning and algorithm. The different types of Machine Learning were also presented and defined. The chapter ended with an explanation of the concept, functionality and uses of the different Machine Learning Algorithms applied in research to predict sleepiness in drivers. In the following chapter, these algorithms will be tested using a dataset obtained in a motion-base driving simulator. The algorithms will predict different levels of sleepiness using physiological indicator of sleepiness and driving behaviour.

Chapter 4:
Identifying markers of fatigue in secondary data

4. Identifying markers of fatigue in secondary data

4.1 Introduction

In the previous chapter, we examined the different type of Machine Learning Algorithms (MLAs) researchers have used to predict sleepiness while driving. These included Neural Networks (NeuroNets), Support Vector Machines (SVM) and Bayesian Network. It was found that NeuroNets and SVM were the most common and the ones to achieve highest accuracy level. Therefore, in the following chapters, these NeuroNets and SMV will be adapted to predict sleepiness using driving behaviour data. The data analysed in the present chapter were obtained from a previously published experiment conducted in the driving simulator at the University of Leeds. This dataset was chosen as the experiment was designed to induce high levels of sleepiness in young drivers. The aim of this study was to use blinking behaviour to define multiple levels of sleepiness. Another aim to be achieved during this study was to test different MLAs to determine the most suitable MLA to predict sleepiness in drivers.

4.2 Participants and data

The data used to train and test the MLAs was obtained from the experiment conducted by Merat & Jamson (2013). This dataset was selected as the experiment was designed to induce high levels of sleepiness in the participants (Merat & Jamson, 2013). The experiment was originally aimed to test the effects of three low-cost engineering treatments on drivers' sleepiness (Merat & Jamson, 2013). Participants were asked to drive in a 55-kilometre 3-lane motorway scenario. The driving task took around 30 minutes to be completed. Participants had to repeat the driving task four times. The motorway scenario contained gentle curves and straight segments. In addition, low traffic and little visual clutter, e.g. road signs, were present during the driving task. The experiment was designed to induce high levels of sleepiness in the participants. The experiment was conducted in a motion-base driving simulator at the University of Leeds. The concept of motion based and static driving simulator will be described in Chapter 5.

Two sets of participants were recruited for this experiment: 16 young drivers under the age of 35 (M=31.41, SD=5.37) and 16 older drivers above the age of 45

(M=53.2, SD=5.48). The young drivers were night shift workers whilst the older drivers did not work during the night, i.e. they had normal night sleeping patterns. Participants had to visit the driving simulator on two occasions. During their first visit, young and older participants undertook the driving task while they were awake, i.e. young drivers drove before their night shift and older drivers drove in the morning after a normal night sleeping pattern. During the first visit, they performed a driving task, which was used to determine the baseline, i.e. the blinking and driving behaviour of the participants while their sleepiness levels were low.

On their second visit to the driving simulator, young and older drivers were tested when their sleepiness levels were higher. The young drivers arrived to the driving simulator at 8:00 in the morning, immediately after their night shift. The older drivers arrived at 1:30 in the afternoon. The older drivers were instructed to have a heavy meal of their choice just before arriving to the driving simulator. Although the older drivers had a normal night sleeping pattern before their second visit, they were tested during the post-lunch dip, which, as explained in Chapter 2, increases sleepiness in people (Monk, 2005; Reyner et al., 2012; Smith & Miles, 1986a,b; Wells & Read, 1996; Wells et al., 1995; Lloyd, Green & Rogers, 1994; Cunliffe, Obeid & Powell-Tuck, 1997).

During their second visit, participants repeated the driving task three times. During each driving task, participants were asked to drive as they normally would drive. A speed limit was set in the driving simulator so that participants could not speed further than the legal speed limit even if they kept pressing the accelerator to avoid the participants finishing too early the experiment and not reaching a high level of sleepiness. For the first 48 kilometres, the road had low traffic and little visual clutter. The three driving tasks differed after the 48th kilometre. At kilometre 48, different low-cost treatments were presented to wake up the driver. During one of the drives, the participant found a 3-kilometre stretch of chevrons, which occurred every 40 metres. During the other drive, a stretch of 3 kilometres of rumble strips was presented to the driver. In the final drive, variable message signs were presented for a stretch of 3 kilometres length. For the purposes of this PhD study, the first 48 kilometres, i.e. before the low-cost treatments were presented, were used to test the different MLAs. Participants 3, 11 and 16 from the 'older' group and participants 1, 4, 9 and 15 from the 'young' group were removed as their data were considered outliers.

4.2.1 Variables recorded

During all the drives, the blinking and driving behaviour of the participants were recorded. The blinking variables recorded were percentage of eye closure (PERCLOS; as presented in chapter 2) blinking frequency and blinking duration. The driving variables recorded were lane position, steering wheel, acceleration and speed. The driving simulator environment was programmed to have a speed limiter, i.e. participants could not drive faster than 80 mph even if they kept pushing the accelerator pedal (Merat & Jamson, 2013). As there was no information on the changes in speed behaviour of the participants once they reached the speed limit, the speed variable was discarded for the training and testing of the different MLAs. One of the problems encountered in this dataset was that data were missing in certain time segments, i.e. the blinking and driving variables were not recorded continuously throughout the driving task. The researchers in charge of this experiment were only interested in certain stages of the driving task so they did not find it necessary to record the whole driving task.

During each driving task, only ten segments were recorded. During those events, blinking and driving behaviour were also recorded. Every event lasted 3 kilometres. The first event was recorded once participants crossed the first 5.8 kilometres. The blinking and driving behaviour were recorded during the following 3 kilometres, i.e. until the participants reached the 8.8 kilometres. After that, during the following 2.5 kilometres, driving and blinking behaviour were not recorded. The second event started around 2.5 kilometres after the ending of the first event, i.e. at kilometre 11.3. This pattern continued until kilometre 47. The last two events (events 9 and 10) were recorded continuously, i.e. event 9 recorded from kilometre 49 to kilometre 52.5 and event 10 recorded from kilometre 52.5 to kilometre 55.5. Table 4-1 presents the initial and final kilometre for each event.

In the following sections, each blinking and driving variable will be described. As explained in chapter 3, the variables used to train and test the MLAs are divided into target variables (the variables that the algorithm is trying to predict) and feature variables (the variables used by the algorithm to predict the target variables) (Harrington, 2012; Bell, 2015; Marsland, 2015; Murphy, 2012; Alpaydin, 2010). In this study, the target variables were the blinking variables and the feature variables are the driving variables.

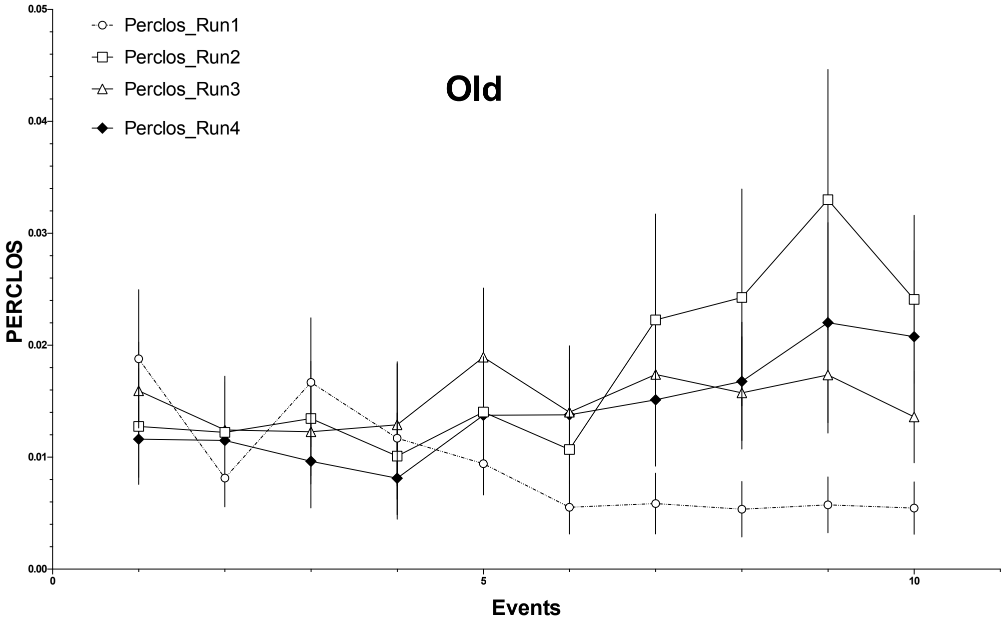
Table 4-1 Initial and final distances of each event recorded (the units are in metres). The driving and physiological variables were recorded only during certain segments throughout the whole driving task.

Event	Initial Dist	Final Dist	Difference	Difference between initial distance and final distance of previous event
1	5817	8840	3023	N/A
2	11361	14384	3023	2521
3	16739	19763	3024	2355
4	22283	25306	3023	2520
5	27662	30685	3023	2356
6	33206	36229	3023	2521
7	38584	41608	3024	2355
8	44129	47152	3023	2521
9	49507	52530	3023	2355
10	52531	55555	3024	1

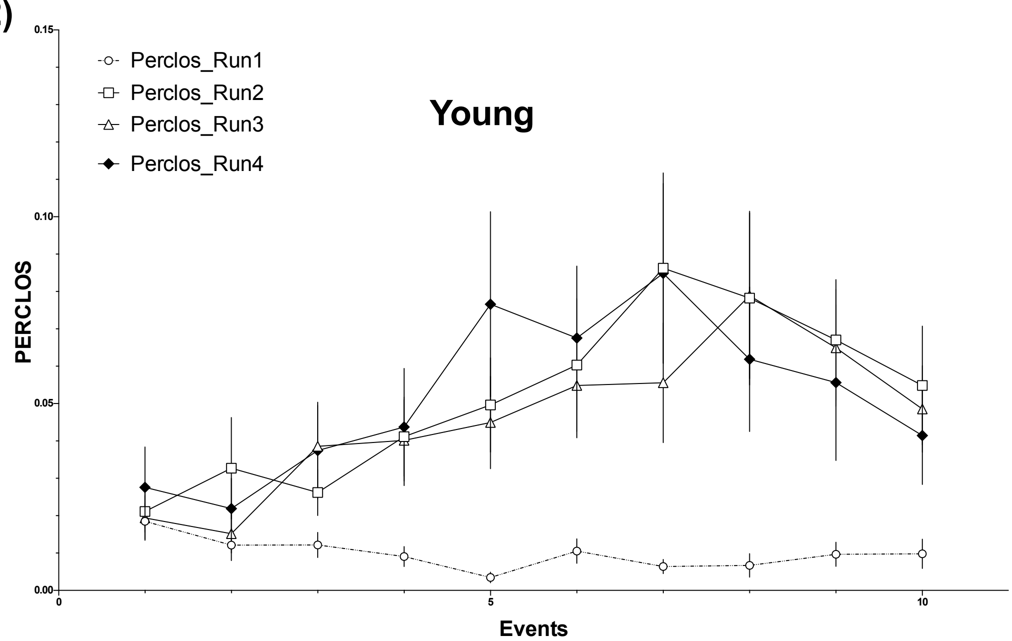
4.2.1.1 Target variables

As discussed in Chapter 3, the variables the MLAs will try to predict, i.e. the outcome of the MLAs, are called the target variables or values. For the experiment presented in this chapter, the target variables were the blinking behaviour. The blinking behaviour was composed of three variables: PERCLOS, blinking duration and blinking frequency. Figure 4-1 shows the blinking behaviour throughout the four different driving tasks for the young and the older drivers.

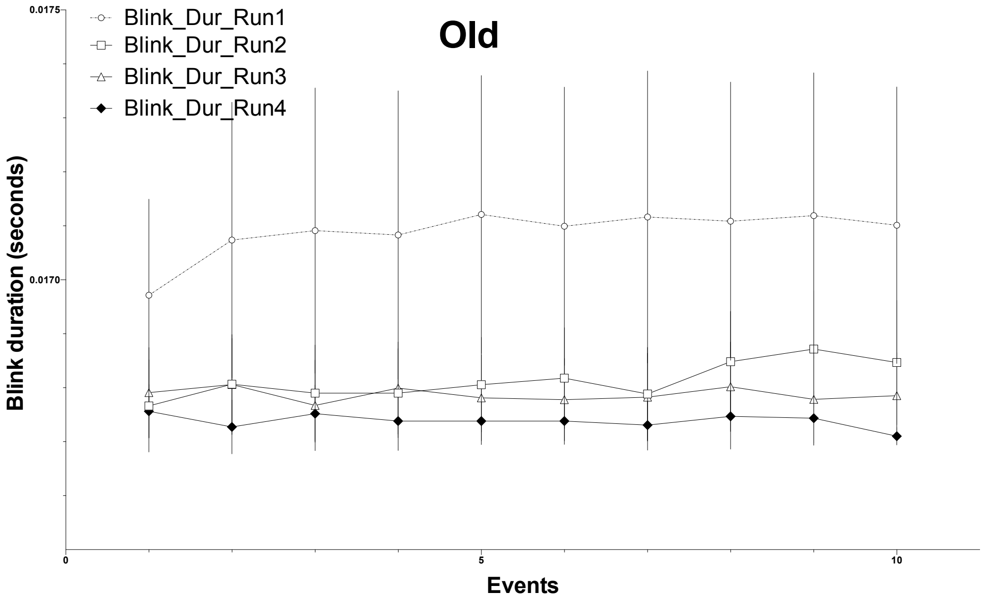
A1) **PERCLOS**



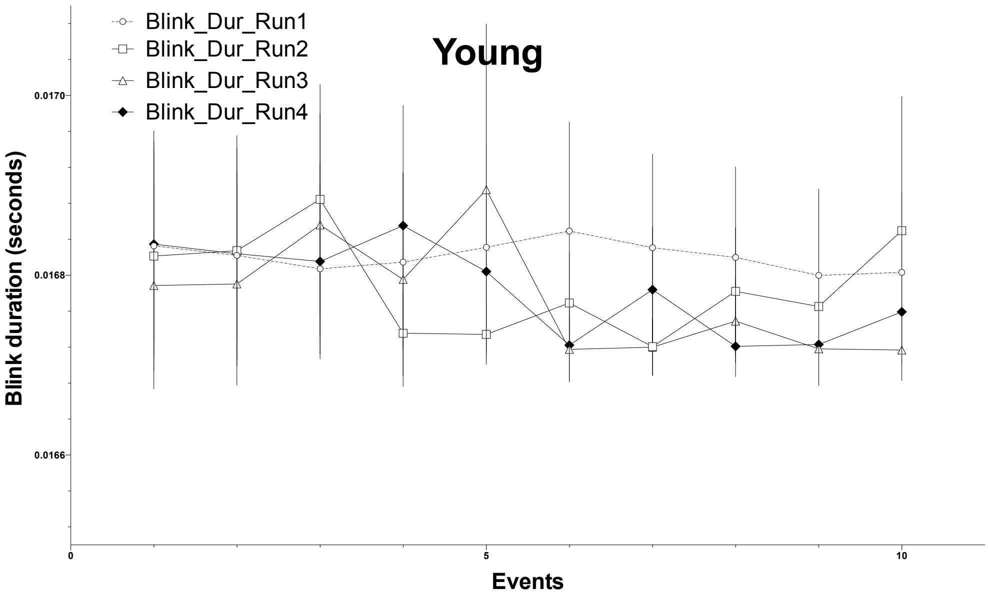
A2)



B1) **BLINK DURATION**

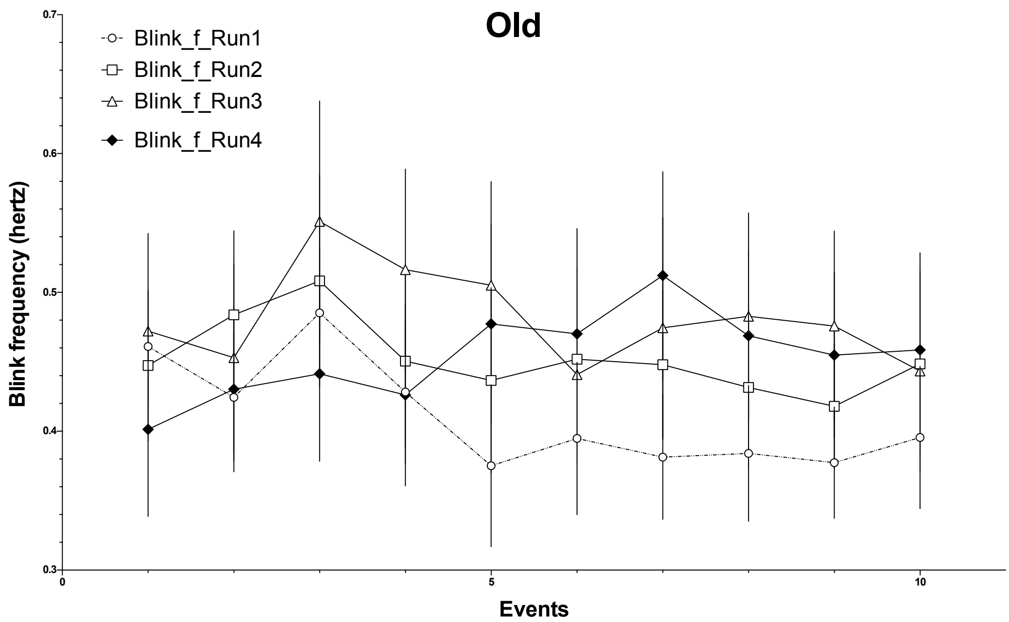


B2)



BLINK FREQUENCY

C1)



C2)

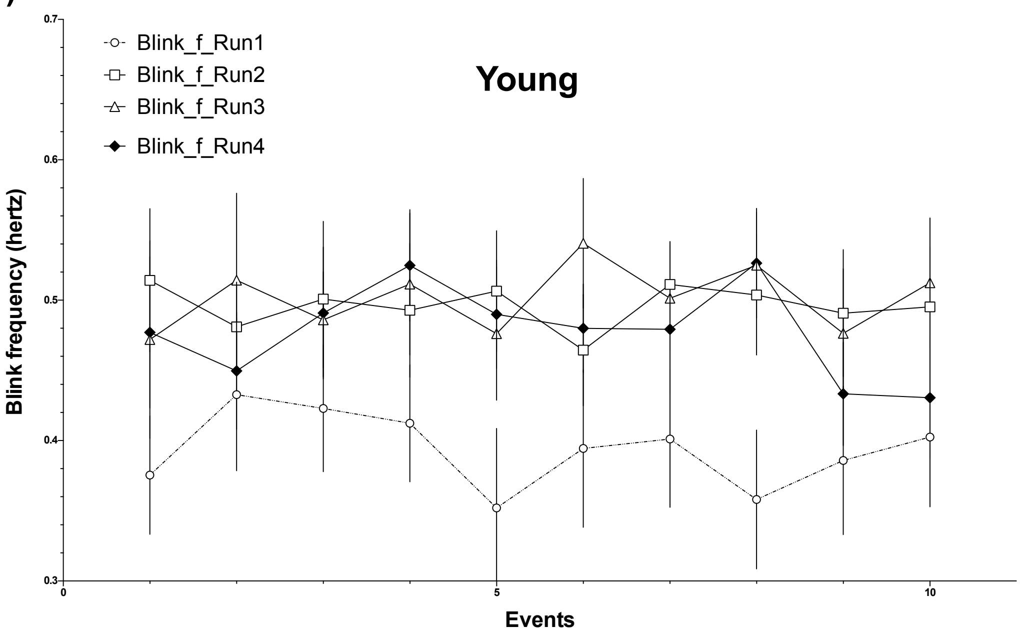


Figure 4-1 Blinking behaviour variables (PERCLOS, blinking duration and blinking frequency) for the older and young group. Each plot contains four lines: the dotted line with circles is the baseline line ("awake" run) and the other three are the experimental runs ("sleep" runs). A) The plots present the PERCLOS (percentage of closure of the eye) values for the older group (top plot) and the younger group (bottom plot). B) The plots present the blinking duration values for the older group (top plot) and the

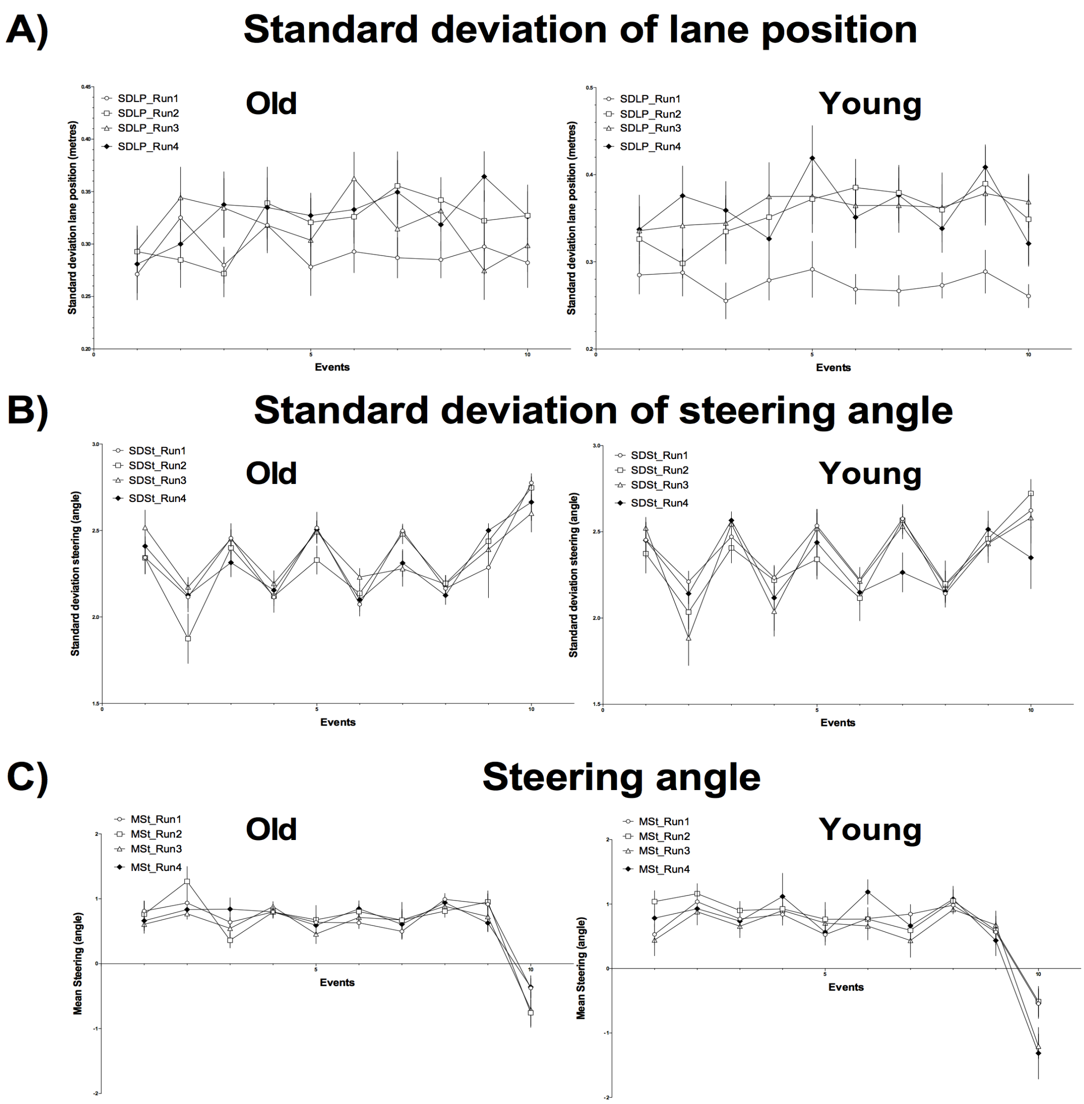
younger group (bottom plot). C) The plots present the blinking frequency values for the older group (top plot) and the younger group (bottom plot).

PERCLOS (percentage of eye closure) was defined and explained in Chapter 2, as the percentage of time the eyes are closed beyond a predefined threshold N during a specific time segment T. For this experiment, the threshold N was .75 and the time segment T was 180 seconds. PERCLOS ranged between zero and one, where zero meant the eyes were opened throughout the whole 3 minutes segment and one meant the eyes were closed throughout the whole 3 minutes segment. The blinking duration referred to the duration of time the eye was completely closed and it was measured in seconds. The blinking frequency referred to rate at which blinking occurred during a time segment and it was measured in Hertz.

4.2.1.2 Feature variables

As discussed in Chapter 3, the variables MLAs use to predict an outcome (target set) are call set. The feature set used in this experiment was the driving variables: lane position, steering wheel and acceleration. Speed was not used, as the speed limiter in the driving simulator did not allow the participants to go faster, i.e. once the limit was reached, the speed did not increase.

Additional variables were derived from the driving data recorded. Using the lane position of the car, the standard deviation of the lane position was calculated and used as a feature variable. From the steering wheel angle, high frequency steering, standard deviation of steering wheel angle and mean steering wheel angle were calculated. Finally, using the acceleration values, the mean and standard deviation of acceleration was calculated. Figure 4-2 show the driving variables per group for the four driving tasks.


A)
Standard deviation of lane position
Old
Young
SDLP_Run1
SDLP_Run2
SDLP_Run3
SDLP_Run4
Standard deviation lane position (metres)
Events
B)
Standard deviation of steering angle
Old
Young
SDSt_Run1
SDSt_Run2
SDSt_Run3
SDSt_Run4
Standard deviation steering (angle)
Events
C)
Steering angle
Old
Young
MSt_Run1
MSt_Run2
MSt_Run3
MSt_Run4
Mean Steering (angle)
Events

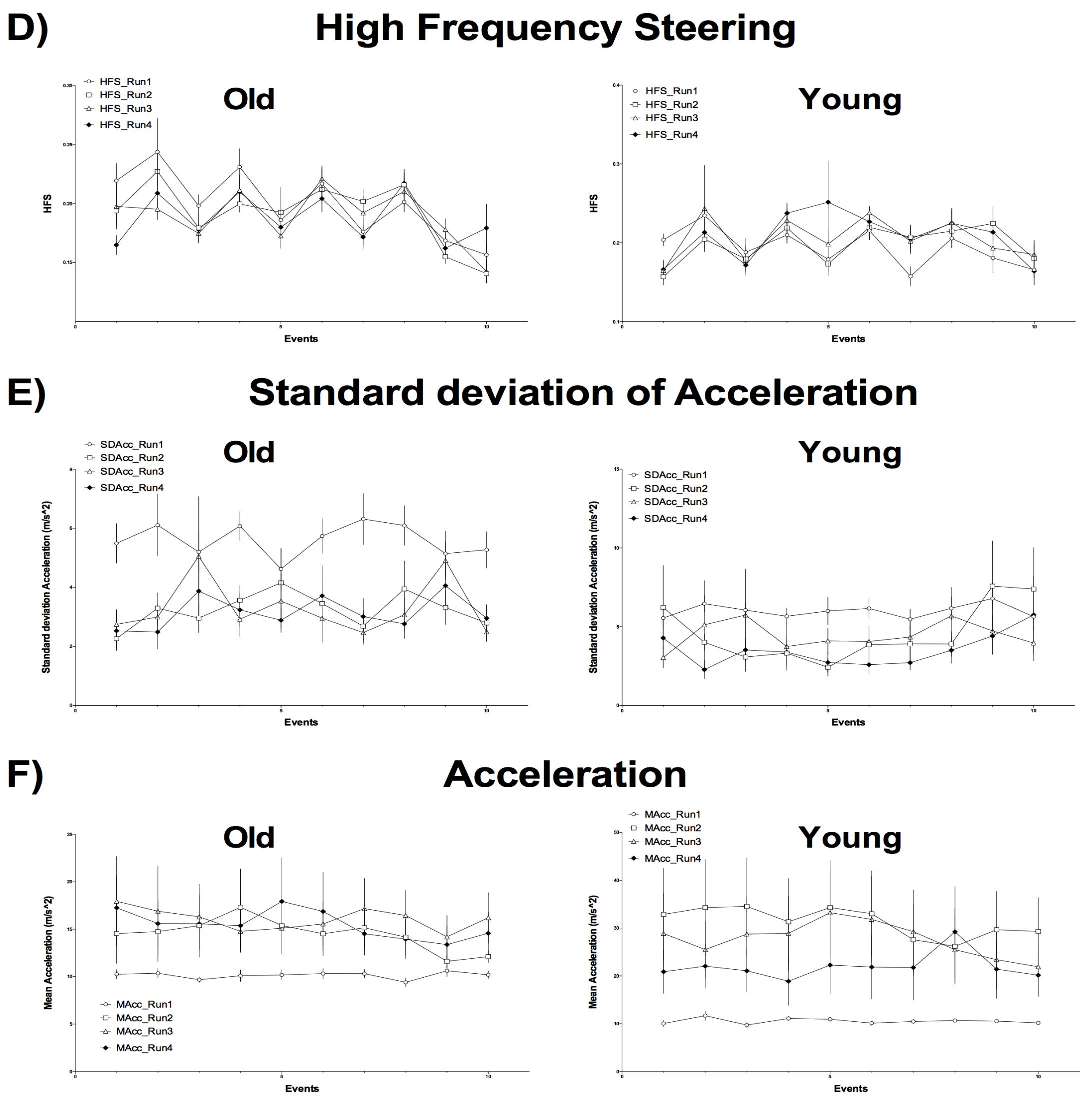


Figure 4-2 Standard deviation lane position, standard deviation steering and mean steering variables of the older and young group. Each plot contains four lines: the dotted line with circles is the baseline line ("awake" run) and the other three are the experimental runs ("sleep" runs).

4.2.2 Statistical Analysis

The results obtained by Merat and Jamson (2013) showed that PERCLOS reduced in young and older drivers after presenting the three low-cost engineering treatments compared to the kilometres before the treatments were presented. It was concluded that sleepiness was affected by the treatments, as the participants became more awake after they experienced the treatments (Merat & Jamson, 2013). This led to conclude that there was an increase in the levels of sleepiness right until the moment the treatments were presented to the participants. Therefore, this dataset (until before the moment the participants experienced the treatments) could be used for the purposes of the present PhD study, i.e. determining sleepiness states in drivers. They also found that even though there was an increase in PERCLOS for both groups (older and younger drivers) from day 1 to day 2. Younger drivers presented higher values of PERCLOS compared to older drivers (Figure 4-1), which are related to higher levels of sleepiness (Merat & Jamson, 2013). Therefore, the algorithms were trained and tested only with the data of the young participants. The young participants were also selected as it has been found that they are the target group most at risk in accidents related to sleeping while driving (Pack et al., 1995; Akerstedt & Kecklund, 2001; Johns, 2000; Horne & Reyner, 1995).

4.3 Predicting sleepiness

As discussed in Chapter 3, MLAs are trained to predict a value or state depending on the inputs given to the algorithm (Harrington, 2012; Bell, 2015; Marsland, 2015; Murphy, 2012; Blum, 2014; Alpaydin, 2010). In this experiment, MLAs were used to predict the state of sleepiness of the participants through their driving behaviour. The different discrete states of sleepiness were determined using the blinking behaviour of the participants. The following section presents the method (k-means clustering) to define the different levels of sleepiness using blinking behaviour. Afterwards, the proposed levels of sleepiness defined were tested using different MLAs.

4.3.1 Discrete targets using k-Means Clustering algorithm

Many researchers have found specific threshold values of blinking behaviour to define different levels of sleepiness (Jimenez-Pinto & Torres-Torriti, 2013; Boverie et al., 2013; Yang et al., 2010; Yeo et al., 2009). One of the most commonly used is PERCLOS (Jimenez-Pinto & Torres-Torriti, 2013; Boverie et al., 2013; Yang et al., 2010). Jimenez-Pinto & Torres-Torriti (2013) determined that during the 'Awake' state when PERCLOS has a mean value of 0.025 and it ranges from 0 to 0.05; driving the 'Drowsy' state has a mean value of 0.09 and ranges from 0.04 to 0.15; and driving the 'Sleepy' state has a mean value of 0.18 and ranges from 0.09 to 0.3. Boverie et al. (2013) defined the 'Awake' state as any PERCLOS values under the value of 0.24; the 'Fatigue' state PERCLOS values between 0.24 and 0.45; and

‘Drowsy’ state any PERCLOS value above 0.45. Finally, Yang et al. (2010) defined ‘Alert’ state when PERCLOS has a value between 0.01 and 0.05 and ‘Fatigue’ state when PERCLOS has a value between 0.05 and 0.94. It was concluded that there is not a clear consensus in literature in the threshold values that separate the different levels of sleepiness.

Due to the lack of consensus in research regarding the specific blinking behaviour values that can be used as threshold to determine the different sleepiness states, the present research used unsupervised MLAs that would cluster the blinking behaviour data into different sleepiness states. The unsupervised algorithm used was the K-means clustering and was run in WEKA software (WEKA 3.6, University of Waikato, New Zeeland). As discussed in Chapter 3, unsupervised algorithms are used when the researcher does not have knowledge of specific desired output from the data (Jain & Dubes, 1988; Davis, 2014; Wagstaf et al., 2001). In the present study, the specific clusters for the blinking behaviour values are unknown; therefore, unsupervised learning algorithms are suitable in this case. The k-means clustering algorithm classifies the data into a specific number of clusters. The researcher defines the number of clusters a priori. For each cluster, a centroid is defined, i.e. if the data wants to be divided into k clusters, there will be k centroids. The position of the centroids is randomised. The algorithm then uses the Euclidean distance between each data point and each centroid (Davis, 2014). Each data point is then associated to the nearest centroid, i.e. the smallest Euclidean distance between a specific data point and each of the centroids (Jain & Dubes, 1988; Davis, 2014; Wagstaf et al., 2001). Once all the data points are associated to a centroid, the position of the centroids is updated as the mean of the position of all the data points contained in the cluster (Davis, 2014). The process is then iterated by calculating the Euclidean distance of each data point to the updated position of the each of the centroids. Once the data points are associated to a centroid, the position of the centroids is updated once more (Jain & Dubes, 1988; Davis, 2014; Wagstaf et al., 2001). These steps are iterated until the centroids are unchanged and the data points are not re-clustered.

4.3.2 Defining a binary levels of sleepiness

The first clustering division to be tested was a binary classification of sleepiness, i.e. the blinking behaviour data would be clustered into ‘Awake’ and

‘Sleep’. This meant that the number of clusters predefined on the k-means clustering algorithm was two. As PERCLOS and blinking frequency showed significant difference between the baseline (awake) and the experiment (sleep) driving tasks (Figure 4-1), these two variables were used in the k-means clustering algorithm. There was no statistically reliable relationship between PERCLOS and blink frequency. The Pearson’s correlation statistic ranged from -.378 to .35 (all p's > .26), reflecting the independence of these two measures. Blanco et al. (2009) also found that there was no correlation between PERCLOS and blinking frequency. Therefore, it is possible to use these two independent variables to determine the clusters.

The k-means clustering algorithm represents the two variables as coordinates where PERCLOS was positioned in the x-axis and blinking frequency was positioned in the y-axis. As the predefined number of clusters to be found by the k-means clustering algorithm is two, there were two centroids. The algorithm started with random positions for each of the centroids: (0.064, 0.633) for one centroid and (0.001,0.425) for the other centroid. As stated previously, the algorithm uses Euclidean distance as the function to improve the position of the centroids. The result of the k-means clustering algorithm is presented in Figure 4-3.

In Figure 4-3, PERCLOS was the variable that had a bigger influence in the clustering of the data points, i.e. high values of PERCLOS were classified in one cluster (‘Sleep’ state) while the low values of PERCLOS were classified in another cluster (‘Awake’ state), being the value of ~0.09 the threshold value between the two clusters. Blinking frequency had small influence in the clustering of the data points. The threshold value of 0.09 matches the ‘Drowsy’ state threshold established by Jimenez-Pinto & Torres-Torriti (2013) and is part of the ‘Fatigue’ range determined by Yang et al. (2010). The data were divided in 335 “awake” segments and 48 “sleep” segments.

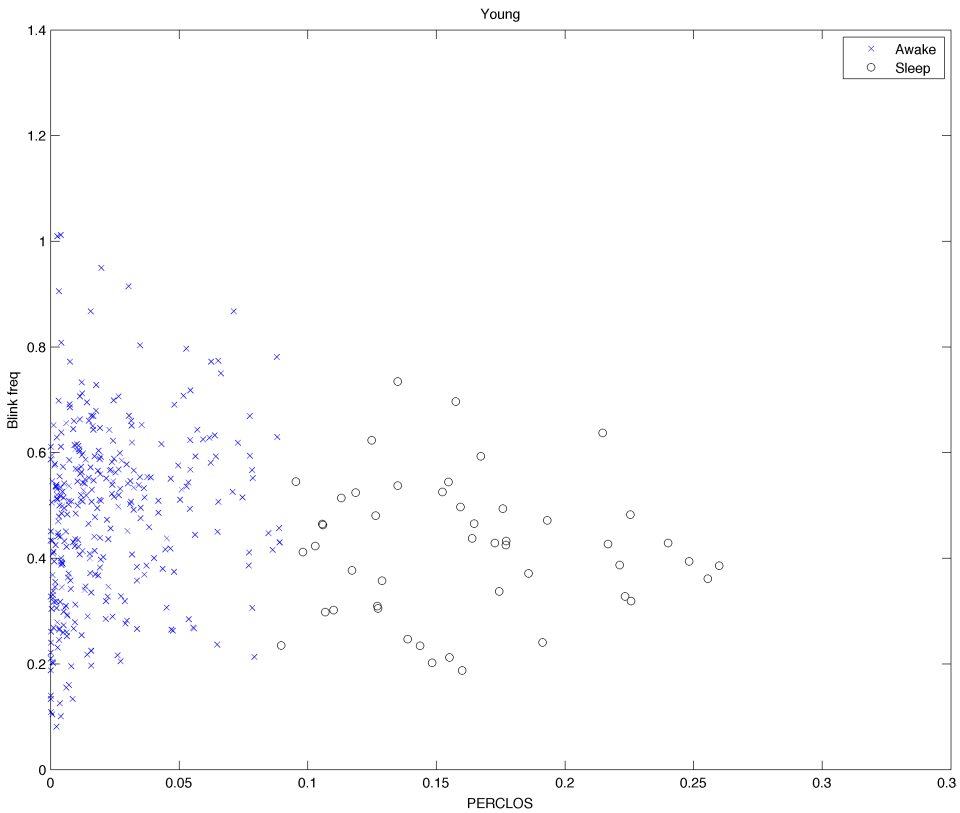


Figure 4-3 Result of the unsupervised k-means algorithm for two clusters using PERCLOS (plotted in the x-axis) and Blinking frequency (plotted in the y-axis). The blue crosses were classified as "awake" segments and the black circles were classified as "sleep" segments.

4.3.3 Defining a ternary levels of sleepiness

The k-means clustering algorithm was run one more time with the same blinking data in search of 3 clusters, i.e. to be classified in 'Awake', 'Drowsy' and 'Sleep'. Figure 4-4 shows the result of a three clustering unsupervised k-means algorithm using PERCLOS and blinking frequency. The initial position of the centroid was randomised: (0.0638, 0.633) for cluster 1, (0.002, 0.425) for cluster two and (0.002, 0.306) for cluster three. In contrast to the 2-cluster k-means algorithm, when defining three clusters, blinking frequency has a higher influence in determining one of the clusters; an intermediate cluster is defined between low values of PERCLOS and low values of blinking frequency, and low values of PERCLOS and high values of blinking frequency. As discussed in Chapter 2, an increase in blinking frequency is correlated to increase in sleepiness level (Yang et al., 2010; Bergasa et al., 2006; Wierwille et al., 1994; Dinges et al., 1998). Therefore, the low values of PERCLOS and high values of blinking frequency cluster was considered as the ‘Drowsy’ state; low values of PERCLOS and low values of blinking frequency was considered as the ‘Awake’ state; and high values of PERCLOS was considered as the ‘Sleep’ state. The data were divided into 135 “awake” segments, 199 “drowsy” segments and 49 “sleep” segments.

It was found that in the two - clustered and the three - clustered datasets, there was a higher number of “awake” and “drowsy” segments (in the 3-clustered dataset) than “sleep” segments. This is referred to as a set of imbalanced data (Elhassan et al., 2016). Imbalance data could lead the MLAs to favour the majority cluster, as more training data of the majority cluster would be available for the MLA to learn, leading to an increase in the false negative rate. Therefore, the following step, before the clustered datasets (2 - clustered dataset and 3 - clustered dataset) can be tested using different supervised MLAs, was to solve the imbalance of the clusters. To solve the imbalance of the clusters, an over-sampling method was used to artificially increase the examples of the under-sample clusters. The over-sampling method used was the Synthetic Minority Oversampling Technique (SMOTE) (Elhassan et al., 2016; Chawla et al., 2002).

SMOTE creates artificial examples of the under-sample cluster using the existing data of the cluster, i.e. it does not replicate the data as this leads to over-fitting of the algorithm. For each dependant, the SMOTE method finds a number of k-nearest neighbours. The number of neighbours to be found depends on the rate of over-sampling required to obtain. Artificial data points are created across the line that connects the data point and its nearest neighbours. After running the SMOTE method across all clusters, the data were divided into 335 “awake” segments and 336 “sleep” segments for the 2-clustered dataset and 270 “awake”, 199 “drowsy” and 245 “sleep” segments for the 3-clustered dataset. Once the imbalance of the data was solved, the following section discusses the MLAs used and the results obtained with each algorithm.

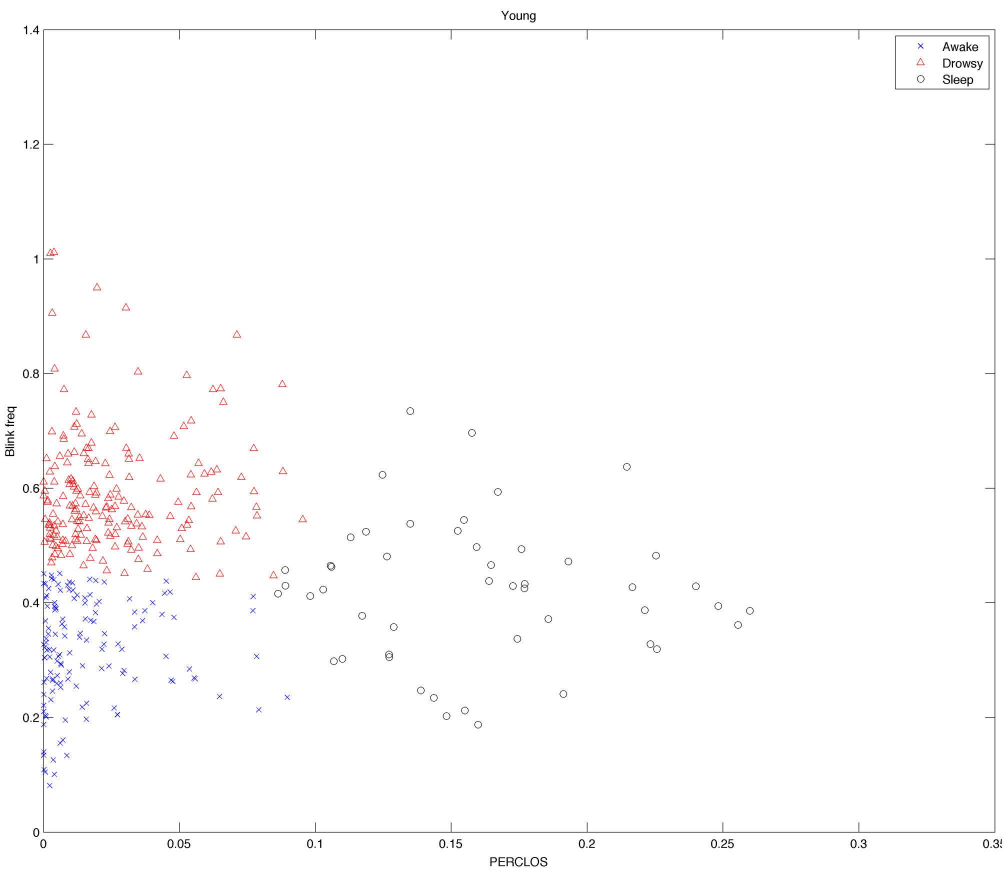


Figure 4-4 Result of the unsupervised k-means algorithm for two clusters using PERCLOS (plotted in the x-axis) and Blinking frequency (plotted in the y-axis). The blue crosses were classified as "awake" segments, the red triangles were classified as "drowsy" segments and the black circles were classified as "sleep" segments.

4.3.4 Predicting levels of sleepiness using Support Vector Machine

As discussed in Chapter 3, Support Vector Machine (SVM) algorithms are able to define a linear function in higher dimensions that will classify and distinguish between two or more categories (Harrington, 2012; Bell, 2015; Marsland, 2015; Murphy, 2012; Alpaydin, 2010). The clustered datasets (obtained with the k-means algorithm) were used as targets and the driving behaviour as features for the SVM to predict the sleepiness state. The first dataset that was used was the 2-clustered dataset, i.e. the target dataset only had two possible outputs ('Awake' or 'Sleep'). To determine the suitability of the SVM algorithm, the accuracy was calculated. As presented in chapter 3, accuracy is calculated using the following formula:

$$\frac{\boldsymbol{True\ Positive + True\ Negative}}{\boldsymbol{True\ Positive + True\ Negative + False\ Positive + False\ Negative}}$$

where “true positive” is the number of times “sleep” was predicted correctly, “true negative” is the number of times “awake” was predicted correctly, “false positive” is the number of times “sleep” was predicted where “awake” was expected and “false negative” is the number of time “awake” was predicted where “sleep” was expected.

The accuracy for the MLA algorithms was calculated using k-fold cross-validation. K-fold cross-validation is an evaluation method where the data are split into a specific number of equal sized sub-samples (the number of equal sub-samples is normally assigned the variable name K, therefore the name K-fold). The algorithm is trained with K-1 subsamples and the remaining subsamples are used for testing. An accuracy value is obtained for the subsample tested. For the next iteration, a different subsample is chosen for testing and the algorithm is trained with the remaining K-1 subsamples. The iterations continue until all subsamples are used for. At the end of the cross-validation, there are K accuracy values, i.e. one for each subsample tested. An average is done to obtained the accuracy of the algorithm. By using k-fold cross-validation, it is possible to assure that the algorithm will not be trained to over-fit the dataset. For the dataset of this study, the k-fold cross-validation was performed by leaving the data of one participant aside for testing and using the data of the rest of the participants for training, i.e. each participant obtained an accuracy value. The accuracy presented for every algorithm is this chapter is the mean value of the accuracy values obtained during the k-fold cross validation process. The accuracy obtained using SVM was 81.19% (SD = 4.17%). Table 4-2 presents the error box of the SVM using driving behaviour to predict the sleepiness states.

Table 4-2 SVM error box using two clustered datasets as targets and driving variables as feature. The columns refer to the target and the rows refer to the prediction

	Awake	Sleep
Awake	78.01%	21.99%
Sleep	15.88%	84.12%

The following step was to determine if creating a baseline for each participant’s driving behaviour would increase the accuracy of the algorithm. The hypothesis was that by knowing the baseline of each participant, i.e. the driving behaviour of each participant when they are in an "awake" state, would reduce the variability of the data and decrease the error rate. The baseline was calculated as the average of each driving behaviour variable for each participant during the first 10 minutes of the 'awake' driving task, i.e. the driving task during his or her first visit to the driving simulator. Once the baseline was obtained, the values for each driving variable were subtracted from the baseline to obtain a relative value in relation to the normal driving behaviour of each participant. These values were used to predict the 2-clustered dataset using SVM. The accuracy of the SVM using the baseline decreased (76.53%, SD=6.03) compared to not using the baseline. This suggested that knowing the driving behaviour of each participant does not increase the accuracy of the prediction. Table 4-3 presents the error box obtained by the SVM algorithm using the relative value of standard deviation of lane position according to baseline as feature to predict the 2-clustered dataset. The results obtained using SVM to predict the 2-clustered ("awake" and "sleep") did not reached the accuracy levels presented by other researchers as high levels of accuracy (Yeo et al., 2009; Shuyan & Gangtie, 2009; Yang et al., 2010; Patel et al., 2011). Therefore, the following step was to test the 2-clustered data with Neural Networks (NeuroNets), another type of MLAs, to determine if better accuracy could be obtained. Using baseline to obtain relative driving behaviour values did not have any effect on SVM. Therefore, relative values of driving behaviour variables according to baseline per participants were not tested in NeuroNets.

Table 4-3 SVM error box using 2 clustered datasets as targets and the relative value of the driving variables according to baseline as feature

	Awake	Sleep
Awake	74.15%	25.85%
Sleep	21.31%	78.69%

4.3.5 Predicting levels of sleepiness using Neural Networks

The second algorithm tested to predict the clustered datasets was the Neural Networks (NeuroNets) algorithm. As discussed in chapter 3, NeuroNets algorithms can define single or multiple linear functions that will classify and distinguish between two or more categories (Harrington, 2012; Bell, 2015; Marsland, 2015; Murphy, 2012; Alpaydin, 2010). Similarly to the previous section, driving behaviour variables were input as features to predict the clustered datasets using NeuroNets. The NeuroNets used was a three layer feed-forward algorithm with 12 neurones in the hidden layer and a learning rate of 0.1, as shown in Figure 4-5. The number of neurons in the hidden layer was decided by running the algorithm with different number of hidden neurons. The number of hidden neurons varied from 6 to 24 with increasing steps of one. The algorithm increase accuracy until it reached 12, afterwards the accuracy did not change or decreased. The value for the learning rate was also decided using the same approach. The learning rate value was varied from 0.05 to 0.2 with increasing interval steps of 0.01. The accuracy of the algorithm increased once it reached 0.1 and then the accuracy decreased with every step increase. The number of iteration for the training phase was also decided heuristically. The number of iterations were increases from 0 until 10,000 in batches of 100. When reaching 1,000, the accuracy of the algorithm did not increase further.

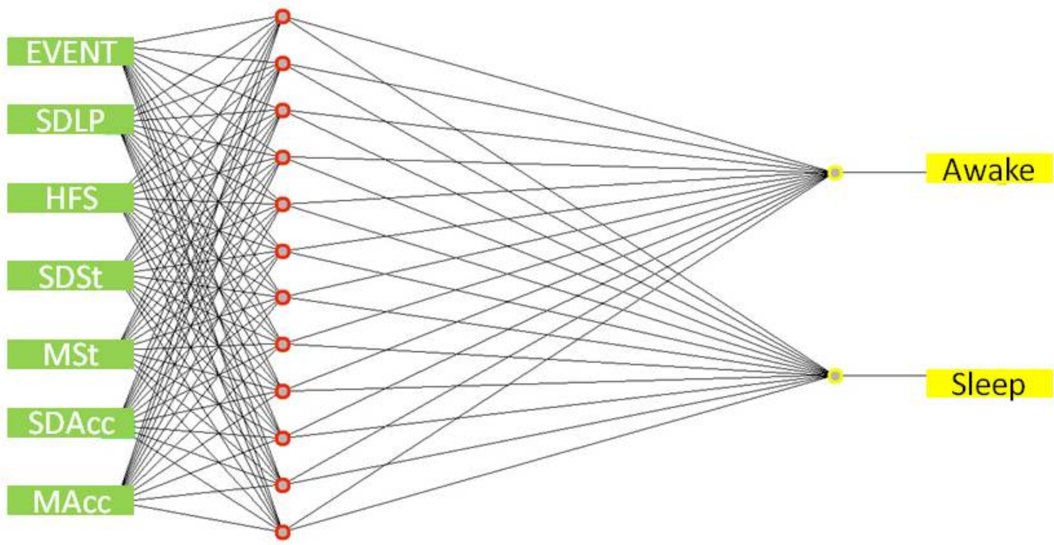


Figure 4-5 NeuroNets 3 layer feed-forward with 12 hidden layers design used to predict a binary classification of sleepiness

For the 2-clustered dataset, i.e. predicting between 'Awake' and 'Sleep', the accuracy of the algorithm was 89.70% (SD=3.55%), higher than the accuracy obtained with SVM. Table 4-4 shows the error box of the NeuroNets classification. An accuracy of 89% has been considered as a high level of accuracy in the literature

(Patel et al., 2011; Sayed & Eskandarian, 2001; Shuyan & Gangtie, 2009). This led to the conclusion that the method used to cluster the blinking behaviour into a binary level of sleepiness, i.e. using unsupervised k-means clustering algorithm, was suitable. To validate the threshold value obtained using the k-means clustering, which separates the data into two levels of sleepiness, a manual analysis was done. For the manual analysis, the threshold value was varied manually from 0.01 to 0.2 with interval jumps of 0.005 and was then run using the NN. The best result was obtained using the threshold value of 0.09, which is similar to the threshold value obtained using the k-means clustering (although the threshold obtained by the k-means clustering is not a constant value, the behaviour of the threshold is almost a vertical line crossing at a PERCLOS value of 0.09).

As a high accuracy level was reached, the next step was to determine if the data could be classified into three levels of sleepiness instead of just two. As discussed previously, the aim of creating multiple levels of sleepiness was to be able to predict pre-states of “sleep”, as the “sleep” state is a high-risk state for the driver. In addition, by defining multiple levels of sleepiness, the actions taken by a safety system will not have high jumps in automation. When the driver presents low levels of sleepiness, a warning signal might be enough to indicate that the driver is in need of a rest. When the driver presents high levels of sleepiness, the driver’s skills are reduced (Merat et al., 2014; Endsley, 1995; Carsten et al., 2012) and the system should take partial or complete control of the driving tasks. High jumps in automation would happen when passing from low levels of sleepiness directly to high levels of sleepiness. Having multiple levels of sleepiness can gradually increase the automation needed as the levels of sleepiness increase.

Table 4-4 NeuroNets error box using 2 clustered datasets as targets and driving variables as feature

	Awake	Sleep
Awake	85.98%	14.02%
Sleep	6.86%	93.14%

For the 3-clustered dataset, i.e. predicting between ‘Awake’, ‘Drowsy’ and ‘Sleep’, the accuracy using NeuroNets to determine three clusters decreased considerably (68.45%, SD=4.26). Table 4-5 shows the error box for NeuroNets when predicting 3-clustered dataset. The reasons for the low accuracy obtained with NeuroNets when classifying the 3-clustered dataset could be that there is not enough differentiation in the blinking behaviour to cluster the data into more than two states. Another hypothesis was that the driving task was not long enough for the participants to achieve higher levels of sleepiness and therefore obtain higher changes in blinking behaviour.

Table 4-5 NeuroNets error box using 3 clustered datasets as targets and driving variables as feature

	Awake	Drowsy	Sleep
Awake	74.16%	19.04%	6.80%
Drowsy	39.48%	42.62%	17.90%
Sleep	8.39%	7.50%	84.11%

4.3.6 Continuous target

The previous sections presented the methods and results when predicting a specific number of clusters representing different sleepiness state. Although this reduces the error due to the limited number of possible outcomes the algorithm can have, it depends widely on the thresholds used to cluster the data into different categories. If the thresholds are incorrect, even if the accuracy of the algorithm is high, the results obtained will not be robust. An alternative is to predict a continuous value and then act accordingly to the specific value obtained. The following section presents the results obtained using the Radial Basis Function Network algorithm (a type of NeuroNets algorithm) for the prediction of a continuous value of PERCLOS and blinking frequency.

4.3.6.1 Radial Basis Function Network

Radial Basis Function Network (RBF) algorithm is a variant of NeuroNets (Orr, 1996). The structure of the RBF is a two layer NeuroNets with a hidden layer, same as the ones discussed in chapter 3. The difference is that each hidden node represents a data point from the training set. Each hidden node is a non-linear activation function. The most common activation function used in RBF is the Gaussian function. Each activation function has a central point, which depends on the data point it represents. When testing a new data point, the responses of each activation node will increase or decrease depending on the proximity of the new data point to the central point of each activation node. Figure 4-6 shows the result of four different RBF with different regularisation parameters (how much the RBF is affected by the weight of each activation functions) using sampled data obtained from a sine wave.

RBF was used in the present study to predict a continuous value of PERCLOS and blinking frequency. The feature variables used were the same driving variables presented in section 4.2.1.2. The best results when predicting PERCLOS using RBF were obtained using standard deviation of lane position. Taking into account that the possible outcome from this algorithm is not categorical, the validation of the accuracy of the algorithm was performed using a calculation of root mean square error (RMSE). The RMSE provides a measure of the variability of the expected value (real) compared to the estimated value (predicted)- see Equation (6).

$$\boldsymbol{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (6)$$

In Equation (6), y_i is the i-th target value and $\hat{y}_i$ is the i-th predicted value. The result obtained by the RBF when predicting PERCLOS using standard deviation of lane position was a RMSE value of 0.0421. The results suggested a low accuracy obtained from the RBF algorithm as an error value of 0.0421 of PERCLOS is very high. Comparing this error value with the value obtained calculating the mean of every other PERCLOS value, it shows that the value predicted with the algorithm is no better than calculating the mean of every PERCLOS value. RBF was also used to predict a continuous value of blinking frequency using the driving variables. In this case, the variable that obtained the best results was high frequency steering, with a RMSE of 0.1383. Similarly to the results obtained when predicting a continuous value of PERCLOS, an error value of 0.1383 is very high for blinking frequency.

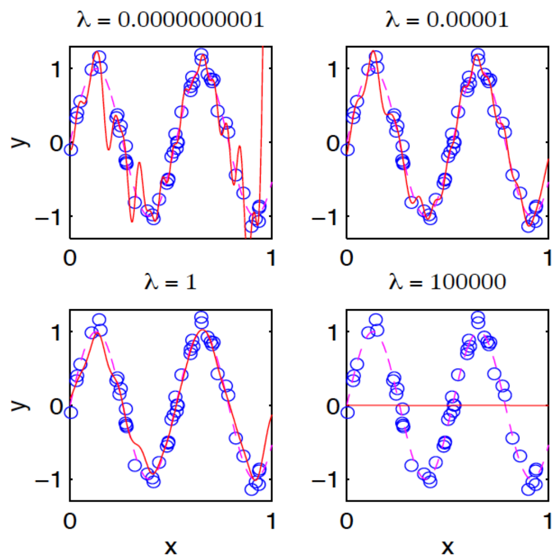


Figure 4-6 Results of using RBF with four different regularisation parameters using sampled data obtained from a sine wave (dotted line is the sine wave, the blue circles are the sampled data and the solid red line is the result of the RBF). (Source: Orr, 1996)

4.4 Conclusion

In the present chapter, a first attempt to determine multiple levels of sleepiness was proposed. Knowing that blinking behaviour is considered a reliable variable to identify changes in sleepiness (Lal & Craig, 2001a,b; Hayami et al., 2002; May & Baldwin, 2009, Wierwille et al., 1994, Dinges and Grace, 1998), PERCLOS and blinking frequency were used to determine the different levels of sleepiness. The aim of this study was to determine the accuracy when blinking behaviour was used to define different levels of sleepiness. To determine the suitability of the defined levels of sleepiness, different MLAs were used to predict the different stages of sleepiness using driving behaviour. The data used during the study was obtained from an experiment previously conducted at the driving simulator of the University of Leeds. The experiment was designed to induce high levels of sleepiness in the participant and the results found by the researchers in charge of the experiment found showed that driving and blinking behaviour were significantly different between the awake and the sleep state (Merat & Jamson, 2013), which suggested that the data contained behavioural data related to different levels of sleepiness. The missing driving and blinking data in the dataset allowed the researcher to test the accuracy of MLAs algorithms when dealing with incomplete data.

Because no consensus could be found in the literature regarding the threshold values of different blinking variables to determine the different levels of sleepiness, an unsupervised MLA was used to cluster the blinking data into different levels of sleepiness. The blinking data were clustered into a binary classification of sleepiness (“awake” and “sleep”) and later into a ternary classification of sleepiness (“awake”, “drowsy” and “sleep”). The suitability of the proposed clustered data were tested using MLAs, specifically SVM and NeuroNets. SVM did not reach high levels of accuracy when using the 2-clustered dataset. NeuroNets achieved high levels of accuracy with the 2-clustered dataset. Therefore, the next step was to test the 3-clustered dataset with NeuroNets. The accuracy obtained was very low. This led to the conclusion that the data did not contain enough changes in sleepiness to classify the data in more than two levels of sleepiness. Another MLA was use to test the accuracy when predicting continuous values of blinking behaviour, which could later be related to a specific level of sleepiness. The results showed that the algorithms obtained better accuracy when the target variables (blinking behaviour) were clustered into discrete sleepiness states instead of predicting a continuous value.

The dataset used in this study was not suitable to define more than two levels of sleepiness. Subsequent empirical work should seek to: (i) define different levels of sleepiness using a variable more susceptible and reliable to changes in sleepiness, such as brain wave activity; (ii) design a longer driving task experiment to be able to record different levels of sleepiness; (iii) additionally to the driving data recorded, record also physiological data to understand the behaviour of the drivers in different levels of sleepiness, and therefore obtain better accuracy when using MLAs. The following chapter explains the design of the experiment conducted during the present PhD study, which attempted to address the issues previously described.

Chapter 5:
Inducing high levels of sleepiness in drivers

5. Inducing high levels of sleepiness in drivers

5.1 Introduction

Results from the previous chapter suggested that the collection of new, additional data might increase prediction accuracy for different levels of sleepiness. The following chapter describes the experiments conducted during the present PhD study. During the experiment, changes in subjective, physiological and driving behaviour due to sleepiness were recorded whilst participants completed a predetermined driving task. EEG activity was recorded for each participant, as it is one of the most reliable indicators of sleepiness. This chapter describes the design, participants, statistical analysis and conclusions for the two sets of experiments. During the first experiment, participants were provided with a high carbohydrate and low protein lunch to the participants to induce high levels of sleepiness. During the analysis, it was found that lunch did not have an effect in sleepiness. In addition, due to the noise and individuality of the EEG recording, it was necessary to conduct a second experiment to obtain more data. The second experiment attempted to induce higher levels of sleepiness than in the first experiment by increasing the length of the driving task.

5.2 Study 1: Effects of lunch on drivers' sleepiness

5.2.1 Aims

The aim of this study was to induce sleepiness in participants to record the changes in driving and physiological behaviour of drivers as sleepiness increases. This would allow the researcher to develop algorithms that could predict the level of sleepiness of the participant. During this study, three key factors to improve prediction accuracy were explored. First, the study assessed the viability of using EEG to record sleepiness in a driving simulator set-up. In chapter 4, blinking behaviour was used as the variable to determine sleepiness. The results obtained in chapter 4 showed that, although there was a significant difference in time in the blinking behaviour indicating an increase in sleepiness, the MLAs could not predict sleepiness, i.e. blinking behaviour using driving behaviour. As presented further in this chapter, the present study examined differences in EEG over time as an index of sleepiness.

The second issue addressed in this study was to explore the effects of lunch in the sleepiness of the participants while driving. Sleepiness may be induced under different artificial conditions approaches (Monk, 2005; Horne & Gibbons, 1991; Merat & Jamson, 2013; Lenne, Triggs & Redman, 1997). One of these approaches to induce sleepiness is lunch. Previous research has found that high calorie meal causes an increase in sleepiness (Reyner et al., 2012). Specially, lunch with high carbohydrate content has been correlated to a decrease in performance (Smith & Miles, 1986a,b; Wells & Read, 1996; Wells et al., 1995; Lloyd, Green & Rogers, 1994; Cunliffe, Obeid & Powell-Tuck, 1997). In the present experiment, participants were given a high carbohydrate and low protein content lunch to explore if it induced higher levels of sleepiness in the participant during the driving task. If this appeared to be successful, it would indicate that for future experiments it would be needed to give lunch to the participants to induce higher levels of sleepiness.

The third issue addressed was to determine the effect of gender. The study aimed to determine if only short hair participants could be recruited for future experiments. The reason to determine this was that EEG in male participants was less noisy than in female participants due to the amount of hair.

5.2.2 Method

5.2.2.1 Driving simulator

Obtaining naturalistic driving data, i.e. driving behaviour data in the real world, presents a challenge due to safety of the driver, e.g. it can be dangerous to obtain naturalistic data of drivers falling asleep at the wheel (Philip et al., 2005; Bos, Bles & Graaf, 2002). This is the main reason why many researchers opt for using an examination of driving under simulation- a safe environment to study such driving behaviour (Philip et al., 2005; Bos, Bles & Graaf, 2002). Evidence indicates that the driving behaviour of participants is similar in the simulator compared to real driving (Philip et al., 2005; Hallvig et al., 2013). Thus, it is possible to extrapolate driving behaviour findings from a driving simulator into the real world. For some variables, such as subjective sleepiness and reaction times, driving simulators appear to have a stronger effect in participants than in the real life (Philip et al., 2005). This makes a driving simulator an ideal context to carry out research on sleepiness.

There are two main types of driving simulators: motion base and static driving simulators, as shown in Figure 5-1. A motion-base driving simulator can move according to the driving behaviour of the participant to give a 'realistic' feel of driving (Greenberg et al., 2003; Merat & Jamson, 2013). As an example, the driving simulator in the University of Leeds consists of a fully equipped Jaguar S-type cab inside a moving 4 metre diameter simulation dome (rigid glass-fibre construction), associated with control loading (steering, brake and accelerator due to its large amplitude eight degree-of-freedom motion). Inside the dome there is a nine-channel 300° field-of-view projection system showing the road (Merat & Jamson, 2013).

The static driving simulator used in the present PhD study (Figure 5-1b) consists of a steering wheel, a set of pedals (accelerator and brake), a wide computer screen and an audio system. Greenberg et al. (2003) found that the lack of motion cues from the static driving simulator affects the lateral and heading control of the vehicle. Despite these differences, the static driving simulator was used in the present study, as it would reduce the noise in the EEG. The effects of movement and electrical equipment in the noise generated in EEG recordings will be explained in detail in further sections.

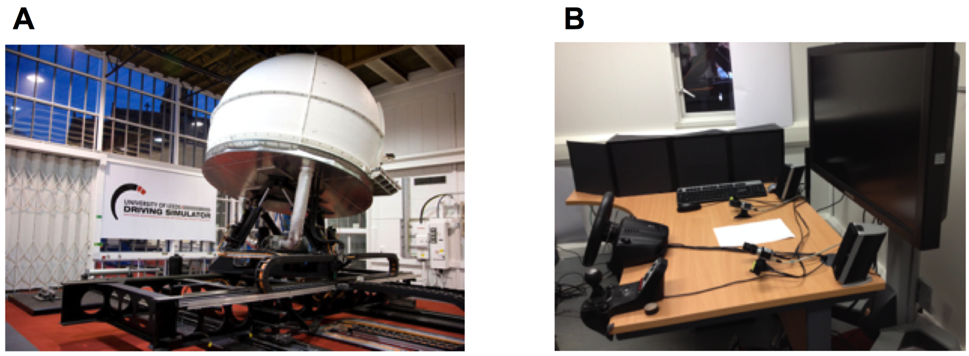


Figure 5-1 a) Motion base driving simulator at the University of Leeds (source: UoLDS, 2012) b) Static driving simulator at the University of Leeds. The second one was the driving simulator used during the present PhD

For this experiment, the participants undertook the driving tests in the static driving simulator in the Physics Research Deck of the University of Leeds (Figure 5-1). As discussed before, the static driving simulator was chosen instead of the motion base driving simulator as the constant movement of the motion base simulator and the electrical equipment around it can increase noise in the recording of physiological variables (Fisch, 2000; Núñez, 2010; Benbadis, 2006). The static driving simulator consisted of a Logitech Steering Racing Wheel with force feedback, accelerator, brake and clutch pedals (clutch was not used during this experiment). The road was displayed in a Samsung 400MX2 monitor (40" size with a resolution of 1920 x 1800) with a vertical field of view of 45 degrees and a horizontal field of view of 80 degrees. A sound system reproduced the sounds of the engine and the environment around the car.

5.2.2.2 Participants

As discussed in chapter 2, the group most at risk for crashes due to falling asleep while driving are young people under the age of 30. This could be due to the many differences between young drivers and old drivers outlined previously (Campagne, Pebayle & Muzet, 2004; Lowden et al., 2009; Filtness et al., 2012). For this experiment, only young participants under the age of 30 with a valid UK license (2 years minimum) were recruited. The number of participants taking part in this experiment was 18 students and faculty members from the University of Leeds (8 males, 10 females) aged between 19 and 29 (M=22.72, SD=3.04). Before being accepted to take part in the experiment, participants had to fill up a screening questionnaire (Appendix E) where they were asked to provide their height and weight information while signing up to the experiment. Body-Mass Index (BMI), a ratio between the height and the weight of an individual, has been related to sleep apnoea (Romero-Corral et al., 2010). Therefore, participants were only accepted to take part in the experiment if their BMI was lower than 30 $\frac{kg}{m^2}$. The BMI of participants ranged between 18.94 and 27.73 $\frac{kg}{m^2}$ (M=21.97, SD=2.5).

Participants were also asked to refrain from alcohol and caffeine during the experiment day. Caffeine has been shown to be an effective countermeasure against sleepiness (Horne & Reyner, 1996; Reyner & Horne, 1997, 2000, 2002). It also has been found that even low levels of alcohol in the blood have an increasing effect in sleepiness, causing a worsening in driving performance (Banks et al., 2004; Wilkinson & Colquhoun, 1968; Huntley & Centybear, 1974; Peeke et al., 1980). Thus, participants were instructed to avoid alcohol for 24 hours before the experiment and caffeine should be avoided on the day of the experiment. Participants were also asked to maintain a normal (7-8 hours) sleep pattern for three days before the experiment days.

Finally, participants completed the Epsworth Sleepiness Scale (ESS) test (Johns, 1991). The ESS test captures an individual's predisposition to sleepiness in daily situations (Appendix F). The scores of the ESS test range from 0 to 15, where a score higher than 10 indicates an individual is more likely to fall asleep in normal daily situations than the average population (Filtness et al., 2012; Yeo et al., 2009). A score close to zero means that a person has a wakeful level higher compared to the average people (Yeo et al., 2009). Previous studies have used a strict approach and excluded participants with a score under eight and over nine. Others have taken a less stringent approach- only excluding participants with scores above 10 (Filtness et al., 2012; Reyner et al., 2012). For the current experiment, only one participant (participant 10) was excluded due to an ESS score above 10. The rest of the participants had an ESS score between 2 and 10 (M=6.72, SD=2.87). Participant 2, 3, 4, 9, 12, 13, 15 and 19 were also removed, as the noise in their EEG recording could not be removed after the artefact correction process. After excluding the participants above mentioned, the number of participants analysed were nine (6 males, 3 females).

5.2.2.3 Design

As presented in chapter 2, sleepiness, i.e. the probability of a person falling asleep, increases due to many different factors. These factors could be environmental, e.g. time of the day affecting circadian rhythm; task related, e.g. monotonous road with low traffic; or physiological, e.g. length of time the person has been awake or amount of food consumed (Curcio et al., 2001; Zhao & Rong, 2013; Thiffault & Bergeron, 2003; May & Baldwin, 2009; Johns, 2000; Lenne et al., 1998; Philip et al., 2005; Vitaterna, Takahashi, & Turek, 2001). In this experiment, participants were asked to undergo two monotonous driving task happening in two different day sessions. During one of the sessions, participants were given a lunch before the driving task whilst in the other session, participants were to abstain from eating before the experiment. The experiment tested the effects that lunch and the monotony of the task have in the increase of sleepiness in the participants and followed a within participants repeated measures design where lunch was the independent variable.

Each participant took part in both conditions and each condition was tested in different days. During each condition, participant underwent a monotonous driving task in a static driving simulator. For the 'Lunch' condition, participants arrived to the driving simulator at 12:00 in the afternoon. This period of the day is called 'Post-lunch dip' and it has been found to increase sleepiness in participants (Reyner et al., 2012; Monk, 2005). The lunch provided during the 'Lunch' condition was 250 grams of Tomato and Herbs Risotto (Uncle's Ben). The reason to select this as the appropriate lunch was that it was high in calories, high in carbohydrate and low in protein, which is known to increase the level of sleepiness in drivers (Reyner et al., 2012). For the 'No Lunch' condition, participants arrived to the driving simulator at 9:00 in the morning. For this condition, participants were asked to refrain from eating anything two hours before the experiment. The order of the conditions ('Lunch' vs. 'No Lunch') was randomised and counter-balanced to reduce the effect of the order.

During both conditions, participants had a practice run (lasting 5-10 minutes) on their arrival at the driving simulator in order to be acquainted to the driving environment. Following the practice run, during the 'No Lunch' condition participants were given an hour break and instructed to avoid food, alcohol and coffee; and during the 'Lunch' condition, participants were taken to an adjacent area where they could eat the aforementioned lunch. From this step onwards, both conditions followed the same design.

At their return to the driving simulator (either from their break or their lunch), participants were asked to determine their level of stress using the Perceived Stress Scale and the Stress and Arousal Checklist (McCormick, Walkey & Taylor, 1987) (Appendix I). They were also asked to rate their subjective sleepiness using the Karolinska Sleepiness Scale (KSS; Appendix D) (Akerstedt & Gillberg, 1990). Following this, an EEG net was positioned on the head of the participant to record the electrophysiological activity during the task (Figure 5-2). After the EEG was positioned, participants were asked again to rate their subjective sleepiness level using the KSS to explore if there were any effects in the sleepiness of the participant due to the length of time while positioning the EEG.

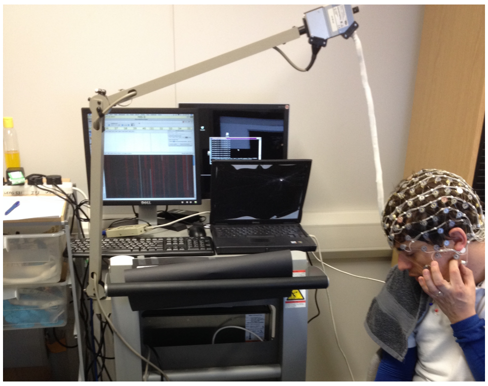


Figure 5-2 EGI system used to record brain wave activity. The EEG net is a 129 electrode cap that is positioned on the head of the participant. The cap is then connected to the EGI system, which consists of an amplifier (not visible) and a computer that records the signal (screen on the left side). The laptop and the right screen were used to synchronise the EGI system with the driving simulator for both systems to start recording at the same time.

Participants underwent a 45-minute driving task in a two-lane motorway with no traffic and few gentle curves. The driving task was conducted in a night driving environment. The driving simulator room was also set up to resemble a night environment by covering the windows to impair light to enter the room and turning off the room lights during the task. Before the driving, ask started, participants were instructed to maintain the same speed (40 miles per hour) and stay in the same lane (left lane) throughout the experiment. The participants were monitored from an adjacent room using a video camera set up in the corner of the room. There was no interaction between the participant and the researcher during the driving task. At the end of the driving session, participants were asked once more to rate their level of sleepiness using the KSS test (Akerstedt & Gillberg, 1990).

5.2.3 Subjective data recording

The KSS test is a 9-point rating scale, where “1” is completely awake and “9” is almost falling asleep. This test has been widely used in many driving experiments to assess sleepiness (Lowden et al., 2009; Reyner et al., 2012; Shuyan & Gangtie, 2009; Filtness et al., 2012). A score of five or below is considered an awake state and a score of six and above is considered a sleepy state.

Unfortunately, KSS cannot be considered the “ground truth” for sleepiness level due to the lack of sensitivity of the measurement (Kaida et al., 2006; Akerstedt & Gillberg, 1990; Kozak et al., 2005). Although some researchers have found the people are capable of assessing correctly their level of sleepiness (Williamsom et al., 2014), it has been found that, due to a low sensitivity in the scale, participants are not able to assess with precision their level of sleep (Kaida et al., 2006; Dinges et al., 1998). Although the KSS lacks high sensitivity and precision, it remains a useful tool to determine subjective changes in the sleepiness of the participant (Lowden et al., 2009; Reyner et al., 2012; Shuyan & Gangtie, 2009; Filtness et al., 2012).

5.2.4 Driving data recording

The following driving variables were measured during the driving task:

- Standard deviation of lane position (SDLP). The participant was asked to maintain the same lane throughout the driving task. If the left tyre or right tyre touches the left or right edge of the lane, respectively, it counts as an ‘out of lane’. ‘Out of the lane’ data are not taken into account until the car comes back inside (the left or right tyre are not longer touching the edge of the lanes or are not longer outside the edges of the lane). The standard deviation of the position of the car is calculated from the mean of the position of the car during a specified segment of time, i.e. not in respect to the centre of the lane. It has been found that as sleepiness increases, the standard deviation of the lane position also increases (Lowden et al., 2009).
- Standard deviation of speed (SDSpeed): Participants were asked to maintain the same speed throughout the driving task. The standard deviation of the speed was calculated with respect to the mean of the speed in a specified segment of time, i.e. not in respect to the specified speed they were asked to maintain. The first segments (first 5 minutes) of SDSpeed were excluded from the analysis. As the participants started from a resting position and had to reach the desired speed, the variability of speed in this first segment was too high. Although there is not a common conclusion regarding if speed increases or decreases in sleepy drivers, it has been found that speed changes as sleepiness increases (Bloomfield, Harder & Chihak, 2009; Hargutt et al., 2000; Oron-Gilad & Shinar, 2000; Riemersma et al., 1977).

- Standard deviation of steering wheel angle (SDSteering): A study conducted in a driving simulator showed that drivers' steering ability was impaired when they became sleepy (Bloomfield, Harder & Chihak, 2009).
- High frequency steering (HFS): number of high frequency movements in the steering wheel. Variations in the steering wheel movements in the frequency range of 0.3 to 0.6 Hz are accounted as steering corrections (Merat et al., 2012; Ostlund et al., 2006). Merat & Jamson (2013) found that the HFS of sleepy drivers differs from the HFS of awake drivers.
- Time to lane crossing (TTLC): time it will take to cross the lane if the car continues with the same direction and speed (Mammar, 2006). It has been found that the TTLC tends to decrease as sleepiness increases.
- Out of lane (OOL): number of times the car goes out of the specified lane. When the left or right tyre touched the edge of the left or right lane, respectively, was accounted as an 'out of lane' event (Reyner et al., 2012). Sleepy drivers tend to have a higher number of 'incidents' (running out of the lane) than awake drivers.

5.2.5 EEG data recording

As presented in chapter 3, EEG was used to record the electrophysiological activity of the participant, which has shown to be strongly correlated with levels of wakefulness and sleep (Lal and Craig, 2002; Eoh, Chung & Kim, 2005; Jap et al., 2009). The EEG system used for this experiment was the 129 channels EGI net cap (Electric Geodesic Inc. [EGI], Eugene OR, USA). After recording the subjective sleepiness of the participant, the EEG acquisition net was positioned on the head of the participant while they sat down (Figure 5-2). The net was soaked in a solution of potassium chloride electrolyte, which increases the conductivity between the scalp and the electrodes of the EEG (Electrical Geodesics Inc, 2007). After the EEG is positioned on the head of the participant, the EEG is connected to the EGI recording station. The EGI recording station consisted of an amplificatory and bespoke software for data acquisition (Electrical Geodesics, Inc., Eugene, OR). The sample rate for the EGI system was set on 500 Hz with a bandwidth of 0.01-100 Hz. The process of positioning the EEG on the participant head took around 20 to 30 minutes.

5.2.5.1 Frequency bands

The EEG variables analysed for this experiment were the magnitude of the theta frequency band (4 to 8 Hz), the magnitude of the alpha frequency band (8 to 13 Hz), the magnitude of the beta frequency band (13 to 20 Hz), the magnitude of the ratio $\frac{theta+alpha}{beta}$ and the magnitude of the ratio $\frac{alpha}{beta}$. The magnitude of each frequency band was calculated by obtaining the mean of the magnitudes of all frequencies in each frequency range, e.g. a mean of the magnitudes of all frequencies falling in the range of 4 to 8 Hz resulted in the magnitude of the theta frequency band. The method used to transform the EEG data into the frequency domain is further explained in the following section.

5.2.5.2 Fast Fourier Transform

To be able to transform the data into the frequency domain it is necessary to perform Fast Fourier Transform on the data (Cohen, 2014). Any signal (electrical, audio, etc.) can be analysed in the time domain or in the frequency domain. The time domain refers to the data recorded in time, i.e. the raw data recorded. For EEG, time domain analysis refers to determining changes in voltage in the electrodes as time passes. The frequency domain analysis refers to determining changes in power in different frequencies.

FFT is a method that determines the power certain frequencies have on the data. It uses the theory of convolution to be able to determine the power of each frequency in the data. Convolution is the weight one signal has over another signal. Convolution is a method used in other fields beside EEG analysis. In statistical analysis, this is used in cross-covariance and in signal analysis, this is used to filter the data at certain frequencies.

In mathematical terms, convolution is the dot product between two vectors. A dot product between two vectors produces a scalar result. One of the vectors represents the original data, which is called the signal, and the other vector represents the signal to be weight, called the kernel. The objective is to know how much of the kernel signal is present in the original signal. Figure 5-3 shows a convolution between a signal (Figure 5-3a) and a kernel (Figure 5-3b). The result of the convolution is shown in Figure 5-3c.

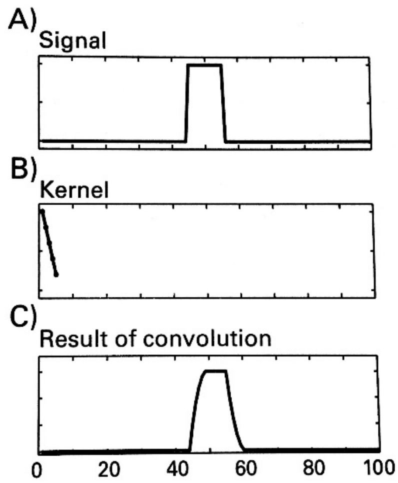


Figure 5-3 Effect of convolution on a signal. a) An example of a square signal that will be used to show the effect of convolution (this is called the original signal) b) The Kernel signal is the signal of interest and its determined by the researcher depending on his/her objective c) Result of the convolution between the signal and the kernel (when the kernel signal passes through the original signal) (Source: Cohen, 2014). Reprinted from "Analyzing Neural Time Series Data: Theory and Practice" by Mike X. Cohen. Copyright © 2014 by Mike X. Cohen. Used by permission of The MIT Press, Cambridge, MA, USA.

The kernel is a sliding signal that goes through all the points of the signal. Figure 5-4 shows the principle of convolution using the points of the signal and the kernel. Figure 5-4a shows that the result of the dot product between the signal and the kernel creates a single value. The kernel signal is then shifted one data point to the right of the signal and repeats the dot product, results in a second single value. This continues until the kernel reaches the last possible data point of the signal, as shown in Figure 5-4b. As it can be concluded, the resulting signal from the convolution is smaller than the original signal. If there is a need for the resulting signal to be the same size as the original signal, zero values (called zero padding) should be added to the beginning and the end of the original signal as presented in Figure 5-4c.

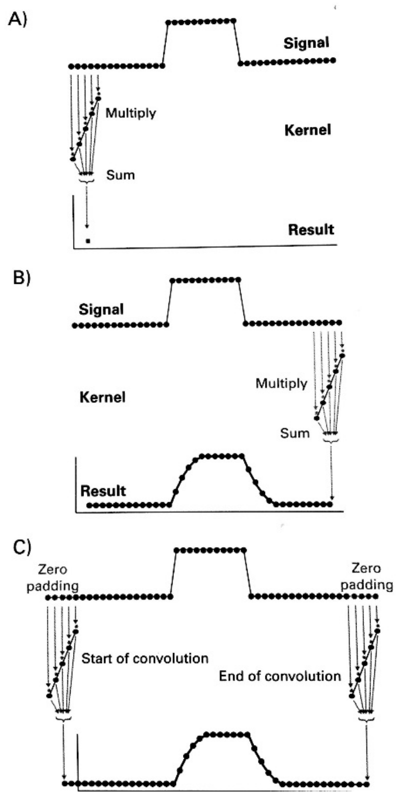


Figure 5-4 Steps to perform a convolution process on a signal. a) When transforming a signal using convolution, it is needed to do a dot product. The dot product of kernel and signal results in a new data point b) The kernel is then moved one data point to the right to obtain a new second point. This process is continued until the end of the original signal Dot product of the kernel and the signal to obtain last convolution value c) For the resulting signal from the convolution process be same length as the original signal, zero values (zero padding) has to be added to the original signal (this means that new data points need to be added to the right and left of the original signal with a value of zero) (Source: Cohen, 2014). Reprinted from "Analyzing Neural Time Series Data: Theory and Practice" by Mike X. Cohen. Copyright © 2014 by Mike X. Cohen. Used by permission of The MIT Press, Cambridge, MA, USA.

In the case of EEG, the EEG signal is the original signal and the kernels are sine waves of different frequencies. By performing a convolution with a sine wave with an specific frequency in the EEG signal, the result will be the weight that specific sine wave has on the EEG signal, i.e. the power that specific frequency has in the EEG signal. By performing convolutions between the EEG signal and sine waves with different frequencies, it is possible to determine that power each frequency has on the EEG signal. This method is called Discrete Fourier Transform and takes substantial computational time, as the process is repeated for each different sine wave. The FFT reduces the computational time by removing the redundant processes found in the Discrete Fourier Transform.

One of the problems encountered with Fourier Transform is the need for the data to be stationary, i.e. does not changes over time. EEG is a non-stationary signal, i.e. the signal changes due to cognitive processes happening inside the brain. When performing Fourier Transform in non-stationary data, noise is created in the frequency domain. Figure 5-5 a shows two signals in the time domain. The one on the left is a stationary signal composed of four sine waves with 3, 5, 7 and 10 Hz of frequency, respectively. The one on the right is a non-stationary signal composed as well of four sine waves with 3, 5, 7 and 10 Hz, respectively. Figure 5-5b shows the results of a Fourier Transform of both signals. For the non-stationary signal, the frequency peaks are not as well defined as in the stationary signal. There is also more noise, i.e. non-zero values in the non-peak frequencies, in the Fourier Transform for the non-stationary signal than in the stationary signal.

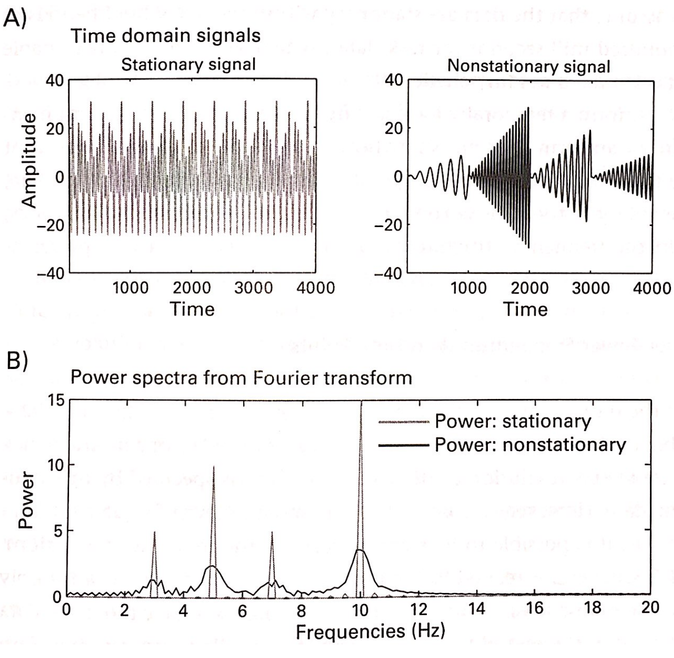


Figure 5-5 Difference between stationary and non-stationary signal. a) On the left is a stationary signal (does not change over time) and on the right a non-stationary signal b) When performing FFT on a stationary signal, the frequency analysis has clearer results than when performing FFT in a non stationary signal (Source: Cohen, 2014). Reprinted from "Analyzing Neural Time Series Data: Theory and Practice" by Mike X. Cohen.

Copyright © 2014 by Mike X. Cohen. Used by permission of The MIT Press, Cambridge, MA, USA.

The non-stationary assumption of the FFT can be solved by a method called the Short FFT. It can be assumed that for short time segments (~200-300 milliseconds) the EEG data are stationary. This means that if the FFT is performed in that brief segment of EEG data, the noise created from non-stationary data will be attenuated. This is the basis of the short FFT, i.e. performing FFT in a short moving window of EEG data. When performing a FFT, the edges of the data, i.e. the beginning and ending of the time series data, create edge artifacts. When performing many FFTs in several short segments, the noise from the edge artifacts can contaminate the data. These reduce the edge artifacts from appearing, hence making it necessary to taper the data. A taper is any function that dampens the edges of the short segment of the signal to a zero value. The most common taper functions are Hann, Hamming and Gaussian. Figure 5-6 shows these three tapers. These tapers allow the data to be gradually dampened to zero on the edges without creating any edge.

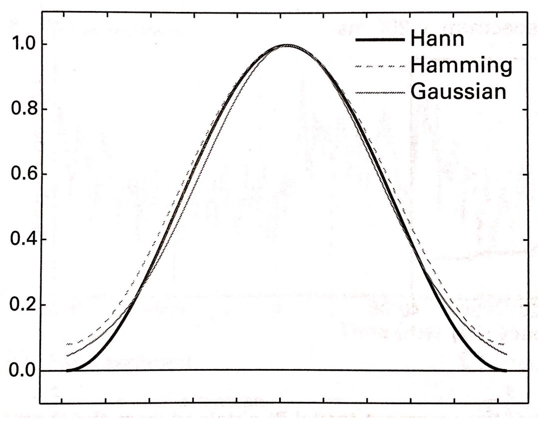


Figure 5-6 Taper functions for segmentation of a EEG dataset used to reduced the non-stationary problem of EEG signals (Source: Cohen, 2014). Reprinted from "Analyzing Neural Time Series Data: Theory and Practice" by Mike X. Cohen. Copyright © 2014 by Mike X. Cohen. Used by permission of The MIT Press, Cambridge, MA, USA.

The next step is to determine the size of the short moving window. The minimum size of the window depends on the Nyquist theorem and the number of unique frequencies needed. The Nyquist theorem states that the maximum frequency that can be obtained from a dataset is half of the sampling rate plus the zero frequency. This means that if a it is needed to obtain information about the 50 Hz signal, the dataset needs to have a sampling rate at least 100 Hz. The reason for the Nyquist theorem is the fact that it is needed to know at least 2 points in a sine wave to determine the frequency of that sine wave. The number of unique frequencies that can be obtained from a dataset is half of the number of points in a dataset or segment of dataset plus the zero-frequency. Thus, if the minimum frequency that is needed for analysis is 0.1 Hz, then the minimum number of points needed in the segment is 20 seconds, i.e. at least two times the period related to that frequency ($\boldsymbol{frequency} = \frac{\mathbf{1}}{\boldsymbol{Period}}$, where period is the length of the segment in seconds). For the present PhD the frequencies band analysed were: theta (4 to 8 Hz), alpha (8 to 13 Hz) and beta (13 to 20 Hz), meaning that minimum time of the segment to be analysed was 0.25 seconds as the minimum frequency was 4 Hz ($\frac{\mathbf{1}}{\mathbf{4}} = \mathbf{0.25}\ \boldsymbol{seconds}$).

5.2.5.3 Electrode clusters

The magnitude of the frequency bands were not obtained for each individual electrode, instead electrodes were joined in 9 blocks, or clusters, depending on their position on the head and a mean value was obtained for the block in order to increase signal strength (Oken & Chiappa, 1986). The clusters were divided in left, middle and right in the horizontal axis (from the left ear to the right ear) and in frontal, central and parietal in the vertical axis (from the nose to the back of head). Figure 5-7 shows the map of the position of the electrodes in the head. The electrodes were divided in the following 9 blocks (not all electrodes were used and some electrodes appear in more than one block):

- Frontal left: Electrodes 18, 19, 23, 24, 25, 26 and 33.
- Frontal middle: Electrodes 9, 10, 15, 16, 18 and 22.
- Frontal right: Electrodes 2, 3, 8, 10, 122, 123 and 124.
- Central left: Electrodes 29, 30, 31, 35, 36, 37, 41 and 42.
- Central middle: Electrodes 7, 31, 55, 80 and 106.
- Central right: Electrodes 80, 87, 93, 103, 104, 105, 110 and 111.
- Parietal left: Electrodes 58, 59, 64, 65, 66, 68, 69 and 70.
- Parietal middle: Electrodes 67, 71, 72, 76 and 77.

- Parietal right: Electrodes 83, 84, 89, 90, 91, 94, 95 and 96.

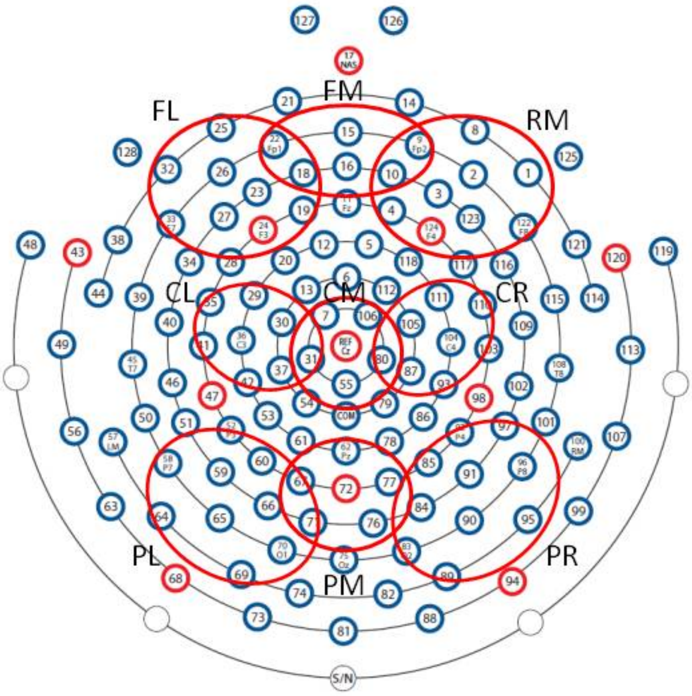


Figure 5-7 The electrodes' labels and locations on the 128 channel EGI sensor net (Hydrocel Geodisic Sensor Net Map). Clusters were created to obtain an average amplitude from different regions of the scalp. Nine regions were determined: Left (L), Middle (M) and Right (R) of the Frontal (F), Central (C) and Parietal (P) region of the scalp. (Source: EGI, 2007). Copyright © 2007 by Electrical Geodesics, Inc. Used by permission of Electrical Geodesics, Inc. Eugene, OR.

5.2.6 Artifacts and automatic cleaning for EEG data

Before the EEG variables were analysed, the data were cleaned to remove artifacts. The EEG data contains noise from muscle movements, the recording device and the environment (Fisch, 2000; Núñez, 2010; Benbadis, 2006; Cohen, 2014). Noise artifacts can be divided in extra physiologic artifacts (due to the environment, machines around the participant and the EEG recording device) or physiologic artifacts (from the physiology of the participant) (Benbadis, 2006). In the following section, these types of artifacts will be explained.

5.2.6.1 Type of artifacts

Extrinsic artifacts are often present in EEG data due to electrical interference produced by power lines and equipment around the EEG system (Fisch, 2000; Núñez, 2010; Benbadis, 2006; Cohen, 2014). It can be identified in the EEG recording, as the noise appears across all electrodes, as seen in Figure 5-8d. This type of artifacts has a frequency of 50 Hz in the United Kingdom (in America the frequency of the electrical noise has a frequency of 60 Hz). Unless it is used a shielded room, i.e. Faraday Cage (Arman, Ahmed & Syed, 2012), it is unavoidable when working with any equipment that uses alternative current. This type of artifact can be introduced either electromagnetically, e.g. the strong current generated in cables by transformers, or electrostatically, e.g. when power cables do not have a proper shield. Any equipment such as televisions, radios, speakers and computers can also introduce this type of artifact. To get rid of this type of artifact, the amount of electrical equipment in the environment should be reduced to a minimum. If electrical equipment is essential to the experiment, a notch filter (filter that removes a specific frequency) with a null frequency of 50 Hz can be applied to the data.

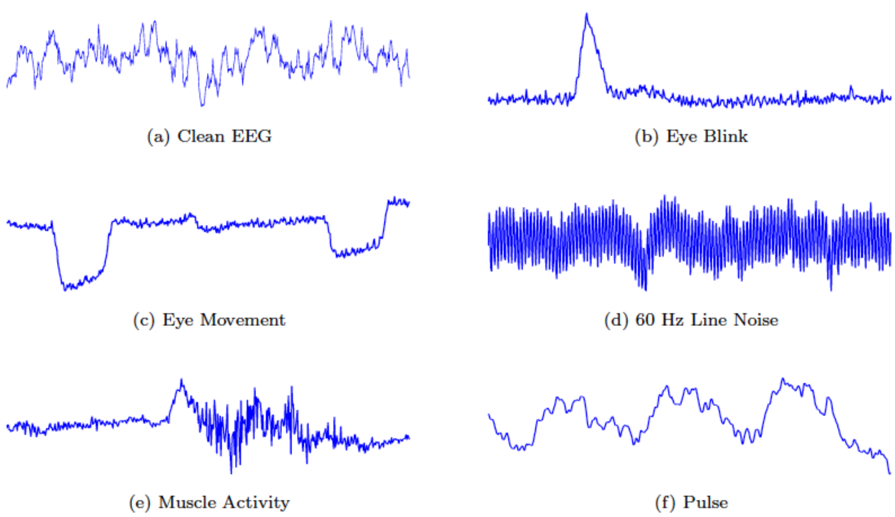


Figure 5-8 A single electrode EEG recording showing different artifacts a) Original EEG recording b) Ocular artefact due to blink c) Ocular artefact due to eye movement d) Extrinsic artefact at 60 Hz (in the UK it is 50 Hz) e) Artifact introduced due to muscle movement f) Artifact introduced due to an external pulse (Source: Knight, 2003)

In order to minimise artifacts due to poor electrode connectivity, one needs to minimise the impedance between the electrodes and the scalp. The graphical user interface of the EGI system presents the impedance of all the electrodes, as shown in Figure 5-9. During the experiment, impedance below 50 ohms was desired and an impedance below 100 ohms was acceptable. The interface allows one to identify faulty electrodes- with their impedance values shown as 3000 ohms. Electrodes with high levels of impedance were noted down so that they may be removed from any subsequent analysis.

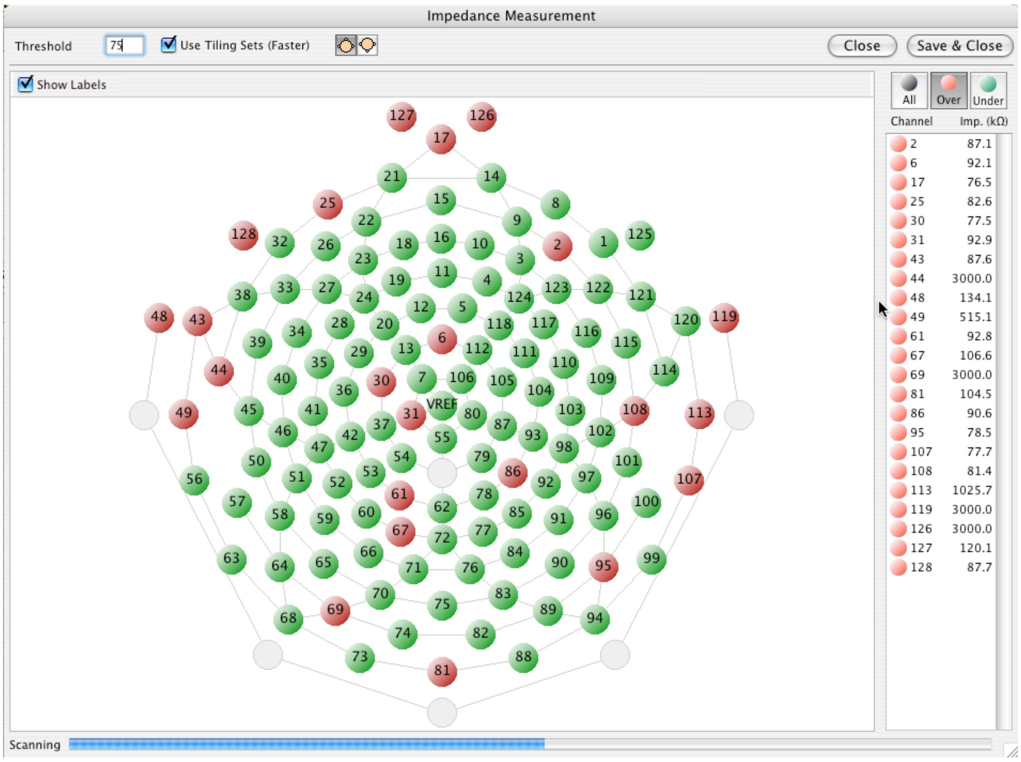


Figure 5-9 EGI interface which allows to identify electrode whose impedance are above certain threshold. The impendence threshold was set to 75kohms. The smaller the impedance, the better the conductivity of that particular electrode. The green electrodes are the ones below the threshold, whilst the red are above. The reason to have electrodes with high impedance may be due to a broken electrode (impedance above 3,000) or not enough solution. The right panel gives a list of the electrode's number and the impedance value.

Physiological artifacts are also common in EEG data- such noise induced from eye movements – ocular atifacts. These are most prominent in the electrodes positioned in the frontal electrodes but can be also detected in central and parietal regions. Lateral and vertical eye movements and as well as blinking can affect the quality of the EEG data. Lateral eye movements can be large voluntary movements, e.g. looking at the side of the screen, or saccades, i.e. small involuntary movement of the eyes.

In the EGI system, there are four electrodes positioned above and below the eye to be able to detect the ocular artifacts. Specialist EEG analysis software such as BESA (Brain Electrical Source Analysis, Grafelfing, Germany) use the information of these electrodes to detect vertical and horizontal eye movements. Blinking artifacts can also be visually detected in the recording due to their electrical magnitude (between 150 and 200 microvolts) and the length of time each blink has (200 to 400 milliseconds (Núñes, 2010). An example of an ocular artifact in an EEG recording can be seen in Figure 5-8b and Figure 5-8c.

Muscle artifacts are introduced due to participant movement (Fisch, 2000; Núñez, 2010; Benbadis, 2006; Cohen, 2014). The most common are introduced due to the movement of frontalis and temporalis muscles, e.g. clenching the jaw or chewing (Benbadis, 2006). This type of artifacts can be identified due to the short outburst of a high magnitude electrical signal (Fisch, 2000; Núñez, 2010; Benbadis, 2006; Cohen, 2014). Muscle artifacts are present generally in low frequency ranges (in the delta frequency range) (Filtness et al., 2012) or in very high frequencies (above 20) (Filtness et al., 2012; Fisch, 2000; Núñez, 2010; Benbadis, 2006; Cohen, 2014). Due to this reason, the delta frequency range was not considered for the analysis and a bandwidth filter with a low cut of 4 Hz and a high cut of 20 Hz was performed on the data. An example of muscle artifacts can be seen in Figure 5-8e.

The other major contributors of physiological artifacts are cephalic and corporeal movements (Fisch, 2000; Núñez, 2010; Benbadis, 2006; Cohen, 2014) such as yawning and head movement. Many researchers opt to ask participants to avoid any movement during the experiment (Fisch, 2000). As this was a driving task, participants were allowed to move freely during the experiment. The artifacts due to cephalic and corporeal movement were removed in a further stage before analyses of the data were done.

Finally, although heart rate, intraoral (tongue movement), perspiration and galvanic skin response and respiration also create artifacts in the EEG recordings, they are not as invasive as the other artifacts explained before and are removed through the application of a high pass filter (Fisch, 2000; Núñez, 2010; Benbadis, 2006; Cohen, 2014).

5.2.6.2 EEG cleaning methods

Before analysing the EEG recording, it is necessary to remove all unwanted artifacts as they might affect the results obtained (Cohen, 2014). Specialised EEG software like BESA (Brain Electrical Source Analysis, Grafelfing, Germany) are capable of recognising and removing ocular artefacts, due to their identifiable pattern (Fisch, 2000; Núñez, 2010; Benbadis, 2006; Cohen, 2014). Unfortunately, other types of artifacts are less stereotyped and have to remove in different ways. Many researchers opt to hire a clinician with EEG experience who can manually identify and reject the artifacts in the data, amongst other type of analysis the clinician can do, e.g. identifying different frequency bands (Vuckovic et al., 2002; Yeo et al., 2009). Although this is a good approach, it requires the time and the expertise of a clinician to clean and analyse the data. An alternative method is the use of independent component analysis (ICA) (Klass, 1995; Cohen, 2014; Jung et al., 2000). Through pattern recognition, ICA determines different elements present in the EEG data, e.g. the small burst of muscle artifacts, the high values from ocular artifacts, blinking behaviour, etc. The researcher then removes the components identified as artifacts based on visual inspection.

5.2.6.3 Automatic cleaning method developed in Matlab

Although some artifacts have a random behaviour, it is possible to discriminate an artefact from EEG data. As shown in the previous figures, artifacts are visually identifiable. This means that an EEG dataset can be manually cleaned, i.e. remove the segments containing noise. For the present experiment, an automatic cleaning process was developed in Matlab (MATLAB 2011b, The MathWorks Inc., Natick, MA, USA) to detect and remove artifacts. The process identified the clean EEG segments (referred hereafter as 'best'), which were used as baseline to compare the rest of the EEG data. Figure 5-10 shows the process of selecting the 'best' EEG segments.

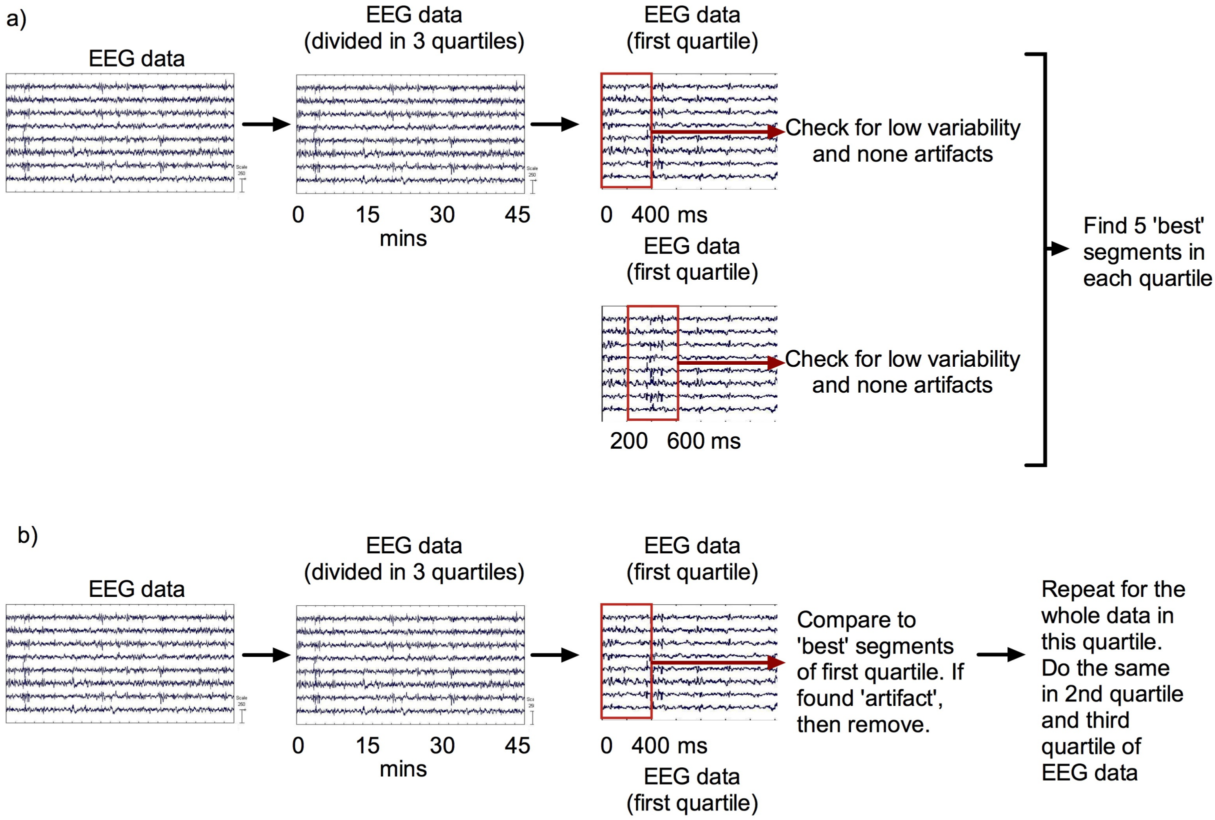


Figure 5-10 Cleaning process to remove artifacts automatically developed in Matlab a) Process of finding segments that correspond to 'good' EEG data. The EEG data are divided in three quartiles and five 'good' segments (low variability and none artifacts) are obtained from each quartile. b) Process of comparing the rest of the data with the 'good' segments to determine if they are artifacts or not. The "good" segments found in the previous process are used to compare the rest of the EEG data. If the difference between the "good" segment and the segment being tested is too high is considered an artefact.

To detect the 'best' EEG segments that could be used as baseline, the 45 minutes EEG recording was segmented in three 15-minutes sections. It was found that as time passes the electrodes of the EEG dried up and more noise was recorded in the EEG. Due to this reason, a set of 'best' segments was selected for the three different 15 minutes time sections. In each 15-minutes section, five 'best' segments were selected as baselines for the specific 15-minutes section. To identify the 'best' segments, a 400 milliseconds moving window with a 50% overlap was used to analyse each 15-minutes segment. The criteria to select the 'best' segments were the following ones:

- Lowest variability
- Lack of artifacts with a voltage magnitude above 20 microvolts

After the 'best' five segments per 15-minutes section were selected, each 15-minutes section was segmented in 400 millisecond segments that were compared to the 'best' segments. A segment was considered an artifact if it did not meet the following requirements:

- More than 5 electrodes have magnitudes above 100 microvolts
- More than 100 electrodes have a higher value than 5 times the standard deviation of the 'best' segment
- There was a 100% error or more between the variability of the 'best' segment and the segment being compared in more than $\frac{2}{3}$ of its electrodes
- There was a 75% error or more between the variability of the 'best' segment and the segment being compared in more than $\frac{1}{3}$ of its electrodes
- More than 5 electrodes have a mean value of 30 microvolts

5.2.6.4 Results of automatic cleaning method

To determine the accuracy of the automatic artefact rejection process developed in Matlab (MATLAB 2011b, The MathWorks Inc., Natick, MA, USA), a comparison was made against an EEG dataset cleaned manually. The results are shown in Figure 5-11. Two random EEG datasets were selected (Participant 9 and 12). From the manual cleaning, the EEG data removed were 27.28% and 14.98% for participant 9 and 12 respectively. The automatic cleaning removed 21.96% and 10.6%, for participant 9 and 12 respectively. The results of the comparison gave the researcher enough evidence to use the automatic cleaning process for all the remaining EEG datasets.

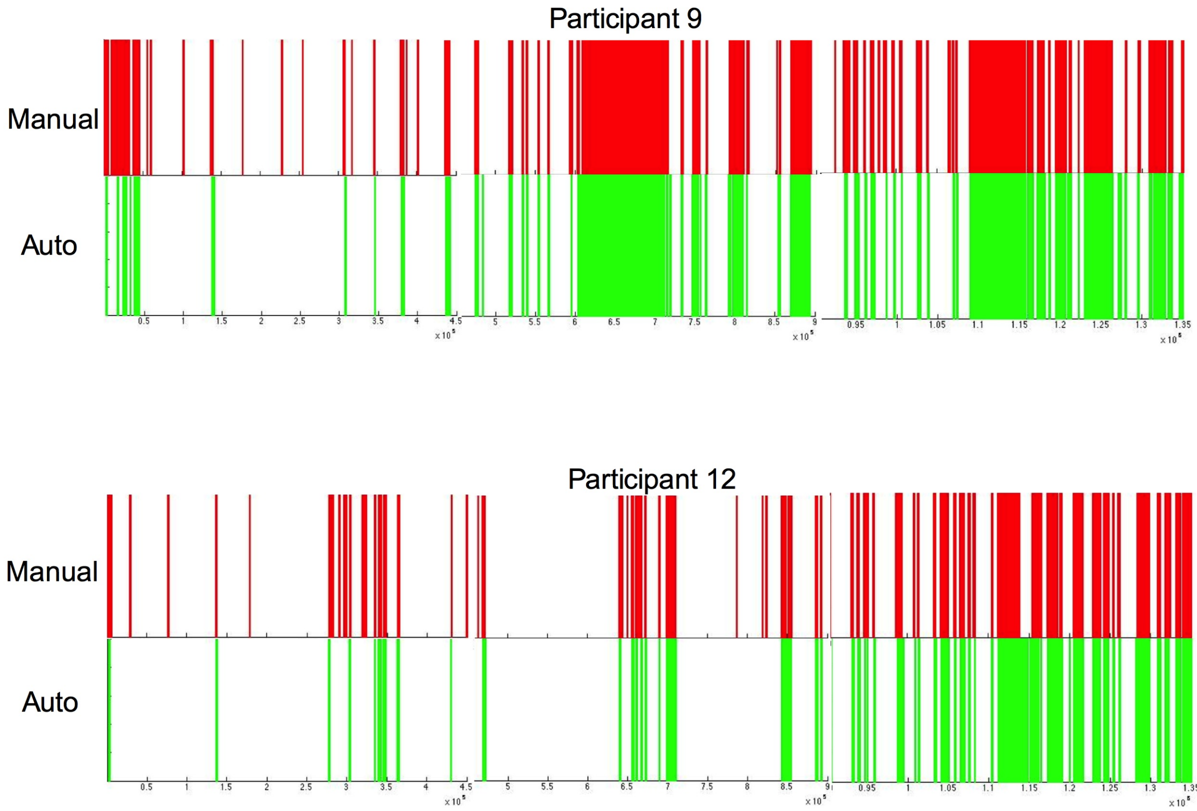


Figure 5-11 EEG data removed using manual (red) and automatic cleaning (green) in two random participants

5.2.7 Statistical analysis

Each driving task lasted 45 minutes. Each 45 minutes driving task was divided in 5 minutes segments (9 equal segments) for analysis. A Lunch condition (Lunch vs. No Lunch) x Gender (Male vs. Female) x Time Segment (9 time segments) repeated measures ANOVA was performed for the subjective variables and the driving variables. For the EEG variables a Lunch condition (Lunch vs. No Lunch) x Gender (Male vs. Female) x Time Segment (9 time segments) x Head Blocks (9 head blocks) repeated measures ANOVA was performed.

5.2.7.1 Subjective sleepiness results

There was no significant effect of time in the subjective sleepiness of the participants before positioning the EEG and after positioning the EEG ($F(1, 7)=0$, $p=1.00$, $\eta^2=0$), i.e. participants did not become more awake or more sleepy while the EEG was being positioned on their heads. From this step onwards, only the KSS scores obtained before the driving task and after the driving task were used. There was no effect of lunch in the subjective sleepiness ratings of the participants (F(1, 7)=0.429, p=.553, η^2=.058). Neither did gender have an effect on the subjective sleepiness of the participants (F(1, 7)=0.628, p=.454, η^2=.082). Nevertheless, time did have a significant effect in KSS scoring before and after the driving task (F(1, 7)=79.398, p<.001, η^2=.919), i.e. the monotonous driving task made participants more sleepy according to their subjective assessment. The mean value for the KSS score before the driving task was 4.083, i.e. the participants were awake before the driving task. The mean value for the KSS score after the driving task was seven, i.e. the participant ended up in a sleepy state after the driving task. This is shown in Figure 5-12.

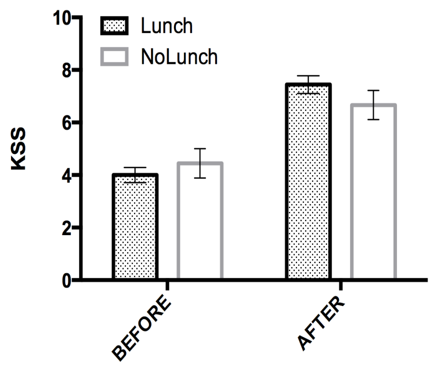


Figure 5-12 The scores of the KSS test before and after the driving task separated by the lunch condition

5.2.7.2 Driving behaviour results

The first hypothesis tested was the effect of lunch in the driving variables. There was no effect of lunch in SDLP (F(1, 7)=1.397, p=.276, η^2=.165), SDSpeed (F(1, 7)=1.136, p=.322, η^2=.140), TTLC (F(1, 7)=2.232, p=.179, η^2=.242) and OOL(F(1, 7)=0.234, p=.644, η^2=.032). In comparison, lunch did have an effect in SDSteering (F(1, 7)=7.133, p=.032, η^2=.505) and HFS (F(1, 7)=27.713, p=.001, η^2=.798).

The second hypothesis tested was the effect of a long monotonous task in the driving variables. The long monotonous driving task did have an effect of all of the driving variables: SDLP (F(8, 56)=5.131, p<.001, η^2 =.423), SDSpeed (F(7, 49)=4.644, $p<.001$, η^2=.399), SDSteering ($F(8, 56)$=2.332, p=.031, η^2=.250), HFS ($F(8, 56)$=5.913, $p<.001$, η^2=.458), TTLC ($F(8, 56)$=5.197, $p<.001$, η^2=.426) and OOL ($F(8, 56)$=5.279, $p<.001$, η^2=.430). This meant that driving on a monotonous road in a night environment for 45 minutes increased the sleepiness of the participants. The results are presented in Figure 5-13. The driving behaviour shown in Figure 5-13 went in line with the conclusions obtained in previous research found in literature. For SDSpeed, the behaviour of the drivers in this experiment went more in line with the results obtained by Bloomfield, Harder & Chihak (2009) where an increase in speed was found.

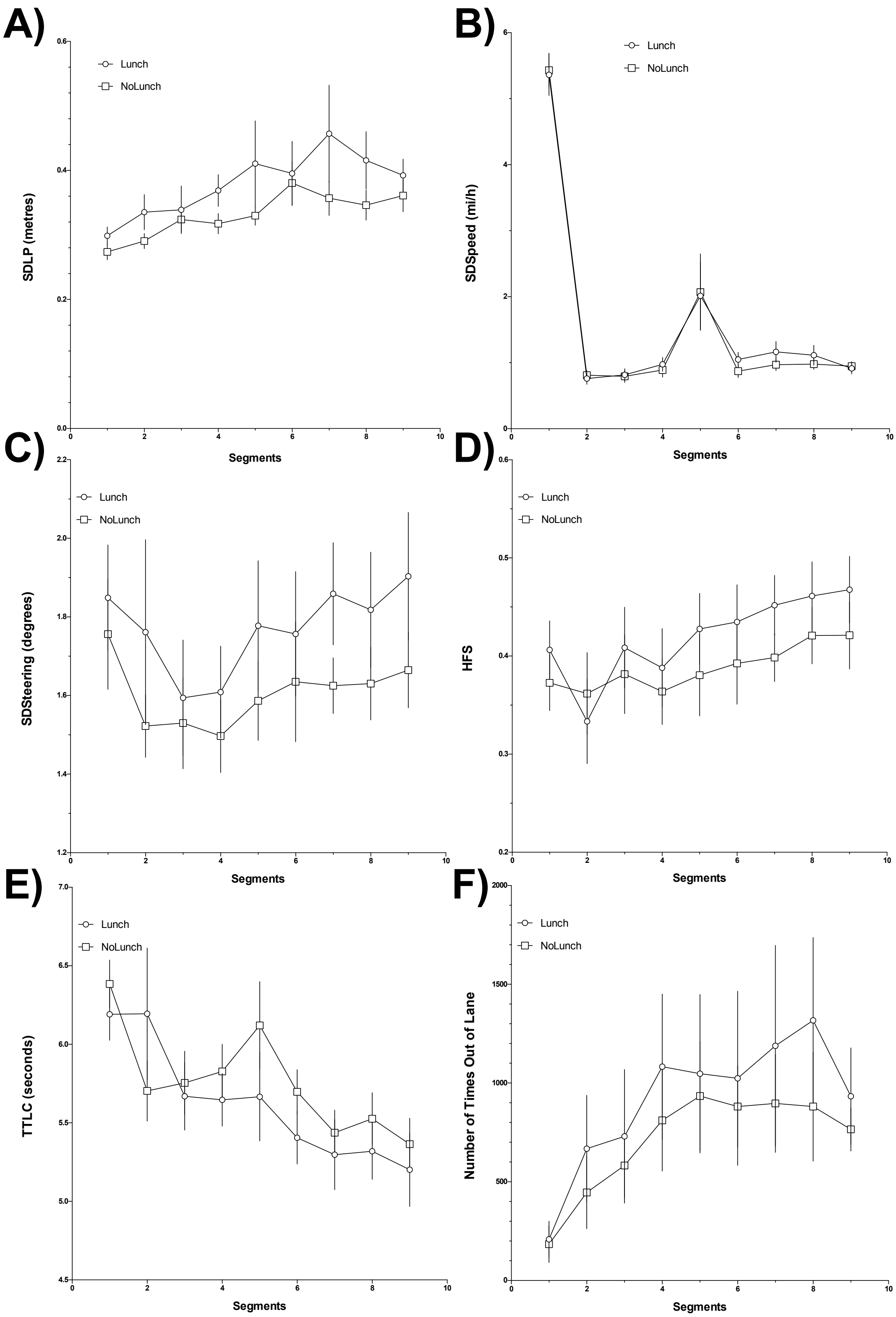


Figure 5-13 Driving variables behaviour separated by lunch condition with error bars representing the standard error. The x-axis shows the nine segments of time (each segment represents 5 minutes of the driving task).

The last hypothesis to be tested is the effect of gender (between subjects analysis) on the driving variables. Gender did not have any effect in any of the driving variables, i.e. female and male participants had a similar driving behaviour in the driving task. For all driving variables the results of the statistical analysis are the following: SDLP (F(1, 7)=0.055, p=.821, $\boldsymbol{\eta}^2$=.008), SDSpeed (F(1, 7)=1.368, p=.280, $\boldsymbol{\eta}^2$=.164), SDSteering (F(1, 7)=0.559, p=.479, $\boldsymbol{\eta}^2$=.074), HFS (F(1, 7)=0.627, p=.455, $\boldsymbol{\eta}^2$=.082), TTLC (F(1, 7)=1.451, p=.268, $\boldsymbol{\eta}^2$=.172) and OOL (F(1, 7)=1.320, p=.288, $\boldsymbol{\eta}^2$=.159).

5.2.7.3 EEG results

Lunch did not have any effect in theta (F(1, 7)=2.173, p=.184, $\boldsymbol{\eta}^2$=.237), alpha (F(1, 7)=2.776, p=.140, $\boldsymbol{\eta}^2$=.284), beta (F(1, 7)=1.991, p=.201, $\boldsymbol{\eta}^2$=.221), $\frac{\boldsymbol{theta+alpha}}{\boldsymbol{beta}}$ (F(1, 7)=0.057, p=.818, $\boldsymbol{\eta}^2$=.008) and $\frac{\boldsymbol{alpha}}{\boldsymbol{beta}}$ (F(1, 7)=0 p=.992, $\boldsymbol{\eta}^2$=0). Gender did not have any effect neither in theta (F(1, 7)=2.654, p=.147, $\boldsymbol{\eta}^2$=.275), alpha (F(1, 7)=1.068, p=.336, $\boldsymbol{\eta}^2$=.132), beta (F(1, 7)=1.251, p=.300, $\boldsymbol{\eta}^2$=.152), $\frac{\boldsymbol{theta+alpha}}{\boldsymbol{beta}}$ (F(1, 7)=0.123, p=.736, $\boldsymbol{\eta}^2$=.017) and $\frac{\boldsymbol{alpha}}{\boldsymbol{beta}}$ (F(1, 7)=0.034, p=.858, $\boldsymbol{\eta}^2$=.005).

The long monotonous driving task did have an effect in all the EEG variables: theta (F(8, 56)=10.624, p<.001, $\boldsymbol{\eta}^2$=.603), alpha (F(8, 56)=7.947, p<.001, $\boldsymbol{\eta}^2$=.532), beta (F(8, 56)=7.058, p<.001, η^2=.502), $\frac{\boldsymbol{theta+alpha}}{\boldsymbol{beta}}$ (F(8, 56)=5.130, p<.001, $\boldsymbol{\eta}^2$=.423) and $\frac{\boldsymbol{alpha}}{\boldsymbol{beta}}$ (F(8, 56)=4.101, p=.001, $\boldsymbol{\eta}^2$=.369). This means that there was a significant difference in the segments of time. It was also found that there was a significant difference between the different blocks of the head: theta (F(8, 56)=4.740, p<.001, $\boldsymbol{\eta}^2$=.404), alpha (F(8, 56)=6.503, p<.001, $\boldsymbol{\eta}^2$=.482), beta (F(8, 56)=6.214, p<.001, $\boldsymbol{\eta}^2$=.470), $\frac{\boldsymbol{theta+alpha}}{\boldsymbol{beta}}$ (F(8, 56)=3.182, p=.005, $\boldsymbol{\eta}^2$=.312) and $\frac{\boldsymbol{alpha}}{\boldsymbol{beta}}$ (F(8, 56)=2.295, p=.033, $\boldsymbol{\eta}^2$=.247). This is in line with other research where the changes in certain frequency bands are identified in particular electrodes positioned in specific areas in the scalp. Figure **5-14** presents the changes in each frequency band and in the EEG ratios over time.

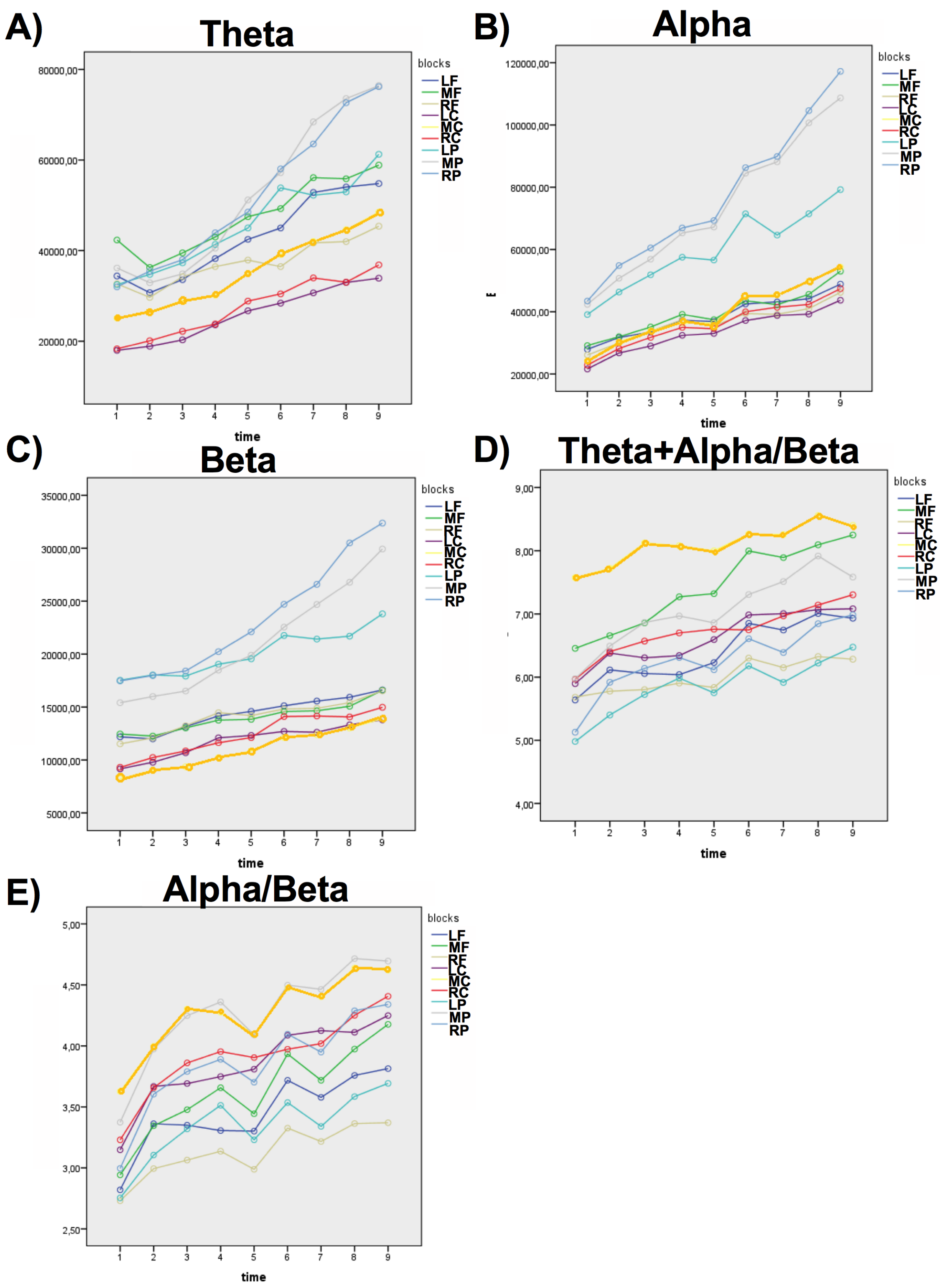


Figure 5-14 Changes in theta, alpha, beta, $\frac{\text{theta+alpha}}{\text{beta}}$ and $\frac{\text{alpha}}{\text{beta}}$ over time for the first study. Each graph represents one of the three frequency bands (A - Theta, B- Alpha and C – Beta) as well as the two EEG ratios (D - $\frac{\text{theta+alpha}}{\text{beta}}$ and E - $\frac{\text{alpha}}{\text{beta}}$) throughout the driving task. The driving task was divided in nine segments of 5 minutes (represented in the y-axis). Each graph contains nine lines, which represent the nine blocks of the head (Left Frontal, Middle Frontal, Right Frontal, Left Central, Middle Central, Right Central, Left Parietal, Middle Parietal and Right Parietal).

Correlation analyses were also performed to determine the relationship between the EEG variables and the driving variables. The number of correlation analyses done was 90 (combination of the 6 driving variables, the 3 frequency bands and 2 EEG ratios, and the 9 different blocks of the head). Examples of the correlation analysis is presented using heatmap tables (only a couple of representative examples are shown). As seen in Figure 5.14, SDLP and $\frac{alpha}{beta}$ in the middle parietal block of the head showed strong significant positive correlations during the initial part of the experiment (from segment 2 to segment 5). Figure 5.15 shows the same analysis for SDLP and $\frac{theta+alpha}{beta}$ in the middle parietal block of the head, where the correlations where not found significant. The same analysis was performed using driving and EEG variable, and no significant correlation was found. (Appendix M shows more examples of the different correlation analysis for other blocks of the head for SDLP and $\frac{alpha}{beta}$, different driving variable with alpha/beta for the middle parietal section and different frequency bands for SDLP and $\frac{alpha}{beta}$ in the middle parietal section, as a comparison with the figures presented here).

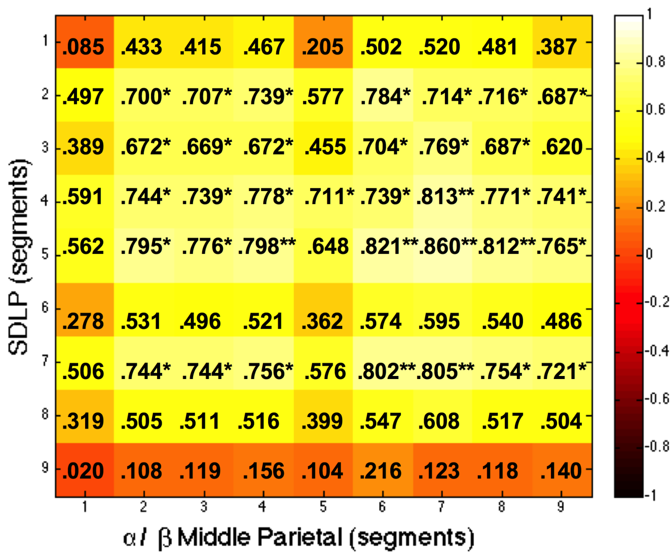


Figure 5-15 Heatmap presenting the correlation analysis between the nine time segments (each time segment represents 5 minutes of the driving task) of SDLP and $\frac{alpha}{beta}$ in the middle parietal region of the head.

As only SDLP and $\frac{alpha}{beta}$ in the middle parietal region of the scalp showed significant correlation in certain time segments it is not sufficient to determine a strong correlation between driving variables and EEG variables. The reason could be to a low number of participants and the need for more variables, specifically physiological behaviour variables, to predict and correlate sleepiness. It was concluded that a new set of experiments had to be conducted to obtain more data. The new experiment also recorded new variables such as eye movement and head movement behaviour of the participants.

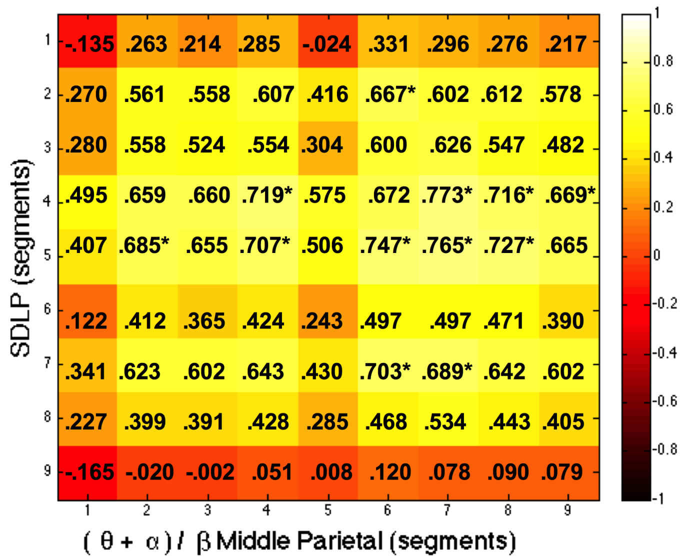


Figure 5-16 Heatmap presenting the correlation analysis between the nine time segments (each time segment represents 5 minutes of the driving task) of SDLP and $\frac{\text{theta+alpha}}{\text{beta}}$ in the middle parietal region of the head.

5.2.8 Study 1: Conclusion

In the present experiment, the main objective was to assess the viability of EEG as an indicator of sleepiness. It was reasoned that this would then provide a fundamental measure for use in follow-on experiments employing Machine Learning Algorithms to predict sleepiness, i.e. based upon changes in EEG signals. The study showed that the main frequency bands and ratios correlated to sleepiness change over time with increases in α *and* θ. This led to the conclusion that the experiment as a whole resulted in participants becoming sleepier over time, and this sleepiness was evident in the EEG signals.

The study also assessed the effect of lunch and gender on level of sleepiness. The results obtained from the present experiment showed that the 'lunch' condition (and time of the day) did not have an effect on the level of sleepiness of the participants, i.e. participants who had lunch behaved in a similar way to participants that did not have lunch. Similarly, gender did not appear to play any role in the level of sleepiness of the participants. From these results, it was concluded that for future experiment it is not necessary to provide lunch to the participants. For practicality purposes, it was identified that long haired individuals produced more noise than the EEG recording obtained from short-haired participants, so for future experiments in this domain was deemed appropriate to only recruit short-haired participants.

Although the EEG behaviour showed an increase in sleepiness, it was expected that the driving variables would be correlated to this increase of sleepiness as well. An interesting strong significant correlation was found between SDLP and $\frac{alpha}{beta}$ in the middle parietal section of the head. This is in line with research in literature were they found that an increase in sleepiness is related to an increase in SDLP (Lowden et al., 2009) and an increase in $\frac{alpha}{beta}$ (Eoh, Chung & Kim, 2005). Unfortunately, due to the high number of correlation analyses done (around 90 analyses) and the fact that the correlation is only found in few time segments during the driving task, it is possible that the results obtained for this correlation results could be a type 1 error. Additionally, the rest of the correlation analyses did not show any strong significant correlation leading stronger to a general impression that the results between SDLP and $\frac{alpha}{beta}$ in the middle parietal section is a type 1 error and there is no correlation between driving and EEG variables. This could have been due to lack of data; therefore, more recorded variables are needed. A new set of experiments was conducted to obtain more data. Specifically, the aim was to record more physiological variables to help predict the level of sleepiness with better accuracy. In addition, the length of the driving task will be extended to assure high levels of sleepiness.

5.3 Study 2: Effects of a one-hour monotonous driving task on drivers' sleepiness

5.3.1 Aims

In the previous section, it was found that lunch and gender did not have an effect in the subjective, driving and EEG variables. In comparison, the long monotonous driving task in the night environment did have an effect in the subjective, driving and EEG variables. This means that participants became sleepier after the driving task. Unfortunately, no correlations were found between the driving and the EEG variables. One hypothesis is that the data obtained is not sufficient and more data are needed. Adding to this hypothesis, the number of physiological features might not be sufficient to predict sleepiness. A second experiment was run focused only in inducing sleepiness by designing a long monotonous driving task in the participants. As lunch did not have any effect in sleepiness, participants were not provided with lunch during any of the driving tasks.

As presented in the analysis of the previous experiment, it was also found that gender did not have an effect in any of the variables, i.e. no difference was found between male and female participants. As the task of positioning the EEG is easier on participants with shorter hair, for this experiment only participants with short hair were targeted for recruitment (male participants with long hair were also excluded). In the previous experiment, seven of the female participants had to be excluded due to bad EEG that was not possible to analyse. On the other hand, only two of the male participants had to be excluded due to the same reason.

One of the aims of the second set of experiments was to explore the effects of undertaking the same driving task described in section 5.2 with a modification in the length of time (the driving task lasted 15 minutes more). Another aim of this study was to record physiological variables (blinking behaviour, gaze behaviour and head movement behaviour) not recorded during the first experiment. It has been found in literature that physiological variables are indicators of changes in sleepiness; therefore, obtaining information of the physiological behaviour of the driver might lead to a better prediction of sleepiness.

5.3.2 Method

5.3.2.1 Participants

The screening process was the same as the screening process for the experiment in the first experiment:

- Participants under the age of 30
- At least 2 years of driving experience
- BMI under 30 $\frac{kg}{m^2}$
- ESS score equal to or under 10

Participants in this experiment were also instructed to avoid alcohol for 24 hours before the experiment and caffeine should be avoided on the day of the experiment and they were asked to maintain a normal (7-8 hours) sleep pattern for three days previous to the experiment days. For this experiment, 15 male participants were recruited aged between 21 and 30 years old (M=24.93, SD=3.1). The BMI of the participants ranged between 17.51 and 26.3 $\frac{kg}{m^2}$ (M=22.77, SD=2.17) and their ESS score ranged between one and nine (M=4.93, SD=2.23). Participants 4, 6, 8, 10, 14 and 15 had to be excluded from the analysis due to unknown source of noise in their EEG data. Participant 3 was considered an outlier when inspecting the participant's driving behaviour and participant 13 was removed due to an ESS score higher than 10. The data of the remaining seven participants was analysed.

5.3.2.2 Design

The experiment follows the same design as the first experiment, except that participants were not provided with lunch and only took part in one driving task, i.e. they only came one day to the driving simulator. Participants were given the choice of two arrivals time to the driving simulator: 9:00 in the morning or 12:00 in the afternoon. Participants undertook the driving experiment in the static driving simulator in the Physics Research Deck of the University of Leeds explained in the previous section. At their arrival, participants were asked to fill the KSS test and their level of stress using the Perceived Stress Scale and the Stress and Arousal Checklist (Appendix J and I). Following that, participants had a 5 to 10 minute practice drive so they could be acquainted to the driving environment. After the practice run, the EEG was position in the head of the participant while they sat down. The process of positioning the EEG on the participant took around 20 to 30 minutes. The design presented here followed the same design further explained in the first experiment.

As previously discussed, one of the aims of this experiment was to record physiological variables which have been found to be indicators of changes in sleepiness. Blinking behaviour is one of the most reliable indicators of sleepiness (Yang et al., 2010; Bergasa et al., 2006; Wierwille et al., 1994; Dinges et al., 1998). It has also been found in literature, that certain changes in body and head position are related to high levels of sleepiness (Hartley et al., 2000; Haworth & Vulcan, 1991; Kaplan et al., 2007; Lal & Craig, 2002). Therefore, an eye tracker device and the head movement sensor were included in the design of this experiment.

The eye tracker device used was the SmartEyePro system (Smart Eye AB, Gothenburg, Sweden). It consisted of two infrared (IR) cameras with 8 mm lenses with a recording sample rate of 60Hz. Figure 5-17 shows the SmartEyePro system. The tracking accuracy of each camera is 0.5 degrees. Each camera has attached a matrix of IR LED's. The IR LED's are used to illuminate the face of the participants and reduce the effect of different lighting conditions. The cameras were positioned 30 centimetres apart from each other. The viewing angle of the cameras was fixed for every participant. A third non-IR camera was attached to the ceiling of the experiment room positioned towards the screen showing the driving environment. This third camera was used to correlate the gaze position of the participant with a point in the screen. The IR cameras, the IR LED's matrices and the non-IR camera were attached to a computer running the SmartEye recording software. The software recorded eye position, eye gaze, pupil diameter, saccades, fixations, blinks and eyelid opening. The eye tracker device was calibrated for each participant. The calibration process required the researcher to present a chessboard in the line of view of the cameras. Once the cameras recognised the chessboard, the focus had to be adjusted according to each participant.

Figure 5-17 SmartEyePro camera and their position in the experiment room (Source: Rogue-Resolution, 2016). The SmartEyePro system consisted on two infrared cameras oriented towards the participant's face. This infrared cameras are able to detect the pupil in dark environments.

After calibrating the eye tracker device, the head movement tracker was positioned on top of the EEG and hold in place with a net cap, as shown in Figure 5-18. The head movement tracker used was the MTx kit (Xsens, AN Enschede, The Netherlands). The MTx is a 3 degrees of freedom inertial orientation tracker. It recorded acceleration and rate of turn in three dimensions. The head movement tracker recorded at a sample frequency of 100 Hz. The MTx device had a roll/pitch accuracy of 0.5 degrees and a heading accuracy of 1 degree. The head tracker device was connected to a computer outside the experiment room through a USB cable. The head movement could be traced in real time from the computer outside the experiment room.

Figure 5-18 MTx device, used to track the head movement, positioned on top of the EEG. The MTx device was then hold using a second net positioned on top of the EEG net. The MTx can recorded the movements in x, y and z axis using accelerometers.

After positioning and calibrating all the devices, participants started the driving task, which took 60 minutes. The scenario was a night environment in a two-lane motorway with no traffic and few gentle curves. The environment inside the driving simulator room was set up to a dark environment by covering the windows to impair light to enter the room and turning off the lights during the task. The participants were instructed to maintain the same speed (40 miles per hour) and maintain the same lane throughout the experiment. During the driving task, no highway sign was shown to the participant. The participants were monitored from an adjacent room using a video camera so there was no interaction between the participant and the researcher during the task. At the end of the driving session, participants were asked again to assess their level of sleepiness using the KSS test (Appendix D). The design of this experiment followed the design of the first experiment. The subjective, driving and EEG variables recorded were the same variables recorded during the first experiment.

5.3.3 Statistical analysis

5.3.3.1 Subjective sleepiness results

Following the design of the experiment in chapter 5, KSS was measured for each participant before and after the driving task. The results of the repeated measures test are shown in

Figure 5-19. The results showed that there is a significant difference ($F(1, 6)=153.60$, $p<.001$, $\eta^2=.962$) between KSS before the driving test and after the driving test. This meant that the driving task had an effect in the subjective sleepiness of the participants. The mean score of KSS before the driving task was 4.857 (a score of 5 or less represents an awake state) and the mean score of KSS after the driving task was 7.143 (a score of 6 or above represents a sleepy state). This means that the participants were awake before the driving task and sleepy after driving 60 minutes in a monotonous road.

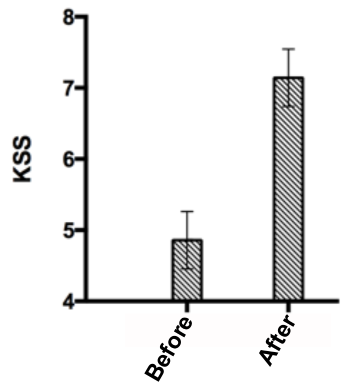


Figure 5-19 The scores of the KSS (subjective sleepiness) test before and after the driving task

The statistical results obtained in the EEG variables and the driving variables were consistent with the results obtained in the previous experiment, i.e. increase in SDLP, TTLC and HFS (as presented in Table 5-1). As presented in Figure 5-19, time had a similar effect in the subjective sleepiness of the participants (similar difference in KSS score between the KSS value before and after the experiment and similar KSS mean value at the end of the driving task). Figure **5-20**, and Figure **5-21** show a similar effect of time in the driving and EEG behaviour of participants, although the increase in SDLP, SDSteering and HFS as well as the decrease in TTLC seems less steep that in the first study. SDSpeed shows the same behaviour as in the first study. In similar way as the driving variables, the EEG behaviour does not increase in the same rate as in the previous study. It was concluded that the low steepness in the driving and EEG variables was due to the low amount of participants (only data of 7 participants was possible to use for the analysis. Nevertheless, results obtained in the second study still showed a significant increase in the levels of sleepiness as the driving task went on.

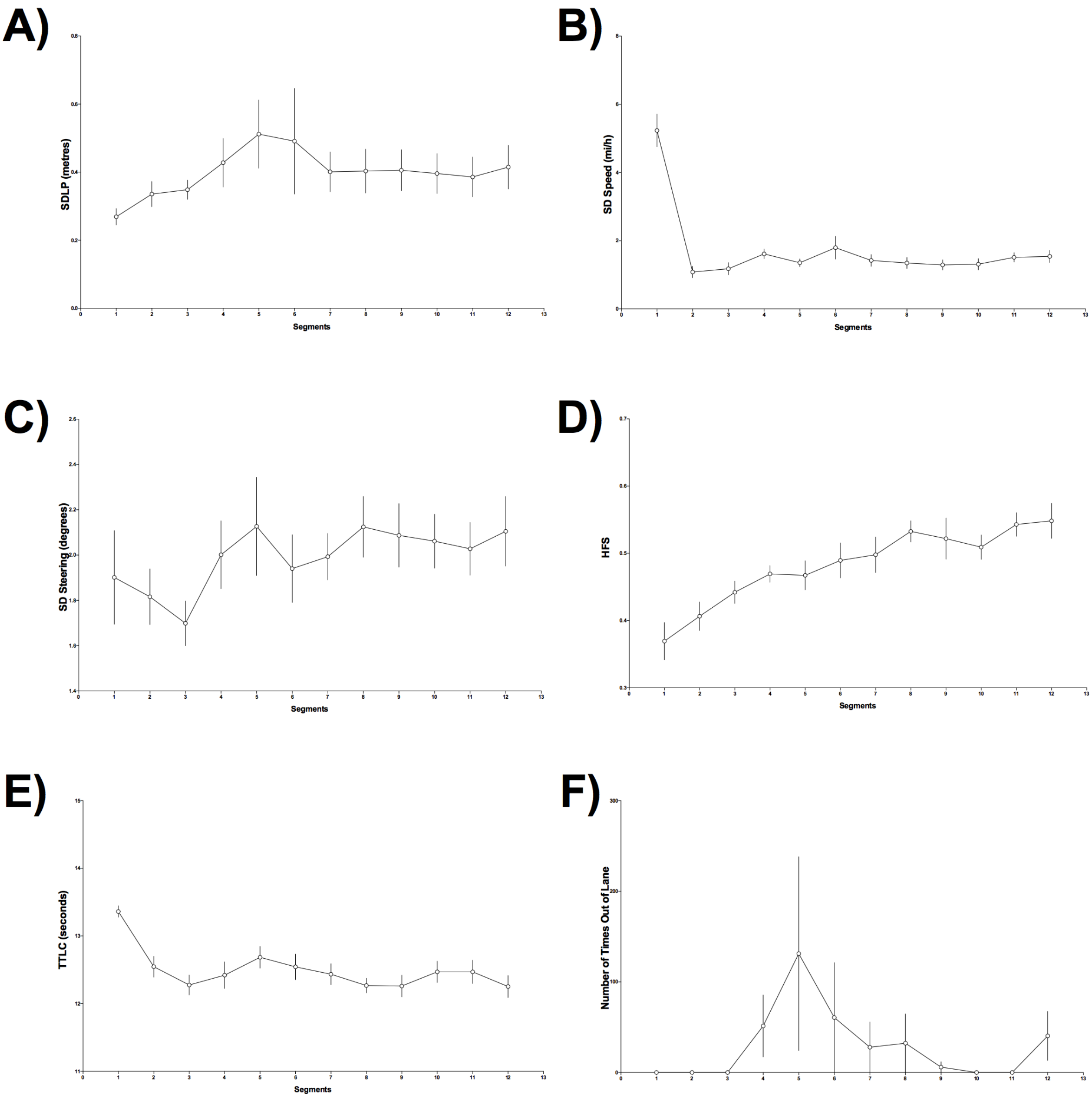


Figure 5-20 Driving variables behaviour with error bars representing the standard error. The x-axis shows the twelve segments of time (each segment represents 5 minutes of the driving task).

As in the first study, correlation analysis was done between the driving variables and the EEG variables (90 correlation analyses were performed). Compared to the first study, SDLP and $\frac{\text{alpha}}{\text{beta}}$ did not show any strong significant correlation (as shown in Figure **5-22**). The rest of the correlation analyses did not show any significant correlation.

Table 5-1 Statistical analysis of the effect of drive time in the driving variables

Measure	Drive Time
SDLP	F(11,66)=2.199, P=.025, $\boldsymbol{\eta}^2$=.268
HFS	F(11,66)=8.592, P<.001, $\boldsymbol{\eta}^2$=.589
TTLC	F(11,66)=3.711, P<.001, $\boldsymbol{\eta}^2$=.382

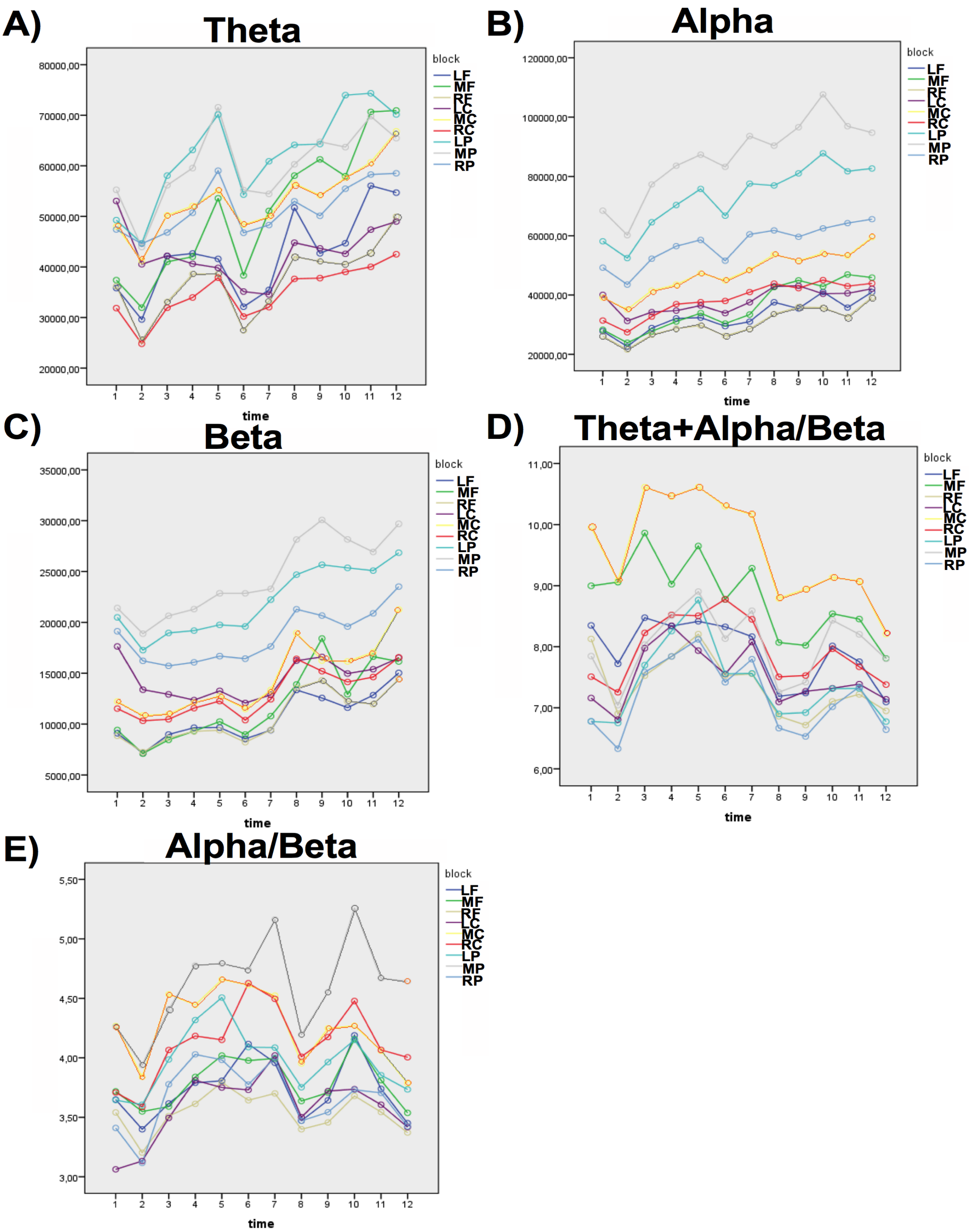


Figure 5-21 Changes in theta, alpha, beta, $\frac{\textbf{theta+alpha}}{\textbf{beta}}$ and $\frac{\textbf{alpha}}{\textbf{beta}}$ over time for the second study. Each graph represents one of the three frequency bands (A - Theta, B- Alpha

and C – Beta) as well as the two EEG ratios (D - $\frac{\textbf{theta+alpha}}{\textbf{beta}}$ and E - $\frac{\textbf{alpha}}{\textbf{beta}}$) throughout the driving task. The driving task was divided in twleve segments of 5 minutes (represented in the y-axis). Each graph contains nine lines, which represent the nine blocks of the head (Left Frontal, Middle Frontal, Right Frontal, Left Central, Middle Central, Right Central, Left Parietal, Middle Parietal and Right Parietal).

5.3.3.2 Physiological behaviour results

The two physiological variables recorded were the participants' blinking behaviour and head movement. Unfortunately, because the system used to record eye and blinking behaviour was not tested before by the department, there was little knowledge regarding the functionality and drawbacks of the system. It was found that the eye and blinking data contain too much noise, therefore could not be accounted during the analysis. Comparatively to the results obtain in the first experiment no correlation were found between driving and EEG variables.

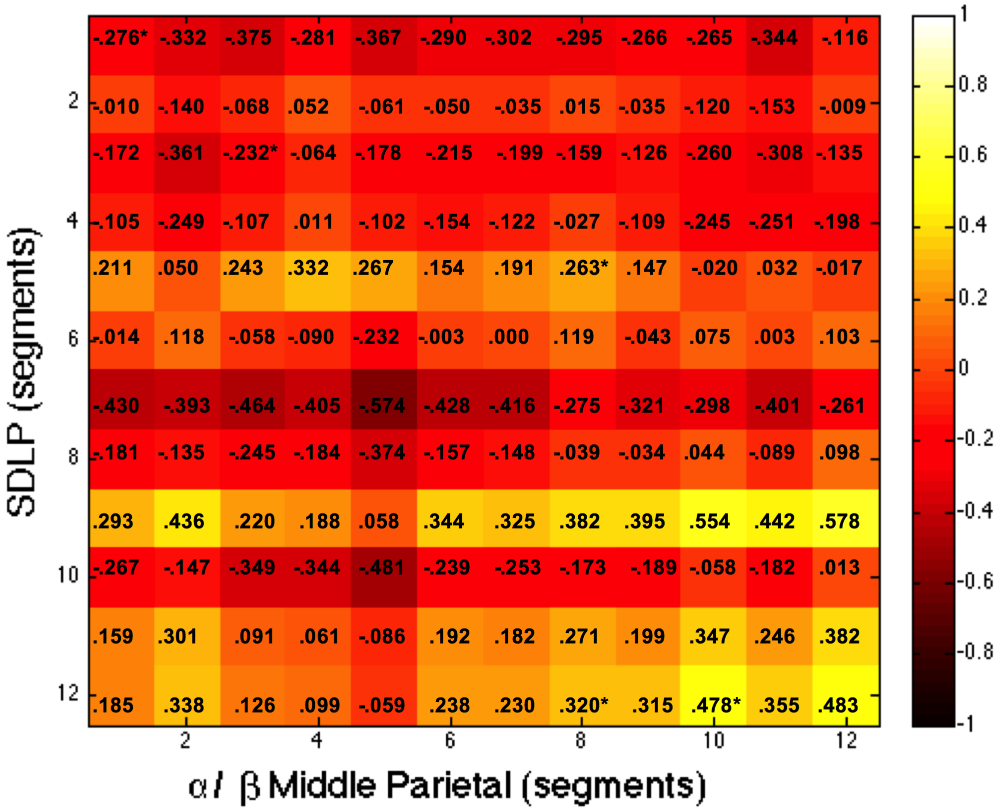


Figure 5-22 Heatmap presenting the correlation analysis between the twelve time segments (each time segment represents 5 minutes of the driving task) of SDLP and $\frac{\textbf{alpha}}{\textbf{beta}}$ in the middle parietal region of the head for the results obtained in study 2.

The second physiological variable recorded was head movement behaviour. The data recorded with the accelerometer was correctly recorded so it was possible to analyse. The main hypothesis to test was to determine the effect of time, i.e. a long monotonous road, had on the head movement behaviour of the participants. The variables recorded were the x, y and z position of the head of the participants. Figure 5-23 presents the axis coordinates direction according to the head of the participant.

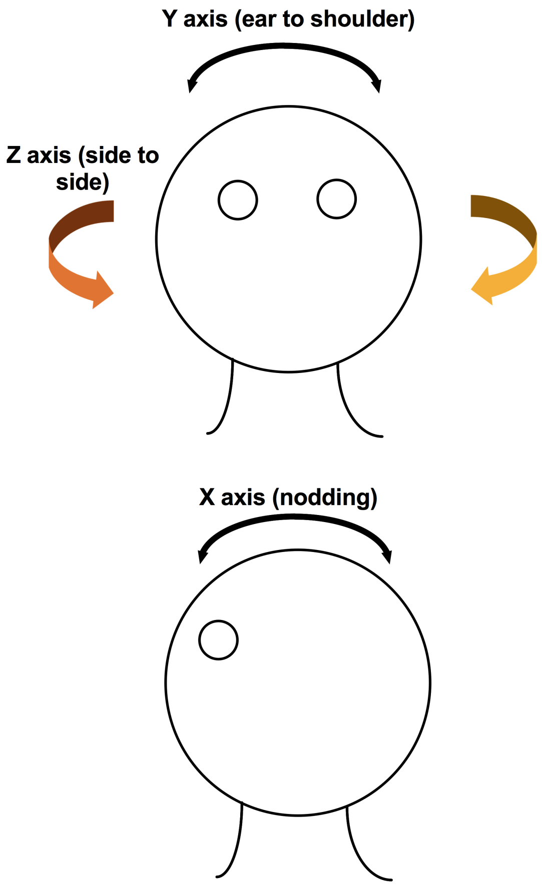


Figure 5-23 Axis coordinates according to head position of the participants

The long monotonous road only had an effect (as presented in Table 5-2) in the standard deviation of the head position in the y axis (ear to shoulder; $F(11, 66)=5.267$, $p<.001$, $\eta^2=.467$) and in the standard deviation of the head position in the z axis (side to side; $F(11, 66)=2.115$, $p=.031$, $\eta^2=.261$): Time did not have an effect in the x axis movement of the head (nodding). The same correlation analysis was done with the head movement and EEG variables. There was no correlation found between the physiological variables and EEG. Figure **5-24** present a heatmap of the correlation analysis between the SD of head movement in y and $\frac{alpha}{beta}$ for the middle parietal section of the head.

Table 5-2 Statistical analysis of the effect of drive time in the head movement variables

Measure	Drive Time
X position of the head (nodding)	F(11,66)=0.517, P=.885, η^2=.079
Y position of the head (ear to shoulder)	F(11,66)=1.001, P=.456, η^2=.143
Z position of the head (side to side)	F(11,66)=0.951, P=.499, η^2=.137
SD in X position of head (nodding)	F(11,66)=1.305, P=.242, η^2=.179
SD in Y position of head (ear to shoulder) *	F(11,66)=5.267, P<.001, η^2=.467
SD in Z position of head (side to side) *	F(11,66)=2.115, P=.031, η^2=.261
Number of nods	F(11,66)=1.409, P=.190, η^2=.190
Inter-nodding time (time between nods)	F(11,66)=1.000, P=.467, η^2=.250

5.3.4 Study 2: Conclusion

The present experiment was aimed to determine the effect of time in physiological variables and if there existed any correlation between physiological variables and EEG variables. Unfortunately, one of the physiological variables could not be used due to the amount of noise contained in the data. The second physiological variable, head movement, could be analysed. The long monotonous driving task had an effect only in two of the three axis of coordinates and no correlations were found between the head movement and the EEG variables. The second aim of the experiment was to reduce the variability of the EEG data by increasing the length of the dataset. No correlations were found between driving and EEG variables for this study. Comparing to the results in study 1, no correlation was found in study 2 between SDLP and $\frac{\text{alpha}}{\text{beta}}$ in the middle parietal section, which strengthens the impression that the correlation found in study 1 should not be considered statistically significant.

One of the hypothesis is that no correlations were found between the driving and EEG variables because the detection of sleepiness is dependant of many driving and physiological variables and is not dependant of any individual variable. A MLA can include the information of all the driving and physiological variables to determine the levels of sleepiness. Another necessary step in achieving this goal is to understand the most appropriate means of quantifying the EEG data to determine the levels of sleepiness. In the following chapter, different EEG variables were analysed to determine the most suitable to define the different levels of sleepiness. Following this, a Neural Networks MLA, using only the driving and physiological variable was trained to be able to determine different levels of sleepiness in drivers.

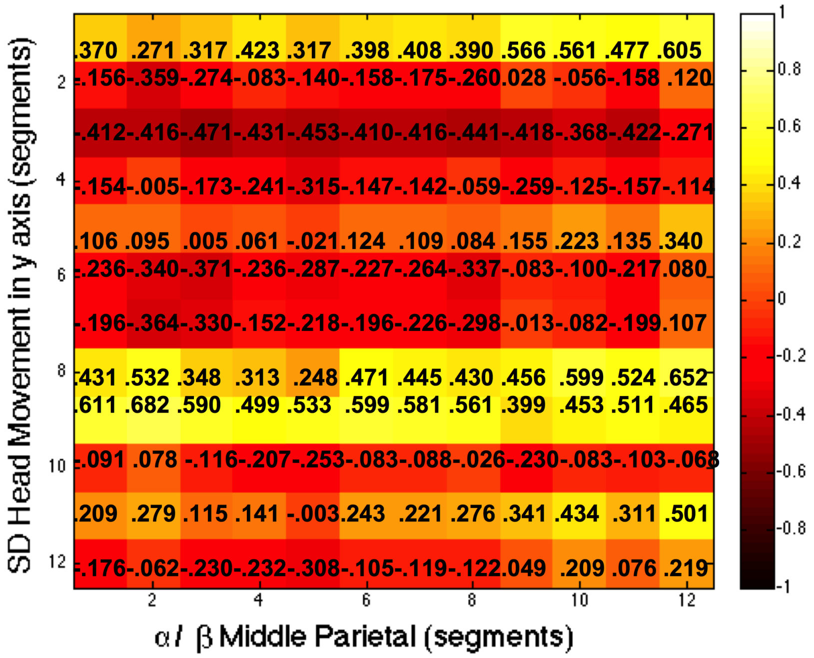


Figure 5-24 Heatmap presenting the correlation analysis between the 9 time segments (each time segment represents 5 minutes of the driving task) of the SD of the head movement in y axis and $\frac{alpha}{beta}$ in the middle parietal region of the head.

Chapter 6:

Identifying markers of sleepiness using EEG

6. Identifying markers of sleepiness using EEG

As discussed in previous chapters, sleepiness is one of the major contributors to driving accidents and driving-related fatalities. One potential approach for addressing this issue is to provide in-vehicle mechanisms that can predict sleepiness and alert the driver before an opportunity for collision arises. In the present chapter, EEG data were used to detect multiple level of sleepiness of drivers. The desired outcome was to distinguish different levels of sleepiness based on the driving and physiological data of the driver. Specifically, Machine Learning Algorithms (MLAs; discussed in chapter 3) were used to predict the different levels of sleepiness using behavioural data. The previous chapter presented the design of the experiment and the driving and physiological variables obtained from participants undertaking a driving task in a monotonous road. The brain activity during the driving task was also recorded using electroencephalogram (EEG). As EEG is the gold-standard measure of fatigue (Artaud et al., 1994; Lal et al., 2003; Jap et al., 2009), the EEG data were used as "ground truth" for sleepiness. The analysis of the different levels of sleepiness and the results obtained from the MLAs are presented in this chapter.

6.1 EEG variables used to determine different levels of sleepiness

Before being able to separate the EEG data into different levels of sleepiness, it was necessary to understand the components of the EEG signal that other researchers have used to identify sleepiness. Researchers have attempted to quantify different degrees of sleepiness using different methods and measures. For example, Lal et al. (2003) and Yang et al. (2010) used the changes in the magnitudes in each frequency band to determine different levels of sleepiness. In this approach, the magnitude was calculated as the sum of the values within a particular frequency band (Lal et al., 2003). Yeo et al. (2009) proposed a different set of EEG features to determine the different levels of sleepiness: dominant frequency (the frequency with the highest power within the particular frequency band); average power of dominant peak (the average power within the full width half maximum band of the dominant frequency); centre of gravity frequency (CGF; contrary to the dominant frequency, the CGF would lay down in the mean of the power spectral density); and frequency variability within the bands.

Other researchers have employed the mean power of each particular frequency band as a way of representing changes in sleepiness over time (Campagne, Pebayle & Muzet, 2004; Otmani et al., 2005; Eoh, Chung & Kim, 2005; Jap et al., 2009; Filtness et al., 2012). In some cases the combination of slow and fast brain waves analysed using EEG ratios have been used (Campagne, Pebayle & Muzet, 2004; Otmani et al., 2005; Eoh, Chung & Kim, 2005; Jap et al., 2009). Several other researchers have demonstrated that changes in alpha and theta bursts correlate with high levels of sleepiness (Knoblauch et al., 2003; Andrillon et al., 2011; Benbadis, 2006; Carskadon & Dement, 2011; Gennaro & Ferrara, 2003; Rivera & Salas, 2013; Cantero et al., 2002; Parekh et al., 2015; Teplan, 2002). For this study, a two-pronged approach was adopted. First, using α, β and θ frequency bands, the two ratios $\frac{\theta+\alpha}{\beta}$ and $\frac{\alpha}{\beta}$ were calculated to capture sleepiness. This approach was adopted as it allows combining fast waves (related to increase in sleep) and slow waves (related to an increase in wakefulness), providing relative measures of changes in sleepiness over time.

Secondly, an adopted Karolinska Drowsiness Scale was used as it measures the instantaneous theta and alpha bursts, which are related to high levels of sleepiness (Shuyan & Gangtie, 2009; Rechtschaffen & Kales, 1968). As alpha and theta bursts account for instantaneous periods of sleep, it was considered a suitable measure of sleepiness. The KDS is a measure to determine the drowsiness in people in an active situation, i.e. situation when the participant is expected to be awake (Shuyan & Gangtie, 2009; Rechtschaffen & Kales, 1968). The KDS is based on the scoring method proposed by Rechtschaffen & Kales (1968). The data are separated into epochs of 20 seconds and each 20 seconds epoch is further divided into 2 seconds epochs (Shuyan & Gangtie, 2009; Rechtschaffen & Kales, 1968). The KDS used in the present study was adapted from Shuyan & Gangtie (2009). For each 2 seconds epoch, if a theta or alpha burst is detected, the epoch is assigned a value of one; otherwise, the epoch is assigned a value of zero. The percentage of epochs with a value of one in the 20 seconds epoch is the KDS value of the 20 seconds epoch. To be able to determine a burst in alpha and theta a "baseline" power value of alpha and theta need to be calculated (Knoblauch et al., 2003; Andrillon et al., 2011; Benbadis, 2006; Carskadon & Dement, 2011; Gennaro & Ferrara, 2003; Rivera & Salas, 2013;

Cantero et al., 2002; Parekh et al., 2015; Teplan, 2002). The baseline power value was obtained by calculating the mean power of alpha and theta during the first 5 minute of the drive. A burst was considered to happen when the mean power of alpha or theta during a specific epoch was twice or higher than the baseline power. By using the KDS and the frequency band ratios, it is possible to account for instantaneous changes in sleepiness and changes in sleepiness over time. After computing these EEG outcome measures (EEG ratios and KDS), the next step was to determine the different levels of sleepiness according to the EEG data.

6.2 Determining different levels of sleepiness using EEG and MLA

As presented in Chapter 3, there have been some successful previous attempts at predicting different levels of sleepiness using driving and physiological data (Yang et al., 2010; Shuyan & Gangtie, 2009; Yeo et al., 2009; Vuckovic, et al., 2002; Patel et al., 2011). It is worth noting the success of these approaches: Vuckovic et al. (2002) and Patel et al. (2011) obtained between 90% and 94% accuracy when predicting two sleepiness states (awake and sleep) using a Neural Networks algorithm. When adding a new sleepiness level to be predicted, i.e. three sleepiness levels (awake, drowsy and sleep), the accuracy some researchers have obtained varies between 90% and 99% using a SVM algorithm (Shuyan & Gangtie, 2009; Yeo et al., 2009). However, to the author's knowledge, there is no evidence in the literature of any successful attempts at predicting more than three levels of sleepiness (awake, drowsy and sleep) in drivers using MLAs. Thus, one of the aims of this PhD was to determine more than three levels of sleepiness. The implications and motivations to develop a method to detect more than three levels of sleepiness are discusses in detail in the discussion chapter.

Whilst the majority of previous researchers have used driving and physiological variables to determine different levels of sleepiness (Shuyan & Gangtie, 2009; Patel et al., 2011), in the present study EEG data were used to determine levels of sleepiness. As discussed in chapter 2, EEG data has been found to be more reliable and less prone to external noise when detecting sleepiness compared to other variables (Lal & Craig, 2002; Eoh, Chung & Kim, 2005; Jap et al., 2009). Researchers have achieved high accuracy when using EEG data to determine the different levels of sleepiness (Vuckovic et al., 2002; Yeo et al., 2009; Lal et al., 2003; Yang et al., 2010).

Typically, in previous research, different levels of sleepiness were decided by a set of expert EEG clinicians following visual inspection of the EEG data (Yeo et al., 2009; Vuckovic et al., 2002). In this method, clinicians search for visual indicators of awake and sleep (Knoblauch et al., 2003; Andrillon et al., 2011; Benbadis, 2006; Carskadon & Dement, 2011; Gennaro & Ferrara, 2003; Devuyst et al., 2010a,b; Rivera & Salsa, 2013; Cantero et al., 2002; Parekh et al., 2015; Teplan, 2002). The indicator of the awake state is an increase in the slow wave activity (beta activity) while indicators of "sleep" state are an increase in fast wave activity (alpha and theta activity), alpha and theta bursts and alpha spindles. Figure 6-1 show examples of the visual indicators used by clinicians to determine the levels of sleepiness. Whilst this approach has the benefit of expert knowledge and is a viable solution to determine the different levels of sleepiness, it is a time-consuming approach and would not be that feasible in an environment where large quantities of data are generated and need to be evaluated on a brief time-scale.

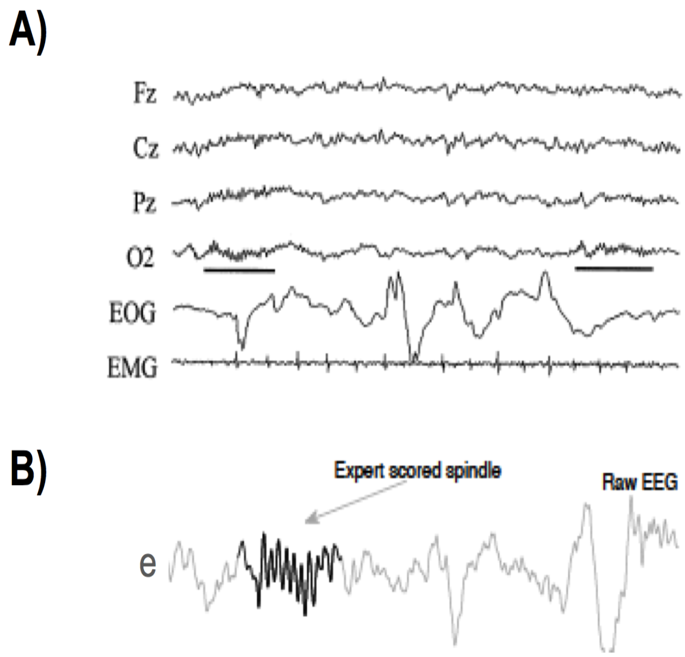


Figure 6-1 Visual indicators of sleep in EEG data used by clinicians to determine the different sleepiness states. Fast frequencies (a) as well as spindles (b) are indicators of high levels of sleepiness (Source: Cantero, Atienza & Salas, 2002). Reprinted from

“Human alpha oscillations in wakefulness, drowsiness period, and REM sleep: different electroencephalographic phenomena within the alpha band” by Jose L Cantero et al. Copyright © 2002 by Jose L Cantero et al. Used by permission of Elsevier.

Furthermore, the assessments of the EEG data are very subjective and vary from clinician to clinician. At least two clinicians are needed to assess the different levels of sleepiness (Yeo et al., 2009; Vuckovic et al., 2002). Studies using this method have found that a number of EEG epochs often need to be removed because the clinicians could not agree in the state of sleepiness. This means that when using clinicians to determine the different levels of sleepiness in EEG data, it is inevitable that a percentage of the data will be lost due to lack of consensus between the clinicians.

Another problem encountered is the lack of automation and repeatability of the EEG assessing process. Compared to a computer algorithm, when hiring clinicians each EEG dataset has to be assessed manually. If each dataset is long and the number of EEG datasets too assess is high, this could lead to a decrease in performance of the clinicians, e.g. as seen in repetitive tasks performed by humans (Thompson et al., 2006; Sheridan, Vámos, and Aida, 1983; Haga, 1984; Dunn &and Williamson, 2011; Oron-Gilad, Ronen, and Shinar, 2000; Weinger, 1999), as well as an increase in time assess all the EEG datasets.

Therefore, an aim of this PhD project was to attempt to determine the different levels of sleepiness through automatized, quantitative analysis of the EEG data. Specifically, a program was developed in Matlab, which used the values $\boldsymbol{\frac{\theta+\alpha}{\beta}}$ and $\frac{\alpha}{\beta}$ $and\ KDS\ of\ \theta\ and\ \alpha$ to determine the different states of sleepiness.

6.2.1 Defining binary clusters of sleepiness using EEG

The first step was to determine the two extreme levels of sleepiness, i.e. awake and sleep. Threshold values were determined for the awake and sleep state using the $\frac{\theta+\alpha}{\beta}$ and $\frac{\alpha}{\beta}$ $ratios\ and\ KDS\ of\ \theta\ and\ \alpha$. To determine the threshold value of the "awake" state, the mean values of the $\frac{\theta+\alpha}{\beta}$ and $\frac{\alpha}{\beta}$ $ratios\ and\ KDS\ of\ \theta\ and\ \alpha$ were obtained for the first 5 minutes of the driving task only for participants with a Karolinska Sleepiness Scale (KSS) score of 4 or less (subjective alert state) (Akerstedt & Gillberg, 1990) before they started the driving task. The length of the epoch (first five minutes of the drive) was derived from previous research using this epoch length to calculate baseline levels for alertness (Shuyan & Gangtie, 2009; Eoh et al., 2005).

The threshold for the sleep state was calculated by obtaining the mean values of the EEG variables before a "complete lane departure" happened, i.e. when the left tyre crossed the right lane boundary or when the right tyre crossed the left lane boundary. Shuyan & Gangtie (2009) used "lane departures" as an indicator of the highest level of sleepiness. They obtained high levels of accuracy when detecting "very sleepy", the state of sleepiness related to "lane departure" (Shuyan & Gangtie, 2009). Therefore, this approach was adopted in the present PhD study.

As the primary interest of the analysis was to identify EEG correlates that precede a "complete lane departure", the mean values of the $\frac{\theta+\alpha}{\beta}$ and $\frac{\alpha}{\beta}$ $and\ KDS\ of\ \theta\ and\ \alpha$ during the 20 seconds before the lane departure event happened were calculated. The reasoning behind the chosen length of the epoch (20 seconds) is related to the Karolinska Drowsiness Scale variable. Once the EEG values for the two extreme levels of sleepiness (awake and sleep) were established, it was then necessary to determine the position in the scalp of the participants where sleepiness has the highest effect.

As discussed in the design experiment conducted during the present PhD (chapter 5), the electrodes of the EEG net were divided into 9 clusters (front left, middle left, right left, central left, central middle, right middle, parietal left, parietal middle and parietal right) according to their position on the scalp of the participants (Oken & Chiappa, 1986). An analysis was needed to be done to determine in which section of occur the highest the changes in sleepiness. A comparison was done using the EEG ratio $\boldsymbol{\frac{\theta+\alpha}{\beta}}$ and $\boldsymbol{\frac{\alpha}{\beta}}$ (indicators of sleepiness) between the values obtained during the 20 seconds before a “complete lane departure” (considered to be highest sleep state) and 20 seconds epochs obtained during the first 5 minutes of the drive (considered to be the highest awake state). The highest changes were found in the middle parietal section, as shown in Figure 6-2 and Figure 6-3. Hence, the EEG values obtained in the middle parietal are the ones being used to determine the different levels of sleepiness in the following section. This is in agreement with the results found in literature. As presented in Chapter 2, the posterior and central sections of the head are the most sensitive to changes in EEG variables related to sleepiness (Cantero et al., 2002).

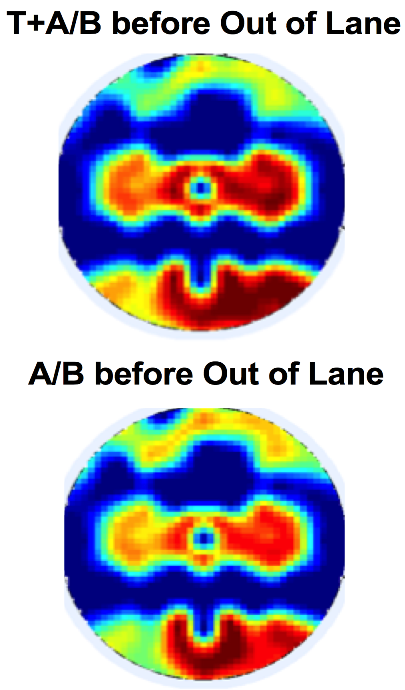
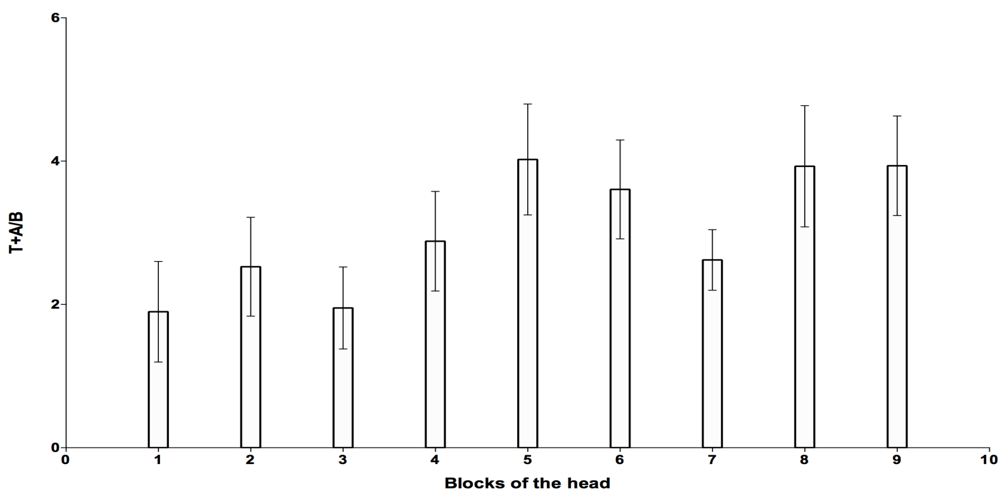


Figure 6-2 Topological plot of the mean power of the EEG ratios a) theta+alpha/beta and b) alpha/beta 20 seconds before the participant drove out of the lane.

Once the middle parietal was selected, an analysis was done to determine the variability of the different EEG variables in the awake and sleep conditions. It was found that in the “awake” state (first 5 minutes of driving for participants with KSS score of 4 or below), the EEG variable with the lowest variability was the EEG ratio $\frac{\alpha}{\beta}$ (M=3.5433, SD=0.9356). Therefore, the mean value was implemented as the maximum threshold for the “awake” state, i.e. every segment where the value of $\frac{\alpha}{\beta}$ was below the mean was clustered as “awake” data. For the “sleep” state (the 20 second segment before the participant went into a “complete out of lane”), it was found that the EEG variable with the lowest variability was KDS of $\boldsymbol{\alpha}$ (M=28.1944, SD=15.6822). The mean value was used as the threshold for the “sleep” state, i.e. every segment where the value of KDS of α was below the mean was clustered as “sleep” data. Every segment that had a $\frac{\alpha}{\beta}$ value higher than the “awake” threshold and had a KDS α value lower than the “sleep” threshold was cluster into the “neither” category.



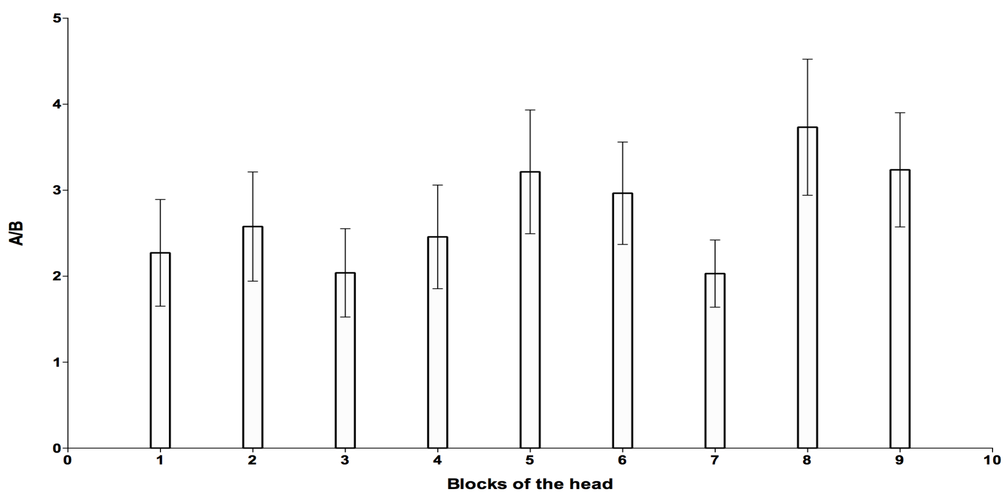


Figure 6-3 Mean power of the EEG ratios a) theta+alpha/beta and b) alpha/beta 20 seconds before the participant drove out of the lane per block of the head (1-3 are the Frontal Left, Middle and Right, respectively. 4-6 are the Central Left, Middle and Right, respectively. 7-9 are the Parietal Left, Middle and Right, respectively)

The data of each participant were collated and clustered (awake, sleep or neither) following the conditions presented previously explained. The data were divided in 80 “awake” segments, 51 “sleep” segments and 121 “neither” segments. Each segment contained a $\mathbb{R}^{\mathbf{14}}$vector, which consisted of the target set (state of sleepiness) and the feature set (temporal and driving variables). The temporal and driving variables contained in the feature set were: segment of time, standard deviation (SD) of lane position, lane position in reference from the centre of the lane, SD of lane position in reference from the centre of the lane, speed of the car, SD of speed of the car, SD of steering angle, high frequency steering, time to lane crossing, number of times the car touch the edge of the lane, number of “complete out of lane”, SD of acceleration and steering entropy. The length of each segment was 5 minutes, same length used by other researchers (Papadelis et al., 2007). The driving and temporal variables (features) were used to train and predict the cluster assigned (target). Unfortunately, there was a higher number of “awake” and “neither” segments compared to “sleep” segments, i.e. imbalanced dataset. An imbalanced dataset could result in learning problems for the MLA.

As presented in chapter 4, imbalanced data could lead the MLAs to favour the majority cluster, as more training data of the majority cluster would be available for the MLA to learn, leading to an increase in the false negative rate (Elhassan et al., 2016). Thus, before the clusters could be tested in the MLAs, it was necessary to solve the imbalance of the clusters. Therefore, the over-sampling method used in chapter 4 to artificially increase the examples of the under-sample clusters (Synthetic Minority Oversampling Technique [SMOTE]) was used (Elhassan et al., 2016; Chawla et al., 2002). After running the SMOTE method across all clusters, the data were divided into 240 “awake” segments, 255 “sleep” segments and 242 “neither” segments.

6.2.2 Predicting binary clusters of sleepiness using driving behaviour

The data obtained after running the SMOTE method was used to train and test the algorithm and hereafter any mention to the data points used refers to the over-sampled data of each cluster. As presented in Chapter 4, Neural Network algorithms obtained better accuracy than Support Vector Machine algorithms when using driving variables to detect levels of sleepiness; therefore, Neural Network algorithms were used in this analysis. A supervised Neural Network algorithm (NeuroNets) was trained to predict the state of sleepiness (awake, sleep or neither) through the driving variables. The NeuroNets design consists of a three layer feed-forward algorithm with seven neurones in the hidden layer and a learning rate of 0.1, as shown in Figure 6-4. The number of iterations that obtained the best accuracy was 2,000. Following the method used in chapter 4, the number of neurons in the hidden layer, the learning rate and the number of training iterations was defined heuristically, i.e. the number of neurons, the learning rate and the number of iterations were increased until the accuracy became stagnant.

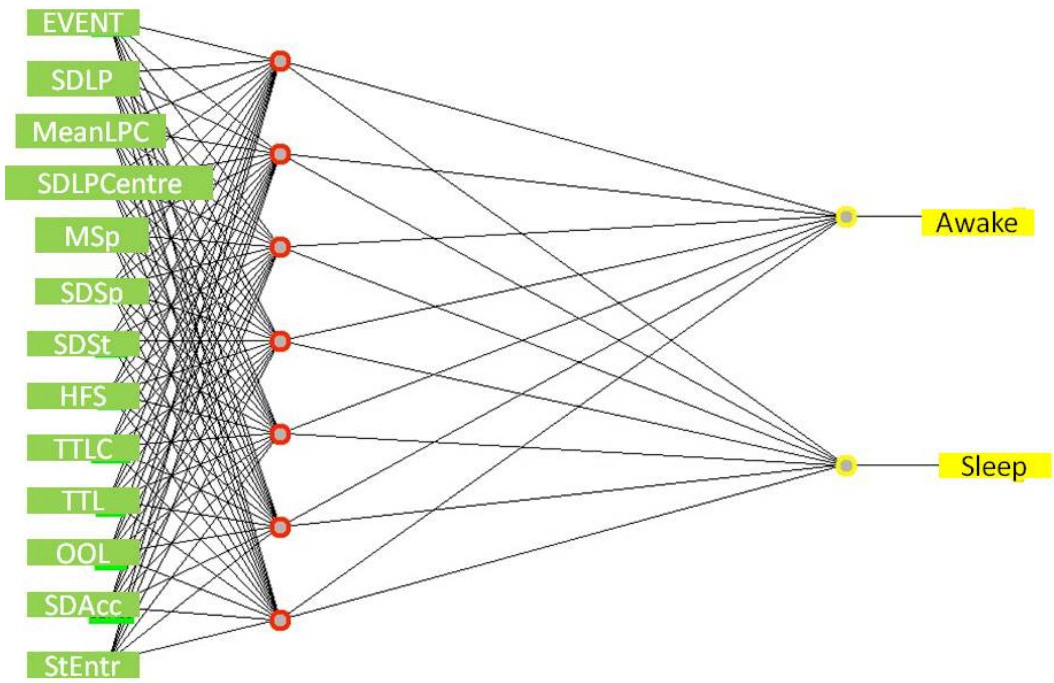


Figure 6-4 NeuroNets 3 layer feed-forward with 7 neurons in the hidden layers design used to predict a binary classification of sleepiness using driving variables as inputs

First, the NeuroNets was trained and tested only with “awake” and “sleep” data, i.e. the “neither” data were discarded. The goal of training and testing the algorithm only with “awake” and “sleep” data were to determine if the clustering thresholds used to separate the data into “awake” and “sleep” were suitable. A high level of prediction accuracy from the NeuroNets would back up the results from the clustering rules used to separate the different levels of sleepiness. As presented in chapter 3, accuracy was used to determine the suitability of the MLA and was calculated using the following formula:

$$\frac{\boldsymbol{True\ Positive + True\ Negative}}{\boldsymbol{True\ Positive + True\ Negative + False\ Positive + False\ Negative}}$$

Using the same criteria as in chapter 4, k-fold cross-validation was performed to evaluate the algorithm. For this dataset, the data of one participant was set aside for testing the algorithm and the data of the rest of the participants was used for training the algorithm. This was iterated for each participant. The accuracy results presented for the algorithms in this chapter are the mean value of the accuracy values obtained in the k-fold cross-validation process. The accuracy obtained when training and testing the NeuroNets with only "awake" and "sleep" data were 95.71% (SD=2.8%) and the error box is shown in Table 6-2. The NeuroNets was also trained with the three levels of sleep (awake, sleep and neither) to determine the effect of adding the "neither" cluster. The accuracy obtained was 85.75% (SD=3.31%) and the error box is shown in Table 6-3.

An ablative analysis was done to determine that all of the features i.e. the temporal and driving variable were contributing towards the accuracy of the overall system. An ablative analysis determines the relative weighting of each of the features in an algorithm, i.e. how much did each of this features affected the accuracy of the overall system (Ng, 2008). To perform an ablative analysis, each feature is removed one by one and it is determined how much the accuracy of the overall system drops once that feature is removed. If a feature is removed and the overall accuracy did not decrease, it means that the feature can be removed, as it does not affect the accuracy of the overall system. Table 6-1 presents the ablative analysis for the NeuroNets algorithm only trained with the "awake" and "sleep" data. As seen in Table 6-1, all features input in the NeuroNets affected the accuracy of the overall system, i.e. all these features were needed to obtain the overall accuracy. In addition, a sensitivity analysis was done to determine the effect of increasing the rate of over-sampling using SMOTE, i.e. how much the accuracy increases according to the amount of data available. Figure 6-5 shows the result of the sensitivity analysis. It starts with low accuracy when using the original imbalanced data. Once the data are over-sampled using SMOTE, the accuracy increases as soon as the number of data points double from the original set. The accuracy does not change once it reaches four times the number of original data points.

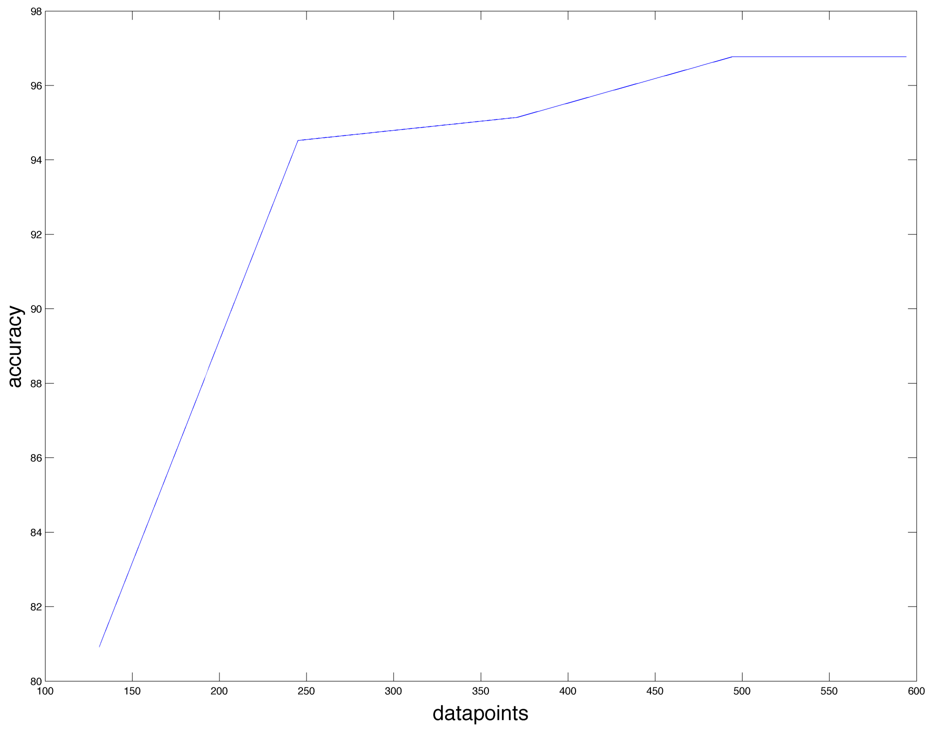


Figure 6-5 Sensitivity analysis showing the changes in accuracy as the number of data points increase. Once the algorithm crosses the 230 data points, the accuracy does not increase further with the increase of data points used.

Table 6-1 Ablative analysis to determine the effect each feature has in the accuracy of the overall system

Overall accuracy	95.71%
Segment of time	90.94%
SD lane position	89.44%
Mean lane position	89.28%
SD lane position from centre	87.49%
Mean Speed	82.73%
SD Speed	82.14%
SD Steering	74.99%

High Frequency Steering	72.01%
Time to Lane Crossing	60.71%
Times touching the lane	58.92%
Number of "out of lane"	48.21%
SD Acceleration	47.82%
Steering Entropy	45.23%

Table 6-2 Error box to determine the accuracy of the NeuroNets when predicting 2 levels of sleepiness

	Awake	Sleep
Awake	94.60%	5.39%
Sleep	3.08%	96.91%

Table 6-3 Error box to determine the accuracy of the NeuroNets when predicting 3 levels of sleepiness

	Awake	Neither	Sleep
Awake	83.21%	10.85%	5.94%
Neither	17.48%	77.10%	5.42%
Sleep	2.21%	1.73%	96.06%

6.2.3 Predicting multiple clusters of sleepiness using driving behaviour

Once the extreme sleepiness levels were determined, the following step was to determine if the “neither” cluster (in between “awake” and “sleep”) could be separated into different levels of sleepiness, i.e. to obtain a “post-awake” and a “pre-sleep” state of sleepiness. An unsupervised k-means clustering algorithm was used to cluster the “neither” data. This algorithm was discussed in Chapter 3. As the EEG ratio $\frac{\boldsymbol{\alpha}}{\boldsymbol{\beta}}$ and KDS of $\boldsymbol{\alpha}$ variables were used to determine the thresholds for the “awake” and the “sleep” states, these two variables were used to determine the clusters in the k-means clustering algorithm.

The k-means clustering algorithm represents the two variables as coordinates where $\frac{\boldsymbol{\alpha}}{\boldsymbol{\beta}}$ was in the x-axis and KDS of $\boldsymbol{\alpha}$ was in the y-axis. The algorithm started with random positions for each of the centroids: (7.9794, 25.333) for one centroid and (6.4708, 19) for the other centroid. The algorithm used Euclidean distance as the function to improve the position of the centroids. The result of the k-means clustering algorithm is presented in Figure 6-6.

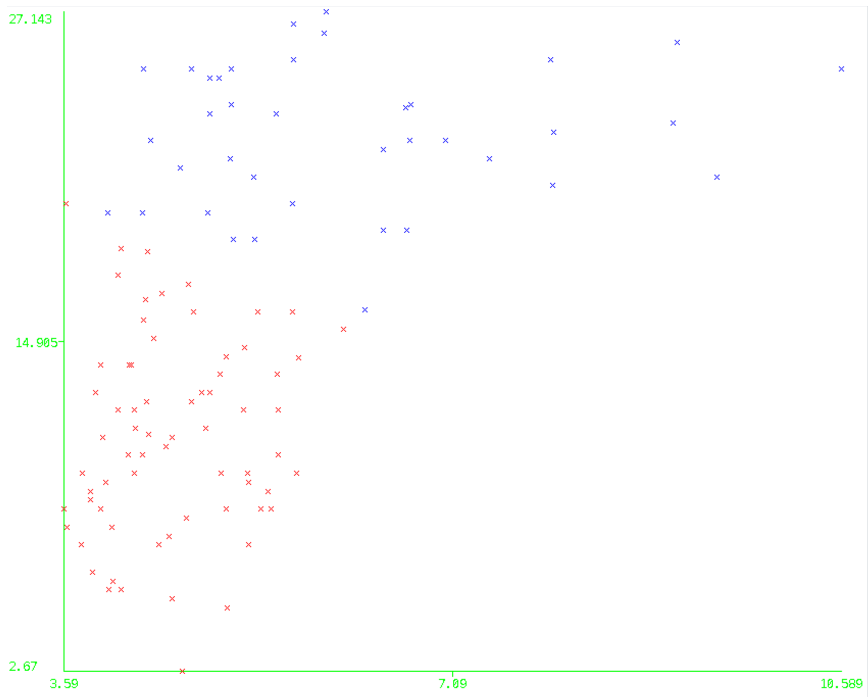


Figure 6-6 Results of the k-means clustering algorithm using the EEG ration alpha/beta and KDS of alpha to determine two levels of sleepiness using the “Neither” section. The red dots were classified as “post-awake” and the blue dots were classified as “pre-sleep”.

Plotting the $\frac{\alpha}{\beta}$ in the x-axis and KDS of $\boldsymbol{\alpha}$ in the y-axis, the position of the centroids for each cluster were (4.5847, 11.4072) for the post-awake and (6.1646, 22.5295) for the pre-sleep. The NeuroNets was trained and tested using only the “post-awake” and the “pre-sleep” segments to determine if the clustering obtained from the unsupervised k-means clustering algorithm was suitable. The accuracy obtained was 92.08% (SD=2.92%) and the error box is shown in Table 6-4. As performed in Chapter 4, different manual thresholds used to cluster the data were selected to determine the accuracy of the clustering obtained through k-means clustering. The α/β value was varied from 5 to 30 with 0.5 increments and the value of KDS α was varied from 3 to 10 with 0.5 increments. The best accuracy was obtained with a threshold set at α / β value of 16.5, similar to the result obtained with the k-means clustering algorithm. The above result confirmed that the thresholds used to cluster the data were suitable.

The NeuroNets was then trained with the four levels of sleep (awake, post-awake, pre-sleep and sleep) to determine the effect of adding the “neither” cluster. The accuracy obtained was 79.04% (SD=4.34%) and the error box is shown in Table 6-5.

Table 6-4 Error box to determine the accuracy of the NeuroNets when predicting 2 levels of sleepiness within the “Neither” cluster

	Awake	Sleep
Awake	93.32%	6.68%
Sleep	10.02%	89.98%

Table 6-5 Error box to determine the accuracy of the NeuroNets when predicting 4 levels of sleepiness

	Awake	Post-Awake	Pre-Sleep	Sleep
Awake	76.51%	10.46%	7.00%	6.03%
Post-Awake	6.61%	80.38%	11.00%	2.01%
Pre-Sleep	6.31%	11.91%	74.43%	7.35%
Sleep	6.23%	1.84%	5.47%	86.46%

6.2.4 Predicting binary clusters of sleepiness using driving and physiological behaviour

As presented in chapter 5, during the first experiment conducted for the present PhD study, only the EEG and driving behaviour was recorded for the participants undertaking the experiment. In the second experiment, for a new set of participants recruited, the head and blinking movements were recorded in addition to the EEG and driving. Unfortunately, the blinking movement recordings were too noisy and had to be discarded. Only head movement could be used as another feature to predict sleepiness.

The NeuroNets algorithm used in the previous section was trained and tested with the data obtained during the second experiment. When adding the head movement variables (number of nods per segment, standard deviation (SD) of the position of the head in the x axis, SD of the position of the head in the y axis and SD of the position of head in the z axis) to the MLA, the accuracy of the algorithm increases by 4%.

6.3 Conclusion

This chapter has presented the analysis performed to determine the multiple levels of sleepiness using an objective analysis of EEG. The EEG ratio $\frac{\boldsymbol{\alpha}}{\boldsymbol{\beta}}$ and KDS of $\boldsymbol{\alpha}$ were considered the most suitable variables to determine the different levels of sleepiness. A combination of objective analysis of the EEG variables as well as the use of unsupervised learning algorithms were used to determine the multiple levels of sleepiness. NeuroNets were used to determine the suitability of the proposed classification of sleepiness levels using driving and physiological behaviour.

Chapter 7: Discussion

7. Discussion

7.1 Contributions to knowledge

The primary contribution to knowledge of this PhD is a proposed new classification criterion for multiple levels of sleepiness using objective analysis of EEG data tested using MLAs. The proposed sleepiness clusters were analysed using NeuroNets, a type of MLA. The multiple levels of sleepiness were predicted using driving behaviour and physiological changes. The results obtained in chapter 6 showed that it achieved some reasonable degree of success.

This work was influenced by the necessity for early warning signals from safety systems in cases where drivers are falling asleep at the wheel. In the literature, sleepiness is mostly divided in two or three levels of sleepiness. This reduces the possibility for a safety system to act before the driver is in a high level of sleepiness and as a consequence being at high risk of having an accident (Lal et al., 2003; Sayed & Eskandarian, 2001; Yeo et al., 2009; Shuyan & Gangtie, 2009; Yang et al., 2010; Patel et al., 2011; Klauer et al., 2006; Lamond & Dawson, 1999). It also leads to the problem of high jumps in automation, e.g. if the actions taken by the safety system go from warning, at low levels of sleepiness, to complete control of the car, at high levels of sleepiness (Merat et al., 2014; Endsley, 1995; Carsten et al., 2012).

A second unique contribution was the development of an objective analysis and classification approach of driving related EEG data. In existing literature, researchers using EEG and MLAs to determine and predict different levels of sleepiness, hire expert EEG clinicians to cluster the data into different levels of sleepiness (Yeo et al., 2009; Vuckovic et al., 2002). The clinicians perform a visual subjective analysis of the EEG to determine different levels of sleepiness (Knoblauch et al., 2003; Andrillon et al., 2011; Benbadis, 2006; Carskadon & Dement, 2011; Gennaro & Ferrara, 2003; Devuyst et al., 2010a,b; Rivera & Salsa, 2013; Cantero et al., 2002; Parekh et al., 2015; Teplan, 2002). By developing an objective analysis of the EEG data to define the different levels of sleepiness there is potential to reduce the time and money that would be needed when hiring a clinician (specially with a high number of participants and long datasets). In addition, this method allows for reproducibility of the results. The following section explains the results obtained in relation to the contribution to knowledge of the present PhD study.

7.2 Overview

Investigating the behavioural aspects related to sleeping while driving has received a large amount of interest in the field of transport safety, from both research and industry. Evidence has indicated that there is a relationship between the rise in sleepiness and the driving and physiological behaviour of drivers. In line with this, this thesis attempted to explore the possibility of detecting different levels of sleepiness in drivers using their driving and physiological behaviour. In order to determine the different levels of sleepiness, brain wave activity, recorded using electroencephalogram (EEG), was used as the classifier to determine the different levels of sleepiness. Neural Networks (NeuroNets), a type of Machine Learning algorithm (MLA), was used to determine the accuracy of the proposed classification of the levels of sleepiness. The driving and physiological data as well as the EEG data were recorded from the experiments conducted in the static driving simulator of the University of Leeds.

During the experiments, participants had to drive in a long monotonous road in a night environment with no other traffic on the road. To compare the accuracy of the algorithm when using EEG against using other physiological variable to determine the levels of sleepiness, the Neural Networks algorithm was tested using driving and physiological dataset obtained by other researcher in 2013. The following chapter summarises the results of the algorithms in light of existing literature and discusses their theoretical implication and limitations alongside recommendations for future research.

7.3 Introduction

Although there is a deep understanding in the literature regarding the effects of sleepiness in the driving and physiological behaviour of drivers, to correlate these variables to different levels of sleepiness has proven to be a difficult challenge in safety and transport for researchers and industries (Lal et al., 2003; Sayed & Eskandarian, 2001; Yeo et al., 2009; Shuyan & Gangtie, 2009; Yang et al., 2010;

Patel et al., 2011). Changes in sleepiness manifest differently in each individual and, due to this individuality, it is difficult to develop an accurate explicit predicting model (Hartley et al., 2000; Haworth & Vulcan, 1991; Kaplan et al., 2007; Lal & Craig, 2003). Using MLAs allows the possibility of obtaining a more accurate detection of changes in sleepiness by training the algorithm with example data of the different stages of sleepiness of a driver (Harrington, 2012; Bell, 2015; Marsland, 2015; Murphy, 2012; Alpaydin, 2010). This means that the algorithm learns from the individual behaviour of each driver and finds patterns in the data, leading to a model that accounts for the individuality problem. MLAs are also useful when the amount of data are large, as it is the case in the present PhD study where the datasets of each participant contains more than 150,000 data points.

Another big challenge researchers in academia and industry are facing when developing systems that can predict sleepiness in drivers, is the lack of consensus on how to determine the different levels of sleepiness (Yeo et al., 2009; Vuckovic et al., 2002; Knoblauch et al., 2003; Andrillon et al., 2011; Benbadis, 2006; Carskadon & Dement, 2011; Gennaro & Ferrara, 2003; Devuyst et al., 2010a,b; Rivera & Salsa, 2013; Cantero et al., 2002; Parekh et al., 2015; Teplan, 2002). As discussed before, the effects of sleepiness in the driving and physiological behaviour vary from participants to participant. Therefore, researchers use different behavioural and/or driving variables to determine the different levels of sleepiness (Yang et al., 2010; Shuyan & Gangtie, 2009; Yeo et al., 2009; Vuckovic, et al., 2002; Patel et al., 2011). For example, Patel et al. (2011) used the heart rate variability of the participants to classify the data in “awake” and “fatigue”; Sayed & Eskandarina (2001) classified the data in “drowsy” and “non-drowsy” using the steering behaviour of the participants; and Shuyan & Gangtie (2009) used a combination of blinking behaviour and EEG to classify the data in “awake”, “sleepy” and “very sleepy”. Even though different driving and physiological variables have been used to classify the levels of sleepiness, it has been found that EEG and blinking behaviour are the most reliable variables to determine sleepiness, although the former is less affected by environmental factors (Artaud et al., 1994; Lal et al., 2003; Jap et al., 2009).

Therefore, this PhD analysed the accuracy of a model using blinking behaviour as a classifier of sleepiness and another model using EEG as a classifier of sleepiness. The model using blinking behaviour as the clustering variable of sleepiness was tested using a binary (“awake” and “sleep”) and a ternary (“awake”, “drowsy” and “sleep”) level of sleepiness. The model using EEG as the clustering variables of sleepiness was tested with two, three and four levels of sleepiness. The following section provides a summary of the results from these analyses.

7.4 Summary of results

7.4.1 Dataset 1: Determining the levels of sleepiness using blinking behaviour

There were two aims of this secondary analysis of a dataset obtained from a prior experiment done in the driving simulator of the University of Leeds in 2013. The first aim was to determine the accuracy of NeuroNets compared to Support Vector Machine (SVM; another type of MLAs). NeuroNets and SVM are the most common algorithms used in literature by researchers predicting different levels of sleepiness in drivers (Shuyan & Gangtie, 2009; Yeo et al., 2009; Vuckovic, et al., 2002; Patel et al., 2011). Previously, researchers using NeuroNets to predict different levels of sleepiness have obtained accuracy levels above 90% (Patel et al., 2011; Sayed & Eskandarian, 2001; Vuckovic et al., 2002). On the other hand, researchers using a SVM also obtained accuracy levels above 90% when predicting different levels of sleepiness (Yeo et al., 2009). The above shows that researchers have obtained high levels of accuracy using different types of MLAs.

Within NeuroNets, there are many possible designs that could lead to many different outcomes (Hagan et al., 2014). For the present PhD study, a three layer feed-forward with 12 neurons in the hidden layer was chosen. The number of iterations used was 10,000 and the learning rate was 0.1. This design has been used to predict sleepiness by many researchers, obtaining around 90% (Vuckovic et al., 2002; Sayed & Eskandarian, 2001; Patel et al., 2011). The number of neurons inside the hidden layer, the number of iterations and the learning rate were determined after a series of iterations using different values for each of the three variables. The values reported in this PhD study were the ones that achieved the highest accuracy.

The learning rate is the parameter that controls the update of the weights of the algorithm (Hagan et al., 2014; Harrington, 2012; Bell, 2015; Marsland, 2015; Murphy, 2012; Alpaydin, 2010). The weights are multiplied by the input, which in this case are the temporal driving variables, to obtain the desired output, the expected level of sleepiness. Depending on the error between the expected level of sleepiness and the predicted level of sleepiness, the weights are modified. The size of the change is controlled by the learning rate. Therefore, a large learning rate would lead the algorithm to diverge, i.e. not learning at all. A small learning rate would lead the algorithm to become slower and there might be a possibility of it being stuck in local minima, leading to a wrong learning. When the algorithm was tested with a lower learning rate (0.01), the accuracy decreased. This may be due to the individuality of the behaviour of the participants, leading the learning to local minima that do not correspond to the correct learning of the data. When using a bigger learning rate (0.3), the accuracy also decreased, meaning that the algorithm was diverging. The number of iterations and the number of hidden neurons was determined heuristically, as there is not an argument that can backup the suitability of the choice of these values. The number of hidden neurons and the iterations were increased until the accuracy did not change. Choosing the small number of iterations and few number of hidden neurons leads to a faster processing time of the NeuroNets algorithm.

The NeuroNets algorithm was tested against a SVM. The SVM was designed using linear kernels. A heuristic approach was followed to determine the best kernel, i.e. linear kernels were the kernel functions that gave the best results. The SVM has the possibility of transforming the dataset into a higher dimension dataset in problems where the data are not linearly separable. This is one of the biggest advantages of a SVM. Another great advantage is that the separation of the dataset is done using just few data points called support vectors. This support vectors create the margin that differentiates the different clusters. The number of support vectors can be any number from two up to the number of data points in the dataset, depending on the complexity and the choice of kernel function. In addition, when the output is not binary, the use of a multi-class SVM is not straightforward. Although, Shuyan & Ganagtie (2009) and Yeo et al. (2009) obtained high accuracies using SVM (around 87% and 99% respectively), for the present dataset used during this section the SVM did not perform better than the NeuroNets algorithm. This could be due to the noise in the data and the over-lapping of the different levels of sleepiness. In Vuckovic et al., (2002) and Yeo et al. (2009), the levels of sleepiness were classified by clinicians and the segments that were not clear or where the clinicians did not reached a consensus were eliminated. As the dataset used in the PhD study during this section was classified according to a more objective approach, every segment was classified, i.e. no segment was eliminated. This could have led to have more noise and overlapping levels of sleepiness amongst the data. Due to all these reasons, the NeuroNets performed better than the SVM and therefore the NeuroNets was used throughout this PhD study when analysing other datasets.

The second aim was to analyse the accuracy of the NeuroNets algorithm when the blinking behaviour data were clustered in two (“awake” vs. “sleep”) and three levels of sleepiness (“awake”, “drowsy” and “sleep”). Although there is no research where blinking behaviour has been used in conjunction with MLAs to determine different levels of sleepiness, many researchers have found deterministic thresholds values of blinking behaviour data that is related to different levels of sleepiness (Jimenez-Pinto & Torres-Torriti, 2013; Yeo at al., 2009; Boverie et al., 2013; Yang et al., 2010). Specifically, PERCLOS (percentage of closure of the eye) values have been related to different levels of sleepiness (Bergasa et al., 2006; Dinges et al., 1998; Mallis, 1999; Jimenez-Pinto & Torres-Torriti, 2013; Yeo at al., 2009; Boverie et al., 2013; Yang et al., 2010). Unfortunately, there is little consensus on the values of the threshold for each level of sleepiness (Jimenez-Pinto & Torres-Torriti, 2013; Yeo at al., 2009; Boverie et al., 2013; Yang et al., 2010). Jimenez-Pinto & Torres-Torriti (2013) determine that the state of “awake” was related with a PERCLOS value of 0.025, the level “drowsy” with a value of 0.09 and “sleep” with a value of 0.18. On the other hand, Boverie et al. (2013) determined that a value of PERCLOS of 0.24 or less was related to “awake”, between 0.24 and 0.45 was related to “fatigue” and above 0.45 was related to “drowsy”. Finally, Yang et al. (2010) concluded that the “alert” state is when PERCLOS has values between 0.01 and 0.05 and “fatigue” happens with PERCLOS is between 0.05 and 0.94.

Due to the discrepancy in the literature regarding the threshold values, it was not possible to use a specific threshold value to separate the blinking data of the current dataset into different levels of sleepiness. Instead, a different approach was followed using unsupervised MLAs using PERCLOS and blinking frequency as the variables. By using a k-means clustering algorithm, it was possible to classify the data without the necessity for the researcher to specify an explicit threshold value. When clustering the data into two levels of sleepiness (awake and sleep), it was found that PERCLOS had a bigger effect than blinking frequency. This was found to coincide with the results found in literature, where PERCLOS is considered to be the variable most highly correlated to sleepiness (Bergasa et al., 2006; Dinges et al., 1998; Mallis, 1999). The threshold value found by the k-means algorithm that separates the 2 levels of sleepiness (0.09) agrees with the value found by Jimenez-Pinto & Torres-Torriti (2013) giving validity to the clustering done by the k-means clustering algorithm. In addition, the accuracy obtained by the NeuroNets (89%) validates the use of k-means clustering algorithm to separate the different levels of sleepiness.

When the k-means clustering algorithm clustered the data into three levels of sleepiness, blinking frequency played a higher role in the division of the levels of sleepiness. The low accuracy obtained when using 3 levels of sleepiness implies that either the experiment was not long enough for participants to present symptoms of different levels of sleepiness or it is not possible to determine more than 2 levels of sleepiness using blinking behaviour data. These results confirmed that the dataset obtained from a previous experiment conducted in the driving simulator in 2013 was not suitable enough to obtain more than two levels of sleepiness. The next step taken was to design an experiment where sleep would be induced for a longer period and different physiological variables would be recorded.

7.4.2 Sleep inducing experiments

The conclusions obtained from the analysis of the first dataset confirmed the need to gather more sleepiness data from drivers. Therefore, two sleep-inducing experiments were conducted where the participants had to drive using a static driving simulator on a monotonous road during a night environment with no traffic around them. It was also concluded from the analysis of the first data set that a different variable needed to be used to classify the levels of sleepiness. The EEG of the participants was recorded during the experiments as it is considered a highly reliable variable to detect changes in sleepiness (Artaud et al., 1994; Lal et al., 2003; Jap et al., 2009).

The first set of experiments tested the effect of lunch as a sleepiness-inducing factor. Reyner et al. (2012) found that a lunch with high calories induced higher levels of sleepiness in participants compared to a lunch with low calories in a driving simulator experiment. Unfortunately, there is no research regarding the effects of lunch compared to the effects of not having lunch. In the experimental set up conducted in the present PhD study, lunch did not show any effect in changes in the sleepiness levels of the drivers compared to when they did not have lunch. The lunch provided might have not contained enough calories to create an effect in the level of sleepiness of the participants. Although a correlation between EEG and driving behaviour has been found in literature, for the first set of experiments no correlations were found between the EEG and the driving variables. The first hypothesis is that there were not enough participants to account for the individuality in the EEG and driving behaviour of each participant. In addition, the EEG recorded for each participant contained a lot of noise produced from blinking and muscle movement from the participant, e.g. yawning and moving the head. As the first set of experiments did not show any correlation between the EEG variables and the driving variables, a second experiment was designed, without the "lunch" condition (as lunch did not show any effect in the driving or EEG variables). For the second set of experiments, head movement and blinking behaviour was recorded for each participant. As mentioned before, noisy EEG data were removed due to blinking and head movement, therefore recording head movement and blinking would allow the researched to have a better understanding of the missing EEG data. It was also found that during the first set of experiments, not every participant experienced a "complete out of lane" (considered to happen during the highest level of sleepiness). Due to this fact, it was concluded that a longer driving task would be presented to the participants. The data obtained during the experiments was used in the following section to determine multiple levels of sleepiness.

7.4.3 Dataset 2: Determining the levels of sleepiness using EEG

The dataset used to determine the levels of sleepiness using EEG was obtained from the sleep inducing experiments conducted in the present PhD study. As discussed in previous sections, the MLA that would be used to analyse the clustering of the levels of sleepiness was NeuroNets. The two main aims to be reached in this section were the following: to define an objective method to analyse EEG, and determine the accuracy of the NeuroNets algorithm when predicting multiple levels of sleepiness (classified according to the EEG data).

As discussed previously, the first aim was to define an objective method to analyse EEG. In Yeo et al. (2009) and Vuckovic et al. (2002), two clinicians classified the data through visual inspection, into different sleepiness' state. It was reported that part of the EEG data were excluded as no consensus were reached between the clinicians (Yeo et al., 2009; Vuckovic et al., 2002). The methods used in the present study allowed to classify the data without the need to exclude any EEG data. Due to this fact, more training data were available for the NeuroNets to learn. A theory of visual classification of EEG was used to determine the variables to be used in the objective method to classify the EEG data. Visual bursts of alpha and theta are indicators of high levels of sleepiness (Knoblauch et al., 2003; Andrillon et al., 2011; Benbadis, 2006; Carskadon & Dement, 2011; Gennaro & Ferrara, 2003; Devuyst et al., 2010a,b; Rivera and Salsa, 2013; Cantero et al., 2002; Parekh et al., 2015; Teplan, 2002). A combination of the alpha and theta bursts and Karolinska Drowsiness Scale (KDS) was used to determine the different levels of sleepiness. By determining the number of alpha and theta burst during a 20 seconds epoch, an objective measure for high levels of sleepiness was developed. During low levels of sleepiness, e.g. awake, there is a high beta activity in combination with low alpha and theta activity (Knoblauch et al., 2003; Andrillon et al., 2011; Benbadis, 2006; Carskadon & Dement, 2011; Gennaro & Ferrara, 2003; Devuyst et al., 2010a,b; Rivera & Salsa, 2013; Cantero et al., 2002; Parekh et al., 2015; Teplan, 2002). Therefore, the power ratio of $\frac{\boldsymbol{\theta}+\boldsymbol{\alpha}}{\boldsymbol{\beta}}$ and $\frac{\boldsymbol{\alpha}}{\boldsymbol{\beta}}$ were a good indicator of low levels of sleepiness. Alpha bursts and the ratio $\frac{\boldsymbol{\alpha}}{\boldsymbol{\beta}}$ presented less variability than theta bursts and the ratio $\frac{\boldsymbol{\theta}+\boldsymbol{\alpha}}{\boldsymbol{\beta}}$, therefore $\frac{\boldsymbol{\alpha}}{\boldsymbol{\beta}}$ and KDS $\boldsymbol{\alpha}$ were used to determine the multiple levels of sleepiness. Muscle movement noise, e.g. yawning, jaw movement, talking, recorded by the EEG during the experiment might have affected theta frequency band (a low frequency band), as this type of noise is normally found in the low frequency range (Filtness et al., 2012). The following step was to determine the accuracy of this method of classification of

EEG data compared to results obtained when clinicians visually classified EEG data. The results were presented in terms of accuracy. As presented in Chapter 3, there are many validation measures for MLAs. In the present PhD study, accuracy was used as the validation measure, as it was important to achieve a high number of true positives (a system that can accurately determine when the driver is asleep) as well as true negatives (a system that does not ring an alarm when the driver is awake).

When the NeuroNets model was trained only with “Awake” and “Sleep” data, the accuracy of the algorithm was very high (95.71%, SD=2.8%). This is similar to excluding EEG data where clinicians do not reach a consensus, i.e. when the data does not clearly presents characteristics of certain cluster. The accuracy obtained by researchers in literature when they hired clinicians to classify the data into different levels of sleepiness was around 90-99% (Yeo et al., 2009; Vuckovic et al., 2002). This leads to the conclusion that the objective method used to classify and analyse the EEG data achieves the same accuracy as the visual classification and analysis done by clinicians. It is important to state that although better accuracy is obtained when the algorithm is presented only with data that can be clearly classified into a specific state of sleepiness, in a real scenario, the driver will present behaviour that does not clearly belongs to a certain state of sleepiness. It is therefore important to account for every type of behaviour, even the segments that cannot be clearly classified.

Once it has been concluded that the objective method developed to classify and analyse the EEG data were suitable, the second main aim was to determine multiple levels of sleepiness. Data that was not classify into “awake” or “sleep” was classified as “neither”. The “neither” data represents the transition between the two extreme levels of sleepiness (awake and sleep). A transition state before reaching the highest level of sleepiness, would allow a safety system to present warnings before the driver reaches a dangerous levels of sleepiness. The accuracy obtained from the NeuroNets model when predicting three levels of sleepiness, i.e. when “awake”, “sleep” and “neither” was used, was moderately high (85.75%, SD=3.31%). In the literature, researchers predicting three levels of sleepiness obtained accuracy of around 80-90% (Shuyan & Gangtie, 2009; Yeo et al., 2009). This leads to the conclusion that the objective method used to classify and analyse the EEG data were suitable. The “neither” data accounted for around 30% of the data and contained data close to the “awake” level of sleepiness as well as data close to the “sleep” level of sleepiness. If a warning is signal when the driver reaches the “neither” level of sleepiness, it might be considered as a false positive if the “neither” segment detected is closer to the “awake” level of sleepiness than to the “sleep” level of sleepiness.

It has been found that in literature there is a lack of knowledge and analysis of this transition state. Shuyan & Gangtie (2009) determined three levels of sleepiness (awake, sleepy and very sleepy). To classify the data into the different levels of sleepiness, Shuyan & Gangtie (2009) used the KDS variable (Rechtschaffen & Kales, 1968). Shuyan & Gangtie (2009) did not use a continuous value of KDS to separate the three states as presented in Figure 7-1. This means that the segments of data that lay in between the levels of sleepiness presented were not accounted for, i.e. the transition segments. Yeo et al. (2009) also separated the sleepiness data into three levels of sleepiness (awake, drowsy and sleep). It was not reported if the “drowsy” state is the transitional state between “awake” and “sleep” or if it is a state immediately prior to “sleep”, leading to a lack of data in the transitional stages.

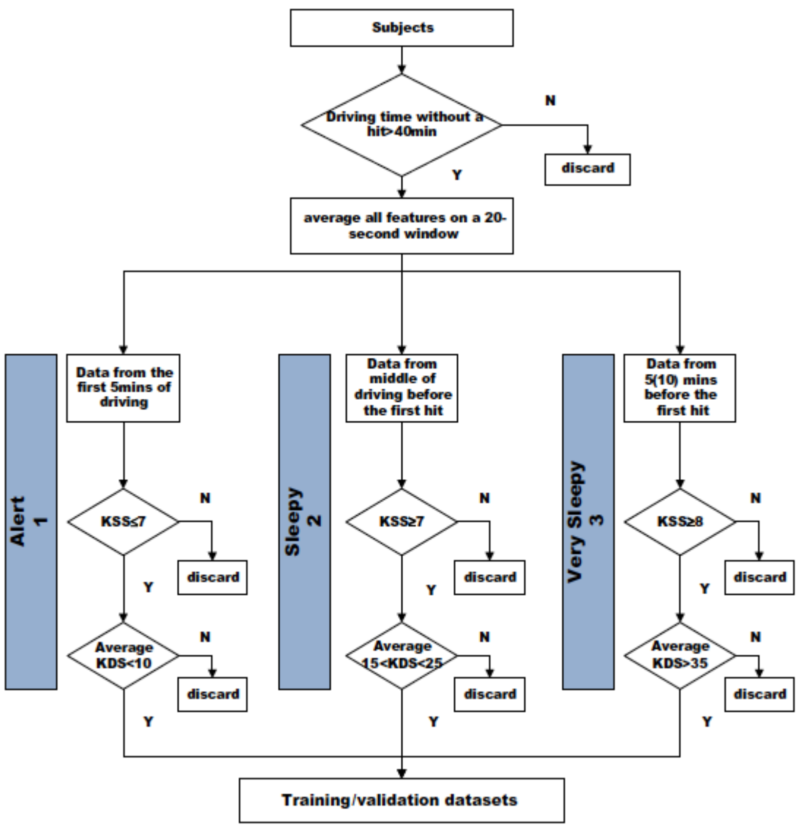


Figure 7-1 Flow chart used by Shuyan & Gangtie (2009) to classify the different levels of sleepiness using KSS and KDS (Source: Shuyan & Gangtie, 2009). Reprinted from "Driver drowsiness detection with eyelid related parameters by Support Vector Machine" by Shuyan Hu and Gangtie Zheng. Copyright © 2009 by Shuyan Hu and Gangtie Zheng. Used by permission of Elsevier.

Therefore, it was concluded that separating the "neither" section into "post-awake" and "pre-sleep" would decrease the possibility of false positives, i.e. premature warning from a safety system, and will account for the transitional data between the extreme levels of sleepiness (awake and sleep). Although in the literature visual indicators were used by clinicians to determine intermediate states of sleepiness between the two extreme of sleepiness (awake and sleep), in the present PhD study a MLA approach was followed. Because high accuracy results were obtained when determining sleepiness using k-means clustering and blinking behaviour, k-means clustering algorithms were used to determine the intermediate states "post-awake" and "pre-sleep". The variables used to determine the clusters were the same ones used to determine the "awake" and "sleep" states: the combined variable of alpha burst and KDS (hereafter called KDS of $\boldsymbol{\alpha}$ and the ratio $\frac{\boldsymbol{\alpha}}{\boldsymbol{\beta}}$). Same as in section 7.4.1, the k-means clustering algorithm was defined to find two clusters with random initial position using Euclidean function. When the NeuroNets model was trained with only the "post-awake" and the "pre-sleep" data, high accuracy (92.08%, SD=2.92%) was obtained. This lead the researcher to conclude that the separation between this two levels obtained using unsupervised k-means clustering algorithm was suitable. When the NeuroNets model was trained with the four levels of sleepiness (awake, post-awake, pre-sleep and sleep), the accuracy reduced to 79.04% (SD=4.34%). This means that the reduction in accuracy of the algorithms is due to the transition between the low levels of sleepiness, i.e. "awake" and "post-awake", and between the high levels of sleepiness, "pre-sleep" and "sleep". Although a high accuracy is preferred, it has been presented that the NeuroNets was capable to differentiate with high accuracy within the proposed levels of sleepiness obtained using objective analysis of EEG. Even when the multiple levels of sleepiness were tested together, a reasonable degree of success was achieved.

Another objective of the present PhD study was to assess the physiological changes in drivers as determinants to predict the different levels of sleepiness. In specific, blinking and eye behaviour as well as head movements by the driver were recorded during the monotonous driving study. The blinking and eye behaviour presented a difficulty for the researcher as the participants were constantly moving their heads outside of the range of the eye trackers (device to record blinking and eye behaviour). It was concluded that during driving simulator studies there is a trade off between the freedom allowed for the participants to behave "naturally" and the need to constraint the movement of the participants to obtain recording with lower amount of noise. The second physiological measure (the head movement of the participants) did not yield any additional information that could increase the accuracy of the algorithm. This could be caused by the fact that the appearance of certain physiological triggers of high levels of sleepiness (e.g. nodding or increase in the movement of the head from side to side) and the frequency appears during different levels of sleepiness depending on the individual. More research has to be done to obtain a better understanding of the appearance and cause of these types of physiological triggers of high levels of sleepiness. The present study could not exploit the information obtain from physiological variables and should be further study in future research.

The fact that in the present PhD study segments that might not be clearly classified into a specific level of sleepiness are not excluded, increases the possibility of incorrect classification from the algorithm. This led to the conclusion that in the dataset there is the existence of segments that could be classified as "awake" or "post-awake" or, in the other extreme, either as "pre-sleep" or "sleep". This means that if every segment of the dataset is to be tested, it will be inevitable to have a reduction of accuracy in the transitional segments. The fact that the accuracy obtained when determining between "awake" and "sleep" as well as when determining between "post-awake" and "pre-sleep" leads to the conclusion that the method developed to determine the different levels of sleepiness has achieved the aims of a suitable system to determine multiple levels of sleepiness in drivers.

7.5 Limitations

7.5.1 Real world driving versus control environment experiment

One of the biggest limitations in research related to sleepiness in driving is the lack of driving data of people falling asleep in the real world (Philip et al., 2005; Bos, Bles & Graaf, 2002). Using a driving simulator, it is possible to record driving behaviour by assuring the safety of the driver (Philip et al., 2005; Bos, Bles & Graaf, 2002). Although it has been found to be a suitable method to research driving behaviour in dangerous situations, real world driving behaviour is different from the driving behaviour in a driving simulator. This means that the MLAs trained using behavioural data obtained during an experiment conducted in a driving simulator may not be exactly adaptable to the real world behaviour of drivers. An understanding of the driving and physiological behaviour of people in real driving situation will improve the safety systems related to drivers falling asleep at the wheel. The above-mentioned problems need further examination in future work.

7.5.2 Use of EEG in a real environment

Although EEG was found to be a reliable estimator of sleepiness (Artaud et al., 1994; Lal et al., 2003; Jap et al., 2009), EEG is susceptible to noise caused by muscle movements from the driver, e.g. yawning, talking or moving, as well as noise in the environment, e.g. electromagnetic noise, sound vibrations, vibration from objects, etc. (Fisch, 2000; Núñez, 2010; Benbadis, 2006, Cohen, 2014). Due to the fact that head movement was also recorded during the experiments, the participants were only instructed to avoid talking or signing without any restriction to body or head movement. The EEG data obtained from the experiments conducted in the driving simulator presented a lot of noise. If the experiments were conducted in a moving base driving simulator, it is highly likely that the EEG signal would have contained more noise from the electrical equipment of the driving simulator as well as from the intrinsic movement of the simulator. This means that the noise-to-data ratio is a limitation when using EEG and the need to obtain real driving and physiological behaviour from the participants.

7.5.3 Individuality of sleepiness patterns

A third factor affecting the accuracy of the algorithm is the individual differences in the level of sleepiness of drivers (Campagne et al., 2009; Lowden et al., 2009; Filtness et al., 2012; Romero-Corral et al., 2010). During the experiments, participants had to undertake a monotonous driving task lasting 45 minutes (for the first experiment) and 60 minutes (for the second experiment). A participant was considered to have reached the maximum level of sleepiness when a "complete out of lane" was detected, as defined in other experiments found in literature (Shuyan & Gangtie, 2009). It was found that not every participant reached a maximum level of sleepiness during the experiment conducted for the present PhD study. A rigorous approach was taken when designing the experiment: only young participants were tested; participants were tested with Epsworth Sleepiness Scale; participant were asked to maintain a normal 8 hours sleep pattern; no alcohol was allowed 24 hours prior the experiment and no caffeine or energised drinks the day of the experiment; participants were asked not to consume large amounts of food before the experiment; and participants with high body mass index were not recruited as it is related to sleeping disorders (Romero-Corral et al., 2010). Despite the above approach, a great difference in the sleepiness pattern was found amongst the participants. Filtness et al. (2012) found that even in a 2 hour driving experiment not every participant fell asleep. This means that, for certain levels of sleepiness, data are hard to obtain regardless of the length of the experiment due to the individuality in the sleepiness pattern of the drivers. A large number of participants are needed to have enough data for every state of sleepiness.

7.6 Future Directions

In order to provide a more complete understanding and prediction of different levels of sleepiness as well as the actions to be taken by a safety system to avoid accidents, one crucial line of future investigation should be the understanding of the effects that different levels of sleepiness have in different driving and physiological behaviour of drivers.

Although correlations have been found between driving and physiological variables and levels of sleepiness, there is not enough research in relation to how each driving and physiological variable is affected in each specific level of sleepiness. By determining the effect of specific sleepiness states in the different driving variables, the actions taken by the safety system can be focused in mitigating the driving variable that is most affected in that specific state of sleepiness. In addition, by determining the behaviour of different physiological variables in each sleepiness state, a better prediction algorithm could be developed. This would reduce even more then need for high jumps in automation, which can lead to accidents (Merat et al., 2014; Endsley, 1995; Carsten et al., 2012). Therefore, more research should be devoted into obtaining a higher amount of behavioural data of drivers as sleepiness increases.

One recent trend in artificial intelligence has been the development of more powerful and faster MLAs (Harrington, 2012; Bell, 2015; Marsland, 2015; Murphy, 2012; Alpaydin, 2010; Jones, 2014; Hinton et al., 2006; Mo, 2012). With the introduction of deep learning neural networks, it has been possible to obtain better and faster learning of very complex data than with previously developed MLAs (Jones, 2014; Hinton et al., 2006; Mo, 2012). SVM and NeuroNets are considered shallow architectures, i.e. containing only fixed feature layer and a weight-combination layer, deep learning are considered deep architectures as it contains a multi-layer network of shallow architectures connected between each other. Better results have been obtained using deep learning compared to SVM and NeuroNets using the same data. As MLAs keep developing, it is more probable that better algorithms that can learn and predict better and faster the different levels of sleepiness using a large amount of driving and physiological data may be found. This could also lead for a MLA to be able to predict future states, i.e. given a particular state (or possibly history of states), what is the next state and how soon will it be reached. In addition, instead of having a deterministic predicted value, a confidence measure (probabilistic) could be achieved, e.g. the output of the present state is “awake” 0.2, “post-awake” 0.5, “pre-sleep” 0.2 and “sleep” 0.1.

Finally, the use of EEG was a constant reminder that methods to record brain wave activity is in need for improvement. The susceptibility to noise and the lack of automated EEG analysis packages for long datasets, are two of the drawbacks when using the majority of research-grade EEG systems. New technologies, such as fNIRS (functional near-infrared spectroscopy), are being developed and, even though it has been found to have drawbacks regarding its temporal accuracy drawbacks could be a method to reduce some of the drawbacks to EEG (Leon-Carrion & Leon-Dominguez, 2012). However, obtaining more EEG data from drivers as their sleepiness levels increase seems to be the only method to account for the noise recorded in the EEG, until new methods for recording brain wave activity suitable for our purposes are developed.

References:

Akerstedt, T., & Gillberg, M. (1990). Subjective and objective sleepiness in the active individual. International Journal of Neuroscience, 52, 29-37.

Akerstedt, T., & Kecklund, G. (2001). Age, gender and early morning highway accidents. Journal of Sleep Research, 10(2), 105-110.

Akerstedt, T., Peters, B., Anund, A., & Kecklund, G. (2005). Impaired alertness and performance driving home from the night shift a driving simulator study. Journal of Sleep Research, 14, 17-20.

Alpaydin, E. (2010). Introduction to Machine Learning (Second Edition ed.). London, England: The MIT Press.

Andrillon, T., Nir, Y., Staba, R. J., Ferrarelli, F., Cirelli, C., Tononi, G., & Fried, I. (2011). Sleep spindles in humans: insights from intracranial EEG and unit recordings. Journal of Neuroscience, 31(49), 17821-17834. doi:10.1523/JNEUROSCI.2604-11.2011

Anand, R., & Saravanan, S. (2016). A Correlative Study of Perturb and Observe Technique and GA-RBF-NN Method Supplying a Brushless DC Motor. Circuits and Systems, 7(8), 1-12.

Anund, A., Kecklund, G., Peters, B., Forsman, A., Lowden, A., & Akerstedt, T. (2008). Driver impairment at night and its relation to physiological sleepiness. Scandinavian Journal of Work, Environment & Health, 34, 142-150.

Apparies, R. J., Riniolo, T. C., & Porges, S. W. (1998). A psychophysiological investigation of the effects of driving longer-combination vehicles. Ergonomics, 41(5), 581-592.

Arman, S. I., Ahmed, A., & Syed, A. (2012). Cost-Effective EEG Signal Acquisition and Recording System. International Journal of Bioscience, Biochemistry and Bioinformatics, 2(5), 301-304.

Arnedt, J. T., Wilde, G. J., Munt, P. W., & MacLean, A. W. (2001). How do prolonged wakefulness and alcohol compare in the decrements they produce on a simulated driving task? Accident Analysis & Prevention, 33, 337-344.

Artaud, P., Planque, S., Lavergne, C., Cara, H., de Lepine, P., Tarriere, C., & Gueguen, B. (1994). An on-board system for detecting lapses of alertness in car driving. Paper presented at the 14th International Technical Conference on Experimental Safety Vehicles, Munich, Germany.

Banks, S., Peter Catcheside, Lack, L., Grunstein, R. R., & McEvoy, R. D. (2004). Low Levels of Alcohol Impair Driving Simulator Performance and Reduce Perception of Crash Risk in Partially Sleep Deprived Subjects. Sleep, 27(6), 1063-1067.

Bartlett, F. C. (1953). Psychological criteria of fatigue. In W. F. Floyd & A. T. Welford (Eds.), Symposium of fatigue. London: Lewsi.

BBC. (2016). Designing an Algorithm. Retrieved from http://www.bbc.co.uk/education/guides/z3bq7ty/revision/3

Bell, J. (2015). Machine Learning Hand-on for developers and technical professionals. Indianapolis, IN: John Wiley and Sons.

Benbadis, S. R. (2006). Introduction To Sleep Electroencephalography. In T. Lee-Chiong (Ed.), Sleep: A Comprehensive Handbook: John Wiley & Sons.

Berg, J. v. d., & Landstrom, U. (2006). Symptoms of sleepiness while driving and their relationship to prior sleep, work and individual characteristics. Transportation Research Part F, 9, 207-226.

Bergasa, L. M., Nuevo, J., Sotelo, M. A., Barea, R., & Lopez, M. E. (2006). Real-time system for monitoring driver vigilance. IEEE Transport Intelligent Transport Systems, 7(1), 63-77.

Bi, R. (2014). Will Deep Learning take over Machine Learning, make other algorithms obsolete? Retrieved May 22, 2015 from http://www.kdnuggets.com/2014/10/deep-learning-make-machine-learning-algorithms-obsolete.html

Billings, C. (1997). Aviation automation: The Search for a Human-Centered Approach. Mahwah, NJ: Lawrence Erlbaum Associates.

Blanco, M., Bocanegra, J. L., Morgan, J. F., Fitch, G. M., Medina, A., Olson, R. L., Green, K. (2009). Assessment of a Drowsy Driver Warning System for Heavy-Vehicle Drivers: Final Report. Retrieved August 6, 2013 from Blacksburg, Virginia, USA: https://www.nhtsa.gov/DOT/NHTSA/NRD/Multimedia/PDFs/.../2009/811117.pdf

Block, M., Bader, M., Tapia, E., Ramírez, M., Gunnarsson, K., Cuevas, E., . . . Rojas, R. (2008). Using Reinforcement Learning in Chess Engines. Journal Research in Computing Science: Special Issue in Electronics and Biomedical Engineering, Computer Science and Informatics, 35, 31-40.

Bloomfield, J., Harder, K. A., & Chihak, B. J. (2009). The Effect of Sleep Deprivation on Driving Performance. Retrieved October 18, 2013 from Minnesota, USA: www.cts.umn.edu/Publications/ResearchReports/pdfdownload.pl?id=1078

Borbély, A. A., Achermann, P., Trachsel, L., & Tobler, I. (1989). Sleep initiation and sleep intensity: interaction of homeostatic and circadian mechanisms. Journal of Biological Rhythms, 4, 149-160.

Bos, J. E., Bles, W. & Graaf, B. D. (2002) Eye movements to yaw, pitch, and roll about vertical and horizontal axes: adaptation and motion sickness. Aviation, Space, and Environmental Medicine, 73(5), 436-44.

Bosch (2012). "Bosch Driver Drowsiness Detection." Retrieved January 24, 2012 from http://www.bosch-presse.de/presseforum/details.htm?txtID=5037&tk_id=108.

Box, E. (2011). Mortality statistics and road traffic accidents in the UK. Retrieved October 30, 2015 from http://www.racfoundation.org/assets/rac_foundation/content/downloadables/road accident casualty comparisons - box - 110511.pdf

Boverie, S., Rodriguez, N., Bande, D., & Saccagno, A. (2013). General driver monitoring module definition SoA. (295364). Deserve Retrieved November 9, 2014 from http://www.deserve-project.eu/wp-content/uploads/2013/04/DESERVE-D32.1-General-Driver-Monitoring-Module-Definition1.pdf.

Boyle, L. N., Tippin, J., Paul, A., & Rizzo, M. (2008). Driver Performance in the Moments Surrounding a Microsleep. Transportation Research Part F: Traffic Psychology and Behaviour, 11(2), 126-136.

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

Broughton, R., & Hasan, J. (1995). Quantitative topographic electroencephalographic mapping during drowsiness and sleep onset. Journal of Clinical Neurophysiology, 12, 372-386.

Brown, I. D. (1994). Driver fatigue. Ergonomics, 36, 298-314.
C., C. F., & R., O. M. (2001). Corticotropin-releasing hormone (CRH) as a regulator of waking. Neuroscience Biobehaviour, 25, 445-453.
Campagne, A., Pebayle, T., & Muzet, A. (2004). Correlation between driving errors and vigilance level influence of the drivers age. Physiology & Behavior, 80, 515-524.
Campbell, M., Hoane, A. J., & Hsu, F.-h. (2002). Deep blue. Artificial Intelligence, 134, 57-83.
Canan, S. (ND). Physiology of Sleep. Retrieved November 8, 2015 from http://www.ybu.edu.tr/sinancanan/contents/files/605sleep.pdf
Cantero, J. L., Atienza, M., & Salas, R. M. (2002). Human alpha oscillations in wakefulness, drowsiness period, and REM sleep. Neurophysiologie Clinique, 32, 54-71.
Carskadon, M. A., & Dement, W. C. (2011). Normal Human Sleep : An Overview. In M. H. Kryger, T. Roth, & W. C. Dement (Eds.), Principles and practice of sleep medicine (5th edition ed.). St. Louis: Elsevier Saunders.
Carsten, O., Lai, F. C. H., Barnard, Y., Jamson, A. H., & Merat, N. (2012). Control Task Substitution in Semiautomated Driving: Does It Matter What Aspects Are Automated? Human Factors: The Journal of the Human Factors and Ergonomics Society, 54(5), 747-761.
Castaño, R., Anderson, R. C., Estlin, T., DeCoste, D., Fisher, F., Gaines, D., . . . Judd, M. (2003). Rover Traverse Science for Increased Mission Science Return. Paper presented at the Proc. 2003 IEEE Aerospace Conf., Big Sky, Montana.
Center-for-Automotive-Research. (2011). The U.S. Automotive Market and Industry in 2025. Retrieved October 24, 2015 from http://www.cargroup.org/assets/files/ami.pdf
Chang, F. C., & Opp, M. R. (2001). Corticotropin-releasing hormone (CRH) as a regulator of waking. Neuroscience Biobehaviour, 25, 445-453.
Chawla, N., Bowyer, K., Hall, L., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. Journal of Artificial Intelligence Research, 16, 321-357.
Chua, E. C., Tan, W. Q., Yeo, S. C., Lau, P., Lee, I., Mien, I. H., . . . Gooley, J. J. (2012). Heart rate variability can be used to estimate sleepiness-related decrements in psychomotor vigilance during total sleep deprivation. Sleep, 35(3), 325-334.
Cohen, M. X. (2014). Analyzing Neural Time Series Data: Theory and Practice. Cambridge: The MIT Press.
Connor, J., Whitlock, G., Norton, R., & Jackson, R. (2001). The role of driver sleepiness in car crashes: a systematic review of epidemiological studies. Accident Analysis & Prevention, 33, 31-41.
Costa, E. P., Lorena, A., Carvalho, A., & Freitas, A. (2007). A Review of Performance Evaluation Measures for Hierarchical Classifiers. Retrieved February 18, 2016 from http://www.aaai.org/Papers/Workshops/2007/WS-07-05/WS07-05-001.pdf
Cunliffe, A., Obeid, O., & Powell-Tuck, J. (2007). Post-prandial changes in measures of fatigue: Effect of a mixed or a pure carbohydrate or pure fat meal. European Journal of Clinical Nutrition, 51, 831-838.
Curcio, G., Casagrande, M., & Bertini, M. (2001). Sleepiness: evaluating and quantifying methods. International Journal of Psychophysiology, 41(3), 251-263.

Davis, N. (2014). From online dating to driverless cars, machine learning is everywhere. Retrieved May 18, 2016 from https://www.theguardian.com/science/2014/sep/18/machine-learning-artificial-intelligence

De Valck, E., & Cluydts, R. (2001). Slow-release caffeine as a countermeasure to driver sleepiness induced by partial sleep deprivation. Journal of Sleep Research, 10(3), 203-209.

Devuyst, S., Dutoit, T., Stenuit, P., & Kerkhofs, M. (2010a). Automatic K-complexes Detection in Sleep EEG Recordings using Likelihood Thresholds. Paper presented at the 32nd Annual International Conference of the IEEE EMBS, Buenos Aires, Argentina.

Devuyst, S., Dutoit, T., Ravet, T., Stenuit, P., Kerkhofs, M. (2010b). Comparison of Visual Sleep Stage Classification according to AASM and Rechtschaffen & Kales', Journal of Sleep Research, 19(2), 358.

Di Stasi, L. L., Renner, R., Catena, A., Cañas, J. J., Velichkovsky, B. M., & Pannasch, S. (2012). Towards a driver fatigue test based on the saccadic main sequence: A partial validation by subjective report data. Transportation Research Part C: Emerging Technologies, 21(1), 122-133.

Dinges, D., & Kribbs, N. (1991). Performing while sleepy: effects of experimentally induced sleepiness. In T. Monk (Ed.), Sleep, sleepiness and performance (pp. 97-128). Chichester: John Wiley and Sons Ltd.

Dinges, D. F. (1989). Napping patterns and effects in human adults. In D. F. Dinges & R. J. Broughton (Eds.), Sleep and alertness: chronobiological, behavioral, and medical aspects of napping (pp. 171-204). New York: Raven Press.

Dinges, D. F., Mallis, M. M., Maislin, G., & Powell, J. W. (1998). Evaluation of Techniques for Ocular Measurement as an Index of Fatigue and the Basis for Alertness Management. Retrieved July 30, 2013 from Washington, DC:NHTSA: http://ntl.bts.gov/lib/21000/21900/21955/PB99150237.pdf

Doughty, M. J. (2002). Further assessment of gender- and blink pattern- related differences in the spontaneous eyeblink activity in primary gaze in young adult humans. Optometry and Vision Science, 79, 439-447.

DTREG. (2014-2016). SVM - Suport Vector Machines. Retrieved June 4, 2016 from https://www.dtreg.com/solution/view/20

Dunn, N., & Williamson, A. (2012). Driving monotonous routes in a train simulator: the effect of task demand on driving performance and subjective experience. Ergonomics, 55(9), 997-1008.

EGI. (2009). Hydrocel Geodesic Sensor Net. Retrieved September 15, 2014 from ftp://ftp.egi.com/pub/support/Documents/net_layouts/hcgsn_128.pdf

Electrical Geodesics, Inc.(2007-2016). EGI: Innovation in neuroscience and neurology. Retrieved August 10, 2016 from https://www.egi.com/

Electronic Tutorials. (2016). Electrical Waveforms. Retrieved April 24, 2016 from http://www.electronics-tutorials.ws/waveforms/waveforms.html

Elhassan, T., Aljurf, M., Al-Mohanna, F., & Shoukri, M. (2016). Classica on of Imbalance Data using Tomek Link (T-Link) Combined with Random Under-sampling (RUS) as a Data Reduc on Method. iMedPub Journals, 1, 2-11.

Elsenbruch, S., Harnish, M. J., & Orr, W. C. (1999). Heart Rate Variability During Waking and Sleep in Healthy Males and Females. Sleep, 22(8), 1067-1071.

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. Human Factors, 37, 65-84.

Endsley, M. R., & Kiris, E. O. (1995). The out-of-the-loop performance problem and the level of control in automation. The journal of the human factors and ergonomics society, 37(2), 381-394.
Eoh, H. J., Chung, M. K., & Kim, S.-H. (2005). Electroencephalographic study of drowsiness in simulated driving with sleep deprivation. International Journal of Industrial Ergonomics, 35(4), 307-320.
Eskandarian, A. (2012). Fundamentals of Driver Assistance. In A. Eskandarian (Ed.), Handbook of Intelligent Vehicles. London, UK: Springer.
EU-OSHA. (2010). A review of accidents and injuries to road transport drivers. Retrieved December 24, 2013 from https://osha.europa.eu/en/publications/literature_reviews/Road-transport-accidents.pdf/view
European-Commission. (2013). Functional Magnetic Resonance Imaging. Retrieved November 2, 2014 from http://ec.europa.eu/research/participants/data/ref/h2020/other/hi/ethics-guide-fmri_en.pdf
Evans, R. J. (2011) Comparing methods for the syntactic simplification of sentences in information extraction. Literary and Linguistic Computing. Vol. 26, Nr. 4, 371-388.
Ferguson, S. A. (2003). Other high-risk factors for young drivers—how graduated licensing does, doesn't or could address them. Annual Proceedings—Association for the Advancement of Automotive Medicine, 47, 539-542.
Filtness, A. J., Reyner, L. A., & Horne, J. A. (2012). Driver sleepiness-comparisons between young and older men during a monotonous afternoon simulated drive. Biological Psychology, 89(3), 580-583.
Fisch, B. J. (2000). Fisch and Spehlmann's EEG Primer: Basic Principles of Digital and Analog EEG (3rd Edition ed.). USA: Elsevier.
Flemisch, F., Adams, C., Conway, S., Goodrich, K., Palmer, M., & Schutte, P. (2003). The H- metaphor as a guideline for vehicle automation and interaction (Technical Memorandum No. NASA/TM—2003-212672). Retrieved January 1, 2014 from Hampton, VA: https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20040031835.pdf
Gennaro, L. D., & Ferrara, M. (2003). Sleep spindles: an overview. Sleep Medicine Reviews, 7(5), 423-440.
George, C. F. P. (2005). Driving and automobile crashes in patients with obstructive sleep apnoea/hypopnoea syndrome. Sleep Diagnosis and Therapy, 1(1), 51-55.
George, N., & Kershaw, K. (2016). Road Use Statistics Great Britain 2016. Retrieved February 20, 2016 from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/514912/road-use-statistics.pdf
Gillberg, M., Keeklund, G., & Akerstedt, T. (1996). Sleepiness and performance of profes- sional drivers in a truck simulator – comparison between day and night driving. Journal of Sleep Research, 5, 12-15.
Goodrich, K., Flemisch, F., Schutte, P., & Williams, R. (2006). Application Of The H-Mode, A Design And Interaction Concept For Highly Automated Vehicles, To Aircraft. Paper presented at the 25th Digital Avionics Systems Conference.
Google. (2016). Google Self-Driving Car Project. Retrieved May 22, 2016 from https://www.google.com/selfdrivingcar/

Graw, P., Krauchi, K., Knoblauch, V., Wirz-Justice, A., & Cajochen, C. (2004). Circadian and wake-dependent modulation of fastest and slowest reaction times during the psychomotor vigilance task. Physiology & Behavior, 80, 695-701.

Greenberg, J., Artz, B., & Cathey, L. (2003, 8 - 9 October). The Effect of Lateral Motion Cues During Simulated Driving. Paper presented at the DSC North America, Dearborn, Michigan.

Gronfier, C., Simon, C., Piquard, F., Ehrhart, J., & Brandenberger, G. (1999). Neuroendocrine processes underlying ultradian sleep regulation in man. Journal of Clinical Endocrinology & Metabolism, 84(8), 2686-2690.

Grove, J. (2015). Vehicle Licensing Statistics: Quarter 4 (Oct - Dec) 2014. Retrieved March 2, 2016 from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/421337/vls-2014.pdf

Gu, H., Q, J., & Zhu, J. W. (2002). Active facial tracking for fatigue detection. Paper presented at the Workshop on Application of Computer Vision, Orlando.

Haga S. (1984). An experimental study of signal vigilance errors in train driving. Ergonomics 27, 755–765.

Hagan, M. T., Demuth, H. B., & Beale, M. H. (2014). Neural Network Design. Colorado, USA: Campus Pub. Service.

Hallvig, D., Anund, A., Fors, C., Kecklund, G., Karlsson, J. G., Wahde, M., & Akerstedt, T. (2013). Sleepy driving on the real road and in the simulator—A comparison. Accident Analysis & Prevention, 50, 44-50.

Hargutt, V., Hoffmann, S., Volrath, M, and Kruger, H.P. (2000). "Compensation for drowsiness and fatigue—a driving simulation study." In: Proceedings of the International Conference on Traffic and Transport Psychology (ICTTP), September 4-7, 2000, Bern, Switzerland.

Harrington, P. (2012). Machine Learning in Action. New York, USA: Manning Publications.

Harrison, Y., & Horne, J. A. (1996). Occurrence of microsleeps' during daytime sleep onset in normal subjects. Electroencephalography and Clinical Neurophysiology, 98(5), 411-416.

Hart, W. M. (1992). Adler's Physiology of the Eye: Clinical Application (Ninth edition ed.). Philadelphia: Mosby.

Hartley, L., Horberry, T., & Mabbott, N. (2000). Review Of Fatigue Detection And Prediction Technologies. Retrieved November 22, 2013 from Virginia, USA: https://www.researchgate.net/profile/Laurence_Hartley2/publication/238308422_REVIEW_OF_FATIGUE_DETECTION_AND_PREDICTION_TECHNOLOGIES/links/00b7d52c7b6bf34a63000000.pdf

Hartley, L. R., Arnold, P. K., Smythe, G., & Hansen, J. (1994). Indicators of fatigue in truck drivers. Applied Ergonomics, 25(3), 143-156.

Haworth, N. L., & Vulcan, P. (1991). Testing of Commercially Available Fatigue Monitors (Report 15. ISBN 0 7326 0015 4). Retrieved December 1, 2015 from http://www.monash.edu/__data/assets/pdf_file/0010/216793/muarc015.pdf

Hayami, T., Matsunaga, K., Shidoji, K., & Matsuki, Y. (2002). Detecting Drowsiness while Driving by Measuring Eye Movement. Paper presented at the The IEEE 5th International Conference on Intelligent Transportation Systems, Singapore.

Hinton, G. E., Osindero, S., & Teh, Y.-W. (2006). A Fast Learning Algorithm for Deep Belief Nets. Neural Computation, 18, 1527-1554.

Horne, J. A., & Gibbons, H. (1991). Effects on vigilance performance and sleepiness of alcohol given in the early afternoon ('post lunch') vs. early evening. Ergonomics, 34(1), 67-77.

Horne, J. A., & Reyner, L. A. (1995). Driver sleepiness. Journal of Sleep Research, 4(S2), 23-29.

Horne, J. A., & Reyner, L. A. (1996). Counteracting driver sleepiness: effects of napping, caffeine, and placebo. Psychophysiology, 33, 306-309.

Horne, J. A., & Reyner, L. A. (1999). Vehicle accidents related to sleep: a review. Occupational and Environmental Medicine, 56, 289-294.

Hsu, F.-h. (1999). IBM'S Deep Blue Chess Grandmaster Chips. Retrieved August 11, 2015 from http://www.csis.pace.edu/~ctappert/dps/pdf/ai-chess-deep.pdf

Hu, S., & Zheng, G. (2009). Driver drowsiness detection with eyelid related parameters by Support Vector Machine. Expert Systems with Applications, 36(4), 7651-7658.

Huntley, M. S. & Centybear, T. M. (1974). Alcohol, sleep deprivation and driving speed effects upon control use during driving. Human Factors, 16, 19.

International-Road-Transport-Union. (2007). A scientific Study 'ETAC' European Truck Accident Causation, Executive Summary and Recommendations. Retrieved June 30, 2015 from http://www.iru.org/index/cms-filesystem-action?file=mix-publications/2007_ETACstudy.pdf

Inagaki, T. (2003). Adaptive Automation Sharing and Trading of Control. In E. Hollnage (Ed.), Handbook of Cognitive Task Design (pp. 147-169): LEA.

Inagaki, T. (2006). Design of human–machine interactions in light of domain-dependence of human-centered automation. Cognition, Technology and Work, 8, 161-167.

Inagaki, T. (2007). Towards monitoring and modelling for situation-adaptive driver assist systems. In P. C. Cacciabue (Ed.), Modelling Driver Behaviour in Automotive Environments (pp. 43-57). London: Springer.

Inagaki, T. (2009). Human Machine Collaboration for Safety and Comfort. Retrieved January 2, 2014 from http://www.enri.go.jp/eiwac/2009/ppts/HumanMachineCollaborationForSafetyAndComfort.pdf

Inagaki, T., & Stahre, J. (2004). Human supervision and control in engineering and music: similarities, dissimilarities, and their implications. Proceedings IEEE, 92(4), 589-600.

IndexMundi. (2016). United Kingdom Age structure. Retrieved July 30, 2016 from http://www.indexmundi.com/united_kingdom/age_structure.html

Ingre, M., Akerstedt, T., Peters, B. R., Anund, A., & Kecklund, G. (2006). Subjective sleepiness, simulated driving performance and blink duration examining individual differences. Journal of Sleep Research, 15, 47-53.

IRTU (2007). A scientific Study 'ETAC' European Truck Accident Causation, Executive Summary and Recommendations. Retrieved August 2, 2013 from http://www.iru.org/index/cms-filesystem-action?file=mix-publications/2007_ETACstudy.pdf

Jain, A. K. & Dubes, R. C. (1988) Algorithms for clustering data. Prentice-Hall, Inc,.

Jap, B. T., Lal, S., & Fischer, P. (2011). Comparing combinations of EEG activity in train drivers during monotonous driving. Expert Systems with Applications, 38(1), 996-1003.

Jap, B. T., Lal, S., Fischer, P., & Bekiaris, E. (2009). Using EEG spectral components to assess algorithms for detecting fatigue. Expert Systems with Applications, 36(2), 2352-2359.

Jimenez-Pinto, J., & Torres-Torriti, M. (2013). Optical Flow and Driver's Kinematics Analysis for State of Alert Sensing. Sensors, 13, 4225-4257.

Johns, M. W. (1991). A new method for measuring daytime sleepiness: the Epworth sleepiness scale. Sleep, 14(6), 540-545.

Johns, M. W. (1998). Rethinking the assessment of sleepiness. Sleep Medicine Reviews, 2, 3-15.

Johns, M. W. (2000). A sleep physiologists view of the drowsy driver. Transportation Research Part F, 3, 241-249.

Johnson, D. H. (2013). Fundamentals of Electrical Engineering I. Retrieved January 23, 2015 from http://www.ece.rice.edu/~dhj/courses/elec241/col10040.pdf

Jones, N. (2014). Computer Science: The learning machines. Retrieved March 10, 2015 from http://www.nature.com/news/computer-science-the-learning-machines

Jung, T.-P., Makeig, S., Humphries, C., Lee, T.-W., Mckeown, M. J., Iragui, V., & Sejnowski, T. J. (2000). Removing electroencephalographic artifacts by blind source separation. Psychophysiology, 37, 163-178.

Kaber, D. B., & Endsley, M. R. (2004). The effects of level of automation and adaptive automation on human performance, situation awareness and workload in a dynamic control task. Theoretical Issues in Ergonomics Science, 5(2), 113-153.

Kaida, K., Takahashi, M., Akerstedt, T., Nakata, A., Otsuka, Y., Haratani, T., & Fukasawa, K. (2006). Validation of the Karolinska sleepiness scale against performance and EEG variables. Clinical Neurophysiology, 117(7), 1574-1581.

Kaplan, K. A., Itoi, A., & Dement, W. C. (2007). Awareness of sleepiness and ability to predict sleep onset: can drivers avoid falling asleep at the wheel? Sleep Medicine, 9(1), 71-79.

Kecklund, G., & Akerstedt, T. (1993). Sleepiness in long distance truck driving: an ambulatory EEG study of night driving. Ergonomics, 36(9), 1007-1017.

Kircher, K. (2001). General information Vitaport II. VTI, Swedish National Road and Transport Research Institute, Linköping

Klauer, S. G., Dingus, T. A., Neale, V. L., Sudweeks, J., & Ramsey, D. J. (2006). The impact of driver inattention on near-crash/crash risk: an analysis using the 100-car Naturalistic Driving Study Data. Retrieved March 10, 2014 from https://vtechworks.lib.vt.edu/bitstream/handle/10919/55090/DriverInattention.pdf?sequence=1

Knight, J. N. (2003). Signal Fraction Analysis and Artifact Removal in EEG. (Master of Science), Colorado State University, Colorado, USA.

Knoblauch, V., Martens, W. L. J., Wirz-Justice, A., & Cajochen, C. (2003). Human sleep spindle characteristics after sleep deprivation. Clinical Neurophysiology, 114, 2258–2267.

Koike, S., Nishimura, Y., Takizawa, R., Yahata, N., & Kasai, K. (2013). Near-infrared spectroscopy in schizophrenia: a possible biomarker for predicting clinical outcome and treatment response, 4(145), 1-12. Retrieved May 1, 2014 from www.frontiersin.org/doi:0.3389/fpsyt.2013.00145

Korb, K. B., & Nicholson, A. E. (2004). Bayesian Artificial Intelligence. London, UK: Chapman & Hall.

Kozak, K., Curry, R., Greenberg, J., Artz, B., Blommer, M., & Cathey, L. (2005). Leading Indicators of Drowsiness in Simulated Driving. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 49(22), 1917-1921.

Lal, S. K., & Craig, A. (2005). Reproducibility of the spectral components of the electroencephalogram during driver fatigue. International Journal of Psychophysiology, 55(2), 137-143.

Lal, S. K. L., & Craig, A. (2001a). A critical review of the psychophysiology of driver fatigue. Biological Psychology, 55, 173–194.

Lal, S. K. L., & Craig, A. (2001b). Electroencephalography activity asso- ciated with driver fatigue: Implications for a fatigue countermeasure device. Journal of Psychophysiology, 15, 183-189.

Lal, S. K. L., & Craig, A. (2002). Driver fatigue Electroencephalography and psychological assessment. Psychophysiology, 39, 313-321.

Lal, S. K. L., Craig, A., Boord, P., Kirkup, L., & Nguyen, H. (2003). Development of an algorithm for an EEG-based driver fatigue countermeasure. Journal of Safety Research, 34(3), 321-328.

Lamond, N., & Dawson, D. (1999). Quantifying the performance impairment associated with fatigue. Journal of Sleep Research, 8(4), 255-262.

Lavesson, N. (2006). Evaluation and Analysis of Supervised Learning Algorithms and Classifiers. Karlskrona, Sweden Blekinge Institute of Technology.

Leibling, D. (2008). Car ownership in Great Britain. Retrieved June 4, 2013 from London, UK: http://www.racfoundation.org/assets/rac_foundation/content/downloadables/car%20ownership%20in%20great%20britain%20-%20leibling%20-%20171008%20-%20report.pdf

Lenne, M. G., Triggs, T. J., & Red, J. R. (1997). Time of Day Variations in Driving Performance. Accident Analysis & Prevention, 29(4), 431-437.

Lenne, M. G., Triggs, T. J., & Redman, J. R. (1998). Interactive effects of sleep deprivation, time of day, and driving experience on a driving task. Sleep, 21(1), 38-44.

León-Carrión, J., & León-Domínguez, U. (2012). Functional near-infrared spectroscopy (fNIRS): principles and neuroscientific applications. In P. Bright (Ed.), Neuroimaging - Methods (pp. 45-75). Rijeka, Croatia: InTech.

Lexus-Europe (2012). "LS Driver Monitoring System." Retrieved September 20, 2015 from http://www.lexus.eu/range/ls/key-features/safety/safety-driver-monitoring-system.aspx.

Li, Z. Y., Jiao, K., Chen, M., & Wang, C. T. (2004). Reducing the effects of driving fatigue with magnitopuncture stimulation. Accident Analysis & Prevention, 36, 501-505.

Liang, S. F., Lin, C. T., Wu, R. C., Chen, Y. C., Huang, T. Y., & Jung, T. P. (2005). Monitoring driver's alertness based on the driving performance estimation and the EEG power spectrum analysis. Paper presented at the IEEE the 27th Annual Conference on Engineering in Medicine and Biology, Shanghai, China.

Linden, G., Smith, B., & York, J. (2003). Amazon.com Recommendations Item-to-Item Collaborative Filtering. IEEE Computer Society. Retrieved February 28, 2013 from https://www.cs.umd.edu/~samir/498/Amazon-Recommendations.pdf

Lindquist, M. A., & Wager, T. D. (2014). Principles of functional Magnetic Resonance Imaging. London: Chapman & Hall.

Liu, C. C., Hosking, S. G., & Lenne, M. G. (2009). Predicting driver drowsiness using vehicle measures: recent insights and future challenges. Journal of Safety Research, 40(4), 239-245.

Lowden, A., Anund, A., Kecklund, G., Peters, B., & Akerstedt, T. (2009). Wakefulness in young and elderly subjects driving at night in a car simulator. Accident Analysis & Prevention, 41(5), 1001-1007.

Lloyd, H. M., Green, M. W., & Rogers, P. J. (1994), 'Mood and cognitive performance effects of isocaloric lunches differing in fat and carbohydrate content', Phisiological Behaviour. 56, 51-57.

LumeWay. (2014). LumeWay: Safer by Design. Retrieved April 23, 2015 from http://www.lumeway.com/Products.htm

MacLean, A. W., Davies, D. R. T., & Thiele, K. (2003). The hazards and prevention of driving while sleepy. Sleep Medicine Reviews, 7(6), 507-521.

Marion, B. (1994). Turing Machines and Computational Complexity. The Computer Science Sampler, Amer. Math. Monthly. MR1542464.

Mallis, M. (1999). Evaluation of Techniques for Drowsiness Detection: Experiment on Performance-Based Validation of Fatigue-Tracking Technologies. Drexel University, Philadelphia, PA.

Mammar, S. (2006). Time to lane crossing for Lane Departure Avoidance: A Theoretical Study and an Experiential Setting. IEEE Transactions on Intelligent Transportation Systems, 7(2), 226-241.

Markoff, J. (2010). "Google Cars Drive Themselves, in Traffic." Retrieved March 10, 2015 from http://www.nytimes.com/2010/10/10/science/10google.html?_r=2.

Marsland, S. (2015). Machine Learning An Algorithm Perspective. Florida, USA: CRC Press.

MathWorks. (2011b). Matlab. Retrieved August 6, 2016 from https://uk.mathworks.com/products/matlab/?requestedDomain=uk.mathworks.com

May, J. F., & Baldwin, C. L. (2009). Driver fatigue: The importance of identifying causal factors of fatigue when considering detection and countermeasure technologies. Transportation Research Part F: Traffic Psychology and Behaviour, 12(3), 218-224.

Maycock, G. (1997). Sleepiness and driving: the experience of heavy goods vehicle drivers in the UK. Journal of Sleep Research, 6(4), 238-244.

McCall, J. C., Trivedi, M. M., Wipf, D., & Rao, B. (2005). Lane change intent analysis using robust operators and sparse bayesian learning. Paper presented at the CVPR ’05: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Washington, DC, USA.

McCartt, A. T., Rohrbaugh, J. W., Hammer, M. C., & Fuller, S. Z. (2000). Factors associated with falling asleep at the wheel among long-distance truck drivers. Accident Analysis & Prevention, 32(4), 493-504.

Mccormick, I. A., Walkey, F. H., & Taylor, A. J. W. (1987). The Stress Arousal Checklist: An Independent Analysis. Educational and Psychological Measurement, 45(1), 143-146.

McGovern, A., & Wagstaff, K. L. (2011). Machine learning in space: extending our reach. Machine Learning, 84(3), 335-340.

McHugh, M. (2015). Tesla's cars now drive themselves, kinda. Wired Gear. Retrieved October 2, 2016 from https://www.wired.com/2015/10/tesla-self-driving-over-air-update-live/

Merat, N., & Jamson, A. H. (2013). The effect of three low-cost engineering treatments on driver fatigue: A driving simulator study. Accident Analysis & Prevention, 50, 8-15.

Merat, N., Jamson, A. H., Lai, F. C. H., & Carsten, O. (2012). Highly Automated Driving, Secondary Task Performance, and Driver State. Human Factors, 54(5), 762-771.

Merat, N., Jamson, A. H., Lai, F. C. H., Daly, M., & Carsten, O. M. J. (2014). Transition to manual: Driver behaviour when resuming control from a highly automated vehicle. Transportation Research Part F: Traffic Psychology and Behaviour, 27, 274-282.

Michon, J. A. (1985). A critical view of driver behavior models: what do we know, what should we do? In L. Evans & R. C. Schwing (Eds.), Human behavior and traffic safety (pp. 485-520). New York: Plenum Press.

Mickiewicz, M. F. (2012). Brainput. Retrieved July 29, 2015 from https://www.prote.in/journal/articles/brainput

Mo, D. (2012). A survey on deep learning: one small step toward AI. Retrieved June 10, 2015 from http://www.cs.unm.edu/index.html

Moller, H. J., Kayumov, L., Bulmash, E. L., Nhan, J., & Shapiro, C. M. (2006). Simulator performance, microsleep episodes, and subjective sleepiness: normative data using convergent methodologies to assess driver drowsiness. Journal of Psychosomatic Research, 61(3), 335-342.

Monk, T. H. (2005). The post-lunch dip in performance. Clinics in Sports Medicine, 24(2), e15-23, xi-xii.

Mulder, M., Abbink, D. A., & Boer, E. R. (2012). Sharing control with haptics: seamless driver support from manual to automatic control. Human Factors, 54(5), 786-798.

Murphy, K. P. (2012). Machine Learning A Probabilistic Perspective. London, England: The MIT Press.

Nap-Zapper. (2008-2016). Nap Zapper. Retrieved May 22, 2016 from http://www.napzapper.com/

National-Aeronautics-and-Space-Administration. (2015). NASA Reaches New Heights in 2015. Retrieved July 6, 2015 from https://www.nasa.gov/press-release/nasa-reaches-new-heights-in-2015

NASA. (2012). High-Resolution Self-Portrait by Curiosity Rover Arm Camera. Retrieved October 31, 2016 from http://mars.jpl.nasa.gov/msl/multimedia/images/?ImageID=4845

NHTSA. (1999). A Preliminary Assessment of Algorithms for Drowsy and Inattentive Driver Detection on the Road. Retrieved November 4, 2014 from https://babel.hathitrust.org/cgi/pt?id=mdp.39015075377047;view=1up;seq=1

NHTSA. (2015). Critical Reasons for Crashes Investigated in the National Motor Vehicle Crash Causation Survey. Retrieved November 14, 2015 from Washington, DC, USA: https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812115

Núñez, I. M. B. (2010). EEG Artifact Detection Department of Cybernetics: Czech Technical University in Prague. Retrieved December 23, 2015 from https://riunet.upv.es/bitstream/handle/10251/10356/Project_Report_IB.pdf

NeuroSky. (2015). Greek Alphabet Soup – Making Sense of EEG Bands. Retrieved January 30, 2016 from http://neurosky.com/2015/05/greek-alphabet-soup-making-sense-of-eeg-bands/

Ng, A. [Stanford]. (2008, July 22). Lecture 11 | Machine Learning (Stanford) [video]. Retrieved June 2, 2015 from https://youtu.be/sQ8T9b-uGVE?list=PL382F7B6C56973EB8

Oken, B. S. & Chiappa, K. H. (1986). Statistical issues concerning computerized analysis of brainwave topography. Annual Neurology. 19, 493-494.

Oron-Gilad, T, and Shinar, D. (2000). Driver fatigue among military truck drivers. Transportation Research, Part F: Traffic Psychology and Behavior, 3, 195-209.

Orr, M. J. L. (1996). Introduction to Radial Basis Function Networks. Retrieved July 23, 2014 from http://www.cc.gatech.edu/~isbell/tutorials/rbf-intro.pdf

Ostlund, J., Nilsson, L., Töonros, J., & Forsman, A. (2006). Effects of cognitive and visual load in real and simulated driving. Retrieved May 21, 2016 from Linköping Sweden: https://www.diva-portal.org/smash/get/diva2:675275/FULLTEXT02.pdf

Otmani, S., Pebayle, T., Roge, J., & Muzet, A. (2005). Effect of driving duration and partial sleep deprivation on subsequent alertness and performance of car drivers. Physiology & Behavior, 84(5), 715-724.

Pack, A. I., Pack, A. M., Rodgman, E., Cucchiara, A., Dinges, D. F., & Schwab, W. (1995). Characteristics of crashes attributed to the driver having fallen asleep. Accident Analysis & Prevention, 27(6), 769-775.

Papadelis, C., Chen, Z., Kourtidou-Papadeli, C., Bamidis, P. D., Chouvarda, I., Bekiaris, E., & Maglaveras, N. (2007). Monitoring sleepiness with on-board electrophysiological recordings for preventing sleep-deprived traffic accidents. Clinical Neurophysiology, 118(9), 1906-1922.

Parasuraman, R., Bhari, T., Molloy, R., & Singh, I. (1991). Effects of shifts in the level of automation on operator performance. Paper presented at the Proceedings of the 6th International Symposium on Aviation Psychology, Columbus, OH.

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. Human Factors, 39(2), 230-253.

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for Types and Levels of Human Interaction with Automation. IEEE Transactions On Systems, Man, And Cybernetics—Part A: Systems And Humans, 30(3), 286-296.

Parekh, A., Selesnick, I. W., Rapoport, D. M., & Ayappa, I. (2015). Detection of K-complexes and Sleep Spindles (DETOKS) using Sparse Optimization. Journal of Neuroscience Methods, 251, 37-46.

Patel, M., Lal, S. K. L., Kavanagh, D., & Rossiter, P. (2011). Applying neural network analysis on heart rate variability data to assess driver fatigue. Expert Systems with Applications, 38(6), 7235-7242.

Paul, A., Boyle, L. N., Tippin, J., & Rizzo, M. (2005). Variability Of Driving Performance During Microsleeps. Paper presented at the Third International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, California, USA.

Peeke, S. C., Callaway, E., Jones, R. T., Stone, G. C. & Doyle, J. (1980). Combined effects of alcohol and sleep deprivation in normal young adults. Phsychopharmachology. 67, 279-87.

Philip, P., & Akerstedt, T. (2006). Transport and industrial safety, how are they affected by sleepiness and sleep restriction? Sleep Medicine Review, 10(5), 347-356.

Philip, P., Sagaspe, P., Taillard, J., Valtat, C. d., Moore, N., Akerstedt, T., . . . Bioulac, B. (2005). Fatigue, Sleepiness, and Performance in Simulated Versus Real Driving Conditions. Sleep, 28(12), 1511-1516.

Philip, P., Taillard, J., Guilleminault, C., Quera Salva, M. A., Bioulac, B., & Ohayon, M. (1999). Long distance driving and self-induced sleep deprivation among automobile drivers. Sleep, 22, 475-480.

Plankermann, K. (2013). Human Factors as Causes for Road Traffic Accidents in the Sultanate of Oman under Consideration of Road Construction Designs. University of Regensburg, Regensburg, Germany.

Rafaely, V., Meyer, J., Zilberman-Sandler, I., & Viener, S. (2006). Perception of traffic risks for older and younger adults. Accident Analysis & Prevention, 38(6), 1231-1236.

Rasmussen, J. (1983). Skills, rules and knowledge: signals, signs and symbols; and other distinctions in human performance model. IEEE-SMC, 13(3), 257-267.

Rechtschaffen, A. & Kales, A. (1968). A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects. US Department of Health, Education and Welfare, Public Health Service, Bethesda, MD.

Reyner, L. A., & Horne, J. A. (1997). Suppression of sleepiness in drivers: combination of caffeine with a short nap. Psychophysiology, 34, 721-725.

Reyner, L. A., & Horne, J. A. (2000). Early morning driver sleepiness: effectiveness of 200 mg caffeine. Psychophysiology, 37, 251-256.

Reyner, L. A., & Horne, J. A. (2002). Efficacy of a ‘functional energy drink’ in counteracting driver sleepiness. Physiology & Behavior, 75, 331-335.

Reyner, L. A., Wells, S. J., Mortlock, V., & Horne, J. A. (2012). 'Post-lunch' sleepiness during prolonged, monotonous driving - effects of meal size. Physiology & Behavior, 105(4), 1088-1091.

Rich, C. K. (2010). Psychology Researcher Uses Technology to Learn More about Memory. Retrieved August 6, 2013 from https://news.gmu.edu/articles/4389

Riemersma, J.B., Sanders, A.F., Wildervanck, C. and Gaillard, A.W. (1977)

“Performance decrement during prolonged night driving.” In: Mackie, R.R. (Editor) Vigilance: theory, operational performance and physiological correlates. New York, NY: Plenum Press, 41-58.

Risser, M. R., Ware, C., & Freeman, F. G. (2000). Driving Simulation with EEG Monitoring in Normal and Obstructive Sleep Apnea Patients. Sleep, 23(3), 1-6.

Rivera, M., & Salas, L. (2013). Monitoring of Micro-sleep and Sleepiness for the Drivers Using EEG Signal. Malardalen University, Vasteras, Sweden.

Rogue-Resolutions (2016). Smart Eye Pro 3d Eye Tracker. Retrieved October 13, 2016 from http://rogue-resolutions.com/catalogue/neuro-sensory/smart-eye-pro-5-10-3d-eye-tracker/

Romero-Corral, A., Sert-Kuniyoshi, F. H., Sierra-Johnson, J., Orban, M., Gami, A., Davison, D., . . . Somers, V. K. (2010). Modest Visceral Fat Gain Causes Endothelial Dysfunction in Healthy Humans. Journal of American College of Cardiology, 56(8), 662-666.

Rouse, W. B. (1994). Twenty years of adaptive aiding: Origins of the concept and lessons learned. In M. Mouloua & R. Parasuraman (Eds.), Human performance in automated systems: Current research and trends (pp. 28-32). Hillsdale, NJ: Lawrence Erlbaum Associates.

Royal, D. (2003). Volume I Findings National Survey of Distracted and Drowsy Driving Attitudes and Behavior 2002. Retrieved November 14, 2015 from Washington DC, USA: http://www.nhtsa.gov/people/injury/drowsy_driving1/survey-distractive03/index.htm

Sadock, B. J., & Sadock, V. A. (2000). Kaplan and Sadock's Comprehensive Textbook of Psychiatry. Philadelphia: Lippincott Williams & Wilkins.

SAE-International (Producer). (2014, 25-11). Automated driving levels of driving automation are defined in new SAE INTERNATIONAL standard J3016. Retrieved November 25, 2014 from http://www.sae.org/misc/pdfs/automated_driving.pdf

Sagaspe, P., Tailard, J., Chaumet, G., Moore, N., Bioulac, B., & Philip, P. (2007). Aging and nocturnal driving: better with coffee or a nap? A randomised study. Sleep, 30(12), 1808-1813.

Sagberg, F. (1999). Road accidents caused by drivers falling asleep. Accident Analysis & Prevention, 31, 639-349.

Sarter, N. B., & Woods, D. D. (1995). How in the world did we ever get into that mode? Mode error and awareness in supervisory control. Human Factors, 37(1), 5-19.

Sarter, N. B., Woods, D. D., & Billings, C. E. (1997). Automation surprises. In G. Salvendy (Ed.), Handbook of Human Factors and Ergonomics (2nd Edition ed., pp. 1926-1943). New York: Wiley.

Sayed, R., & Eskandarian, A. (2001). Unobtrusive drowsiness detection by neural network learning of driver steering. Proceedings of the Institution of Mechanical Engineers, 215(D), 969-975.

Sgambati, F. (2012). Driver Drowsiness Detection: SAE International.

Sheridan, T. B. (1992a). Musings on telepresence and virtual presence. Presence: Teleoperators and Virtual Environments, 1(1), 120-126.

Sheridan, T. B. (1992b). Telerobotics, automation, and human supervisory control. Cambirdge: MIT Press.

Sheridan, T. B., Vámos, T., & Aida, S. (1983). Adapting automation to man, culture and society. Automatica, 19, 605-612.

Shuyan, H. & Gangtie, Z. (2009). Driver drowsiness detection with eyelid related parameters by Support Vector Machine. Expert Systems with Applications, 36, 7651-7658.

Smith, A. P. & C. Miles (1986a). "The effect of lunch on cognitive vigilence tasks", Ergonomics, 29(10), 1251-1261.

Smith, A. P. & C. Miles (1986b). "Effects of lunch on selective and sustained attention", Neuropsychobiology, 16, 117-120.

Steiger, A. (2002). Sleep and the hypothalamo–pituitary–adrenocortical system. Sleep Medicine Reviews, 6(2), 125-138.

Stern, J. M., & Engel, J. (2005). Atlas of EEG patterns. USA: Lippincott Williams and Wilkins.

Sousanis, J. (2011). World Vehicle Population Tops 1 Billion Units. Retrieved January 3, 2016 from http://wardsauto.com/news-analysis/world-vehicle-population-tops-1-billion-units

Takei, Y., & Furukawa, Y. (2005). Estimate of driver's fatigue through steering motion. IEEE International Conference on Systems, Man and Cybernetics, 2, 1765-1770.

Tamatsu, Y., & Nitanda, N. (2014). Application of Image Recognition Technology to Vehicles. Encyclopedia of Automotive Engineering, 1, 1-8.

Tanaka, H., Hayashi, M., & Hori, T. (1996). Statistical features of hypnagogic EEG measured by a new scoring system. Sleep, 19, 731-738.

Teplan, M. (2002). Fundamentals Of Eeg Measurement. Measurement Science Review, 2(2), 1-11.

The-International-Council-on-Clean-Transportation. (2014). European Vehicle Market Statistics. Retrieved December 31, 2015 from http://www.theicct.org/sites/default/files/publications/EU_pocketbook_2014.pdf

Thiffault, P., & Bergeron, J. (2003). Monotony of road environment and driver fatigue: a simulator study. Accident Analysis & Prevention, 35(3), 381-391.

Thomas, M., Thorne, D., Sing, N., Redmond, T., Balkin, T., Wesensten, N., . . . Belenky, G. (1998). The relationship between driving accidents and microsleep during cumulative sleep deprivation. Journal of Sleep Research, 7(2), 275.

Thompson, W. T., Lopez, N., Hickey, P., DaLuz, C., Caldwell, J. L., & Tvaryanas, A. P. (2006). Effects of shift work and sustained operations: Operator performance in remotely piloted aircraft (OPREPAIR).

Thorpy, M., & Yager, J. (1991). The encyclopedia of sleep and sleep disorders. New York: Facts on File.

Tune, G. S. (1969). Sleep and wakefulness in 509 normal human adults. British Journal of Medical Psychology, 42, 75-80.

UoLDS. (2012). University of Leeds Driving Simulator. Retrieved February 4, 2012 from http://www.uolds.leeds.ac.uk/facility/facility-images/

Vitaterna, M. H., Takahashi, J. S., & Turek, F. W. (2001). Overview of circadian rhythms. Alcohol Research Health, 25(2), 85-93.

Volvo-Car-Group. (2014). Volvo Car Group's first self-driving Autopilot cars test on public roads around Gothenburg. Retrieved May 30, 2014 from https://www.media.volvocars.com/global/en-gb/media/pressreleases/145619/volvo-car-groups-first-self-driving-autopilot-cars-test-on-public-roads-around-gothenburg

Vuckovic, A., Radivojevic, V., Chen, A. C. N., & Popovic, D. (2002). Automatic recognition of alertness and drowsiness from EEG by an artificial neural network. Medical Engineering & Physics, 24, 349-360.

Wagstaff, K., Cardie, C., Rogers, S., & Schroedl, S. (2001). Constrained K-means Clustering with Background Knowledge. Paper presented at the Eighteenth International Conference on Machine Learning.

Wakita, T., Ozawa, K., Miyajima, C., Igarashi, K., Itou, K., Takeda, K., & Itakura, F. (2006). Driver identification using driving behavior signals. IEICE Transactions on Information and Systems, E89-D(3), 1188-1194.

Weinger, M. B. (1999). Vigilance, Boredom, and Sleepiness. [journal article]. Journal of Clinical Monitoring and Computing, 15(7), 549-552.

Wells, A. S., & Read N. W. (1996). Influences of fat, energy and time of day on mood and performance. Physiological Behaviour, 59, 1069–1076.

Wells, A. S., Read N. W., and Craig, A.(1995). Influences of dietary and intraduodenal lipid on alertness, mood, and sustained concentration. British Journal of Nutrition, 74, 115–123.

Wickens, C. D. (1994). Designing for situation awareness and trust in automation. Paper presented at the Proceedings of IFAC Integrated Systems Engineering, London.
Wiener, E. L. (1998). Cockpit Automation. In E. L. Wiener & D. C. Nagel (Eds.), Human Factors In Aviation (pp. 433-461). San Diego, CA: Academic Press.
Wierwille, W. W., Lewin, M. G., & Fairbanks, R. J. (1996). Research on Vehicle-Based Driver StatusPerformance Monitoring PART III. Retrieved October 16, 2014 from Blacksburg, Virginia, USA: http://ntl.bts.gov/lib/5000/5900/5912/827.pdf
Wierwille, W. W., Wreggit, S. S., Kirn, C. L., Ellsworth, L. A., & Fairbanks, R. J. (1994). Research on Vehicle-Based Driver Status/Performance Monitoring; Development, Validation, and Refinement of Algorithms for Detection of Driver Drowsiness (Report No. ISE 94-04, NHTSA Report No. DOT HS 808 247). Retrieved September 20, 2014 from http://ntl.bts.gov/lib/jpodocs/repts_te/9006.pdf
Wilkinson, R. T. & Colquhoun, W. P. (1968). Interaction of Alcohol with incentive and sleep deprivation. Journal of experimental psychology, 76, 623-629.
Williamson, A., Friswell, R., Olivier, J., & Grzebieta, R. (2014). Are drivers aware of sleepiness and increasing crash risk while driving? Accident Analysis & Prevention, 70, 225-234.
Woods, D. (1989). The effects of automation on human's role: Experience from non-aviation industries. In S. Norman & H. Orlady (Eds.), Flight deck automation: Promises and realities (pp. 61-85). Moffet Field, CA: NASA-Ames Research Center.
World-Health-Organization. (2009). Global Status Report On Road Safety
Time For Action. Retrieved March 1, 2013 from Switzerland: http://apps.who.int/iris/bitstream/10665/44122/1/9789241563840_eng.pdf
Wu, J., & Chen, T. (2008). Development of a drowsiness warning system based on the fuzzy logic images analysis. Expert Systems with Applications, 34(2), 1556-1561.
XSens. (ND). Overview of all Xsens Products. Retrieved January 11, 2016 from https://www.xsens.com/products/
Yang, C. M., Han, H. Y., Yang, M. H., Su, W. C., & Lane, T. (2010). What subjective experiences determine the perception of falling asleep during sleep onset period? Consciousness and Cognition, 19(4), 1084-1092.
Yang, G., Lin, Y., & Bhattacharya, P. (2010). A driver fatigue recognition model based on information fusion and dynamic Bayesian network. Information Sciences, 180, 1942-1954.
Yeo, M. V., Li, X., & Wilder-Smith, E. P. (2007). Characteristic EEG differences between voluntary recumbent sleep onset in bed and involuntary sleep onset in a driving simulator. Clinical Neurophysiology, 118(6), 1315-1323.
Yeo, M. V. M., Li, X., Shen, K., & Wilder-Smith, E. P. V. (2009). Can SVM be used for automatic EEG detection of drowsiness during car driving? Safety Science, 47(1), 115-124.
Young, M. S., & Stanton, N. A. (2002). Malleable Attentional Resources Theory: A new explanation for the effects of mental underload on performance. Human Factors, 44(3), 365-375.
Zhao, C., Zhao, M., Liu, J., & Zheng, C. (2012). Electroencephalogram and electrocardiograph assessment of mental fatigue in a driving simulator. Accident Analysis & Prevention, 45, 83-90.

Zhao, X., & Rong, J. (2013). Computational Intelligence for Traffic and Mobility (Vol. 8): Atlantis Press.

Zilberg, E., Xu, Z. M., Burton, D., Karrar, M., & Lal, S. (2007). Methodology and initial analysis results for development of non-invasive and hybrid driver drowsiness detection systems. Paper presented at the The 2nd International Conference on Wireless Broadband and Ultra Wideband Communications.

Zilberg, E., Xu, Z. M., Burton, D., Karrar, M., & Lal, S. (2009). Statistical validation of physiological indicators for noninvasive and hybrid drowsiness detection system. African Journal of Information and Communication Technology, 5(2), 75-83.

Bibliography

Abe, G., & Richardson, J. (2005). The influence of alarm timing on braking response and driver trust in low speed driving. Safety Science, 43(9), 639-654. doi:10.1016/j.ssci.2005.04.006

Abe, T., Nonomura, T., Komada, Y., Asaoka, S., Sasai, T., Ueno, A., & Inoue, Y. (2011). Detecting deteriorated vigilance using percentage of eyelid closure time during behavioral maintenance of wakefulness tests. Int J Psychophysiol, 82(3), 269-274. doi:10.1016/j.ijpsycho.2011.09.012

Ahlstrom, C., Kircher, K., Fors, C., Dukic, T., Patten, C., & Anund, A. (2012). Measuring driver impairments: sleepiness, distraction, and workload. IEEE Pulse, 3(2), 22-30. doi:10.1109/MPUL.2011.2181020

Ahlstrom, C., Nystrom, M., Holmqvist, K., Fors, C., Sandberg, D., Anund, A., . . . Akerstedt, T. (2013). Fit-for-duty test for estimation of drivers' sleepiness level:Eye movements improve the sleep/wake predictor. Transportation Research Part C, 26, 20-32.

Ahlstrom, C., Nyström, M., Holmqvist, K., Fors, C., Sandberg, D., Anund, A., . . . Åkerstedt, T. (2013). Fit-for-duty test for estimation of drivers' sleepiness level: Eye movements improve the sleep/wake predictor. Transportation Research Part C: Emerging Technologies, 26, 20-32. doi:10.1016/j.trc.2012.07.008

Akella, M., Bang, C., Beutner, R., Delmelle, E., Batta, R., Blatt, A., . . . Wilson, G. (2003). Evaluating the reliability of automated collision notification systems. Accident Analysis & Prevention, 35, 349-360. doi:0.1016/S0001-4575(02)00010-6

Akella, M. R., Bang, C., Beutner, R., Delmelle, E. M., Batta, R., Blatt, A., . . . Wilson, G. (2003). Evaluating the reliability of automated collision notification systems. Accident Analysis & Prevention, 35, 349-360. doi:0.1016/S0001-4575(02)00010-6

Akerstedt, T., & Landstrom, U. (1998). Work place countermeasures of night shift fatigue. Industrial Ergonomics, 21, 167-178.

Amditis, A., Andreone, L., Pagle, K., Markkula, G., Deregibus, E., Romera Rue, M., . . . De Gloria, A. (2010). Towards the Automotive HMI of the Future: Overview of the AIDE-Integrated Project Results. IEEE Transactions on Intelligent Transportation Systems, 11(3), 567-578. doi:10.1109/tits.2010.2048751

Anund, A. (2010). Perception of sleepiness before falling asleep. Sleep Medicine, 11, 743-744. doi:0.1016/j.sleep.2010.06.001

Anund, A., & Akerstedt, T. (2010). Perception of sleepiness before falling asleep. Sleep Med, 11(8), 743-744. doi:10.1016/j.sleep.2010.06.001

Bagdadi, O., & Varhelyi, A. (2013). Development of a method for detecting jerks in safety critical events. Accid Anal Prev, 50, 83-91. doi:10.1016/j.aap.2012.03.032

Baulk, S. D., Biggs, S. N., Reid, K. J., van den Heuvel, C. J., & Dawson, D. (2008). Chasing the silver bullet: measuring driver fatigue using simple and complex tasks. Accid Anal Prev, 40(1), 396-402. doi:10.1016/j.aap.2007.07.008

Bellet, T., Mayenobe, P., Bornard, J.-C., Gruyer, D., & Claverie, B. (2012). A computational model of the car driver interfaced with a simulation platform for future Virtual Human Centred Design applications: COSMO-SIVIC.

Engineering Applications of Artificial Intelligence, 25(7), 1488-1504. doi:10.1016/j.engappai.2012.05.010

Benderius, O., & Markkula, G. (2014). Evidence for a fundamental property of steering. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 58(1), 884-888. doi:10.1177/1541931214581186

Benedetto, S., Pedrotti, M., Minin, L., Baccino, T., Re, A., & Montanari, R. (2011). Driver workload and eye blink duration. Transportation Research Part F: Traffic Psychology and Behaviour, 14(3), 199-208. doi:10.1016/j.trf.2010.12.001

Berg, J. V. D. (2006). Sleepiness and Head Movements. Industrial Health, 44, 564-576.

Bishop, C. M. (2006). Pattern Recognition and Machine Learning. Cambridge, UK: Springer.

Blais, A., & Mertz, D. (2001). An introduction to neural networks Pattern learning with the back-propagation algorithm. Retrieved from https://www.ibm.com/developerworks/library/l-neural/l-neural-pdf.pdf

Blanco, S., Quian Quiroga, R., Rosso, O. A., & Kochen, S. (1995). Time-frequency analysis of electroencephalogram series. Physical Review E, 51(3), 2624-2631.

Boer, E. R., Rakauskas, M. E., Ward, N. J., & Goodrich, M. A. Steering Entropy Revisited. Paper presented at the Third International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design.

Bomquist, G. (1986). A Utility Maximization Model Of Driver Traffic Safety Behavior. Accident Analysis & Prevention, 18(5), 371-375.

Bonzani, I., & Mussone, L. (2004). Modeling the driver's behavior on second-order macroscopic models of vehicular traffic flow. Mathematical and Computer Modelling, 40(9-10), 1065-1073. doi:10.1016/j.mcm.2003.09.042

Borghini, G., Astolfi, L., Vecchiato, G., Mattia, D., & Babiloni, F. (2014). Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness. Neurosci Biobehav Rev, 44, 58-75. doi:10.1016/j.neubiorev.2012.10.003

Brookhuis, K. A., & Waard, D. d. (2010). Monitoring drivers mental workload in driving simulators using physiological measures. Accident Analysis & Prevention, 42, 898-903. doi:0.1016/j.aap.2009.06.001

Cacciabue, P. C., & Carsten, O. (2010). A simple model of driver behaviour to sustain design and safety assessment of automated systems in automotive environments. Appl Ergon, 41(2), 187-197. doi:10.1016/j.apergo.2009.03.008

Caldwell, J. A., Prazinko, B., & Caldwell, J. L. (2003). Body posture affects electroencephalographic activity and psychomotor vigilance task performance in sleep-deprived subjects. Clinical Neurophysiology, 114, 23-31.

Carsten, O. M., & Tate, F. N. (2005). Intelligent speed adaptation: accident savings and cost-benefit analysis. Accid Anal Prev, 37(3), 407-416. doi:10.1016/j.aap.2004.02.007

Carsten, O. M. J., & Tate, F. N. (2005). Intelligent speed adaptation accident savings and cost–benefit analysis. Accident Analysis and Prevention, 37, 407-416. doi:0.1016/j.aap.2004.02.007

Casucci, M., Marchitto, M., & Cacciabue, P. C. (2010). A numerical tool for reproducing driver behaviour Experiments and predictive simulations. Applied Ergonomics, 41, 198-210. doi:10.1016/j.apergo.20

Chong, L., Abbas, M. M., Medina Flintsch, A., & Higgs, B. (2013). A rule-based neural network approach to model driver naturalistic behavior in traffic. Transportation Research Part C: Emerging Technologies, 32, 207-223. doi:10.1016/j.trc.2012.09.011

Colquhoun, D. (2014). An investigation of the false discovery rate and the misinterpretation of p-values. R Soc Open Sci, 1(3), 140216. doi:10.1098/rsos.140216

National-Transport-Commission. (2016). Heavy vehicle driver fatigue data. Retrieved from Melbourne, Australia: http://www.ntc.gov.au/Media/Reports/(792A30B5-8CE0-420A-A8F6-79723FE802F6).pdfCommittee, O.-R. A. V. S. (2014). Taxonomy and definitions for terms related to on-Road motor vehicle automated driving systems. Retrieved from http://standards.sae.org/j3016_201401/

Corfitsen, M. T. (1999a). Fatigue among young male night-time car drivers is there a risk-taking group. Safety Science, 33, 47-57.

Corfitsen, M. T. (1999b). 'Fatigue' among young male night-time car drivers: is there a risk-taking group? Safety Science, 33, 47-57.

D'Orazio, T., Leo, M., Guaragnella, C., & Distante, A. (2007). A visual approach for driver inattention detection. Pattern Recognition, 40(8), 2341-2355. doi:10.1016/j.patcog.2007.01.018

Dahlgren, A., Kecklund, G., & Åkerstedt, T. (2006). Overtime work and its effects on sleep, sleepiness, cortisol and blood pressure in an experimental field study. Scandinavian Journal of Work, Environment & Health, 32(4), 318-327. doi:10.5271/sjweh.1016

Dargay, J., Gately, D., & Sommer, M. (2007). Vehicle Ownership and Income Growth, Worldwide: 1960-2030. The Energy Journal, 28(4), 143-170.

Demir, M., & Çavuşoğlu, A. (2012). A new driver behavior model to create realistic urban traffic environment. Transportation Research Part F: Traffic Psychology and Behaviour, 15(3), 289-296. doi:10.1016/j.trf.2012.01.004

Dittner, A. J., Wessely, S. C., & Brown, R. G. (2004). The assessment of fatigue. Journal of Psychosomatic Research, 56(2), 157-170. doi:10.1016/s0022-3999(03)00371-4

Endsley, M. R., & Kaber, D. B. (1999). Level of automation effects on performance, situation awareness and workload in a dynamic control task. Ergonomics, 42(3), 462-492.

Eriksson, M., & Papanikolopoulos, N. P. (2001). Driver fatigue: a vision-based approach to automatic diagnosis. Transportation Research Part C, 9, 399-413.

Flores, M. J., Armingol, J. M., & de la Escalera, A. (2009). Real-Time Warning System for Driver Drowsiness Detection Using Visual Information. Journal of Intelligent & Robotic Systems, 59(2), 103-125. doi:10.1007/s10846-009-9391-1

Forsman, P. M., Vila, B. J., Short, R. A., Mott, C. G., & Van Dongen, H. P. (2013). Efficient driver drowsiness detection at moderate levels of drowsiness. Accid Anal Prev, 50, 341-350. doi:10.1016/j.aap.2012.05.005

Fort, A., Martin, R., Jacquet-Andrieu, A., Combe-Pangaud, C., Foliot, G., Daligault, S., & Delpuech, C. (2010). Attentional demand and processing of relevant visual information during simulated driving: a MEG study. Brain Res, 1363, 117-127. doi:10.1016/j.brainres.2010.09.094

Fuller, R. (2005). Towards a general theory of driver behaviour. Accident Analysis & Prevention, 37, 461-472. doi:0.1016/j.aap.2004.11.003

Garcia, I., Bronte, S., Bergasa, L. M., Hernandez, N., Delgado, B., & Sevillano, M. (2010). Vision-based drowsiness detector for a Realistic Driving Simulator. Paper presented at the 13th International IEEE Annual Conference on Intelligent Transportation Systems, Madeira Island, Portugal.

Gast, H., Schindler, K., Rummel, C., Herrmann, U. S., Roth, C., Hess, C. W., & Mathis, J. (2011). EEG correlation and power during maintenance of wakefulness test after sleep-deprivation. Clin Neurophysiol, 122(10), 2025-2031. doi:10.1016/j.clinph.2011.03.003

Ge, H. X., Dai, S. Q., & Dong, L. Y. (2006). An extended car-following model based on intelligent transportation system application. Physica A: Statistical Mechanics and its Applications, 365(2), 543-548. doi:10.1016/j.physa.2005.08.050

Gennaro, L. D., Ferrara, M., & Bertini, M. (2001). The Boundary Between Wakefulness And Sleep: Quantitative Electroencephalographic Changes During The Sleep Onset Period. Neuroscience, 107(1), 1-11.

Guan, Y., Zhang, N., Zhu, J., & Yang, X. (2010). Modeling On-ramp Capacity with Driver Behavior Variation. Journal of Transportation Systems Engineering and Information Technology, 10(1), 122-127. doi:10.1016/s1570-6672(09)60028-3

Gunzelmann, G., Richard Moore, L., Salvucci, D. D., & Gluck, K. A. (2011). Sleep loss and driver performance: Quantitative predictions with zero free parameters. Cognitive Systems Research, 12(2), 154-163. doi:10.1016/j.cogsys.2010.07.009

Häkkänen, H., Summala, H., Partinen, M., Tiihonen, M., & Silvo, J. (1999). Blink Duration as an Indicator of Driver Sleepiness in Professional Bus Drivers. Sleep, 22(6), 798-802.

Hancock, P. A., & Verwey, W. B. (1997). Fatigue, Workload And Adaptive Driver Systems'. Accident Analysis & Prevention, 29(4), 495-506.

Hassan, H. M., & Abdel-Aty, M. A. (2011). Analysis of drivers behavior under reduced visibility conditions using a Structural Equation Modeling approach. Transportation Research Part F, 14, 614-625. doi:0.1016/j.trf.2011.07.002

Heitmann, A., Guttkuhn, R., Aguirre, A., Trutsche, U., & Moore-Ede, M. Technologies For The Monitoring And Prevention Of Driver Fatigue. Paper presented at the First International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design.

Herrmann, U. S., Hess, C. W., Guggisberg, A. G., Roth, C., Gugger, M., & Mathis, J. (2010). Sleepiness is not always perceived before falling asleep in healthy, sleep-deprived subjects. Sleep Med, 11(8), 747-751. doi:10.1016/j.sleep.2010.03.015

Hoogendoorn, R., Hoogendoorn, S., Brookhuis, K., & Daamen, W. (2011). Adaptation Longitudinal Driving Behavior, Mental Workload, and Psycho-Spacing Models in Fog. Transportation Research Record: Journal of the Transportation Research Board, 2249, 20-28. doi:10.3141/2249-04

Hoogendoorn, R. G., van Arem, B., & Brookhuis, K. A. (2013). Longitudinal Driving Behavior in Case of Emergency Situations: An Empirically Underpinned Theoretical Framework. Procedia - Social and Behavioral Sciences, 80, 341-369. doi:10.1016/j.sbspro.2013.05.020

Hoogendoorn, R. G., van Arem, B., & Hoogendoorn, S. P. (2012). A Neurofuzzy Approach to Modeling Longitudinal Driving Behavior and Driving Task

Complexity. International Journal of Vehicular Technology, 2012, 1-12. doi:10.1155/2012/807805
Horne, J. A., & Baulk, S. D. (2004). Awareness of sleepiness when driving. Psychophysiology, 41(1), 161-165. doi:10.1046/j.1469-8986.2003.00130.x
Hurwitz, D. S., Wang, H., Knodler, M. A., Ni, D., & Moore, D. (2012). Fuzzy sets to describe driver behavior in the dilemma zone of high-speed signalized intersections. Transportation Research Part F: Traffic Psychology and Behaviour, 15(2), 132-143. doi:10.1016/j.trf.2011.11.003
Itoh, M., Sakami, D., & Tanaka, K. (2000). Dependence of Human Adaptation and Risk Compensation on Modification in Level of Automation for System Safety. IEEE.
Jamson, A. H., Merat, N., Carsten, O. M. J., & Lai, F. C. H. (2013). Behavioural changes in drivers experiencing highly-automated vehicle control in varying traffic conditions. Transportation Research Part C: Emerging Technologies, 30, 116-125. doi:10.1016/j.trc.2013.02.008
Jap, B. T., Lal, S., & Fischer, P. (2010). Inter-hemispheric electroencephalography coherence analysis: assessing brain activity during monotonous driving. Int J Psychophysiol, 76(3), 169-173. doi:10.1016/j.ijpsycho.2010.03.007
Jentsch, F., Barnett, J., Bowers, C. A., & Salas, E. (1999). Who Is Flying This Plane Anyway? What Mishaps Tell Us about Crew Member Role Assignment and Air Crew Situation Awareness. Human Factors: The Journal of the Human Factors and Ergonomics Society, 41(1), 1-14. doi:10.1518/001872099779577237
Ji Hyun, Y., Zhi-Hong, M., Tijerina, L., Pilutti, T., Coughlin, J. F., & Feron, E. (2009). Detection of Driver Fatigue Caused by Sleep Deprivation. IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans, 39(4), 694-705. doi:10.1109/tsmca.2009.2018634
Ji, Q. (2002). Real-Time Eye, Gaze, and Face Pose Tracking for Monitoring Driver Vigilance. Real-Time Imaging, 8(5), 357-377. doi:10.1006/rtim.2002.0279
Jun, L., Zhiqiang, S., & Long, C. (2012). Dynamic Neural Network-Based Integrated Learning Algorithm for Driver Behavior. Journal Of Transportation
Systems Engineering And Information Technology, 12(2), 34-40. doi:0.1016/S1570-6672(11)60192-X
Kar, S., Bhagat, M., & Routray, A. (2010). EEG signal analysis for the assessment and quantification of driver's fatigue. Transportation Research Part F: Traffic Psychology and Behaviour, 13(5), 297-306. doi:10.1016/j.trf.2010.06.006
Kassis, O., Katz, N., Ravid, S., & Pillar, G. (2013). double-Blind Placebo and active (caffeine) controlled study to examine the effects of the Herbal nutritional supplement Beverage "wake up" on vigilance and Function after lunch. IMAJ, 15, 487-491.
King, L. M., Nguyen, H. T., & Lal, S. K. L. (2006). Early Driver Fatigue Detection from Electroencephalography Signals using Artificial Neural Networks. Paper presented at the 28th IEEE EMBS Annual International Conference, New York City, USA.
Knapper, A., Christoph, M., Hgenzieker, M., & Brookhuis, K. A. (2015). Comparing a driving simulator to the real road regarding distracted driving speed. EJTIR, 15(2), 205-225.
Kyriakidis, M., Happee, R., & de Winter, J. C. F. (2015). Public opinion on automated driving: Results of an international questionnaire among 5000

respondents. Transportation Research Part F: Traffic Psychology and Behaviour, 32, 127-140. doi:10.1016/j.trf.2015.04.014

Lam, L. T. (2003). Factors associated with fatal and injurious car crash among learner drivers in New South Wales, Australia. Accident Analysis & Prevention, 35(3), 333-340. doi:10.1016/s0001-4575(02)00008-8

Lappi, O., Pekkanen, J., & Itkonen, T. H. (2013). Pursuit eye-movements in curve driving differentiate between future path and tangent point models. PLoS One, 8(7), e68326. doi:10.1371/journal.pone.0068326

Larue, G. S., Rakotonirainy, A., & Pettitt, A. N. (2011). Driving performance impairments due to hypovigilance on monotonous roads. Accid Anal Prev, 43(6), 2037-2046. doi:10.1016/j.aap.2011.05.023

Leandro, M. (2012). Young drivers and speed selection: A model guided by the Theory of Planned Behavior. Transportation Research Part F: Traffic Psychology and Behaviour, 15(3), 219-232. doi:10.1016/j.trf.2011.12.011

Lehtonen, E., Lappi, O., Koirikivi, I., & Summala, H. (2014). Effect of driving experience on anticipatory look-ahead fixations in real curve driving. Accid Anal Prev, 70, 195-208. doi:10.1016/j.aap.2014.04.002

Lehtonen, E., Lappi, O., Kotkanen, H., & Summala, H. (2013). Look-ahead fixations in curve driving. Ergonomics, 56(1), 34-44. doi:10.1080/00140139.2012.739205

Lemke, M. (1992). Correlation Between Eeg And Drivers Actions During Prolonged Driving Under Monotonous Conditions. Accident Analysis & Prevention, 14(1), 7-17.

Lewandowski, A., Rosipal, R., & Dorffner, G. (2013). On the Individuality of Sleep EEG Spectra. Federation of European Psychophysiology Societies, 27(3), 105-112. doi:10.1027/a000001

Lin, Y., Leng, H., Yang, G., & Cai, H. (2007). An Intelligent Noninvasive Sensor for Driver Pulse Wave Measurement. IEEE Sensors Journal, 7(5), 790-799. doi:10.1109/jsen.2007.894923

Liu, Z. (2007). Characterisation of optimal human driver model and stability of a tractor-semitrailer vehicle system with time delay. Mechanical Systems and Signal Processing, 21(5), 2080-2098. doi:10.1016/j.ymssp.2006.06.007

Lucidi, F., Russo, P. M., Mallia, L., Devoto, A., Lauriola, M., & Violani, C. (2006). Sleep-related car crashes: risk perception and decision-making processes in young drivers. Accid Anal Prev, 38(2), 302-309. doi:10.1016/j.aap.2005.09.013

Ma, R., & Kaber, D. B. (2007). Situation awareness and driving performance in a simulated navigation task. Ergonomics, 50(8), 1351-1364. doi:10.1080/00140130701318913

MacLean, A. W., Davies, D. R. T., & Thiele, K. (2002). The hazards and prevention of driving while sleepy. Sleep Medicine Reviews, 7(6), 507-521.

Mahachandra, M., Yassierli, Sutalaksana, I. Z., & Suryadi, K. (2011). Sleepiness Pattern of Indonesian Professional Driver Based on Subjective Scale and Eye Closure Activity. International Journal of Basic & Applied Sciences, 11(6), 87-96.

Maltz, M., & Shinar, D. (1999). Eye Movements of Younger and Older Drivers. Human Factors: The Journal of the Human Factors and Ergonomics Society, 41(1), 15-25. doi:10.1518/001872099779577282

Markkula, G., Benderius, O., & Wahde, M. (2014). Comparing and validating models of driver steering behaviour in collision avoidance and vehicle stabilisation.

Vehicle System Dynamics, 52(12), 1658-1680. doi:10.1080/00423114.2014.954589

Martín de Diego, I., S. Siordia, O., Crespo, R., Conde, C., & Cabello, E. (2013). Analysis of hands activity for automatic driving risk detection. Transportation Research Part C: Emerging Technologies, 26, 380-395. doi:10.1016/j.trc.2012.10.006

Masland, S. (2009). Machine Learning An Algorithm Perspective. Palmeston North, New Zealand: CRC Press.

Matthews, G., & Desmond, P. A. (2002). Task-induced fatigue states and simulated driving performance. Q J Exp Psychol A, 55(2), 659-686. doi:10.1080/02724980143000505

McBain, W. (1970). Arousal, monotony, and accidents in line driving. Jounral of Applied Psychology, 54, 509-519.

McCartt, A. T., Shabanova, V. I., & Leaf, W. A. (2003). Driving experience, crashes and traffic citations of teenage beginning drivers. Accident Analysis & Prevention, 35, 311-320. doi:0.1016/S0001-4575(02)00006-4

McCauley, P., Kalachev, L. V., Smith, A. D., Belenky, G., Dinges, D. F., & Van Dongen, H. P. (2009). A new mathematical model for the homeostatic effects of sleep loss on neurobehavioral performance. J Theor Biol, 256(2), 227-239. doi:10.1016/j.jtbi.2008.09.012

Meech, J., & Parreira, J. (2011). An interactive simulation model of human drivers to study autonomous haulage trucks. Procedia Computer Science, 6, 118-123. doi:10.1016/j.procs.2011.08.023

Merica, H., & Fortune, R. D. (2004). State transitions between wake and sleep, and within the ultradian cycle, with focus on the link to neuronal activity. Sleep Med Rev, 8(6), 473-485. doi:10.1016/j.smrv.2004.06.006

Murphy, T., Richard, M., Masaki, H., & Segalowitz, S. (2006). The effect of sleepiness on performance monitoring I know what I am doing, but do I care? Journal of Sleep Research, 15, 15-21.

Mushtaq, F. (2012). Electrophysiological Correlates of Affective Context and Risk-Taking in Human Decision-Making. (PhD), University of Leeds, Leeds, UK.

Mushtaq, F., Stoet, G., Bland, A. R., & Schaefer, A. (2013). Relative changes from prior reward contingencies can constrain brain correlates of outcome monitoring. PLoS One, 8(6), e66350. doi:10.1371/journal.pone.0066350

Mynttinen, S., Gatscha, M., Koivukoski, M., Hakuli, K., & Keskinen, E. (2010). Two-phase driver education models applied in Finland and in Austria – Do we have evidence to support the two phase models? Transportation Research Part F: Traffic Psychology and Behaviour, 13(1), 63-70. doi:10.1016/j.trf.2009.11.002

Naus, G. J. L., Ploeg, J., Van de Molengraft, M. J. G., Heemels, W. P. M. H., & Steinbuch, M. (2010). Design and implementation of parameterized adaptive cruise control: An explicit model predictive control approach. Control Engineering Practice, 18(8), 882-892. doi:10.1016/j.conengprac.2010.03.012

Niedermeye, E. (1999). The Normal EEG of the Waking Adult. In E. Niedermeyer & F. L. d. Silva (Eds.), Electroencephalography: Basic Principles, Clinical Applications and Related Fields (pp. 149-173). Baltimore MD: Lippincott Williams & Wilkins.

Nixon, M., & Aguado, A. S. (2012). Appendix 3: Principal components analysis. 525-540. doi:10.1016/b978-0-12-396549-3.00018-5

Nixon, M. S., & Aguado, A. S. (2012a). Appendix 1: Camera geometry fundamentals. 489-518. doi:10.1016/b978-0-12-396549-3.00016-1
Nixon, M. S., & Aguado, A. S. (2012b). Appendix 2: Least squares analysis. 519-523. doi:10.1016/b978-0-12-396549-3.00017-3
Nixon, M. S., & Aguado, A. S. (2012c). Introduction. 1-36. doi:10.1016/b978-0-12-396549-3.00001-x
Nixon, M. S., & Aguado, A. S. (2012d). Introduction to texture description, segmentation, and classification. 399-434. doi:10.1016/b978-0-12-396549-3.00008-2
Nordbakke, S., & Sagberg, F. (2007). Sleepy at the wheel: Knowledge, symptoms and behaviour among car drivers. Transportation Research Part F: Traffic Psychology and Behaviour, 10(1), 1-10. doi:10.1016/j.trf.2006.03.003
Ogilvie, R. D. (2001). The process of falling asleep. Sleep Med Rev, 5(3), 247-270. doi:10.1053/smrv.2001.0145
Ohno, H. (2001). Analysis and modeling of human driving behaviors using adaptive cruise control. Applied Soft Computing, 1, 237-243.
Özkan, T., & Lajunen, T. (2005). A new addition to DBQ: Positive Driver Behaviours Scale. Transportation Research Part F: Traffic Psychology and Behaviour, 8(4-5), 355-368. doi:10.1016/j.trf.2005.04.018
Papadelis, C., Kourtidou-Papadeli, C., Bamidis, P. D., Chouvarda, I., Koufogiannis, D., Bekiaris, E., & Maglaveras, N. (2006). Indicators of Sleepiness in an ambulatory EEG study of night driving. Paper presented at the 28th IEEE EMBS Annual International Conference, New York City, USA.
Paul, A., Boyle, L. N., Boer, E. R., Tippin, J., & Rizzo, M. (2005). Steering Entropy Changes As A Function Of Microsleeps. Paper presented at the Third International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design, California, USA.
Paz, A., & Peeta, S. (2009). Information-based network control strategies consistent with estimated driver behavior. Transportation Research Part B: Methodological, 43(1), 73-96. doi:10.1016/j.trb.2008.06.007
Philip, P., Sagaspe, P., Lagarde, E., Leger, D., Ohayon, M. M., Bioulac, B., . . . Taillard, J. (2010). Sleep disorders and accidental risk in a large group of regular registered highway drivers. Sleep Med, 11(10), 973-979. doi:10.1016/j.sleep.2010.07.010
Philip, P., Taillard, J., Klein, E., Sagaspe, P., Charles, A., Davies, W. L., . . . Bioulac, B. (2003). Effect of fatigue on performance measured by a driving simulator in automobile drivers. Journal of Psychosomatic Research, 55(3), 197-200. doi:10.1016/s0022-3999(02)00496-8
Prato, C. G., Toledo, T., Lotan, T., & Taubman-Ben-Ari, O. (2010). Modeling the behavior of novice young drivers during the first year after licensure. Accid Anal Prev, 42(2), 480-486. doi:10.1016/j.aap.2009.09.011
Principe, J., Gala, S., & Chang, T. (1989). Sleep Staging Automaton Based on the Theory of Evidence. IEEE Transactions on Biomedical Engineering, 36, 503-509.
Putilov, A. A., & Donskaya, O. G. (2013). Construction and validation of the EEG analogues of the Karolinska sleepiness scale based on the Karolinska drowsiness test. Clin Neurophysiol, 124(7), 1346-1352. doi:10.1016/j.clinph.2013.01.018
Ranney, T. (1994). Models Of Driving Behavior A Review Of Their Evolution. Accident Analysis & Prevention, 26(6), 733-750.

Raw, R. K., Kountouriotis, G. K., Mon-Williams, M., & Wilkie, R. M. (2012). Movement control in older adults: does old age mean middle of the road? J Exp Psychol Hum Percept Perform, 38(3), 735-745. doi:10.1037/a0026568

Raw, R. K., Wilkie, R. M., Culmer, P. R., & Mon-Williams, M. (2012). Reduced motor asymmetry in older adults when manually tracing paths. Exp Brain Res, 217(1), 35-41. doi:10.1007/s00221-011-2971-x

Rogado, E., García, J. L., Barea, R., Bergasa, L. M., & López, E. (2009). Driver Fatigue Detection System. Paper presented at the 2008 IEEE International Conference on Robotics and Biomimetics, Bangkok, Thailand.

Ropper, A., & Brown, R. (2005). Adams and Victor's Principles of Neurology (8th ed.). USA: McGraw-Hill.

ROSPA. (2001). Driver fatigue and road accidents Statistics. Retrieved from

Salvucci, D. D., & Liu, A. (2002). The time course of a lane change Driver control and eye-movement behavior. Transportation Research Part F, 5, 123-132.

Sandberg, D., Akerstedt, T., Anund, A., Kecklund, G., & Wahde, M. (2011). Detecting Driver Sleepiness Using Optimized Nonlinear Combinations of Sleepiness Indicators. IEEE Transactions on Intelligent Transportation Systems, 12(1), 97-108. doi:10.1109/tits.2010.2077281

Sanei, S., & Chambers, J. A. (2007). EEG Signal Processing. West Sussex, England: John Wiley & Sons.

Sato, T., & Akamatsu, M. (2008). Modeling and prediction of driver preparations for making a right turn based on vehicle velocity and traffic conditions while approaching an intersection. Transportation Research Part F: Traffic Psychology and Behaviour, 11(4), 242-258. doi:10.1016/j.trf.2007.11.002

Schier, M. A. (2000). Changes in EEG alpha power during simulated driving a demonstration. International Journal of Psychophysiology, 37, 155-162.

Schmidt, E. A., Schrauf, M., Simon, M., Fritzsche, M., Buchner, A., & Kincses, W. E. (2009). Drivers misjudgement of vigilance state during prolonged monotonous daytime driving. Accident Analysis & Prevention, 41, 1087-1093. doi:0.1016/j.aap.2009.06.007

Schubert, R., Tangermann, M., Haufe, S., Sannelli, C., Simon, M., Schmidt, E. A., . . . Curio, G. (2008). Parieto-occipital alpha power indexes distraction during simulated car driving. International Journal of Psychophysiology, 69(3), 214. doi:10.1016/j.ijpsycho.2008.05.033

Shakouri, P., Ordys, A., & Askari, M. R. (2012). Adaptive cruise control with stop&go function using the state-dependent nonlinear model predictive control approach. ISA Trans, 51(5), 622-631. doi:10.1016/j.isatra.2012.05.001

Shen, K.-Q., Li, X.-P., Ong, C.-J., Shao, S.-Y., & Wilder-Smith, E. P. V. (2008). EEG-based mental fatigue measurement using multi-class support vector machines with confidence estimate. Clinical Neurophysiology, 119, 1524-1533. doi:0.1016/j.clinph.2008.03.012

Sheridan, T. B., & Parasuraman, R. (2005). Human-Automation Interaction. Reviews of Human Factors and Ergonomics, 1(1), 89-129. doi:10.1518/155723405783703082

Sheu, J.-B. (2008). A quantum mechanics-based approach to model incident-induced dynamic driver behavior. Physica D: Nonlinear Phenomena, 237(13), 1800-1814. doi:10.1016/j.physd.2008.01.023

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Driessche, G. v. d., . . . Hassabis, T. G. D. (2016). Mastering the game of Go with deep neural networks and tree search. NATURE, 529, 484-489. doi:0.1038/nature16961

Simon, M., Schmidt, E. A., Kincses, W. E., Fritzsche, M., Bruns, A., Aufmuth, C., . . . Schrauf, M. (2011). EEG alpha spindle measures as indicators of driver fatigue under real traffic conditions. Clin Neurophysiol, 122(6), 1168-1178. doi:10.1016/j.clinph.2010.10.044

Smilek, D., Carriere, S. A., & Cheyne, J. A. (2010). Out of Mind, Out of Sight Eye Blinking as Indicator and Embodiment of Mind Wandering. Psychology Science, 21(6), 786-9. Doi: 10.1177/0956797610368063.

Smith, A., Leekam, S., Ralph, A., & McNeill, G. (1988). The influence of Meal Composition on Post-Lunch Changes in Performance Efficiency and Modd. Appetite, 10, 195-203.

Smith, S., Carrington, M., & Trinder, J. (2005). Subjective and predicted sleepiness while driving in young adults. Accid Anal Prev, 37(6), 1066-1073. doi:10.1016/j.aap.2005.06.008

Sommer, D., & Golz, M. (2010). Evaluation of PERCLOS based Current Fatigue Monitoring Technologies. Paper presented at the 32nd Annual International Conference of the IEEE EMBS, Buenos Aires, Argentina.

Sonnleitner, A., Simon, M., Kincses, W. E., Buchner, A., & Schrauf, M. (2012). Alpha spindles as neurophysiological correlates indicating attentional shift in a simulated driving task. Int J Psychophysiol, 83(1), 110-118. doi:10.1016/j.ijpsycho.2011.10.013

Sperling, D., & Gordon, D. (2010). Two Billion Cars: Oxford University Press.

St Hilaire, M. A., Sullivan, J. P., Anderson, C., Cohen, D. A., Barger, L. K., Lockley, S. W., & Klerman, E. B. (2013). Classifying performance impairment in response to sleep loss using pattern recognition algorithms on single session testing. Accid Anal Prev, 50, 992-1002. doi:10.1016/j.aap.2012.08.003

Stutts, J. C., Wilkins, J. W., Scott Osberg, J., & Vaughn, B. V. (2003). Driver risk factors for sleep-related crashes. Accident Analysis & Prevention, 35(3), 321-331. doi:10.1016/s0001-4575(02)00007-6

Stutts, J. C., Wilkins, J. W., & Vaughn, B. V. (1999). Why Do People Have Drowsy Driving Crashes?

Sugiyama, K., Nakano, T., Yamamoto, S., Ishihara, T., Fujii, H., & Akutsu, E. (1996). Method of detecting drowsiness level by utilizing blinking duration. JSAE REVIEW, 17, 159-163.

Summala, H. (1996). Accident Risk And Driver Behaviour. Safety Science, 22(1-3), 103-117.

Sung, E. J., Min, B. C., Kim, S. C., & Kim, C. J. (2005). Effects of oxygen concentrations on driver fatigue during simulated driving. Appl Ergon, 36(1), 25-31. doi:10.1016/j.apergo.2004.09.003

Susmakova, K. (2004). Human Sleep and Sleep EEG. Measurement Science Review, 4(2), 59-74.

Svensson, U. (2004). Blink behaviour based drowsiness detection – method development and validation. (Master of Science), Linköping University, Linköping.

Tang, T. Q., Li, C. Y., & Huang, H. J. (2010). A new car-following model with the consideration of the driver's forecast effect. Physics Letters A, 374(38), 3951-3956. doi:10.1016/j.physleta.2010.07.062

Tango, F., Minin, L., Tesauri, F., & Montanari, R. (2010). Field tests and machine learning approaches for refining algorithms and correlations of driver's model parameters. Appl Ergon, 41(2), 211-224. doi:10.1016/j.apergo.2009.01.010

Tatum, W., Husain, A., Bendadis, S., & Kaplan, P. (2008). Handbook of EEG Interpretation. USA: Demos.

Tirunahari, V. L., Zaidi, S. A., Sharma, R., Skurnick, J., & Ashtyani, H. (2003). Microsleep and sleepiness: a comparison of multiple sleep latency test and scoring of microsleep as a diagnostic test for excessive daytime sleepiness. Sleep Medicine, 4(1), 63-67. doi:10.1016/s1389-9457(02)00250-2

Tran, C., Doshi, A., & Trivedi, M. M. (2012). Modeling and prediction of driver behavior by foot gesture analysis. Computer Vision and Image Understanding, 116(3), 435-445. doi:10.1016/j.cviu.2011.09.008

Trutschel, U., Sirois, B., Sommer, D., Golz, M., & Edwards, D. Perclos An Alertness Measure Of The Past. Paper presented at the Sixth International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design.

Vansteensel, M. J., Hermes, D., Aarnoutse, E. J., Bleichner, M. G., Schalk, G., van Rijen, P. C., . . . Ramsey, N. F. (2010). Brain-computer interfacing based on cognitive control. Ann Neurol, 67(6), 809-816. doi:10.1002/ana.21985

Verwey, W. B., & Zaidel, D. M. (1999). Preventing drowsiness accidents by an alertness maintenance device. Accident Analysis & Prevention, 31, 199-211.

Vogel, K. (2003). A comparison of headway and time to collision as safety indicators. Accident Analysis & Prevention, 35(3), 427-433. doi:10.1016/s0001-4575(02)00022-2

Watling, C. N. (2014). Sleepy driving and pulling over for a rest: Investigating individual factors that contribute to these driving behaviours. Personality and Individual Differences, 56, 105-110. doi:10.1016/j.paid.2013.08.031

Watling, C. N., Armstrong, K. A., Obst, P. L., & Smith, S. S. (2014). Continuing to drive while sleepy: the influence of sleepiness countermeasures, motivation for driving sleepy, and risk perception. Accid Anal Prev, 73, 262-268. doi:10.1016/j.aap.2014.09.021

Wenzel, T. P., & Ross, M. (2005). The effects of vehicle model and driver behavior on risk. Accid Anal Prev, 37(3), 479-494. doi:10.1016/j.aap.2004.08.002

Wilkie, R., & Wann, J. (2003). Controlling steering and judging heading: Retinal flow, visual direction, and extraretinal information. Journal of Experimental Psychology: Human Perception and Performance, 29(2), 363-378. doi:10.1037/0096-1523.29.2.363

Wilkie, R. M., Johnson, R. L., Culmer, P. R., Allen, R., & Mon-Williams, M. (2012). Looking at the task in hand impairs motor learning. J Neurophysiol, 108(11), 3043-3048. doi:10.1152/jn.00440.2012

Wilkie, R. M., Kountouriotis, G. K., Merat, N., & Wann, J. P. (2010). Using vision to control locomotion: looking where you want to go. Exp Brain Res, 204(4), 539-547. doi:10.1007/s00221-010-2321-4

Wilkie, R. M., & Wann, J. P. (2003). Eye-movements aid the control of locomotion. J Vis, 3(11), 677-684. doi:10.1167/3.11.3

Xin, W., Hourdos, J., Michalopoulos, P., & Davis, G. (2008). The Less-Than-Perfect Driver: A Model of Collision-Inclusive Car-Following Behavior. Transportation Research Record: Journal of the Transportation Research Board, 2088, 126-137. doi:10.3141/2088-14

Yamakoshi, T., Rolfe, P., Yamakoshi, Y., & Hirose, H. (2009). A novel physiological index for Driver's Activation State derived from simulated monotonous driving studies. Transportation Research Part C: Emerging Technologies, 17(1), 69-80. doi:10.1016/j.trc.2008.09.002

Yu, L., & Shi, Z. (2008). Nonlinear analysis of an extended traffic flow model in ITS environment. Chaos, Solitons & Fractals, 36(3), 550-558. doi:10.1016/j.chaos.2007.07.076

Yu, X. (2009). Real-time Nonintrusive Detection of Driver Drowsiness. Retrieved from Minnesota, USA: http://conservancy.umn.edu/bitstream/handle/11299/97650/CTS%2009-15.pdf?sequence=1

Zhao, C., Zheng, C., Zhao, M., Tu, Y., & Liu, J. (2011). Multivariate autoregressive models and kernel learning algorithms for classifying driving mental fatigue based on electroencephalographic. Expert Systems with Applications, 38(3), 1859-1865. doi:10.1016/j.eswa.2010.07.115

Zhu, H. B., & Dai, S. Q. (2008). Analysis of car-following model considering driver's physical delay in sensing headway. Physica A: Statistical Mechanics and its Applications, 387(13), 3290-3298. doi:10.1016/j.physa.2008.01.103

Appendix A Consent Form given to the participants during the experiments

Study Title: Performance while driving in young drivers before and after lunch

Principal Investigator (s): Pablo Puente Guillen

Prof. Anthony Cohn (supervisor) a.g.cohn@leeds.ac.uk
Prof. Oliver Carsten (supervisor) O.M.J.Carsten@its.leeds.ac.uk
Dr. Richard Wilkie (supervisor) r.m.wilkie@leeds.ac.uk
Dr. Faisal Mushtaq (supervisor) f.mushtaq@leeds.ac.uk

The purpose of this form is to provide you with information regarding the confidentiality of data and your right to withdraw from this study at any time. This form will also ensure that you are happy to participate in this study and fully understand what is involved.

Confidentiality
Any information you provide that can be traced back to you will remain strictly confidential, and will be disclosed only with your permission or as required by law. If information collected in this study is published in scientific journals, where necessary, participants will be referred to by an anonymous code only. The terms of the data protection Act 1988 will be adhered to and information will be securely stored.

Right to Withdraw
You are free to withdraw from the study at any point in time without consequences. You may stop participation during the testing period, or contact a member of the research team to request that your data be destroyed at a later date (you will be provided with contact details before the experiment).

Understanding and Consent Participant code (leave blank):

1. Have you read and understood the implications of this study and agree to participate? (YES/NO)
2. Have you had the opportunity to ask any questions? (YES/NO)
3. Do you understand that you have the right to withdraw at any time? (YES/NO)
4. Do you understand that all information gathered will be kept confidential? (YES/NO)
5. Do you grant permission for your data to be used in research reports on the understanding that your anonymity will be maintained? If you indicate NO your data will not be used outside of the practical. (YES/NO)

Name (print): ___________________ Signed: ________________________

Date: __________________________

Appendix B Debrief Information Sheet (post experiment – after both driving tasks have been completed)

Study Title: Performance while driving in young drivers before and after lunch

Principal Investigator (s): Pablo Puente Guillen

Email: scppg@leeds.ac.uk

Thank you for taking part in this study. Your participation is greatly appreciated.

The aims of this study were:

- To investigate the increase in sleepiness through analysis of the changes in brain wave activity in two different conditions: before and after lunch.
- Investigate changes in physiological behaviour during the driving tasks through analysis of the video monitoring of the participants.

The data you provided will contribute towards a wider investigation of crashes when people fall asleep while driving. If you are unsure about ay aspect of the study, or wish to withdraw your data from the experiment at a later data, please do not hesitate to contact the principle investigator. If you have any questions about the experiment the researcher can answer those questions or in a later moment using the details below:

Prof. Anthony Cohn (supervisor) a.g.cohn@leeds.ac.uk
Prof. Oliver Carsten (supervisor) O.M.J.Carsten@its.leeds.ac.uk
Dr. Richard Wilkie (supervisor) r.m.wilkie@leeds.ac.uk
Dr. Faisal Mushtaq (supervisor) f.mushtaq@leeds.ac.uk

Appendix C Instruction Sheet for the participants

Please read this instruction sheet to familiarise yourself with what will be required of you during the experiment.

You will complete a driving task in a night scenario in the static driving simulator in the Physics Research Deck.

During the task you will have to maintain the same lane. You will have to maintain the same speed during the whole driving task as well.

You will have a 5 minutes training session to familiarize yourself with the driving simulator environment.

Thank you!

Appendix D Karolinska Sleepiness Scale test

How sleepy do you feel at this moment? (Circle your answer)
1 = extremely alert
2 = very alert
3 = alert
4 = rather alert
5 = neither alert nor sleepy
6 = some signs of sleepiness
7 = sleepy, but no effort to remain awake
8 = sleepy, some effort to keep alert
9 = very sleepy, great effort to keep alert, fighting sleep

Appendix E Screening Questionnaire for participants

Please complete the following questions for screening purposes. Your responses will be kept confidential and your identity will not be disclose in any case whatsoever.

Age: ____________________

Gender: Male Female

Weight: ________________

Height: _________________

Are you Right-handed Left handed

To the best of your knowledge, do you have any sleep disorder?

Yes No

To the best of you knowledge, do you have any visual disorder?

Yes No

To the best of you knowledge, do you have any motor disorder?

Yes No

To the best of you knowledge, do you have any auditory disorder?

Yes No

Normally, how many hours do you sleep every night? _________________________

How often you take naps during the afternoon?

Never Rarely Sometimes Frequently Every day

How many cups of coffee you take per day:

Do you regularly drink alcohol?

Never Rarely Sometimes Frequently Every day

Years of driving experience: ____________________

Last time you drove: ______________________

How many kilometres or miles you drive per week? ___________________

How often do you drive in a week?

1 day 2-3 days 4-5 days More than 5 days

What type of driver would you consider yourself to be?

Passive Aggressive Cautious Adventurous

What you ate for lunch (only during the post lunch driving session)? ______

Appendix F The Epworth Sleepiness Scale (ESS)

How likely are you to doze off or fall asleep in the following situations, in contrast to feeling just tired? This refers to your usual way of life in recent times. Even if you have not done some of these things recently try to work out how they would have affected you. Use the following scale to choose the most appropriate number for each situation:

0 = would never doze

1 = slight chance of dozing

2 = moderate chance of dozing

3 = high chance of dozing

SITUATION CHANCE OF DOZING (0–3)

Sitting and reading

Watching television

Sitting inactive in a public place (e.g. a theatre or meeting)

As a passenger in a car for an hour without a break

Lying down to rest in the afternoon when circumstances permit

Sitting and talking to someone

Sitting quietly after a lunch without alcohol

In a car, while stopped for a few minutes in the traffic

Appendix G Big 5 Personality Scale

Instructions: Respond to each item as if it were the only item. That is, don't worry about being 'consistent' in your responses. Choose from the following five response options:

1 = very accurate
2 = somewhat accurate
3 = neither accurate or inaccurate
4 = somewhat inaccurate
5 = very inaccurate

Am the life of the party.	1	2	3	4	5
Am quiet around strangers.	1	2	3	4	5
Don't like to draw attention to myself.	1	2	3	4	5
Don't mind being the center of attention.	1	2	3	4	5
Don't talk a lot.	1	2	3	4	5
Feel comfortable around people.	1	2	3	4	5
Have little to say.	1	2	3	4	5
Keep in the background.	1	2	3	4	5
Start conversations.	1	2	3	4	5
Talk to a lot of different people at parties.	1	2	3	4	5
Am interested in people.	1	2	3	4	5
Am not really interested in others.	1	2	3	4	5
Have a soft heart.	1	2	3	4	5
Am not interested in other people's problems.	1	2	3	4	5
Feel little concern for others.	1	2	3	4	5
Sympathize with others' feelings.	1	2	3	4	5
Insult people.	1	2	3	4	5
Take time out for others.	1	2	3	4	5
Feel others' emotions.	1	2	3	4	5
Make people feel at ease.	1	2	3	4	5
Am always prepared.	1	2	3	4	5
Leave my belongings around.	1	2	3	4	5
Make a mess of things.	1	2	3	4	5
Like order.	1	2	3	4	5
Shirk my duties.	1	2	3	4	5
Pay attention to details.	1	2	3	4	5
Get chores done right away.	1	2	3	4	5
Often forget to put things back in their proper place.	1	2	3	4	5
Follow a schedule.	1	2	3	4	5
Am exacting in my work.	1	2	3	4	5
Am relaxed most of the time.	1	2	3	4	5

Get stressed out easily.	1	2	3	4	5
Worry about things.	1	2	3	4	5
Am easily disturbed.	1	2	3	4	5
Get upset easily.	1	2	3	4	5
Change my mood a lot.	1	2	3	4	5
Have frequent mood swings.	1	2	3	4	5
Get irritated easily.	1	2	3	4	5
Often feel blue.	1	2	3	4	5
Seldom feel blue.	1	2	3	4	5
Have a rich vocabulary.	1	2	3	4	5
Use difficult words.	1	2	3	4	5
Am not interested in abstract ideas.	1	2	3	4	5
Do not have a good imagination.	1	2	3	4	5
Have a vivid imagination.	1	2	3	4	5
Have excellent ideas.	1	2	3	4	5
Am quick to understand things.	1	2	3	4	5
Spend time reflecting on things.	1	2	3	4	5
Am full of ideas.	1	2	3	4	5

Appendix H BIS-BAS questions

Instructions: Respond to each item as if it were the only item. That is, don't worry about being 'consistent' in your responses. Choose from the following four response options:

1 = very true for me
2 = somewhat true for me
3 = somewhat false for me
4 = very false for me

1. A person's family is the most important thing in life.	1	2	3	4
2. Even if something bad is about to happen to me, I rarely experience fear or nervousness.	1	2	3	4
3. I go out of my way to get things I want.	1	2	3	4
4. When I'm doing well at something I love to keep at it.	1	2	3	4
5. I'm always willing to try something new if I think it will be fun.	1	2	3	4
6. How I dress is important to me.	1	2	3	4
7. When I get something I want, I feel excited and energized.	1	2	3	4
8. Criticism or scolding hurts me quite a bit.	1	2	3	4
9. When I want something I usually go all-out to get it.	1	2	3	4
10. I will often do things for no other reason than that they might be fun.	1	2	3	4
11. It's hard for me to find the time to do things such as get a haircut.	1	2	3	4
12. If I see a chance to get something I want I move on it right away.	1	2	3	4
13. I feel pretty worried or upset when I think or know somebody is angry at me.	1	2	3	4
14. When I see an opportunity for something I like I get excited right away.	1	2	3	4
15. I often act on the spur of the moment.	1	2	3	4
16. If I think something unpleasant is going to happen I usually get pretty worked up.	1	2	3	4
17. I often wonder why people act the way they do.	1	2	3	4
18. When good things happen to me, it affects me strongly.	1	2	3	4
19. I feel worried when I think I have done poorly at something important.	1	2	3	4
20. I crave excitement and new sensations.	1	2	3	4
21. When I go after something I use a 'no holds barred' approach.	1	2	3	4
22. I have very few fears compared to my friends.	1	2	3	4
23. It would excite me to win a contest.	1	2	3	4

24. I worry about making mistakes.	1	2	3	4

Appendix I Stress and Arousal Checklist

The adjectives shown below describe different feelings and moods. Please use this list to describe your feelings at this moment in time.

If the adjective definitely describes your feelings circle the:

++ + ? -

If the adjective more or less describes your feelings circle the:

++ + ? -

If you do not understand the adjective, or you cannot decide whether it describes how you feel circle the:

++ + ? -

If the adjective does not describe the way you feel circle the:

++ + ? -

Your first reactions will be the most reliable, therefore do not spend too long thinking about each adjective. Please be as honest and accurate as possible.

Tense	++	+	?	-	Tired	++	+	?	-
Relaxed	++	+	?	-	Idle	++	+	?	-
Restful	++	+	?	-	Up tight	++	+	?	-
Active	++	+	?	-	Alert	++	+	?	-
Apprehensive	++	+	?	-	Lively	++	+	?	-
Worried	++	+	?	-	Cheerful	++	+	?	-
Energetic	++	+	?	-	Contented	++	+	?	-
Drowsy	++	+	?	-	Jittery	++	+	?	-
Bothered	++	+	?	-	Sluggish	++	+	?	-
Uneasy	++	+	?	-	Pleasant	++	+	?	-
Dejected	++	+	?	-	Sleepy	++	+	?	-
Nervous	++	+	?	-	Comfortable	++	+	?	-
Distressed	++	+	?	-	Calm	++	+	?	-
Vigorous	++	+	?	-	Stimulated	++	+	?	-
Peaceful	++	+	?	-	Activated	++	+	?	-

Appendix J Perceived Stress Scale

	Never	Almost never	Sometimes	Fairly often	Very often
1. In the last month, how often have you been upset about something that has happened unexpectedly?	0	1	2	3	4
2. In the last month, how often have you felt that you were unable to control the important things in your life?	0	1	2	3	4
3. In the last month, how often have you felt nervous and "stressed"?	0	1	2	3	4
4. In the last month, how often have you felt confident about your ability to handle your personal problems?	0	1	2	3	4
5. In the last month, how often have you felt things were going your way?	0	1	2	3	4
6. In the last month, how often have you found that you could not cope with all the things that you have had to do?	0	1	2	3	4
7. In the last month, how often have you been able to control the irritations in your life?	0	1	2	3	4
8. In the last month, how often have you felt that you were on top of things?	0	1	2	3	4
9. In the last month, how often have you been angered because of things that were outside your control?	0	1	2	3	4
10. In the last month, how often have you felt difficulties were piling up so high that you could not overcome them?	0	1	2	3	4

Appendix K Recruitment poster for Study 1

Interested in taking part in a Driving Simulator experiment?

We are looking for volunteers to take part in the study "Effects of lunch in young people while driving".

As a participant in this study, you would be asked to undergo a driving task for 45 minutes in the static driving simulator in the Physics Research Deck in the University of Leeds. During the driving task your brain wave activity will be recorded with an EGI 128 channel system. You will also be video monitored during the whole driving task. Finally you will be asked to fill personality, sleepiness and stress questionnaires.

Your participation would involve in 2 sessions in two different days, each of which is approximately 70 minutes (45 minutes driving simulator task and 25 minutes positioning the EGI system to record brain wave activity).

To be able to be part of the study you need to be between 20 and 30 years old and hold a valid UK passport for at least 2 years. In case you have a visual, auditory, motor and/or mental disability we would not be able to recruit you for the experiment. People with Body Mass Index (BMI) above 28 will not be able to be recruited for this experiment ($\boldsymbol{BMI} = \frac{\boldsymbol{Weight}(\boldsymbol{kg})}{\boldsymbol{Height}(\boldsymbol{meters})^2}$).

Before being recruited you need to fill up a sleepiness questionnaire to determine if you have any sleep disorder.

For more information about this study please contact

Pablo Puente Guillen
(Principal researcher)

at

email: scppg@leeds.ac.uk

or

Prof. Anthony Cohn (supervisor) a.g.cohn@leeds.ac.uk
Prof. Oliver Carsten (supervisor) O.M.J.Carsten@its.leeds.ac.uk
Dr. Richard Wilkie (supervisor) r.m.wilkie@leeds.ac.uk
Dr. Faisal Mushtaq (supervisor) f.mushtaq@leeds.ac.uk

Your participation is truly appreciated.

Appendix L Recruitment poster for Study 2

Interested in taking part in a Driving Simulator experiment?

We are looking for volunteers to take part in the study "Effects of lunch in young people while driving".

As a participant in this study, you would be asked to undergo a driving task for 45 minutes in the static driving simulator in the Physics Research Deck in the University of Leeds. During the driving task your brain wave activity will be recorded with an EGI 128 channel system. You will also be video monitored during the whole driving task. Finally you will be asked to fill personality, sleepiness and stress questionnaires.

Your participation would involve in 1 session. The experiment will take approximately 85 minutes (60 minutes driving simulator task and 25 minutes positioning the EGI system to record brain wave activity).

To be able to be part of the study you need to be male participant (with short hair) between 20 and 30 years old and hold a valid UK passport for at least 2 years. In case you have a visual, auditory, motor and/or mental disability we would not be able to recruit you for the experiment. People with Body Mass Index (BMI) above 28 will not be able to be recruited for this experiment ($\boldsymbol{BMI = \frac{Weight(kg)}{Height(meters)^2}}$).

Before being recruited you need to fill up a sleepiness questionnaire to determine if you have any sleep disorder.

For more information about this study please contact

Pablo Puente Guillen
(Principal researcher)

at

email: scppg@leeds.ac.uk

or

Prof. Anthony Cohn (supervisor) a.g.cohn@leeds.ac.uk
Prof. Oliver Carsten (supervisor) O.M.J.Carsten@its.leeds.ac.uk
Dr. Richard Wilkie (supervisor) r.m.wilkie@leeds.ac.uk
Dr. Faisal Mushtaq (supervisor) f.mushtaq@leeds.ac.uk

Your participation is truly appreciated.

Appendix M Examples of correlation analysis of Driving variables and EEG variables for study 1

SDLP Alpha (Middle Parietal) / SDLP (segments)	Alpha (segments) 1	2	3	4	5	6	7	8	9
1	0.135	0.394	0.359	0.334	0.157	0.221	0.249	0.125	0.078
2	0.1	0.393	0.404	0.344	0.196	0.334	0.379	0.31	0.231
3	-0.107	0.221	0.231	0.146	0.02	0.157	0.199	0.133	0.056
4	0.048	0.353	0.347	0.311	0.206	0.245	0.365	0.267	0.187
5	-0.149	0.198	0.182	0.114	-0.041	0.059	0.146	0.042	-0.059
6	-0.33	-0.039	-0.037	-0.082	-0.222	-0.101	-0.055	-0.128	-0.215
7	-0.112	0.216	0.228	0.141	0.004	0.156	0.2	0.136	0.052
8	-0.209	0.041	0.075	0.002	-0.058	0.076	0.09	0.062	0.01
9	-0.028	0.058	0.109	0.101	0.067	0.192	0.149	0.162	0.149

SDLP Beta (Middle Parietal) / SDLP (segments)	Beta (segments) 1	2	3	4	5	6	7	8	9
1	0.055	0.059	0.055	0.098	0.027	0.01	-0.072	-0.101	-0.078
2	-0.041	0.021	0.057	0.095	0.022	0.082	0.059	0.077	0.064
3	-0.168	-0.13	-0.102	-0.049	-0.081	-0.052	-0.112	-0.105	-0.056
4	-0.132	-0.074	-0.023	0.041	-0.032	0.002	-0.003	-0.003	0.006
5	-0.317	-0.26	-0.228	-0.168	-0.24	-0.207	-0.239	-0.23	-0.231
6	-0.401	-0.349	-0.315	-0.265	-0.319	-0.286	-0.316	-0.293	-0.308
7	-0.216	-0.165	-0.135	-0.092	-0.138	-0.086	-0.128	-0.11	-0.085
8	-0.214	-0.174	-0.148	-0.123	-0.122	-0.071	-0.118	-0.096	-0.045
9	0.013	0.056	0.082	0.071	0.06	0.129	0.119	0.155	0.14

SDLP Theta	Theta (segments)

(Middle Parietal) SDLP (segments)	1	2	3	4	5	6	7	8	9
1	-0.223	-0.12	-0.083	0.116	-0.087	0.039	-0.206	-0.202	-0.105
2	-0.178	-0.117	-0.028	0.124	-0.071	0.123	0.035	0.094	0.077
3	-0.182	-0.196	-0.18	0.026	-0.079	0.031	-0.191	-0.202	-0.126
4	-0.149	-0.156	-0.072	0.108	-0.08	0.017	-0.014	-0.069	-0.035
5	-0.365	-0.338	-0.277	-0.074	-0.248	-0.102	-0.262	-0.239	-0.254
6	-0.455	-0.424	-0.372	-0.194	-0.297	-0.148	-0.331	-0.25	-0.316
7	-0.266	-0.263	-0.215	-0.034	-0.172	-0.012	-0.17	-0.141	-0.117
8	-0.18	-0.257	-0.239	-0.087	-0.126	0.004	-0.144	-0.148	-0.094
9	-0.099	-0.098	-0.031	0.01	-0.032	0.151	0.14	0.234	0.185

HFS Alpha/Beta (Middle Parietal) Alpha/Beta (segments)

HFS (segments)	1	2	3	4	5	6	7	8	9
1	-0.014	0.247	0.2	0.174	-0.029	0.257	0.281	0.211	0.122
2	0.237	0.424	0.409	0.371	0.194	0.438	0.433	0.404	0.305
3	0.07	0.325	0.3	0.269	0.045	0.364	0.346	0.293	0.206
4	-0.186	0.024	-0.001	-0.041	-0.224	0.018	0.078	0.008	-0.11
5	0.055	0.295	0.288	0.25	0.045	0.317	0.382	0.29	0.201
6	0.025	0.164	0.159	0.105	-0.026	0.17	0.221	0.146	0.074
7	-0.093	0.034	0.032	-0.023	-0.145	0.051	0.085	0.004	-0.049
8	0.056	0.144	0.151	0.089	-0.002	0.167	0.206	0.126	0.098
9	0.004	0.174	0.18	0.122	-0.073	0.226	0.185	0.143	0.07

SDLP Alpha/Beta (Middle Front) Alpha/Beta (segments)

SDLP (segments)	1	2	3	4	5	6	7	8	9
1	-0.223	0.219	0.218	0.29	-0.207	0.189	0.395	0.163	-0.033
2	0.48	,737*	,741*	,791*	0.391	,757*	,754*	0.654	0.566
3	-0.001	0.457	0.465	0.413	0.013	0.43	,705*	0.33	0.155
4	0.282	0.601	0.58	0.63	0.465	0.645	,859**	0.603	0.546
5	0.198	0.612	0.606	0.613	0.279	0.601	,837**	0.539	0.38
6	0.099	0.442	0.419	0.445	0.091	0.434	0.603	0.325	0.162
7	0.297	0.653	0.659	0.633	0.246	0.637	,792*	0.517	0.388
8	0.176	0.423	0.419	0.381	0.133	0.431	0.619	0.271	0.212
9	0.472	0.38	0.363	0.454	0.238	0.446	0.256	0.263	0.323

SDLP Alpha/Beta (Left Parietal)	Alpha/Beta (segments)								
SDLP (segments)	1	2	3	4	5	6	7	8	9
1	-0.386	-0.145	-0.189	-0.207	-0.351	-0.037	0.156	-0.176	-0.272
2	0.213	0.453	0.46	0.403	0.183	0.542	0.611	0.392	0.304
3	0.041	0.307	0.235	0.135	-0.032	0.343	0.575	0.222	0.036
4	0.505	,705*	0.648	0.621	0.488	,709*	,872**	0.612	0.522
5	0.157	0.426	0.354	0.291	0.133	0.385	0.635	0.27	0.154
6	-0.052	0.172	0.112	0.068	-0.07	0.137	0.343	-0.004	-0.083
7	0.176	0.432	0.391	0.302	0.104	0.458	0.642	0.323	0.184
8	0.281	0.435	0.389	0.323	0.18	0.456	0.615	0.333	0.223
9	0.216	0.246	0.309	0.322	0.174	0.37	0.262	0.215	0.255