STATUS OF THESIS

Title of thesis: IMPROVED SLEEPINESS DETECTION METHOD USING FUSION OF MULTIPLE SLEEPINESS INDICATORS FOR SHIFT WORKERS

I RODNEY PETRUS BALANDONG

hereby allow my thesis to be placed at the Information Resource Center (IRC) of Universiti Teknologi PETRONAS (UTP) with the following conditions:

1. The thesis becomes the property of UTP.
2. The IRC of UTP may make copies of the thesis for academic purposes only.
3. This thesis is classified as

- [ ] Confidential
- [x] Non-confidential

If this thesis is confidential, please state the reason:

____

The contents of the thesis will remain confidential for ________ years.

Remarks on disclosure:

____

Endorsed by:

____
Signature of Supervisor
Name of Supervisor:

AP. Dr. M. NAUFAL BIN M. SAAD

____
Signature of Author

Permanent Address:
P.O BOX, 519
PAPAR, 89608,
SABAH, MALAYSIA

Date: ____________________

Date: ____________________

UNIVERSITI TEKNOLOGI PETRONAS

IMPROVED SLEEPINESS DETECTION METHOD USING FUSION OF MULTIPLE SLEEPINESS INDICATORS FOR SHIFT WORKERS

by

RODNEY PETRUS BALANDONG

The undersigned certify that they have read, and recommend to the Postgraduate Studies Programme for acceptance of this thesis for the fulfilment of the requirements for the degree stated.

Signature:	
Main Supervisor:	AP. Dr. M. NAUFAL BIN M. SAAD
Signature:	
Co-Supervisor:	AP. Dr. TANG TONG BOON
Signature:	
Chair of Department:	AP. Ir. Dr. NURSYARIZAL MOHD NOR
Date:	

IMPROVED SLEEPINESS DETECTION METHOD USING FUSION OF MULTIPLE SLEEPINESS INDICATORS FOR SHIFT WORKERS

by

RODNEY PETRUS BALANDONG

A Thesis
Submitted to the Postgraduate Studies Programme
as a Requirement for the Degree of

DOCTOR OF PHILOSOPHY

DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING
UNIVERSITI TEKNOLOGI PETRONAS
BANDAR SERI ISKANDAR
PERAK

FEBRUARY 2020

DECLARATION OF THESIS

Title of thesis	IMPROVED SLEEPINESS DETECTION METHOD USING FUSION OF MULTIPLE SLEEPINESS INDICATORS FOR SHIFT WORKERS

I RODNEY PETRUS BALANDONG

hereby declare that the thesis is based on my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously or concurrently submitted for any other degree at UTP or other institutions.

Witnessed by

Signature of Author

Permanent Address:
P.O BOX, 519
PAPAR, 89608,
SABAH, MALAYSIA

Date:

Signature of Supervisor
Name of Supervisor:

AP. Dr. M. NAUFAL BIN M. SAAD

Date:

DEDICATION

I would like to dedicate this thesis to you whom one day will read the manuscript full heartily.

ACKNOWLEDGEMENTS

Firstly I would like to thank AP. Dr. Tang Tong Boon for supervising this work. I thank him, especially for his willingness to accept me at the mid of this thesis project, where things not so neatly organised but magically through his guidance leading to the submission of this thesis. Feedback on manuscripts was always rapidly delivered and helpful. I have learnt a lot about research science and the processes involved in doing an excellent job of it from completing this project. I would also like to thank AP. Dr. Naufal M. Saad for given me the chance to be part of HICOE project. His chair but 3000 baby steps (figuratively since he was at France at the time of writing) from mine, but Dr Naufal was always providing genuine caring and concern, which tremendously help me when it comes to deal with management related problem. This project also will not be possible without the help from international researcher whom I had only the chance to communicate via email but tremendously boost my confidence as a researcher. Especially, Dr. Svetlana Postnova (The University of Sydney) for her valuable advice and support during the construction of the Phillips-Robinson Model and St Hillary Model in the MATLAB environment. Her recommendation to use SourceTree when collaborating on code was particularly high-five worthy and revolutionised in the way I organise my code-related work. To Prof. Rongrong Fu (Yanshan University) who passionately teach me about how he utilises Bayesian Network for his work on fatigue prediction. I will keep all the hand-written note that you provided to me during the whole course of our discussion. Furthermore, I am grateful to one of my co-authors, Dr. Michelle Short (Flinders University) who have generously provides the data used to answer all the research problem and all the idea that constitutes my thesis. Also, thanks to Dr. Sarat Dass, and Dr. Raheel Zafar and all the anonymous reviewer for their helpful advice on various technical issues. To CISIR friends and you people of monopoly-board game maniac, thanks for making this journey stress free yet memorable. My ever supporting, loving, and inspiring mother Annie, and father, Petrus. You have always let me take my time to find my way, but still at all times offered guidance when needed. To Felicity and Rhys, I promise to spend more times with both of you. Last, but not least, I'd like to thank myself for writing such a comprehensive set of acknowledgements.

ABSTRACT

Error committed by human due to sleepiness is one of the most common reasons that has led to glitches within the maritime sphere. Thus, development of a sleepiness detection system (SDS) to monitor sleepiness level, is highly desirable. Past studies depicted that the implementation of subjective sleepiness assessment was hampered due to response biases. The SDS with higher sleepiness level resolution reveals the specific arousal state of a shift worker to notify the seafarers sleepiness at an early stage when intervention procedure is most effective. However, most SDSs have some limitations, including constrained prediction horizon length and low sleepiness level resolution. Besides, the existing SDS approaches tend to meet failure upon insufficient or lost information.

Instead of thresholding the raw self-reported sleepiness value (vSRS) directly, this study propose an improved self-reported sleepiness value (IvSRS), a new measure of subjective sleepiness estimation improved by using the likelihood ratio test and kernel density estimation technique. This study further seek to increase the aspects of robustness, prediction horizon, and sleepiness level resolution by embedding multiple contextual factors (CFs) and IvSRS under the Bayesian Network (BN). The BN generated a single probability estimate that was calculated based on prior and posterior probabilities of CF and IvSRS, respectively.

As a result, the IvSRS achieved greater average F_1-measure which reflecting its superiority in classifying a data point into binary and ternary sleepiness states, in comparison to other non-modified vSRS. Upon comparing each CF and IvSRS as a standalone indicator, the integration of all information under BN resulted in greater average F_1-measure and the improvement compared to the other stand alone indicator were all significant as confirmed by the statistical analysis. Apart from functioning well in the event of missing vSRS, the proposed system displayed prediction horizon of 12 h, and F_1-measure exceeding 78%. Finally, comparison of the study outcomes with those reported in other studies indicated that the proposed method can have a superior performance. Due to the simplicity of the proposed SDS, practical deployment of mariner sleepiness detection can be easily proliferated without installing any intricate equipment or making major modification to the ship.

ABSTRAK

Mengemudi kapal dalam keadaan mengantuk merupakan antara punca peningkatan kes kemalangan maut melibatkan maritim. Oleh itu, adalah penting untuk menilai semula risiko secara berkala dan tahap keberkesanan aktiviti mitigasi untuk meningkatkan keselamatan navigasi perkapalan. Kajian terdahulu menunjukkan bahawa penggunaan penilaian mengantuk secara subjektif telah terganggu disebabkan maklumbalas yang berat sebelah. Tambahan pula, kebanyakan Sistem Pengesanan Mengantuk (SDS) mempunyai limitasi dimana memberikan ramalan yang kurang tepat dan tahapresolusi mengantuk yang terhad. Selain itu, SDS semasa cenderung mengalami kegagalan disebabkan kehilangan atau kekurangan maklumat.

Kajian ini menyiasat samada laporan kendiri penilaian tahap mengantuk (vSRS) ditambah baik (IvSRS) dengan mengubahnya menjadi pengiraan kepadatan dan seterusnya dinilai menggunakan ujian nisbah kebolehan. Tesis ini juga telah berjaya meningkatkan aspek keteguhan, ramalan ufuk dan resolusi tahap mengantuk dengan menggabungkan pelbagai pelbagai faktor konteks (CF) dan IvSRS yang diintegrasikan kepada model menggunakan Bayesian Network (BN). BN menghasilkan satu andaian keberangkalian yang dikira berdasarkan kepada keberangkalian sebelum dan selepas bagi CF dan IvSRS.

Hasil kajian menunjukkan IvSRS mempunyai kebolehan yang lebih baik dalam mengklasifikasikan rasa mengantuk kepada dua dan tiga tahap berbanding dengan vSRS. Dengan membandingkan setiap CF dan IvSRS sebagai penunjuk tunggal, integrasi semua maklumat dalam BN dapat mempertingkatkan prestasi sistem. Kajian juga menunjukkan teknik sistem yang dicadangkan ini mempunyai kebolehan yang lebih baik dalam mengklasifikasikan rasa mengantuk berbanding pengklasifikasian asas. Selain dapat berfungsi dengan baik dengan ketiadaan vSRS, sistem yang dicadangkan mempunyai kebolehan ramalan 12 h dengan ukuran F_1-measure melebihi 78%. Akhir sekali, perbandingan dengan hasil kajian yang telah dilaporkan menunjukkan kaedah yang dicadangkan mempunyai prestasi yang lebih unggul.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACRONYMS

AVT	Auditory Vigilance Task
BMM	Bio-Mathematical Models
BN	Bayesian Network
CBT	Core Body Temperature
CF	Contextual Factors
ECG	Electrocardiogram
EEG	Electroencephalography
EMG	Eectromyography
EOG	Electrooculogram
FE	Facial Expression
ISI	Inter-Stimulus Intervals
IvSRS	Improved Self-Reported Sleepiness Value
KDE	Kernel Density Estimation
KSS	Karolinska Sleepiness Scale
LR	Likelihood Ratio
MAD	Model of Arousal Dynamics
ML	Machine Learning
NREM	Non-REM
PERCLOS	Percentage Closure of Eyelids
PSG	Polysomnography
PVT	Psychomotor Vigilance Test
REM	Rapid Eye Movement
RMSE	Root Means Square Error
RT	Response Time
SDS	Sleepiness Detection System
SE	Sleep Efficiency
SQ	Sleep Quality

SW	Shift Worker
SWA	Steering Wheel Activity
SWM	Steering Wheel Movement
SWS	Slow Wave Sleep
TPM	Three-Process-Model of Alertness
TSD	Total Sleep Drive
vSRS	Self-Reported Sleepiness Value

CHAPTER 1

INTRODUCTION

1.1 Background

1.1.1 Sleepiness

Sleepiness is defined as the increased sleep propensity due to decrease in physiological arousal [1, 2]. Sleepiness is a term that indicates the perceived need or readiness to sleep [3]. Elevated level of sleep propensity leads to shorter time taken to fall asleep (i.e., sleep latency) [4]. While occasional sleepiness is a standard daily episode that is generally experienced at the end of the day, one who falls asleep inadvertently during active operation can cause undesired events [3].

Often in the literature, several terms, namely 'fatigue', 'drowsiness', 'sleepiness', and 'sleep propensity' have been used somewhat loosely [5]. The reason to apply these terms interchangeably is attributed to their overlapping features and semiotic between them [6], including compliance with regulatory requirement [7]. That being highlighted, the terms 'fatigue', 'drowsiness', 'sleepiness', and 'sleep propensity' are used interchangeably in this study. As a final note, 'alert' and 'alertness' are treated as opposed to the states of sleepy and sleepiness, respectively.

1.1.2 Split Schedule In The Maritime Industry

The standard working hours normally involve working five days a week between 09:00 h and 17:00 h. Specific industries implement shift schedule to meet 24-hour operation demands. As there is no common definition of shift work, the term in this study refers to a system of scheduling working time, where multiple panels of workers replace one another to extend operation hours [8].

Shift work is a particularly compelling characteristic within the maritime industry to maintain 24-hour operation while accommodating both navigational (tide) and market (time) pressure [9]. The maritime industry applies a split schedule, whereby a single day is fragmented into several work/rest cycles. In general, three split schedules are employed within this sector [10–12]. The first system refers to four-hour work, and followed by eight-hour off (shift$_{4/8}$). The three panels of watchkeepers are scheduled in rotation for daily operation: the first panel covers the operation from 00:00 to 04:00 h and 12:00 to 16:00 h, followed by the second panel that replaces the initial panel from 04:00 to 08:00 h and from 16:00 to 20:00 h, and finally, the third panel takes over the previous panel from 08:00 to 12:00 h and from 20:00 to 00:00 h. This shift system, which is also called the three-watch system, is often implemented in large ships. The more common watch systems employed in the maritime industry refer to the schedule of six-hour work, followed by six-hour off (shift$_{6/6}$), and repeated twice within 24-hour period. In this schedule, the first panel covers the operation from 00:00 to 06:00 h and from 12:00 to 18:00 h, while the second panel is on duty from 06:00 to 12:00 h and from 18:00 to 00:00 h. The shift$_{6/6}$ system is commonly implemented in vessels with moderate size that operate in narrow waters. Finally, the eight-hour on and eight-hour off (shift$_{8/8}$) refers to a schedule that is commonly applied in the rail industry, but has recently gained interest for implementation in the maritime domain [13–16]. In this schedule, two crews of employees alternately work and rest for eight hours each.

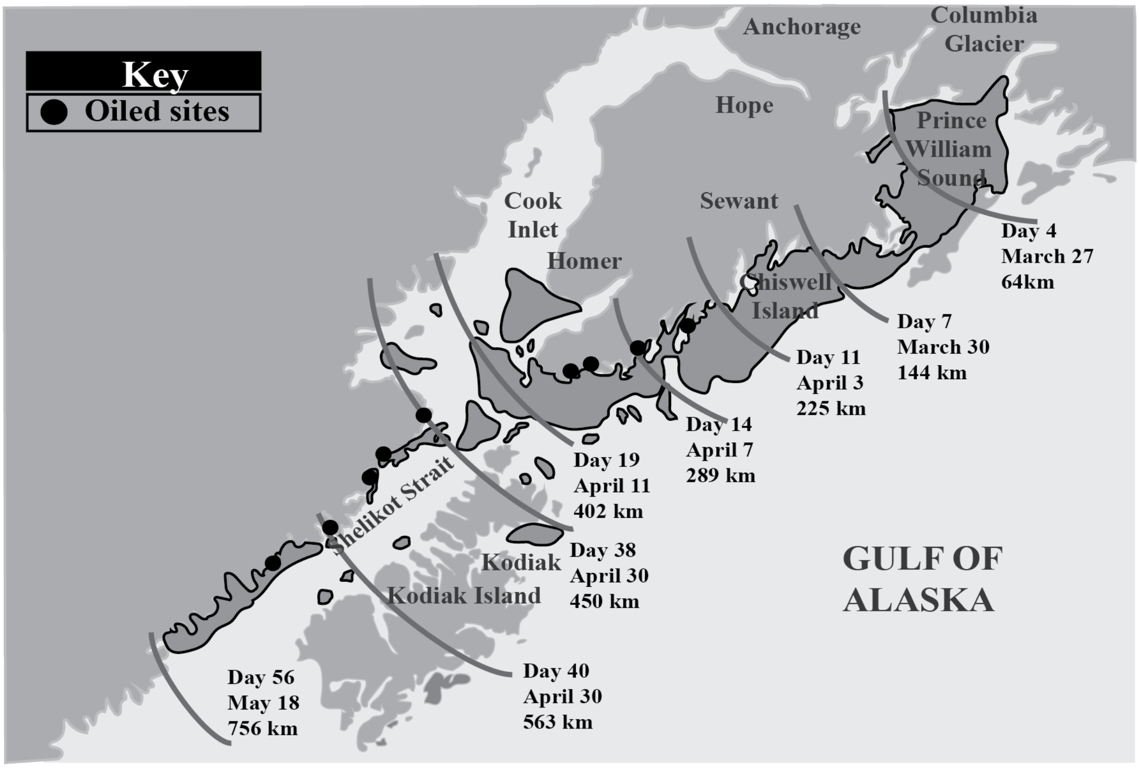


Figure 1.1: The spread of crude oil from the Exxon Valdez for the first two months. Figure adapted from [23].

1.2 Accident Due To Sleepiness

It has been reckoned that the marine industry has a reasonably good safety record, in comparison to other transportation domains [17]. Regardless of the magnitude of accidents and incidences reported within the marine domain, evidence points out that such unpleasant events may indirectly or directly affect the environment, apart from resulting in human casualties [9, 18]. The main impetus for occurrences of marine incidences in the past was attributable to environmental and technical factors, while recent reports point out human sleepiness to upend the role [19, 20]. An infamous case of the grounding of Exxon Valdez in 1989 had been linked with excessive sleepiness by one of the operators whilst on duty [21]. Damages made to the super tanker led to the leakage of 11 million gallons of crude oil that adversely affected nearly 1990 km of shoreline (see Figure 1.1). Due to the environmental damages incurred, Exxon was required to pay a whopping $100 million criminal fine [22].

Several investigation bodies at national level have discovered a substantial number of collision and grounding due to negligent officers falling asleep or feeling sleepy during navigation. A report by the Great Britain's Department of Transportation concluded that about one-third of 1647 near collision, collision, and grounding cases that had occurred in the United Kingdom between year 1994 and 2004 mostly involved a sleepy officer and those who worked alone on the bridge [21]. The Japan Maritime Research Institute (cited in [24]) reported that 53% and 38% of casualties from vessel grounding and collision, respectively, resulted due to seafarers falling asleep during navigation. Similarly, in Norwegian waters, 9% out of 88 grounded vessels in year 2006 had been due to navigation officers had fallen asleep inadvertently [25]. Likewise, the Finland Accident Investigation Board had analysed 10 maritime accidents that occurred between year 1997 and 2003 that displayed either clear or potential link with sleepiness [26]. Five cases showed that the officers had involuntarily fallen asleep while the vessels continued moving until the crash, whereas the other five cases had been due to excessive sleepiness experienced by the officers involved. Phillips, in a study that involved 44 maritime accidents investigated by the Australian Transport Safety Bureau from year 1991 to 2001, concluded that sleepiness was a primary contributor in about 15.9% of the accidents [9]. In a review of crash report, Smith and colleague revealed that the pattern of sleepiness-related accidents was most prevalent during the first week of tour [27]. The number of sleepiness-related incidences in the maritime industry investigated by the Australian Transport Safety Bureau and the Marine Accident Investigation Branch (MAIB) was at an alarming level [28, 29]. The literature highlights that sleepiness among mariners is common and the number of those affected by this occupational problem might have increased progressively throughout the year [30]. Some studies concluded that most seafarers experience sleepiness or involuntarily sleep during night-shift work, which is uncommon for day-oriented schedule [31–33].

Some investigations have identified falling asleep while operating a moving vehicle is a crucial factor in land transportation accidents [34, 35]. The National Highway Traffic Safety Administration (NHTS) reported that drowsy driving claimed about 824 lives in year 2015 [36]. The same report by NHTS also estimated that 4,121 crashes due to drowsy driving had occurred between 2011 and 2015 [36]. A survey conducted by Ontario Canada revealed that out of the 750 randomly polled participants, 58.6% claimed that they occasionally drove while feeling sleepy, while 14.5% slept off unintentionally while driving [37]. The destructive consequence of sleepiness is also noted in the aviation industry. The National Transportation Safety Board (NTSB) in the United States investigated aviation accidents that occurred from 1983 to 1986, and found that 69 cases were resulted by sleepy pilots [38]. The secondary analysis by Caldwell reported that drowsiness might be involved in 4-8% of civil aviation accidents [39].

Sleepiness is also a contributing factor for accidents that take place in the railway industry. An investigation into a coal train collision at Beresford located in Australia occurred in 1997 revealed sleepiness as the major causal contributor [40]. The railway accident analyses in Japan, China, and the United States had identified sleepiness as the main reason in many rail accidents [40]. The detrimental impact of sleepiness is not limited to the transportation industry, but it also affects mining, air traffic controls, and healthcare professionals, to name a few [41–43]. While the consequences from sleepiness-related accidents in these line of work may not cause loss of life, but the direct healthcare expenditures and indirect costs, such as lost employment, lower productivity, and machinery damage, should be of considerable concern amidst these industries.

A number of studies highlighted the impairment in cognitive ability after a period of sleep deprivation, which is equivalent to upon one assessed under the influence of alcohol. In a seminal study, Dawson and Reid [44] revealed that 17-hour forced wakefulness decreased one's vigilance performance, which is equivalent to those observed with blood alcohol concentration of 0.05% (the blood alcohol concentration limit for driving in most nation is 0.05%), whereas the performance after 24-hour sustained wakefulness was comparable to 0.10% level of blood alcohol concentration. Besides, the Swedish merchant ships showed variance in reaction time performance, especially for night shift, which revealed a subject with blood alcohol content of 0.4 parts per thousand (the legal driving limit under the influence of alcohol in Sweden is 0.2 parts per thousand) [45, 46].

1.3 Motivation

1.3.1 Sleepiness Detection System

As the shipping domain continues to play a crucial role in the global economy, accidents are bound to occur despite the best countermeasures taken. Despite the regulation to control the hours of rest a mariner should take, statistical data reflecting accidents indicate that regulation alone is ineffective in addressing issues related to sleepiness. The visible presence of risk factors due to sleepiness and the potential adverse effects highlight the need for an on-board system that monitors the level of sleepiness among seafarers. Such a system is denoted as sleepiness detection system (SDS), which can monitor and warn if one's sleepiness level is incompatible with safety-sensitive operations, thus offering assistance in deciding if it is unsafe for the officer to resume operation. The SDS, as a way to minimise sleepiness-related accidents, has for years been in the top most wanted list for safety improvement sought by the United States' National transportation safety. Meanwhile, PETRONAS under the PETRONAS Technical Guidelines for Fatigue Management at Work Place recommended a periodic risk of fatigue evaluation as a measure to address issues related to fatigue at workplace. The SDS can be applied to collect data concerning sleepy navigation in large-scale naturalistic marine studies in order to add to the knowledge on the pattern between sleepiness and crash risk. Such system may be implemented to gather data regarding one's sleepiness state to enhance work scheduling. These four needs have partly inspired this thesis.

1.3.2 Bayesian Network

Bayesian Network (BN) has been gaining considerable attention with promising avenues for modelling multi-component systems [47]. The BN is a probabilistic graphical model that integrates graph and probability theories [48]. The development of BN is composed of two stages [49]. The initial stage involves specifying the graph structure for a certain issue. The expert's knowledge and experience of the cause and effect are applied to associate or disassociate the correlation between random variables. Next, the second stage elicits prior and conditional probability values. The values of these probabilities are obtained from an expert's judgment or the statistical analysis of data gathered from large-scale surveys [48]. With adequate training data, the construction of graph structure or/and probabilities values can be generated automatically via structure learning (e.g., junction tree algorithms) and parameter learning (e.g., expectation maximisation), thus leaving the experts' input at a minimal or even unnecessary [50].

The BNs possess some distinct advantages. First, the BN architecture offers a ready mechanism to address vast amounts of uncertainties inherent in the problem domain, as reported in building SDS [51]. This is indeed probable as BN defines the strength of the correlation between pairs of variables (causal relationship) in a probabilistic manner. Second, the BN offers the flexibility to update knowledge in real-time as new information and experiences become available, thus commonly known as evidence updating (or belief revision). This process recalculates prior estimates, and hence, always reflects the latest knowledge in the problem domain. Third, since it is common for the graphical representation to mimic the causal structure, it eases knowledge acquisition and comprehension of the model structure. Thus, while developing a network, the domain expert can effortlessly assess the framework and if necessary, modify the model structure to gain better predictive models. The graphical interface based on nodes and arrows indirectly makes BN more transparent and intuitive. Fourth, the explicit representation of independencies that effectively minimises the connectivity in the graph makes BN suitable for domains with multiple sources of variables. This offers more compact networks, and efficient overall computation, mainly because only a subset of all possible connections has to be calculated. Fifth, BN can avoid data over-fitting by embedding prior probabilities during the rule-discovery process. Finally, BN allows 'what-if' notions so that a modeller can change the belief of various variables and assess the outcomes.

1.3.3 Kernel Density Estimation Supplement With Likelihood Ratio Test

The Kernel Density Estimate (KDE) refers to a non-parametric method that estimates the probability density of any random variable. An identified density function (the kernel) is overlaid across the training data to generate a smooth histograms that preserve important density features at multiple scales [52]. Some advantages of KDE are that its estimation can be made independently across all data points, and it has no fixed structure [53]. Apart from insensitivities to the shape of kernel, KDE also can deal with the non-Gaussian data in a better way and offer a flexible way to estimate the densities.

Likelihood ratio (LR) test refers to a hypothesis test that compares the assumptions of varied models. Based on the density estimate obtained from the KDE, LR test can be used to determine which of the classes to assign for any given match.

1.4 Problem Formulation

1.4.1 Research Gap

To this end, numerous attempts have been dedicated to develop an effective SDS. Despite the vastly proposed SDS in many studies, some gaps have been determined from the literature review. For instance, the application of subjective sleepiness assessment as a universal tool for sleepiness indicator is challenging due to the potential occurrence of response biases that may result from the inability of one to accurately interpret perception or varyingly define each scale. Such response biases may be manifested as spikes or noise thus required a post processing to produce smoother distribution.

Most of the proposed SDSs have constrained prediction horizon length which is in the range of 0.2 seconds to 10 minutes. From the accident prevention perspective, a short prediction horizon may limit the time for a suitable intervention measure to be ready. The correlation between varied sleepiness indicators and a different level of sleepiness is far from perfect. The element of uncertainty displayed by each related component may generate an ambiguous system if applied as a standalone decision-making tool. The vast literature has proposed integration of one or more of sleepiness indicators to compensate for the advantage and disadvantage of each indicator. The existing multi-modalities fusion-based SDSs mainly adopt a model with a single-level structure framework. In such framework, all features from multiple sleepiness indicators are fused simultaneously into a mathematical model. The reliability of the proposed frameworks may be affected upon unavailable or insufficient data. More importantly, each SDS is application-driven. This is because SDS for a particular application/ occupation cannot be generalised to other applications. In precise, the applicability and practicality of each indicator are subjected to the location and the type of intended work.

1.4.2 Problem Statement

Based on the research gap, the following problems have been identified:

1. Subjective sleepiness assessment is susceptible to response biases as a result from the inability of an individual to properly interpret their own perception or may define each of the scale differently. Such response

biases may be manifested as spikes or noise thus required a post processing to produce smoother distribution.

2. Most of the proposed SDSs have constrained prediction horizon length.

3. The correlation between varied sleepiness indicators and a different level of sleepiness is limited.

4. The reliability of the proposed frameworks may be affected upon unavailable or insufficient data.

1.5 Hypotheses

Based on the depicted problem statement, the following have been hypothesised:

1. Post-process of subjective sleepiness estimation by using the Likelihood Ratio test and Kernel Density Estimation technique may dampen the effect of response biases.

2. Systematically-fused multiple sleepiness indicators using the Bayesian Network can result in;

   (a) SDS with better classification ability compared to stand alone sleepiness indicator.

   (b) SDS that able to function in the event of missing observational information.

   (c) SDS that capable in lengthening prediction horizon.

1.6 Study Objectives

Based on the hypotheses, the following research objectives have been formulated:

1. To investigate the benefit of post-processing of subjective sleepiness estimation by using the Likelihood Ratio test and Kernel Density Estimation technique in reducing the effect of response biases.

2. To design a ternary SDS based on Bayesian Network that will result in:

   (a) SDS with better classification ability compared to stand alone sleepiness indicator.

   (b) SDS that able to function in the event of missing observational information.

   (c) SDS that capable in lengthening prediction horizon.

1.7 Scope Of Work

The scope of work for the thesis are:

1. The significance of a controlled laboratory experiment is beyond doubt to establish reliable and valid measures of the related constructs. The data that had been employed to develop the proposed system was retrieved from the controlled laboratory experiment carried out by Short *et al* [54]. This particular dataset was chosen mainly because the experiment mimicked the schedule applied in the shift industry. The experiment looked into the impact of naps and varying shift systems implemented in the maritime industry on the aspects of neurobehavioral performance, sleep, and sleepiness. Thus, this thesis focus mainly on data processing and the sample has been done by others.

1.8 Organisation Of Thesis

The organisation of this thesis is elaborated in the following:

Chapter 1 serves as an introduction to this study and highlights the significance of developing SDS by incorporating statistical data and visual pieces of evidence. Chapter 2 elaborates relevant information pertaining to the significant causes of sleepiness during working. The second part of Chapter 2 describes the existing sleepiness indicators and algorithms, as well as an overview of the practical requirement for constructing the SDS.

Chapter 3 is devoted to the development of the proposed framework. The initial part of the Chapter 3 explains the improvement made by using two algorithms, namely Kernel Density Estimate and Likelihood Ratio test, while the second part of the chapter introduces the Bayesian Network paradigm, and followed by the construction of SDS based on Bayesian Network. The final part of Chapter 3 depicts in detail the criteria for validation metric, types of validation, and the aspects considered to select the most viable evaluation metric. Chapter 4 presents the experimental analysis and validation, as well as comparison with other state-of-the-art methods. Finally, Chapter 5 presents the highlights for each chapter, the contributions made within this dissertation, and several suggestions for the future research direction of the thesis work.

Chapter List Of Publication list down all the journal article published arise from the development of this thesis. Chapter Appendix first explained the biomathematical model that capture the interaction between the process homeostatic and circadian. The second part of the Chapter Appendix gives a detailed account of the dataset applied to develop and evaluate the proposed system. The final part of chapter Appendix give the detail mathematical expansion regarding the interaction between dependent and independent variables.

CHAPTER 2

LITERATURE REVIEW

This chapter is composed of two main sections. The first section describes the main factors that control sleepiness. Essentially, the general agreement is that sleepiness is influenced by three main factors, namely time-of-the-day, length of time at work, and quality of the last sleep. The first two factors are associated with the human endogenous process, namely circadian and homeostasis processes, in which misalignment between them may aggravate one's sleepiness while working. Deteriorated quality and duration of sleep may increase one's sleepiness during the next waking bout and performance deficits in terms of neuro-behavioural measures. Hence, it is vital to comprehend these factors as the prerequisite for constructing an effective SDS. The second part of this chapter elaborates the practical requirement to develop SDS. This chapter is continued by reviewing the three group of sleepiness indicators that are used to quantify sleepiness, including vehicle-centred, operator-centred, and hybrid of varied sleepiness indicators. Next, this thesis discusses several mathematical algorithms to uncover patterns in data, and subsequently, to automatically predict sleepiness upon being fed with a new set of data. Finally, the four main issues that revolve around the SDS design are discussed. The initial issue is related to missing or corrupted data in the real environment, while the second issue is the limited length of sleepiness prediction horizon. Next, the third issue is regarding low sleepiness resolution in existing SDSs, and lastly, the generalizability of each sleepiness indicator to varying applications and field of work.

2.1 Significant Causes For Sleepiness During Working

2.1.1 The Circadian And Homeostasis Processes

Based on the two-process concept initiated by Borbély, the timing of human sleep propensity is regulated by homeostatic and circadian processes [55]. The homeostatic process is associated with the build-up and the dissipation of sleep propensity during the time spent awake and asleep, respectively. The regulation of sleep pressure by the homeostatic process has been hypothesised to be linked with the concentration of adenosine, a by-product of cell metabolism in the basal forebrain. This adenosine concentration increases and decreases during wakefulness and sleep, respectively. On the other hand, the circadian process, which is mainly driven by the master circadian clock located in the suprachiasmatic nuclei (SCN) of the anterior hypothalamus, regulates the approximately 24-hour rhythm of sleep propensity [56]. The SCN activity is synchronised by zeitgebers (time givers), predominantly the light-dark cycle [57]. In fact, other zeitgebers are inclusive of feeding time, temperature cycle, humidity cycle, and social interactions [58, 59].

The most commonly used indicator for circadian rhythm refers to the dim light melatonin onset and core body temperature (CBT). When one is exposed to natural light-dark cycle, the daily oscillation of CBT adheres to a near sinusoidal pattern with minimum CBT (CBT_{min}) that usually occurs around two hours prior to habitual wake time. Entrained sleep is commonly initiated about five to six hours prior to CBT_{min} with the peak of circadian sleep propensity typically observed close to CBT_{min}. Meanwhile, melatonin build-up that starts in the late evening serves as a cue to the body that 'biological night' is about to begin [60]. At this stage, the propensity to fall asleep is high. Melatonin increases and hits its maximum around 04:00 h before it decreases to a minimum level during the day.

For the natural sleep-wake cycle, one is usually awake and asleep during the solar day and night, respectively [61–63]. While being awake between

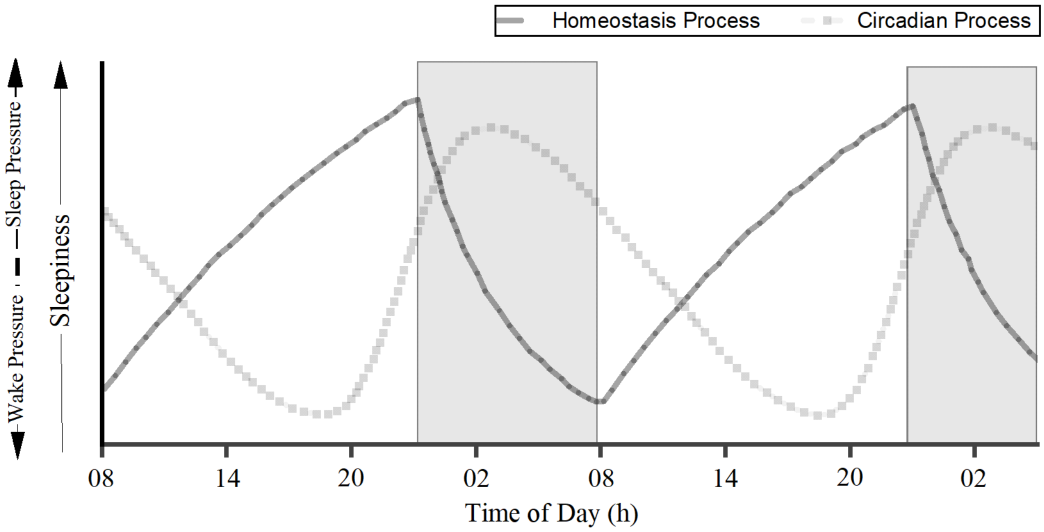


Figure 2.1: The interaction between circadian and homeostatic process. The grey rectangular box indicated sleep period. Figure adapted from [62].

morning and afternoon, the homeostatic sleep propensity accumulates, while the wake-promoting circadian signal increases but drops marginally around 14:00-17:00 h. Both homeostatic and circadian processes ensure a stable net drive for alert wakefulness, with a minor dip in the afternoon [62]. The wake-promoting circadian signal reaches its peak in the late evening (19:00-22:00 h), whereby sleep is usually difficult to initiate and maintain [62]. After that, the wake-promoting circadian signal decreases, while the homeostatic sleep propensity continues to increase with wakefulness. It is within this period that the net drive for alert wakefulness reduces and sleep state sets in.

At solar night, the homeostatic sleep propensity dissipates with the time spent sleeping, while the wake-promoting circadian signal continues to fall until it attains its nadir at 04:00-05:00 h. The correlation between both processes leads to a stable net drive to sleep [62]. Upon reaching its nadir, the wake-promoting circadian signal begins to rise, while the homeostatic sleep propensity continues to drop as one is in slumber. The net drive to sleep gradually diminishes in the morning, which leads to awakening and the beginning of a new sleep-wake cycle. Figure 2.1 shows the interaction between homeostatic and circadian sleep propensities.

The temporal synchronisation between the homeostatic and circadian processes becomes distorted during shift work. At night shift, for instance, the combination of the weakening wake-promoting circadian signal and the increasing sleep pressure of homeostatic process results in increased net pressure for sleep during working period [62, 64]. Simply put, some shift workers are forced to be operationally active during night shift when their circadian rhythm is still aligned to their day-oriented rhythm. This is especially a critical concern when shift workers need to have sustained attention during midnight shift, which is close to the circadian peak in sleepiness [65].

Nevertheless, the SCN can adapt to a new light-dark cycle imposed by atypical shift work by shifting the rhythmic biological function. As such, inter-individual variability in rate, direction, and ability to phase shift circadian rhythm needs to be weighed in [66]. Heterogeneous responses can be manifested by the variance in the variables predominantly by latitude [67], seasonal [68], operational setting [69], social setting, chronotype [70–72], and shift pattern [68,73,74].

Theoretically, split duty schedules tend to minimise the homeostatic sleep pressure while at work, although influences of circadian still exist [75]. Thus, if rest period is fully utilised for recuperation activity, frequent rest period may reduce the accumulation of sleep pressure, especially those that derive mainly from the homeostasis process. The impacts of both homeostasis and circadian processes on human sleepiness level and performance vary based on shift type.

The most commonly applied split duty in the maritime industry are $\text{shift}_{4/8}$, $\text{shift}_{6/6}$, and $\text{shift}_{8/8}$ (see Section 1.1.2). Theoretically, the $\text{shift}_{4/8}$ is the most favourable shift system in terms of minimising sleepiness, since the work time is only four hours and followed by a long resting period. Nonetheless, a series of shipboard studies carried out in the 1980s revealed that watchkeepers under the three-watch system failed to obtain adequate amount of sleep, and even worse, most of the subjects had no circadian synchronisation to the shift schedule [76–78]. The 00:00-04:00 h shift has been linked with high levels of sleepiness [45, 77–80] and increased likelihood of falling asleep [81]. The outcomes suggest that circadian rhythms have an impact on alertness during a shift, even under conditions of limited wakefulness. A major consideration for shift$_{4/8}$ is that it requires three panels of crew for 24-hour operation, when compared to only two panels of crew for shift$_{6/6}$ and shift$_{8/8}$, signifying cost ineffectiveness [82]. The additional changeover has emerged as an issue because shift handovers are related to increased accidents and errors [69].

In a laboratory simulation study, Short and colleague investigated two types of shift$_{6/6}$ schedules (6 h early: awake time 03:00-08:00 h and 15:00-20:00 h, or 6 h late: awake time 09:00-14:00 h and 21:00-02:00 h) [54]. Despite the fact that both shift systems displayed circadian variations in performance outputs, the 6h-late group demonstrated high-performance deficit towards the end of night shift (07:30 h) relative to at any other time during their shifts. This observation is in line with a study that assessed the fluctuation of subjective sleepiness while performing an activity in a ship simulator. Participants who worked under shift$_{6/6}$ system exhibited higher subjective sleepiness on the night watch, when compared to at any other time during their shifts [82]. The observation from these two studies had been expected, based on the homeostatic factor, whereby most participants reported increasing subjective sleepiness and performance impairment as a function of time, but the worst during night watch (shift from midnight to early morning) due to the coupling effect, along with the circadian factor.

The question if shift$_{6/6}$ schedule is better in terms of having lower magnitude of sleepiness and performance deficit than the shift$_{4/8}$ schedules is paramount [10]. This question has been partially motivated by the investigation initiated by MAIB, which concluded that slightly more accidents occurred under shift$_{6/6}$ schedule than those under shift$_{4/8}$ schedule [21]. It was found that most of the accidents occurred between midnight and 06:00 h on ships following the $\text{shift}_{6/6}$ schedule [21]. Two on-board studies [45, 80] investigated the impacts of $\text{shift}_{6/6}$ and $\text{shift}_{4/8}$ schedule on sleepiness among the seafarer. As expected, higher level of sleepiness and higher frequency of falling asleep were reported while on the watch for $\text{shift}_{6/6}$, when compared to the $\text{shift}_{4/8}$ schedule. Two additional observation made from these two studies were that the fluctuation of the subjective sleepiness was depended to the time of the day, and most seafarers whom were working under $\text{shift}_{6/6}$ schedule reported the highest subjective sleepiness scores toward the end of night watch (00:00 to 06:00h).

Nevertheless, a concern was raised that the conclusion made from these shipboard studies was influenced by multiple confounding factors, such as weather conditions, differing workload, and varying journey durations [81]. These issues were addressed in a seven-day laboratory simulation of a typical shipboard environment, with the seafarers grouped into either $\text{shift}_{4/8}$ or $\text{shift}_{6/6}$ systems. Despite sufficient rest period, as well as an ideal environment to rest and sleep, high incidences of nodding off were noted while on duty. Incidences of seafarers falling asleep while on watch seemed higher for those working under the $\text{shift}_{6/6}$ schedule compared to $\text{shift}_{4/8}$ schedule. Similar to the on-board studies, seafarers presented the symptom of sleepiness during afternoon watches and reported to be most sleepy during night duty. Collectively, these outcomes suggest that the $\text{shift}_{6/6}$ system induce more sleepiness than the $\text{shift}_{4/8}$ system, which can more likely cause sleepiness-related accident due to the shorter time for sleep and recovery [83].

A number of studies have reported that train drivers working on $\text{shift}_{8/8}$ system tended to have better sleep and subjective sleepiness outcomes despite getting less sleep than the baseline sleep schedule [13, 14, 84, 85]. While the $\text{shift}_{8/8}$ system schedule is the standard schedule implemented in the rail domain, there has been striking interest for this schedule for implementation within the maritime domain [15]. A shipboard study conducted by the UK Maritime and Coastguard Agency assessed the effectiveness of the $\text{shift}_{8/8}$

schedule amongst seafarers [15]. For comparison purpose, the $\text{shift}_{8/8}$ schedule was benchmarked against the $\text{shift}_{6/6}$ system schedule. When compared to the $\text{shift}_{6/6}$ schedule, seafarers who worked in the $\text{shift}_{8/8}$ system reported good subjective sleep quality (SQ) and greater sleep quantity, as those working in the $\text{shift}_{8/8}$ system gained additional two hours for rest and sleep. This observation was made in light of circadian and homeostatic factors. It was suggested that not only $\text{shift}_{8/8}$ system was able to balance the level of sleepiness at all time during the day, but also offers an opportunity to get full and restorative sleep that can indirectly reduce the accumulation of performance deficits across days. A review concluded that schedules with rotating start times (e.g. $\text{shift}_{8/8}$) were more likely to generate poorer sleepiness outcome, when compared to schedules with fixed shift start times across 24-hour period (e.g. $\text{shift}_{4/8}$ or $\text{shift}_{6/6}$) [10]. Nevertheless, schedule comparison in the review omitted variances in the working environment, such that, the $\text{shift}_{8/8}$ schedule involved railway industries, while $\text{shift}_{4/8}$ and $\text{shift}_{6/6}$ schedules involved the maritime industry.

2.1.2 Sleep

The knowledge of sleep architecture is a prerequisite when assessing the correlation of sleep stages with specific brain and bodily functions. Apart from sleep architecture, sleep has been generally evaluated in terms of quantity (duration) and quality. These sleep parameters can be retrieved from subjective measures, including sleep log and diaries, or more objective measures, for example, polysomnography and actigraphy.

2.1.2.1 Sleep Monitoring

Polysomnography

Polysomnography (PSG), also known as sleep study, is a multi-parametric test that examines the dynamic physiological changes during sleep. It involves

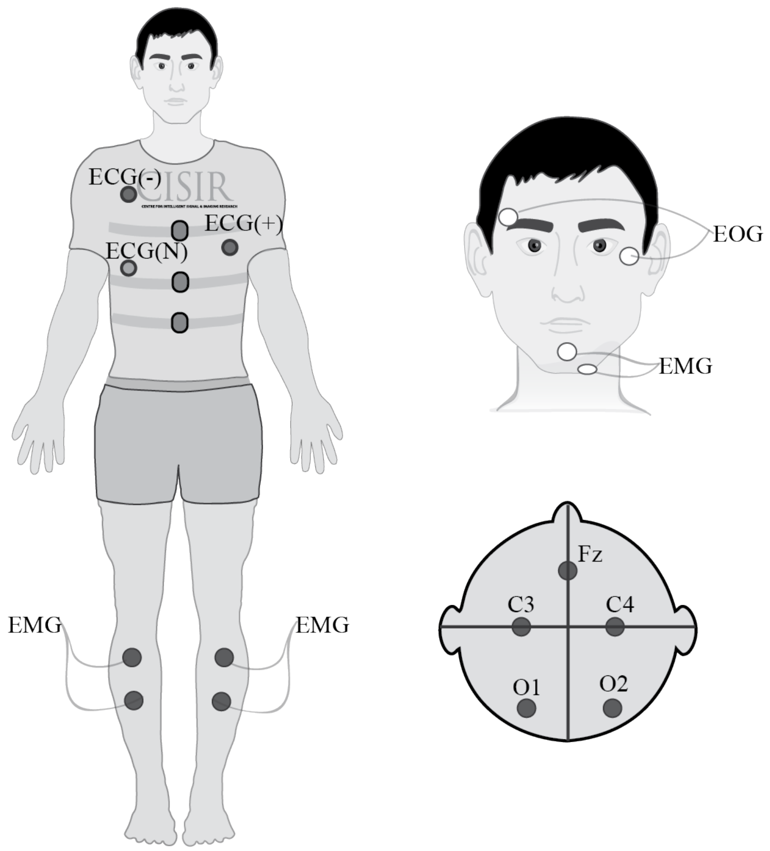


Figure 2.2: Typical PSG setup.

collective evaluation of the brain activity that can be gained via electroencephalography (EEG) recording, heart rate variation that is commonly measured using electrocardiography (ECG) technique, muscle tone that is frequently observed through electromyography (EMG) recording, and eye movements acquired by electroocoulography (EOG) recording. These EOG, EMG, ECG, and EEG record the electrical activities generated by the movement of the eyes, skeletal muscle activation, electrical potential initiation by the heart, and neuronal firing of the brain, respectively. As illustrated in Figure 2.2, electrical activities can be measured by positioning the electrodes at the right and left outer canthus, chin and limb, as well as scalp.

Of all the available alternatives, PSG has been considered a gold standard to assess sleep. Nevertheless, some drawbacks of using PSG are the prohibitive cost of the equipment, discomfort due to placement of many electrodes and wires on the body, time-consuming scoring of biophysiological changes, and requirement of practice prior to usage [86]. It is worth to mention that many studies have addressed the listed issues by proposing sleep monitoring systems that can automatically detect the sleep stages, minimised number of electrodes, and wireless transfer of biophysiological information to a designated processing unit [87].

Sleep log and diaries

Sleep logs and diaries (e.g., Pittsburgh's sleep quality Index [88]) usually require individuals to record several aspects of their sleep/wake habits, including bed and wake times, sleep onset latency (duration needed to fall asleep), both frequency and duration of awakenings, as well as if they were deliberate [89]. Despite the inexpensive and straightforward use of sleep log and diaries, validation studies against PSG test revealed that it common for people to overestimate their sleep duration, as well as underestimate both the number of awakening and sleep onset latencies.

Actigraphy

Actigraphy refers to a technique where sleep-wake patterns are recorded using actigraphs, which reflects a motion-sensitive device with memory storage. In several occasions, actigraphy includes the combination of sleep diaries that is used to identify the time in bed (TIB). These actigraphs are usually worn on the non-dominant wrist, although at times worn on the non-dominant wrist, ankle, or trunk. Regardless of the device location, the epoch (typically 1-min epoch) of inactivity is a gross indicator of the sleep duration. Relative to PSG, actigraphy devices are inexpensive (approximately RM 400 per unit) and free from first-night effect [90]. The unobtrusive actigraph enables one to assess the sleep-wake patterns for a longer duration. Actigraphy has high correlation with PSG-measured sleep parameters, and perhaps, a reliable measure of some sleep issues, such as insomnia or excessive daytime sleepiness, but without the capability of distinguishing between the sleep architectures.

2.1.2.2 Sleep Architecture

Some PSG studies have reported that normal human sleep can be divided into two distinct states; rapid eye movement (REM) and non-REM (NREM). The NREM sleep is further divided into three or four distinct stages based on the scoring system. The conventional standards initiated by Rechtschaffen and Kales [91] classified NREM into four stages, while the more recent American Academy of Sleep Medicine rule [92] has re-grouped NREM into three new categories; N1, N2, and N3 with N1 equivalent to Stage 1, N2 to Stage 2, while N3 is the combination of Stages 3 and 4. As most studies related to sleep to date seem to use terminology and scoring rules prescribed by Rechtschaffen and Kales, likewise, this thesis referred to the nomenclature and the sleep stages according to the standards introduced by Rechtschaffen and Kales, unless stated otherwise.

Stage 1 refers to the initial and the lightest of the four stages that is often seen as a transitional stage between wakefulness and sleep. Stage 1 is characterised by slow and predominantly horizontal eye movements, as well as limb twitches. The electrical activity in Stage 1 sleep is composed of low-voltage waves (amplitudes 50 – 70 µV peak-to-peak), mixed frequency EEG with eminence in the theta band, while the alpha activity reduces and diminished towards the end. The arousal threshold in Stage 1 is low, and consequently, sleep can be discontinued easily due to any sound deriving from a normal conversation or due to the sound made by gentle closing of the door. Those awakened from Stage 1 sleep may feel as if they have not fallen asleep yet.

Next, Stage 2 is determined by the presence of sleep spindles and K-complexes in the EEG recording, besides the minimised eye movement and reduced muscle tone. Similar stimuli that awaken one in Stage 1 may result in evoking K-complexes, but weak to trigger awakening from the Stage 2 sleep.

Stages 3 and 4 occur when the delta wave occupies about 20% of the EEG signals. These stages are collectively known as slow wave sleep (SWS) due to massive delta activities. If more than 20% but less than 50% of the EEG is composed of high-voltage (amplitudes >75 μV peak-to-peak) delta waves, the sleep falls in Stage 3, or else, classified as Stage 4. Limited eye movements and low muscle activity are some features of SWS. The SWS reflects the deepest sleep among all the NREM phases. A stronger stimulus, such as a deafening and startling noise, can cause arousal from SWS, when compared to those in Stages 1 and 2. Nonetheless, should sleepers are awaken, they may feel disorientated and perplexed.

If sleep is uninterrupted at the SWS stage, it will advance to the REM stage. This REM sleep is also known as "paradoxical sleep", mainly because the EEG pattern resembles that in alert waking, but behaviourally one remains asleep and unresponsive. The distinct variance between REM and NREM sleep is the burst of REM under closed eyelids and skeletal muscle, which become paralysed or quiescent. The EEG pattern in REM sleep displays relatively low-voltage with theta predominance, a portion of slow alpha activity (8-12 Hz), and sawtooth waves. REM sleep with high levels of EOG activity is termed "phasic REM," while "tonic REM" in the absence of eye movement [93]. Although internal arousal is typical during REM sleep [94, 95], it is difficult to awake a person from this sleep stage. The most vivid and prolonged dream occurs during REM sleep.

Both timing and arrangement of sleep stages are called "sleep architecture" that can be plotted visually in hypnogram, as illustrated in Figure 2.3. Healthy adults exhibit a cyclic fashion between REM and NREM sleep throughout a typical night sleep period [97, 98]. In one sleep cycle, sleep is initiated at Stage 1, followed by Stage 2, later descending further to the 'deeper' Stages 3 and 4, thus ending with the completion of the first REM stage [99]. During a regular nocturnal sleep episode, both REM and NREM cycles occur alternately for about four to six times with each cycle lasting approximately between 90 and 120 minutes to complete [97]. The proportion of REM and NREM in each

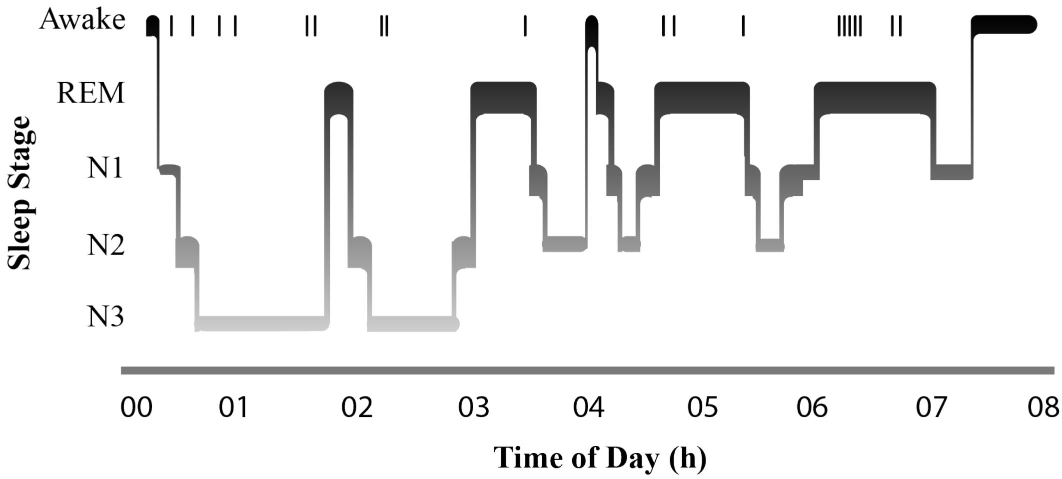


Figure 2.3: Hypnogram showing the typical adult sleep architecture during a normal nocturnal sleep episode. It can be seen that SWS predominate the first half of the sleep period, while stage 2 NREM and REM sleep predominate the last third of the night. Figure adapted from [96].

cycle alters across nocturnal sleep episode. The SWS predominates during the first and second cycles, while the episodes of REM sleep, which alternates with Stage 2 sleep, is extended towards the final cycles. A healthy young adult spends about 20-25%, 2-5%, 45-55%, and 18-23% out of the total sleep time in REM, N1, N2, and N3 sleep states, respectively.

2.1.2.3 Functions Of Sleep

Despite the decades of studies, there is no clear consensus regarding why a human needs to sleep and its functions are yet to be completely understood [100]. Determining the distinct sleep stages has yielded in-depth comprehension pertaining to the role of sleep. One of the earliest and longest-standing theories concerning functions of sleep is regarding their role in body restoration and energy conservation [101, 102]. Benington-Heller asserted that glycogen, which is the single largest energy reserve for neurons, is progressively diminished during wakefulness and it is replenished during NREM sleep [103–105]. The high concentration of growth-hormone promotes tissue growth and repair, especially following the onset of SWS that is in line with this theory [106–110]. Evidence also supports the notion that SWS has an energy-saving function, due to concomitant reduction in body temperature, pulse rate, respiratory rates, metabolic rate, and blood pressure [110]. The SWS can enhance one's memory during wakefulness; a process called declarative memory consolidation [111, 112].

Earlier studies depicted that REM sleep has a role in memory consolidation and learning [2], while recent evidence reveals that such a role is debatable [112, 113]. The concentration of noradrenaline (also known as stress hormone) [114], during REM sleep is lower than that during NREM or wake stage [115]. This suggests that REM sleep is involved in memory consolidation of emotional events during waking, since noradrenaline may have a role in arousal-related emotion processes. Thus, impairment in its regulation is linked with psychiatric disorder, such as that noted in patients diagnosed with post-traumatic stress disorder and major depression [115, 116]. Cerebral blood flow and brain temperature rise rapidly during REM sleep, when compared to NREM sleep. Such physiological changes may be a mechanism that prevents sleep from becoming too deep, thus allowing behavioural activation to be triggered quickly if required during REM sleep [117, 118].

Sleep evaluation goes beyond hygienic recommendation that an adult should obtain, which is at least eight-hour sleep per day (sleep quantity) for maximum performance [119, 120]. It is essential to complete the entire alternating cycles in order to gain the restorative effect from each sleep state. Interruption to any of the NREM stages (except stage 1) and REM sleep may cause the sleep to restart from Stage 1 Sleep. Alteration to sleep architecture can adversely affect SQ, which later affects one's alertness during the next wakefulness bout [97, 121].

2.1.2.4 Sleep Quality

Sleep quality (SQ) contributes to the recuperative benefits of certain sleep episodes that influences the cognitive performance during the next period of wakefulness [122]. The sleep architecture noted during PSG can be divided into REM and NREM [91]. NREM is composed of the following sleep stages: Stage I is the transition from wakefulness to sleep (Stage N1), Stage II refers to light sleep (Stage N2), and Stage III reflects deep sleep (Stage N3) [91]. For healthy individuals, sleep progresses in a cyclic alternating pattern between Stages N1, N2, and N3, later followed by REM sleep. Interruption or truncation sleep due to inadequate sleep duration, environmental interruptions from noise and light, consumption of alcohol or medication, or sleep disorders resulting from circadian misalignment, may result in adverse impact on SQ [97, 121].

SQ can be quantified in many ways which include sleep latency (how fast the time taken from the stage alert to Stage N1), frequency of waking up per night and sleep efficiency (SE) [123]. While there exist many ways to quantify SQ, SE is the most commonly being referred to in the literature of sleep studies. This is because SE covers the whole stage of sleep (i.e., cycling through non-REM and REM) [124].

In this work, SE was applied as an indicator to measure SQ, as given in Equation 2.1. SE is the percentage of total sleep time to the actual time in bed (TIB) [125]. The total sleep time is the addition of the total time spent in Stages N1, N2, and N3, as well as REM sleep, while TIB refers to the period that starts with light-out and ends with light-off [99].

$$SE = \left(\frac{N1 + N2 + N3 + REM}{TIB}\right) \tag{2.1}$$

A subject is deemed to have poor SQ when SE $\leq$ 85% [126, 127]. Long sleep onset latency, wake after sleep onset, and early sleep offset may lead to low SE [99], which suggests that one has spent a large proportion of sleep opportunity awake [124].

2.1.2.5 Sleep Among Mariner

The magnitude in quality and quantity of sleep varies with the shift system. In fact, two field studies [45, 80] reported a similar trend concerning shorter sleep episode for those working on shift$_{6/6}$ than in shift$_{4/8}$. In these studies, the mean period of total sleep time in each episode was approximately 3 hr 14 min and 4 hr in shift$_{6/6}$ and shift$_{4/8}$ systems, respectively. Another study found that the shift$_{6/6}$ system was associated with shorter sleep for 24-hour period, when compared to the shift$_{8/8/4/4}$ system (5.35±101 min vs 6.16±58 min) [128]. These results signify that despite the sufficient sleep time, most participants in these three shift systems spent most of their time being awake. Hence, SE was below 85% for shift$_{6/6}$, shift$_{4/8}$ and shift$_{8/8/4/4}$ system [45, 80, 128]. A variation was noted in SE between fixed- and rotating-shift schedules. The pattern of poor SE (lower than 80%) was similar for seafarers working under fixed (i.e., shift$_{4/8}$) and rotating (work in sequence between the shift$_{4/8}$ system and normal day time from 09:00 to 17:00 h) systems, with rotating shift displaying more significant adverse effect on SE [129].

Truncated and poor SQ is not unique to the maritime industry, but is in line with studies that have probed into other industries that apply the shift system. Typically, morning sleep after night work and an early night sleep prior to morning shift are linked with reduced sleep duration and poor SQ than a night sleep between 00:00 and 06:00 h [33, 82]. The reason for truncated early night sleep before morning shift is due to the inability to advance bedtime, along with the need to be awake very early in the morning [130]. While this failure may be partly social, it is mainly determined by the difficulty to initiate sleep during the forbidden zone for sleep. It is during this time of the day when the wake-promoting circadian signal is at its peak [130, 131], thus increasing sleep latency. The impact of weak wake-promoting circadian signal (circadian nadir) is evident by the difficulty to arise from bed, along with an unpleasant feeling during the early morning awakening [132]. Similarly, early awakening and truncated morning sleep after night work are the results of increasing wake-promoting circadian signal, as well as decreasing homeostatic drive for sleep [130, 133]. This has been verified in a bridge study that utilised sleep parameters derived from polysomnographic recording. It was found that the seafarers obtained their main recuperative sleep during the night time, and a short nap when rest period was available [82]. As a whole, the outcomes showed that most seafarers had difficulty adapting their circadian rhythm to the new light-dark schedule, thus resulting in poor sleep construct [134].

Seafarers work in unique condition as they work and rest in the isolated and confined ship setting. Inevitably, sleeping occurs in the presence of random noise within the vessel, ship motion caused by the waves, vibration induced by the engine, and exposure to organic solvents (e.g., diesel) [135, 136]. Therefore, it is imminent to assess the effects of these interruptions on sleep. A survey study revealed that noise did affect the sleep of seafarers and oil installation workers [137]. In two controlled laboratory studies [138, 139], Tamura and colleague discovered that exposure to ship engine noise at 60 and 65 dBA affected sleep architecture (Stage 2 and REM), thus deteriorating SQ. Another simulator study using the shift$_{6/6}$ system revealed that subjects exposed to motion environment experienced sleep disturbance (a study by [140] and cited by [141]).

2.2 Sleepiness Detection System

2.2.1 Sleepiness Indicators

In numerous safety-critical operations, it is common for an employee to undergo periodic fit-for-work tests, including an evaluation of their sleepiness level [142]. This, in turn, proliferates the demand for a formal process that periodically assesses the sleepiness level. In the last decades, various sleepiness indicators have been proposed, which could be classified into three types, namely (a) vehicle-centred, (b) operator-centred, and (c) hybrid. Here, the taxonomy was adapted slightly from that originally presented by [143] and [144]. In particular, the subjective sleepiness assessment, which was neglected previously, is now parked under the operator-centred sub-group. This adaptation is presumably valid given that the information about subjective sleepiness, in fact, was captured from the human. Figure 2.4

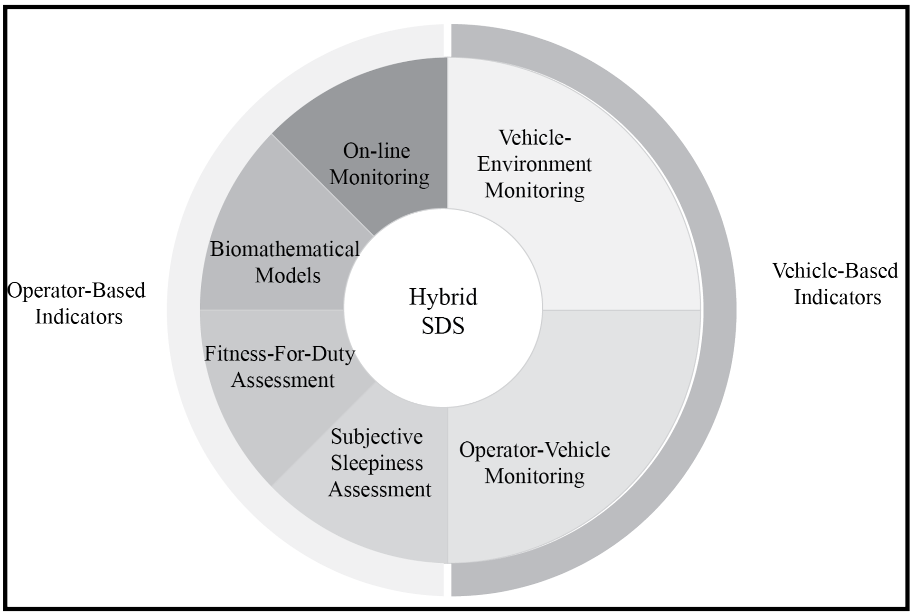


Figure 2.4: Sketch of the SDS taxonomy.

illustrates the three main types of sleepiness indicators, along with their sub-type categories. This subsection elaborates the types of sleepiness indicators, except for hybrid sleepiness indicators that are described in Section 2.2.2. It is worth noting that some of these indicators have often been applied as the benchmark to develop SDS.

2.2.1.1 Vehicle-based Indicators

Vehicle-based indicators operate based on the dynamic states of the vehicle or the position of the vehicle relative to varied landmarks in its surrounding. Vehicle-based indicators can be further sub-divided into two main categories depending if the behaviour displayed by the driver was evaluated based on his interaction with either operator-vehicle or vehicle-environment. Measurements commonly applied to determine operator-vehicle interaction are braking, gear changes, steering wheel movement (SWM), and grip pressure on the steering wheel [145]. Meanwhile, instances of vehicle-environment interactions include the standard deviation of lateral position (SDLP) and the distances between the vehicles [48]. The variability in SWM and SDLP appears to be the most predictive in detecting sleepiness-related driving events, which is vastly applied in commercial Advance-Driver-Assistance-System. For a sleepy driver, the rate of car weaving increases when measured using the SDLP, while the frequency of micro-correction on the steering wheel decreases during alert driving condition [146].

The SWM can be assessed directly through the steering angle sensor fixed on the steering column, whereas the lateral position can be evaluated by using a lane tracker that applies a camera that is mounted on the rear-view mirror and the windshield. A machine vision system is utilised to identify the driving lane boundaries, and subsequently, estimate the lane-departure of the travelling vehicle according to the perpendicular distance between lane tracker and driving lane boundaries. Performance bottleneck may derive from such a topology. The performance of SWM-based systems is mostly associated with their dependence on road geometry, including lane width, cross-section dimension, horizontal and vertical curves, as well as to a lesser extent, several dynamic features of the vehicle [147]. The visibility of road marking affects the validity of SDLP-based systems performance, lighting condition, and weather (snow or fog). Additionally, several other factors, such as varying traffic patterns, vehicle type, as well as driving habit and tendencies, may further complicate or weaken the correlation between sleepiness and vehicle behaviour.

2.2.1.2 Operator-based Indicators

The operator-based indicators reveal information about sleepiness directly from human prior or during the operation. Operator-based indicators can be grouped into four categories, namely (1) fitness-for-duty assessments to evaluate vigilance, (2) on-line monitoring of facial feature and physiological signal, (3) subjective sleepiness assessment, and (4) biomathematical models based on the interaction between the circadian and homeostasis processes.

(1) Fitness-for-duty assessment

Attention refers to a process that is applied to numerous pieces of competing information while selectively bias the selection to one option, and concurrently, suppressing interference from other competing alternatives at any one moment [148]. For example, visual attention can bias selection to only a subset of the incoming visual information, such as a particular feature, orientation, and coordinate [149]. Attention is believed to be a requirement for a more complex cognitive task [150]. A range of studies have shown that both vigilant and sustained attention can deteriorate due to sleep deprivation [151]. The Psychomotor Vigilance Test (PVT) [152] is the most commonly applied assessment to examine one's ability to maintain vigilance and sustain attention in studies concerning shift work in both laboratory and field settings. The PVT records response time (RT) to visual (or auditory) stimuli that appear at random inter-stimulus intervals (ISI). During a standard 10-min PVT, the subject is instructed to press a designated button (see Figure 2.5) as quickly as possible when a stimulus appears at random between 1000 and 2000 ms ISI over 10 minutes. The procedure commonly generates about 80 RTs per trial [153].

The PVT is often used as a 'gold standard' measure to track changes in vigilance or sustained attention due to its simplicity, coupled with minimal aptitude and training effect [155]. The training effect refers to a condition where the performance gradually improves with repeated assessment of cognitive test [155]. The PVT performance can be applied as proxy for "real-world" functioning (i.e., ecological validity), especially for tasks that demand continuous observation and immediate responses. For instance, some tasks, such as operating a vehicle, security-related tasks or operating and maintaining any complex instrument, demand stable and vigilant attention; otherwise, may end up in accidents. Due to its simplicity and portability, PVT has frequently been used as a sleepiness indicator in vast sleep-related studies. A number of PVT performance parameters can be extracted from each PVT

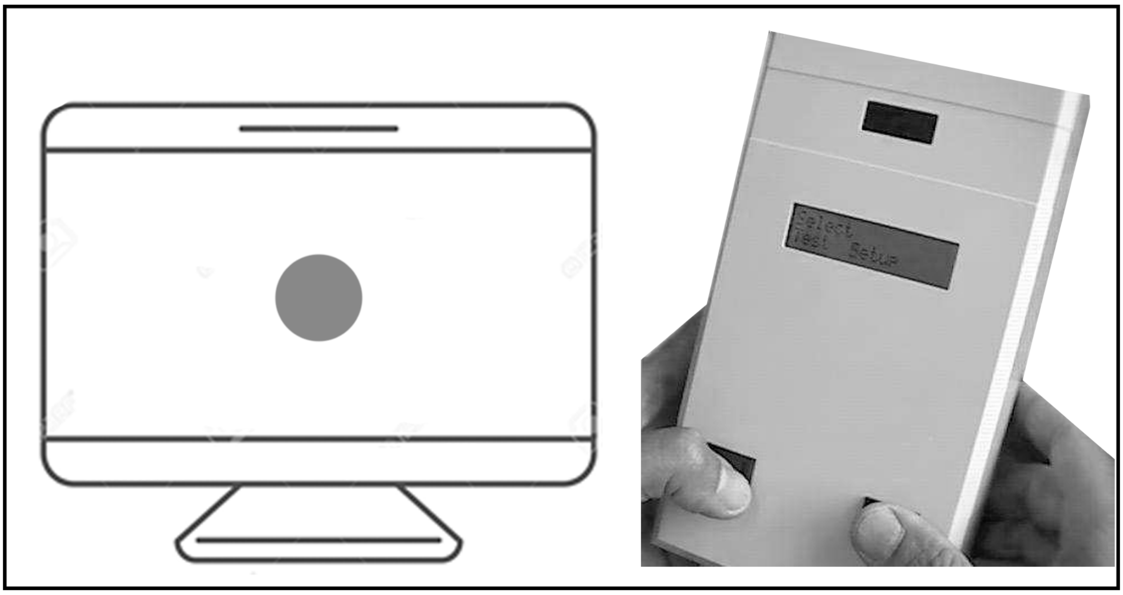


Figure 2.5: Typical setup using handheld-based PVT for short alertness tests. The subject is required to press a button in response to a visual stimulus presented on the monitor. The handheld-based PVT image was adapted from the product supplier's (Ambulatory Monitoring) website [154].

trial, wherein one of them is the fastest RT ($RT_{10\,\text{fast}}$) defined as the average RT of the fastest 10% of all responses, excluding false starts and lapses. Lapse and false starts are trials with RTs below 100 ms and exceeding 500 ms, respectively. The $RT_{10\,\text{fast}}$ has become arguably one of the most commonly used measures as it reflects the optimal alertness level [156], and the occurrence of lapse does not skew its magnitude [54]. Changes in $RT_{10\,\text{fast}}$ performance have been linked with brain activation, such as frontoparietal sustained-attention network, cortical and subcortical motor systems, basal ganglia, and visual cortices [156, 156, 157].

The $RT_{10\,\text{fast}}$ has been demonstrated as sensitive to sleepiness during chronic and prolonged wakefulness. Lim and Dinges, in their review, reported that sleep deprivation increased the $RT_{10\,\text{fast}}$ [158]. Chronic sleep restriction to about 33% of the baseline sleep period resulted in a higher mean value of the $RT_{10\,\text{fast}}$ during sleep restriction, when compared to baseline week [159]. In fact, even restricting sleep from the baseline condition of eight to five hours can lead to a significant increase in the average fastest 10% of RTs [160]. The

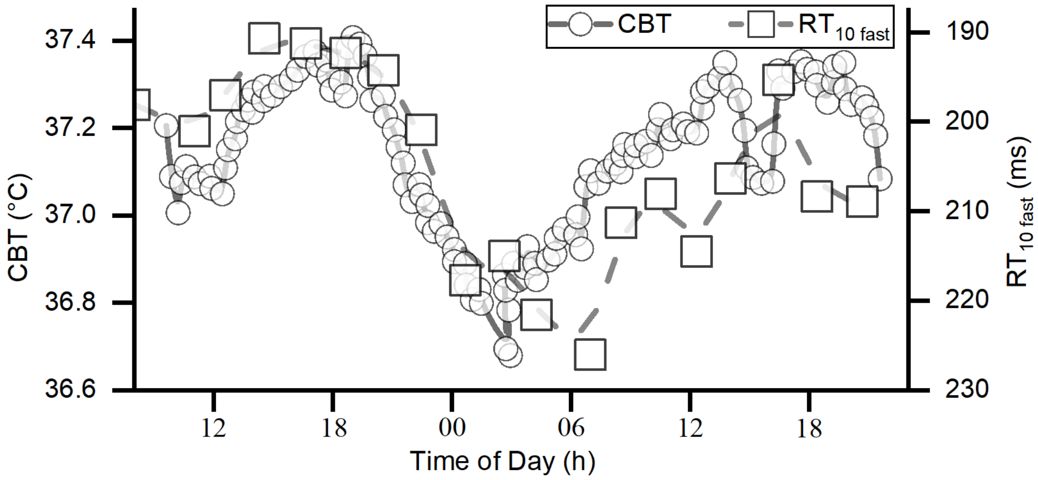


Figure 2.6: Circadian variation across a 40-h total sleep deprivation for the $RT_{10\,\text{fast}}$ (□) and CBT (○) as assessed by a rectal thermistor. Figure adapted from [150].

fluctuation of this measure follows a time-of-day effect. In a 40-hour total sleep deprivation protocol, Goel and colleague [150] discovered that the dynamics of $RT_{10\,\text{fast}}$ was phase-locked to the CBT fluctuation, as portrayed Figure 2.6. They observed that the best and fastest $RT_{10\,\text{fast}}$ occurred in the afternoon, which was closer to the CBT maximum and dip in the performance in the early morning that was closer to the CBT minimum [150].

(2) On-line monitoring

It has been reckoned that deterioration in cognitive performance due to accumulating sleep propensity is accompanied by notable alteration in facial features and physiological measures [161]. These indicators can be monitored by using direct sensor contact or non-contact (e.g., image acquisition) techniques. Those feeling sleepy may exhibit subtle changes in their facial features, particularly facial expressions, head movement, mouth opening, and eyelid movement. As the head faces straight when driving, it tends to face down or to the side as the muscle tone decreases with drowsiness. Yawning reflects an involuntary action of opening the mouth wide, while repeated yawning is a tell-tale sign that one is feeling sleepy. The recent interest in developing SDS has been intensified through the use of variables derived from eyelid movements, especially PERcentage CLOSure of eyelids (PERCLOS). PERCLOS, which is defined as the proportion of time in a minute that the eyes are at least 80% closed over the pupil, increases in its magnitude with accumulating level of sleepiness. The changes in facial features may be determined by using image acquisition technology that can be mounted on the windshield or partially behind the windshield mirror facing the driver, as it is assumed that the driver sits in the same position.

A number of physiological measures have been commonly applied as the input signal to the SDS, whereby brain activities are identified via EEG and fNIR recordings, heart rate variation [162, 163] is measured via ECG technique, and eye movements are determined from EOG recording. Except for fNIR [164] that records the hemodynamic response of the cortical brain activity; ECG, EOG, and EEG record the electrical activities generated in the heart, movement of the eyes in the sockets, and brain regions, respectively. Both electrical activity and hemodynamic response are measured by placing the electrodes (known as "optode" for fNIR) to the chest, at the right and left outer canthus, as well as the scalp, during ECG, EOG, EEG, and fNIR procedures, respectively. Amongst these physiological indicators, the EEG signal appears to be the most reliable and predictive, as it reflects the regulation of sleep-wake activities in the regional brain area directly.

(3) Subjective sleepiness assessment

Humans can experience and subjectively express their perception of sleepiness [165, 166]. This ability is viewed as a precautionary action by the body, so that reasonable actions can be taken to avert risks [166, 167]. In an operational situation, immediate information about one's level of sleepiness is inaccessible. Here, self-subjective-introspection is the only way to gain feedback [168, 169]. For instance, in the case of a mariner who operates the vessel alone for a long duration, self-monitoring seems to be the sole sleepiness-assessment technique. This type of self-assessment assists one to decide whether to continue or to halt temporarily the task he is engaged in or to use a countermeasure, such as pharmaceutical agents, or simply sleep. Therefore, the ability amidst workers to gauge their level of sleepiness has numerous implications to safety critical activity [165].

A subjective frame that is well designed assists one to express to the best his subjective feelings. In fact, Epworth Sleepiness Scale (ESS), Visual Analogue Scale (VAS), Stanford Sleepiness Scale (SSS), and Karolinska sleepiness scale (KSS) are some of the commonly applied sleepiness scale. Although all these scales are purely based on psychological estimate by the subjects, each measure has varying sleepiness characteristics. The ESS is a self-administered questionnaire, in which the subject has to imagine and rate the likelihood of falling asleep based on eight hypothetical scenarios, namely "sitting and reading," "watching the television," or "sitting and talking to someone," with scores scaling from zero ("would never doze") to three ("high chance of dozing"). The final score refers to the total score of the responses, which ranges between zero and twenty-four, with values exceeding eleven being linked with excessive daytime sleepiness. A VAS reflects a horizontal line that is 100-mm long, with each end of the line anchored with a word descriptor that corresponds to the perception of the extreme, such as, "very sleepy" and "very alert". The subject is required to mark at a point along the continuum that best suits his perceived level of sleepiness. The SSS is composed of 10 items that range from 'feeling active, vital, alert, or wide awake' to 'no longer fighting sleep, sleep onset soon, or having dream-like thoughts', wherein some items are far from being related to sleepiness, but more likely boredom [170].

To this end, the KSS [171] appears to be the most popular subjective sleepiness assessment implemented in numerous studies pertaining to sleep. Based on the studies reviewed in this study, KSS had been applied throughout. The KSS refers to a nine-point scale of sleepiness that ranges from "extremely

Table 2.1: The interpretation of each level for the Karolinska Sleepiness Scale.

Level	Interpretation
1	Extremely alert
2	Very alert
3	Alert
4	Rather alert
5	Neither alert nor sleepy
6	Some sign of sleepiness
7	Sleepy, but no effort to keep awake
8	Sleepy, some effort to keep awake
9	Very sleepy, great effort to keep awake, fighting sleep

alert" to "very sleepy, great effort to keep awake, and fight sleep" (see Table 2.1). This scale measures the propensity to fall asleep [170], as it reflects the sleepiness level in absolute terms and easy to comprehend [165]. The behavioural criteria, especially in the intermediate condition, is thought to clearly separate the milder forms ("sleepy, but no effort to keep awake") from absence of sleepiness ("neither sleepy nor alert") [170]. Some studies looked into the relationships between KSS and alterations in cognitive performance [165, 170, 172], driving [166, 173–177], and other subjective sleepiness scale [172, 178]. Increment in subjective sleepiness is in line with increased levels of energy for alpha and theta bands with open eyes condition [170, 171, 174, 178–181]. Interestingly, a recent review concluded that KSS is a robust estimator of drowsy driving in relation to EEG [182]. Moreover, the KSS has been used to evaluate sleepiness in field shift work studies [15, 82, 183–186]. More importantly, KSS is entirely consistent between differing persons. This notion is critical, mainly because variance between people is an essential concern to construct an effective SDS [187].

KSS of eight and more were co-varied strongly with accidents in driving simulator studies [166, 169, 180, 188, 189]. Also, eight or more scores for KSS associated with an extended duration of eyelid closure, and increased brain electrical activity which reflect some efforts in maintaining wakefulness as the propensity to fall asleep increases [171]. Åkerstedt [171] found that increment in subjective sleepiness (measured using KSS) was accompanied by enhancement in eyes-open alpha and theta brain activities. Such observation is of particular interest as high levels of theta energy signify the onset of sleep [190–192], while increased alpha activities means increment in mental effort to maintain vigilance [193]. This particular observation has motivated the researcher to group the driver performance into two or three classification based on their subjective sleepiness (i.e., obtained from KSS). In precise, the driving performances noted during the driving period with ≤ 6 and KSS ≥ 8 were designated as alert and sleepy, respectively. As for the ternary problem, the driving performances recorded for driving durations with ≤ 5, $6 \leq$ KSS ≤ 7, and KSS ≥ 8 were designated as alert, mild drowsy, and sleepy, respectively [175, 176, 194–197]. Such technique to group the data point into either binary or ternary state are denote as simple thresholding technique binary ($\mathrm{STT_b}$) and simple thresholding technique ternary ($\mathrm{STT_t}$), respectively.

(4) Bio-mathematical Models

As elaborated in Section 2.1.1, sleepiness is predominantly regulated by the interaction between circadian and homeostasis processes. This serves as a pillar in building bio-mathematical models (BMM) of sleepiness. The BMM is classified into two groups, namely (1) phenomenological type that is derived from empirical observation, and (2) physiological type that is based on physiological principle. While the requirements for both mathematical abstraction and inputs (e.g., sleep data, work/sleep schedules, light intensity) tend to differ, all BMM regardless of their different groups similarly applies the knowledge of coexistence between circadian and homeostasis mechanisms to predict one's future output (or metrics). The standard outputs produced by BMM are inclusive of subjective alertness, neurobehavioural performance (e.g., reaction time, number of lapse), and metrics that are associated with fatigue-related risk of operational accidents [198, 199]. The BMM has served a number of functions, especially in aiding the scheduler to plan a schedule with improved sleepiness risk management. As part of their methodology in predicting the alertness/performance of the operator at varied duty schedules, preparing optimal forward scheduling, and assessing the level of sleepiness during an unplanned change to the original schedule [7, 200, 201]. Several investigation bodies have implemented the BMM to support their probe by examining the level of sleepiness amidst operators prior to undesired incident [202]. As part of the Fatigue Risk Management System, BMM serves as a tool meant to educate shift workers, including but not limited to factors that cause extreme sleepiness or the appropriate use of countermeasure in minimising sleepiness-related risk [202]. The alternative rosters analysis carried out by BMM has been employed as supporting evidence submitted to the regulator upon request to work exceeding the prescribed hours-of-service regulations [202].

(a) Phenomenological-type

The earlier version of BMM was phenomenological (hence the name) in nature, as its parameters and mathematical structures were initiated by curve-fitting the model output to the dynamic changes of CBT, melatonin, EEG, and cognitive performance for a range of sleep and sleep-deprivation empirical protocols [203–205]. The present use of BMM by the military and commercial transportation industries includes the Three-Process-Model of Alertness (TPM), Circadian Alertness Simulator, and Boeing Alertness Model, to name a few [198, 206, 207]. Nevertheless, as these models seemed to lack in physiological framework, their behavioural predictions also appeared to lack in clear physiological interpretation [208].

(b) Physiological-type

Since past few years, the integration of multiple imaging modalities has advanced the chronobiology field, especially in elucidating the physiology and the anatomy aspects, which regulate the sleep-wake rhythm in human. Upon determining the sleep-wake regulatory in the brain network, which is stratified by its role in promoting sleep or wake stage, a range of physiology-types have been proposed. A physiology-type reflects the sleep-wake behaviour within the context of a neuronal population that is responsible for sleep-wake transitions. As a result, the physiologically-type becomes more versatile and has the ability to relate behavioural prediction with underlying neural, physiological [209, 210].

In the context of physiology-type, the model of arousal dynamics (MAD) has been widely employed in studies concerning shift work [211]. The model refers to an integration of two earlier models, namely Phillips and Robinson's model (PRM) [212] and the model of human circadian pacemaker [213]. Figure 2.7 illustrates a schematic of the model. The PRM portrays the correlation between two vital neuronal groups, namely sleep-promoting neurons found in ventrolateral preoptic region of the hypothalamus (VLPO), and wake-active monoaminergic (MA) neurons found in brainstem, along with circadian and homeostatic drive. Both VLPO and MA reciprocally inhibit each other to generate sleep-wake transitions. The MA receives combined inputs from orexinergic and cholinergic, as well as other neuronal populations, which are simplified as constant input, D_m.

The circadian (C) and homeostatic (H) drive regulate the timing of sleep-wake transitions with the former and latter drive each inhibiting and promoting, respectively, the VLPO firing activity. The drive to VLPO is denoted by D_v. The homeostatic drive is related to the level of adenosine, a by-product of cell metabolism, which increases during wakefulness and decreases as net clearance exceed production during sleep [214]. Adenosine is thought to contribute to promoting the VLPO firing activity partly [167]. The model shows that the contribution of homeostatic drive to VLPO increases and decreases when the MA mean firing rate is high (wake state) and low (sleep), respectively [215]. The circadian drive is modelled as the standard oscillator that generates self-sustaining oscillation by using the model initiated by St. Hilaire et al. [213]. The model illustrates the interaction of varied light intensity and other non-photic variables that could influence the circadian

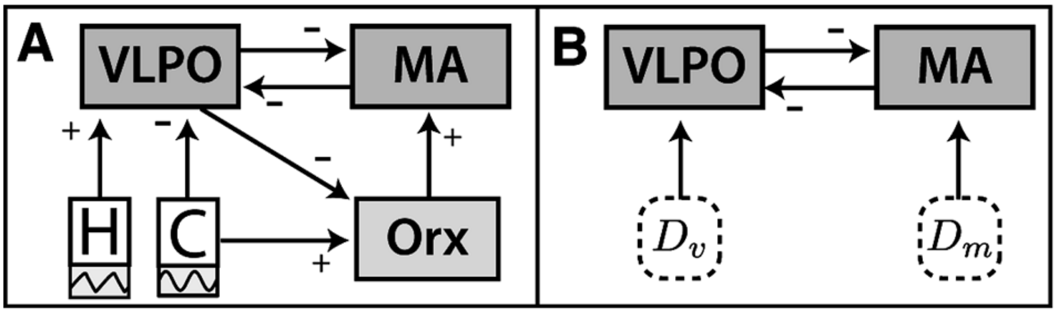


Figure 2.7: Illustration of dynamic interaction between sleep-wake networks regulated by the interaction of circadian(C) process and homeostatic (H) processes. The ventrolateral preoptic nucleus (VLPO) nuclei and monoaminergic (MA) are activated by D_v and D_m, respectively, and consequently, influence the sleep-wake transition. D_v is a combination of homeostasis and circadian processes. Figure adapted from [208].

dynamic.

Considering both homeostatic and circadian drive, the total sleep drive (TSD) to the VLPO is described by the following formula:

$$TSD = V_{vh}H + V_{vc}C + D_0 \tag{2.2}$$

where V_{vh} and V_{vc} are the promoting and inhibiting characteristics of homeostatic and circadian processes, respectively, to the VLPO firing activity. Meanwhile, D_0 represents the baseline sleepiness level that was assumed equal for all subjects [216, 217]. The detailed description of MAD algorithm is given in Appendix A. Figure 2.8 displays the correlation between homeostatic and circadian that results in TSD.

Although the phenomenological value of TSD is not widely established, it has been recently indexed as equivalent to subjective sleepiness (i.e., KSS) and cognitive performance (i.e., PVT) [211]. The TSD is fit to subjective fatigue data gathered during the total sleep deprivation protocol [208]. Interestingly, the subjective fatigue pattern was subjectively similar to other EEG slow wave activity pattern retrieved for total sleep deprivation [208]. This observation displays that increased suppression to MA may increase EEG slow wave activity [212]. Therefore, this study assumed that sleepiness can be determined by TSD [215].

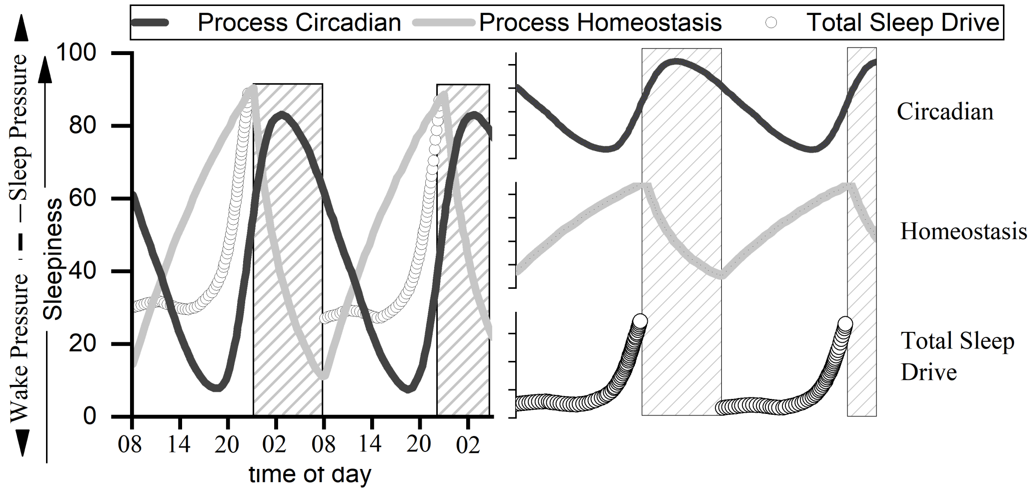


Figure 2.8: The interaction between the circadian and homeostatic process as modeled by MAD.

Additionally, Postnova and colleague have spearheaded a number of exploration studies by including the MAD. For instance, they have looked into the impact on sleepiness when commencing a permanent shift schedule at various time of the day [215], identified the best interval days to allow fast adaptation during forward rotating three-shift system [217], and determined the need for enforced wakefulness for entrainment to permanent shift [216]. Interestingly, the empirical assessment in these three studies was conducted upon appropriate parameter fitting of the MAD according to a single experimental protocol observation.

Factors that affect the successful implementation of operator-based indicator

One critical drawback of the electrophysiological-based techniques refers to the easily-contaminated informative signal from physiologic and extra-physiologic artefacts. Physiological artefacts are signals that derive from unwanted eye and head movements, heart, scalp muscle, tongue, skin stretch, and sweat gland activity. Meanwhile, extra-physiologic artefacts are generated from electromagnetic field interference [218], poor electrode attachment, and movement of others around the vicinity of the recording area [219]. Such artefacts can strongly affect the electrophysiological recordings, despite in a perfect laboratory setup [219–221]. Failing to treat these artefacts properly may lead to very low signal-to-noise ratio, and eventually, unreliable classification [222–224]. In a similar vein, the ability of the facial feature image to distinguish varying levels of sleepiness highly relies on image quality. Despite the improvements attained these recent years, gaining good image quality may be problematical in bright or poor light conditions [225]. The accuracy of the eyelid closure measurement may be influenced by glare reflection from the glasses worn by the subjects or their face being intermittently outside of the detection angle of the recorder during image acquisition [222]. Although subjective sleepiness assessment has been acknowledged as a strong predictor of performance [172, 226]; one cannot say that its correlation is perfect [184, 227]. In precise, the rating may be confounded when the subject is unable, or the subject could define each item from a different stance, thus resulting in response biases [228].

Despite the wide usage of commercial BMM across a range of operational setting, its reported performance is rather scarce in the literature, coupled with no recent updates on its review. It is well documented that individuals differ markedly to their vulnerability to sleep loss. Historically, BMM was parametrised by using group-average (mostly young and healthy subjects) metrics, thus does not lend itself to the assessment of individual level of sleepiness. As a result, the performance of BMM may differ substantially between subjects since the prediction is based on group, instead of individual-specific prediction. Besides, some studies have focused on individuals by implementing a different extension of phenomenological-type BMM. Despite the promising outcomes, the individual-specific BMM has neither been commercially released nor validated in a field-based setting. A recent report involving TPM, Circadian Alertness Simulator, and Boeing Alertness Model had weighed in the issue of individual variances by introducing specific individual input, for instance, the type of chronotype and habitual sleep length [199]. Nonetheless, none has validated the ability of these models in addressing the effect of inter-individual difference.

2.2.2 Fusion Of Sleepiness Indicators

Since each reviewed indicator has a different set of benefits and drawbacks, as well as the imperfect nature of the algorithm applied to extract informative features [161], an ambiguous, incomplete, and uncertain SDS may be resulted if used singly. Hence, the best way to overcome this issue is by combining a number of indicators that can enhance the performance of SDS. Typically, a system that integrates the vector of D feature values, $x = (x_1, x_2, x_3, \cdots, x_D)$, from multiple modalities is known as 'data fusion' or 'hybrid system'. In any classification method, the data obtained either from single or various sources must be comprehended via data mining and fusion process [229]. These two processes are complementary to each other, and may determine the success of an automated classification process across a range of application domains. Data mining refers to the search for a model via abduction, and followed by induction. Abduction reflects the generation of a model hypothesis that is employed to depict a particular data set. Let $c_1, c_2, c_3, \cdots, c_N$ denote varying classes, while z stands for the categorical variable linked with each pattern that represents the class membership. Here, if $z = k$, then the feature vector is belongs to c_k, k $\in \{1, 2, 3, \cdots, N\}$. The primary objective is to produce a model hypothesis that can identify the class to which the feature vector belongs to if it falls in the region of Ω_k, $k = 1, 2, 3, \cdots, N$. Thus, a feature vector is assumed to belong to c_k, $k \in \{1, 2, 3, \cdots, N\}$ if it falls in the region of Ω_k. The line or plane between the regions of Ω_k is the decision boundaries or decision surfaces, respectively. As for the induction, the model hypothesis for the data sets is extended to make a general assertion [229]. In the sleepiness detection problem, the sleepiness indicator (see Section 2.2.1) have been studied and reckoned, thus subsequently applied to establish the model hypothesis. Upon establishing a model, the learned hypothesis is fused and applied to multiple data sources to automatically infer the presence of a target.

Machine learning (ML) refers to an intelligent learning approach that allows computers to acquire knowledge automatically. The ML has been implemented to solve a number of complex and real-world issues that lurk in a range of applications. To date, abundant efforts have been devoted into constructing an SDS by adopting varied MLs and a combination of the varying sleepiness-related indicators. The reason behind the overwhelming response is due to its efficiency and ability to develop classifiers/hypotheses that may depict high dimensional and noisy sleepiness-related data. The proposed algorithms in relation to this area can be classified into four groups, namely threshold-based approach, knowledge-based approach, statistical-based approach, and probability theory-based approach. As the locus of this study is to combine varied indicators, the following discussion highlights studies that have merged heterogeneous information in detecting sleepiness. A thorough review of studies that have exploited individual sleep-related signals is found in [230].

2.2.2.1 Thresholding

A non-intricate and fast method, yet computationally inexpensive in classifying the level of sleepiness is known as 'thresholding'. The key for optimal classification performance in this method refers to the selection of an appropriate cut-off value for a signal. A study proposed a driver drowsiness detector, whereby signals extracted from EEG and PERCLOS were assessed in parallel [231]. The sign of drowsiness extracted from these signals was converted into scalar value and compared with a pre-assigned threshold. A scalar value that exceeds the final rejection threshold reflects sleepy condition. Similarly, a recent investigation determined the specific threshold values of eye and head movements, as well as brain alpha activities for several levels of sleepinessy [232]. While this thresholding method may be viable for most of the time, it could perform poorly if the data are not linearly separable, if the signal is too intricate, or if too many dimensions are involved.

2.2.2.2 *Statistical*

Support vector machine

Khushaba and colleague built a drowsy-driver detection system using Support Vector Machine (SVM) as the classifier, whereby the data were gathered from EEG, EOG, and ECG [233, 234]. Despite using three EEG channels alone may generate satisfactory classification performance, when compared to using either ECG or EOG. Also, the classification accuracy was enhanced with the fusion from either EOG or ECG with the combination of EEG and ECG, which appeared better than EEG combined with EOG. As EEG recordings have been proven to be predictive and a reliable tool to detect transitions from the alert to drowsy state, some studies have incorporated this feature, along with heart rate [235], respiration [236], PERCLOS [237], and road condition [238] by employing the paradigm of SVM to detect drowsy driving in a simulated driving environment.

In a recent study, the classification confidence of SVM for each feature data (facial, vehicle behaviour, sleep condition, driving time, and environmental temperature) was turned into a probability score [239], which was then fused using the Dempster-Shafer theory to address conflicts during multi-source information fusion. Another interesting method that embedded multiple sleepiness indicators using SVM is found in [196], whereby five ML algorithms (random forest, Adaptive Boosting, k-Nearest Neighbor, linear SVM, and Gaussian kernel SVM) were compared with their input data deriving from EEG, EOG, TPM, and a range of driving performances (i.e., steering behaviour and lane positioning) recorded during real road driving. Although the performance of the classifier was near similar for random forest, Adaptive Boosting, and Gaussian kernel SVM; the results were concluded based on the outputs derived from random forest classifier that displayed better success rate after integrating TPM feature with physiological and driving behaviour, in comparison to using only the physiological and driving aspects [196].

Fisher's linear discriminant analysis (FLDA)

Recently, Nguyen and colleague looked into the potential of combining EEG and fNIR features into the Fisher's linear discriminant analysis (FLDA) to detect drowsiness [240] investigate the potential of combining EEG and fNIR feature into the FLDA to detect drowsiness. Interestingly, they revealed that the classification performance was higher when both EEG and fNIR signals were used together, as compared to EEG or fNIR alone [240].

Artificial neural network (ANN)

Some studies have employed Artificial neural network (ANN) to classify a two-class problem (alert and drowsy) during simulated driving based on the features extracted from driving behaviour, eye features, and head motion [195, 241–243].

The computation of ANN is rapid, and hence, may be applied to detect real-time drowsiness. In a study, irregularity analysis of wheel angle and yaw angle was carried out to develop an ANN-based drowsiness classification in a naturalistic setup [244]. The ANN in the application of drowsy driving detection has been employed to integrate a number of physiological features, such as EEG, EOG, and EMG [245, 246], as well as the combination of features derived from EEG, head nodding angle, eye-tracking signal, time-of-day, and time-on-task [247]. Another interesting work in this direction assessed a variety of driving behaviour indicators, wherein these signals were combined with the TPM [175].

Neuro-fuzzy (or fuzzy-neuro) is the fusion of ANN and fuzzy logic, such as the Takagi-Sugeno fuzzy neural network (TSFNN). The integration of these classifiers generates a more intelligent system that mimics the human reasoning process, apart from having the ability to learn and connect the structure of NNs. A study used several vehicle behaviour and facial features, along with TSFNN as the classifier [248]. In a drowsy driving classification study, some features retrieved from EEG, ECG, eye movement, and SQ were passed to the TSFNN system for parameter optimisation and classification [249]. Extreme Learning machine (ELM) refers to a fast and novel model linked with ANN. Unlike the conventional training procedure in ANN, the correlations in ELM between input layer and hidden neurons are randomly assigned at the learning phase. The computational burden of ELM is significantly minimised as the random hidden neuron parameters are retained at the learning phase. Chen introduced an automatic method to identify the drowsy status in EEG and EOG records, which were then fed to an ELM classifier [250].

Logistic regression (LR)

Initially applied in the field of statistics as a method to assess the correlation between two or more independent variables [251], the logistic regression analysis has found its application in the ML domain. The logistic regression (LR) model has been used to estimate the subjective sleepiness level based on physiological (EEG, ECG, EOG) and behavioural (neck bending angle) signals [252]. Another fascinating use of the logistic regression is presented in [194], in which the level of sleepiness of driver was predicted by linearly combining the features extracted from eye movement and TPM.

2.2.2.3 Knowledge-based Approach.

The knowledge-based approach refers to a computing framework that revolves around the idea of linguistic variables and fuzzy if-then rules to represent the qualitative elements in human knowledge and reasoning [253]. The critical process in this approach reflects identifying the fuzzy logic membership functions that map the degree of membership for varied linguistic variables in the interval between zero and one [254]. As a result, the fuzzy method generates more flexibility in determining a diverse aspect of incompleteness and vagueness for an event. In fact, a study proposed a driver sleepiness detection system by fusing multiple cues, including PERCLOS, head nodding, slouching, and postural movement, by applying the linguistic concept modelling of the fuzzy technique [255]. In a similar vein, an online drowsiness detection system, which employed the linguistic variable concept and approximate reasoning methods, was developed by embedding features extracted only from EEG and EOG [256]. It was reported that merging EEG and EOG information resulted in a system that is more robust to inter-individual variances [256].

2.2.2.4 Bayesian Network

While SVM and NN remain as popular approaches, an alternative family of ML methods known as Bayesian Network (BN) has been gaining considerable attention with promising avenues for modelling multi-component systems [47]. The BN is a probabilistic graphical model that integrates graph and probability theories [48]. The idea of presenting a set of sleepiness indicators into a unified framework has been extended in order to address the drowsy driving recognition model based on BN. The trend was initiated by Ji and colleague [51], who highlighted the importance of embedding contextual information, such as working environment, sleep environment, quantity and quality of sleep, time of day, and total wake time.

Nevertheless, as a Static-BN was employed, the system did not weigh in the dynamic features of sleepiness. Sleep propensity accumulates with awakening time and relies on the former state. Some studies have proposed Dynamic-BN (DBN) and fused a range of contextual features with varied sleepiness indicators. In a study, the DBN served as a classifier to discriminate the two-state (alert and drowsy) problem by using sleep quantity, driving conditions, and time-of-day as the contextual factors (CFs), whereas eye movement, EEG, and ECG as the observable physiological data [257]. As a result, the decision maker could perform better when physiological and contextual information was integrated, whereas the absence of EEG and ECG data (considering only eye movement + contextual information) affected the accuracy of predicting sleepiness. A BN model was proposed to detect drowsy driving by considering physiological observations (EEG, EMG, and respiration) and contextual knowledge (SQ, time of day, and driving conditions) [258]. The fusion of the variables resulted in a larger Area-Under-Curve than the other single features (contextual information + physiological observation (RESP, EMG, EEG) > RESP > EMG or EEG).

A study proposed a technique to detect drowsiness by applying BN through the incorporation of EEG and head-based indicators (HB) as the observable features, while time of day and time on task (TOT) as the contextual information [224]. The combination of all variables enhanced the drowsiness state discrimination, whereby the classification outcome as (EEG + HB + time of day + TOT) was better than (EEG + HB), which was better than EEG, while the combination of (EEG + HB + CR + TOT) was better than (EEG + CR + TOT), which was better than EEG data alone. Information linked with driving behaviour and road context was encoded for a BN-based drowsy driving recognition system [230]. Other studies assessed the feasibility of fusing contextual information with varied features extracted from physiological indicators, image acquisition technology, and vehicle-based estimator with BN serving as the classifier [259, 260].

2.2.2.5 Other Technique

Other types of classifiers have been experimented to fuse multiple indicators, such as Dempster–Shafer theory [225], decision tree model [261], random forest [196], Adaptive Boosting [196] and k-nearest neighbour [196]. As a subset of ML, deep learning in recent years has gained considerable attention within the sleepiness detection domain. Since deep learning is based on learning representations of data, this approach can automatically map raw data into a more optimal representation via consecutive non-linear transformations. This process omits the need to perform feature selection beforehand as generally performed in the conventional ML method. Similar to the conventional NN, a deep NN is composed of an input layer, two or multiple hidden layers, many units in each layer, and an output layer. Nonetheless, conventional NN and deep learning differ in the parameterisation of the deep NN structures. Some of the deep learning architectures include long short-term memory. In a recent study, Long-Short Time Memory was developed to classified driver's drowsiness from raw EEG and forehead EOG signals [262]. Nevertheless, it is known that training of deep learning requires a high number of training data set.

2.2.3 Application-specific Consideration

The indicator used for a particular occupation or application cannot be generalised to others. In precise, the applicability and practicality of each indicator are subjected to the dynamic and the mobility of the work system [263]. The dynamic of working place is linked with the response latency to gain information about the surrounding environment. One who operates in a low system dynamic needs no immediate intervention to information from the system, while those working in a highly dynamic system need to frequently detect alterations in their surroundings and respond accordingly in a continuous manner. Example of an operator who works in a low system dynamic can be the controller work in a factory that fits with automatic level control equipment. Despite missing some relevant system information due to increasing sleepiness, these do not necessarily trigger any immediate negative consequence to the person or the surrounding due to the build-in multi-stage fault warning, wherein the operator still can react appropriately based on the warning.

On the contrary, a lookout officer who works in the high dynamic environment has to be vigilant constantly and needs to respond immediately to unexpected changes and events around their immediate surroundings, such as the movements of other ships. A slight disruption or a moment of inattention could lead to severe incidences. Accordingly, SDS is not only

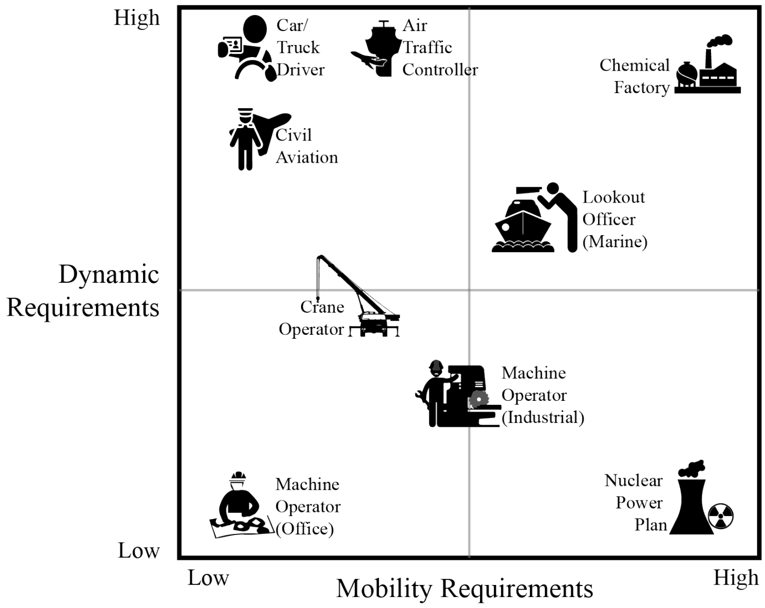


Figure 2.9: Mobility-dynamic system matrix. Figure adapted from [263].

meant to detect the state of sleepiness, but also to predict such states in advance in handling the risks posed by increasing sleepiness. If an operator falls asleep inadvertently during watch-out, it may be already too late to prevent an accident. It is even more critical for the design of sleepiness monitoring system to reflect the mobility of the operator in relation to the work task. The mobility of a working place could be either stationary or mobile. An example is a pilot who is immobile and is fixed to his seat, i.e. the position of the body is precisely marked, while the line of sight is mainly directed to the deck and towards the sky. On the other hand, a lookout officer can move freely and is not fixed to a specific geometric position or orientation. Nevertheless, different sets of setups may be present between the dynamic and the mobility matrix for varied applications/occupations (see Figure 2.9).

Among the various indicators applied in order to detect sleepiness, physiological signals appear to be the most reliable and accurate. Despite a plethora of studies validating this notion, some technical challenges seem to impede the realisation of the physiological-based drowsiness detection system within the marine industry. Every physiological-based drowsiness detection system demands placement of electrodes on the subject's body and connected via wire to a bulky processing hardware system [218]. This is unsuitable and limits the movement of the subject from accomplishing tasks and job demand. Nonetheless, the advent of cutting-edge technology has made feasible the deployment of a less intrusive manner to retrieve electrophysiological data. For instance, wireless system can be used to transmit physiological signals to the processing unit, such as that applied in Zigbee or Bluetooth. Most of the existing electroencephalography-based SDSs are based on multiple EEG electrode (>16 electrodes) settings. Apart from time consumption for preparation (depending on electrodes density), the subjects are bound to feel discomfort with repetitive recordings over extended duration. Despite the effort to minimise the quantity of EEG electrodes (< 3 electrodes) in classifying varied levels of sleepiness (in the SDS domain for classifying drowsy driving), a question rises if this minimisation could maintain the performance [256, 264]. Additionally, one cannot avoid for post-maintenance in the field, such as smearing of gel and re-attachment of electrodes. Furthermore, requirement for a subject to sit for a few minutes to obtain useful electrophysiological data may eventually disrupt the operational work-flow.

Great interest has been noted in replacing the acquisition of physiologically-based indicators with image acquisition methods, such as PERCLOS. While this method seems to perform well for driving scenarios, they are not suitable for subjects engaged in constant movement. This technique is not feasible as a subject has to stay reasonably still during the image-capturing process. Most of such systems face difficulty in functioning properly, especially when the head rotation exceeds 30° in azimuth [265]. Although the image acquisition technology is viable in small areas, such as cockpits (e.g., aeroplane, lorry, small vehicle), it is cost-ineffective to install a huge number of sleepiness monitoring acquisition devices in a large workspace, such as mines, factories, and large cockpits (e.g., ship). More importantly, most subjects are concerned about their privacy when having an image acquisition technology focused on their bodies at all times [266, 267].

One drawback of the vehicle-based indicators is that their design is often catered for land transportation, thus the need to install additional devices to the vehicle. As this system demands the installation of lane-tracking cameras and intricate video signal processing software, they are unsuitable for the marine industry. Moreover, it is impossible to control the navigation course in precise, particularly with random waves excited by the environment, ships, and floating structures.

The KSS is easy to administer, fast to complete (10-15 s), and allows repetitive administration [170], which is unlike lengthier objective measures of sleepiness, such that seen in physiological-based approach and PVT that can consume up to 10-minute duration. Furthermore, the KSS can be easily programmed in a portable device [268]. Despite its practicality in the field, like other self-reported sleepiness, KSS suffer from methodological constrictions. For instance, the correctness of the subjective assessment relies heavily on the quality, as well as accurate interpretation and understanding of each item [228]. Due to age and social diversity among subjects, it might be impossible to design a questionnaire to accommodate each potential problem. The ability of the subjects to accurately interpret and judge their sleepiness perception is crucial in maintaining data quality. To date, the KSS is widely used to assess subjective sleepiness in field studies by monitoring the level of sleepiness amongst the subjects in a range of sleep studies protocol. This is because; a researcher usually develops a strict rating protocol and the participants are trained to use the scale.

2.3 Kernel Density Estimation Supplement With Likelihood Ratio Test

2.3.1 Kernel Density Estimation

The Kernel Density Estimate (KDE) refers to a non-parametric method that estimates the probability density of any random variable. An identified density function (the kernel) is overlaid across the training data to generate a smooth density approximation.

Let the probability density of a random variable to be estimated in region R^D. R^D refers to a D-dimensional hypercube centred at point x. The total number of sample points within a hypercube is exemplified in Equation 2.3, as follows:

$$f_s(x_\mathbf{0}) = \sum_{i=\mathbf{1}}^{N} k\left(\frac{x_i - x_\mathbf{0}}{h}\right) \tag{2.3}$$

where function $k(u)$ is defined as the kernel function that controls the weight assigned to observation x_i at every point x_o, while h is the edge length of the hypercube. Here, s denotes different sleepiness state, and x denote the vSRS.

Despite the vast range of kernel types, the Gaussian has been the most commonly used due to its symmetric probability distribution [269]. As such, the Gaussian kernel was selected in this study for window smoothening.

Let V_d be the volume of hypercube in d dimensions. The probability density estimate at each point x_o reflects the outcome of the overall contribution for each sample x_i as given in Equation 2.4,

$$f_s(x_0) = \frac{1}{N}\sum_{i=1}^{N} \frac{1}{V_\mathrm{d}} k\left(\frac{x_i - x_0}{h}\right) \tag{2.4}$$

The Gaussian kernel was applied by overlaying a Gaussian along each data sample and then, summing up the overall density along the dataset. The term $\frac{1}{N}$ serves as a density normalising factor. Equation 2.4 can be rewritten as Equation 2.5,

$$f_s(x_0) = \frac{1}{N}\sum_{i=1}^{N} \frac{1}{(2\pi h^2)^{1/2}} \exp\left\{-\frac{\|x_i - x_0\|^2}{2h^2}\right\} \tag{2.5}$$

The resulting density estimate depends less on the kernel selection than its bandwidth value h [269]. The selection of h is integral because the value of h that was extremely large or small can hinder small-scale variation from being observed (over smoothing) or result in spiky estimate, respectively. It is common to estimate the optimum value of h by using the Silverman's rule, as given below in Equation 2.6,

$$h = \left(\frac{4\sigma^5}{3n}\right)^{1/5} \tag{2.6}$$

Where standard deviation and total number of samples are denoted by σ and n, respectively. The median absolute deviation was applied in this study to estimate the sample standard deviation.

2.3.2 Likelihood Ratio Test

Likelihood ratio (LR) test refers to a hypothesis test that compares the assumptions of varied models [53].

Let $X = [X_1, X_2, X_3, \cdots, X_K]$ denote the numerical value of K different matcher, where X_k is the numerical value for the K^{th} random variable, $k = 1, 2, \cdots, K$. Let c_1, c_2,$\cdots$, c_m be the varying sleepiness levels. Let $\hat{f}(x)$ reflect the likelihood score, where $x = [x_1, x_2, x_3, \cdots, x_K]$. Here $\hat{f}_k(x_k)$ denote the density estimate obtained from the KDE technique. The likelihood score for each level can be determined through the application of the following:

$$\hat{f}(x) = \prod_{k=1}^{K} \hat{f}_k(x_k) \tag{2.7}$$

According to the logarithmic properties, Equation 2.7 can be represented as

$$\log \hat{f}(x) = \sum_{k=1}^{K} \log \hat{f}_k(x_k) = S_{c_m} \tag{2.8}$$

where S_{c_m} is the likelihood score for x belong to c_m. Then, the match score is assigned to a state where its likelihood score is the highest.

2.4 Treatment Of Missing Data

Based on numerous classification/predictive applications, it is typical to experience issues related to missing useful attribute values (features) [270, 271]. In precise, this applies for sensor-based modalities that may experience faulty and communication error, thus leading to corrupted or missing data from certain modalities [272, 273]. For instance, the electrophysiological-based techniques suffer from unavoidable pervasive motion artefacts [256], noise due to electromagnetic field interference [218], and poor electrode attachment, which all may result in unreliable classification. The image or video-based techniques are affected by the illumination of the surrounding area [274]. The accuracy of eyelid closure measurement, for example, is affected by glare reflection from a subject's glasses or when their face is intermittently outside of the detection angle of the recorder during the image-capturing process [222]. Cases of missing subjective self-report measures data due to non-compliance by subjects have been reported as well [185, 275]. Upon temporary unavailability of certain indicators, a multi-modal system fusion can still offer classification decision by applying the information that implicitly contained in the remaining indicators [276]. The other principal potential advantage of fusion refers to its improved reliability.

The issue of missing data has to be appropriately addressed, since the performance of any predictive model relies heavily on the way missing data are treated. In the context of missing data, the data may be unavailable during induction time (i.e., classifiers training phase) or at prediction time [271]. There are three standard methods one can apply to address a missing feature in the test data. Although the simplest method for handling missing data is to discard instances with missing values [271], it may lead to loss of data, especially when the missing data are of a significant size [277]. Second is via imputation, whereby the missing value is substituted with a meaningful estimate from the available data [271]. Another approach is to allow the classifier to accept and address the incomplete observation during training phase [270, 278].

The issue of missing or corrupted data is common in practical SDS applications. The existing multi-modalities fusion-based SDS are mainly based on ANN and SVM, to name a few. Most studies have adopted a model with a single-level structure framework, in which the feature from multiple sleepiness indicators is fused simultaneously. Hence, their system (i.e. classifier) could only make prediction upon availability of all data. To the best

Table 2.2: Reported results of the prediction horizon in previous studies.

No	Study	Sleepiness Indicator	Prediction Horizon
1	[279]	Vehicle variables.	0.2 s, 0.4 s & 0.6 s
2	[280]	Vehicle variables.	0.2 s & 0.5 s
3	[281]	Steering wheel angle.	6-60 s
4	[252]	Physiological and behavioral.	20 s
5	[282]	Driving factor, eyelid measure, and driving performance.	0-10 min
6	[283]	ECG, eye activity, head movement, electrodermal activity, car and environment variable	0-10 min
7	[284]	Voice feature	3 h
8	[285]	Three process model of alertness	2 h, 6 h and 10 h

of the author's knowledge, no literature has explicitly addressed the issue related to missing or corrupted data within the domain of SDS.

2.5 Prediction Horizon

In light of SDS, the system is considered to operationally anticipate changes in the state (e.g., from alert to sleepiness state) within certain prediction horizon or time window prior to sleepiness state onset. In the stance of hindering accident, if the prediction horizon is sufficiently short, it appears too short to initiate a coordinated precaution in a practical manner [286]. On the contrary, a longer prediction horizon may offer adequate lead time for a stakeholder to implement the necessary intervention measure. Despite being a crucial criterion, the prediction horizon appears to be a topic that has been commonly omitted in the SDS literature. There were several studies that address this issue as summarise in Table 2.2.

In two studies, various vehicle variables have been applied as input for SVM [279] and NN [280] to predict unintentional lane departure. Based on these studies, the prediction horizon was varied between 0.2 and 0.6 s. The outcomes displayed that not only the system consistently had high and medium magnitudes of recall and precision, respectively, but the deterioration of both matrices may turn worse with the prediction horizon length. Nevertheless, such low prediction is somewhat short for a potential coordinated avoidance manoeuvre. For example, the distance is 13.3 m between a vehicle and a front object if the prediction horizon is fixed to 0.6 s and a car travelling at a speed of 80 km/h. McDonald and colleague employed steering wheel angle data as a feature for the Random Forest classifier [281]. In that study, the prediction horizon ranged from 6 to 60 s with 1 s increment as a strategy to assess up to what extent into the future the prediction was reliable. Their proposed system resulted in high false positive and false negative when the prediction horizon was set for 45-60 s and 10-30 s. Based on these two observations, the proposed system (i.e., using vehicle behaviour) performed well only for prediction horizon up to 10 s [281]. It is also interesting to ensure if physiological-based indicators will make a difference in terms of increasing the prediction horizon. Baykaner [284] initiated a method to detect sleepiness based on a linear regression model. The system successfully predicted the state of sleepiness using voice feature obtained 3 h prior to Psychophysiological Tests. Despite this claim, the correlation between the voice feature and the level of sleepiness was very low (r=.49 .58); making the claim somewhat questionable.

Another approach that extends the prediction horizon is the mathematical modelling. BMM is widely used to predict the timescale of minutes [175], hours [211], days [287], and months [215]. Schedulers applied BMM to predict at hourly range the likely level of sleepiness amongst those on a given duty schedule [7]. In a typical BMM validation, the goodness of fit (i.e., the variance between model predictions and actual neurobehavioural data) is usually applied to examine the forecasting ability of a model via measurement of root means square error (RMSE). Often in the literature, RMSE is viewed as a contribution of all data points, thus making it difficult to assess a model's predictive performance at each data point (i.e. prediction horizon). The lone exception is a study that assessed the capability of TPM in estimating PVT performance in advance of 2 h, 6 h, and 10 h [285]. As predicted, the RMSE increased monotonically with the prediction horizon length.

The recent focus is on predicting as accurate and early as possible by taking advantage of fusing multiple sensors. The drowsiness estimation system constructed by Murata was based on the correlation and regression methods to link the changes of physiological (EEG, ECG, EOG) with behavioural (neck bending angles, foot pressure, back pressure) information in predicting the level of sleepiness [252]. The system performed considerably well with 20 s of prediction horizon. In a similar vein, Liang [282] applied the regression method and extracted features from driving factor, eyelid measure, and driving performance measure as independent variables to predict the event of microsleep and lane crossing. The statistical outputs revealed the trend of decreasing sensitivity and specificity as the prediction horizon increased from zero to ten minutes for both scenarios.

Prior studies seemed to omit in ascertaining if the poor performance while extending the prediction horizon is due to poor selection of classifier or as a result of insufficient number of sleepiness indicators. Despite the inexplicit attempt to address the said issue, Larue demonstrated that ANN performed better, when compared to other ML, such as LR, BN, SVM, and decision tree [283]. In the study, several features, such as ECG, eye activity, head movement, electrodermal activity, car, and environment variables, were applied to predict alertness amidst drivers for prediction horizon set between zero and ten minutes with one-minute increment. The ANN-based system obtained above 85% and around 78% accuracy when predicting up to five and ten minutes in advance, respectively. They added that use of more heterogeneous features had improved the reliability of their system upon predicting a longer prediction horizon.

2.6 Multilevel Sleepiness Detection Systems

Sleepiness can be considered as a transient and cumulative process that evolves in a foreseeable manner. This has led to the assumption regarding the existence of the intermediate stages between the sleep and wake continuum. This opens the possibilities for one to track his arousal level in incremental steps, apart from allowing ample lead time for appropriate mitigation procedures [288]. Therefore, an SDS with higher sleepiness level resolution can reflects the specific arousal state of a shift worker and, subsequently notifying the SW's at its early stage when procedure might be the most effective [222]. Take for an example a system with ternary states of sleepiness (awake, drowsy, and sleep): let the system classify the immediate level of sleepiness of the operator to deliver appropriate warning or intervention reflective of the extent of sleepiness level. Upon detecting a low level of sleepiness (drowsy), the system may suggest the operator to take a short break or consume alertness-enhancing compounds. Meanwhile, if the system detects a high level of sleepiness that is incompatible with the safety threshold criteria of the task, strategies to reduce or to remove the operator from the operation can hinder the likelihood of an accident. To date, most SDS studies are intended for binary (e.g., alert or sleepy) classification problem with simplified experimental setting and parameter tuning. Relatively, only a handful has developed multi-level SDSs in the literature.

Several techniques have been proposed to track the progressive arousal changes at five levels. Khushaba and colleague employed a novel feature extractor method to retrieve relevant information from EEG, EOG, and ECG in distinguishing driver sleepiness to a maximum of five levels [233, 234]. Another study employed an extension of SVM to perform a maximum of five-level fatigue detection with features extracted from EEG [289]. In the study, the system was built using auditory vigilance task (AVT) as the ground truth. All three mentioned studies evaluated their proposed method by differentiating the drivers' sleepiness into two, three, four, and five levels [233, 234, 289]. The results exemplified a significant drop in classification accuracy with increment in detection resolution. One reason for such poor performance when classifying multilevel sleepiness problem is due to the selection of validation measurement of sleepiness. In the studies by [233, 234], the system was built using a driving task in a driving simulator; hence the subject might assume that navigation mistakes would not cause any harm, thus compromising the objective assessment [264]. While the author [289] argued the division of the ground truth into different level might be affected by the fluctuation in AVT performance. In other cases, some subjects did not display all levels of sleepiness during the course of data collection. In a study, only six out of thirty-one subjects exhibited all five drowsiness levels [234]. Another reason could be the class being ill-defined or overlapping as the resolution increased; sparking confusion amidst the neighbouring classes.

The difficulties faced from the five level problem has motivated several recent studies constrained their proposed driver SDSs to a maximum of three levels as summarised in Table 2.3. Picot et al., [256] grouped sleepiness into three levels using the Cascaded Decision Rule (CDR) based on electrophysiological signals, inclusive of EOG and EEG features. Meanwhile, in tackling the ternary sleepiness states detection, some features retrieved from statistical and frequency analyses of EEG, EOG, and TPM served as inputs to the SVM classifier [197]. The reports in [290, 291] classified sleepiness to three levels using Naïve Bayes (NB), which was fed with features gained from the driver's facial expression (FE). A study investigated the reliability of a new set of steering wheel activity (SWA) to classify three-level sleepiness states by applying the SVM model [292]. A number of ternary SDSs have proposed a method based on PERCLOS combined with eye movement or SWA using ANN [293] or Multilevel logit order (MLO) [294], respectively, as the classifier.

Table 2.3: Reported results of the ternary sleepiness detection method in previous studies.

No	Study	Parameter(s)	Classifier
1	[197]	EEG, EOG, CF	SVM
2	[292]	SWA	SVM
3	[294]	PERCLOS, SWA	MLO
4	[290]	FE	NB
5	[291]	FE	NB
6	[293]	PERCLOS, EM	ANN
7	[256]	EEG, EOG	CDR

2.7 Research Gap

Following the literature review, the existing sleepiness indicator may be classified into vehicle-based, operator-based, and a combination thereof. Although the proponents had claimed their proposed standalone indicator as strong predictors of performance, it has been reckoned that the correlation, upon used in isolation, is far from demonstrating a perfect correlation. Another issue refers to the prediction horizon length. Depending on the indicator, the prediction horizon can be varied from the scale of seconds, hours, and months. In fact, the performance was found to reduce monotonically with increasing prediction horizon. In order to overcome this issue, some developers have fused multiple approaches to complement the strength of an indicator over another, while minimising the drawback when they are useful in isolation to the extent possible. A number of studies related to SDS displayed that by fusing information from certain indicators enhanced the overall performance and minimised the variance for longer prediction horizon. Ideally, fusion-based SDS demands the availability of all features during induction and prediction phases. Unfortunately, this is always not the case. Based on the system architecture, not only such an event can decrease the performance, but to the extent where the overall system may meet failure. Despite this critical issue, the topic of missing data gained little attention in developing SDS. The existing approaches in handling missing/corrupted data are either by keeping the amount of missing data to a certain low percentage or by simply discarding the affected subject from the analysis. The final limitation noted in the literature review refers to the low sleepiness level resolution. Only a handful has devoted their efforts to multiclass SDSs. Most studies are meant for the binary classification problem, although they are incapable of warning the SW prior to accident.

2.8 Chapter Summary

Sleepiness seems to emerge as an integral contributor to maritime disaster and poor work performance. Hence, establishing a profile of factors linked with sleepiness is of utmost important. The initial part of the literature review offers not only the fundamental comprehension of the aetiology of sleepiness, but also the critical contextual information for constructing an effective SDS. In the interest of this study, selection of suitable indicators is the most critical stage. The literature review, hence, revealed the operator-based indicator, in precise the BMM, and subjective sleepiness assessment, which is the most suitable indicator for direct integration within the constraint of maritime operational environments. Subjective assessment, particularly KSS, not only easy to administer, but can be completed quickly and is able to indicate sleepiness at the time of assessment. Of the four classifiers applied in fusing the sleepiness indicators, BN is the most promising algorithm to address the intricate interaction between a range of contextual information and sleepiness indicators. More importantly, if the domain experts are incorporated into the process of specifying the graph structure, they are more intuitive to comprehend and modify the model structure in obtaining better predictive models, if required. The benefit of representing the dependencies between variables in a graphical manner enables BN to handle missing or lost data entries. The BN can be extended easily to handle multi-class issues and dismisses huge training sets to attain sufficient conditional probability estimate.

CHAPTER 3

METHODOLOGY

This chapter is composed of five main sections. Section 3.1 present an overview about the approach taken in this thesis in developing the SDS. Section 3.2 presents a detailed account of the derivation of the sleepiness indicator that will be used as input into the machine learning. Next, Section 3.3 presents the framework of the proposed SDS. Then, Section 3.4 depicts the approach adopted to minimise the dependecy between the variable embedded in the proposed system as well as to increase the prediction horizon. Lastly, the section 3.5 explains in detail the criteria for validation metric, the types of validation, and the considerations taken in selecting evaluation metric.

3.1 Overview Of The Proposed Methodology

This chapter presents the complete methodology for the fulfilment of research objectives formulated in Chapter 1. Therefore, the design of a SDS, as shown in Figure 3.1, require carefully design methodology to achieve the objective. The Figure 3.1 shows the flowchart of the data-driven approach taken in this work. The approach is (1) to extract relevant data from each of the activities conducted during the laboratory experiment; then (2) to derive the data into some meaningful sleepiness indicator; and finally (3) to construct a machine learning that can monitor the sleepiness level based on these sleepiness indicators. The details about the data collection and data extraction are discussed in detail in Appendix B.

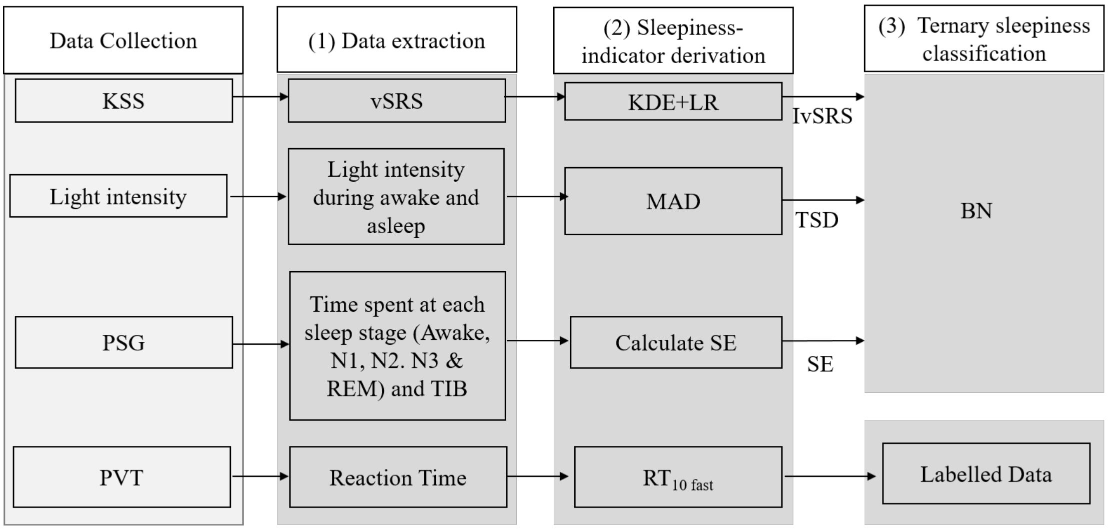


Figure 3.1: Flowchart of the experimental protocol.

We first discussed the approach for the fulfilment of Objective 1. In this study, the decision about the SW sleepiness state, solely by their own self-reported sleepiness value (vSRS), was improved by using the likelihood ratio (LR) test and kernel density estimation (KDE) technique. The new measure of subjective sleepiness estimation is denoted as improved self-reported sleepiness value (IvSRS). The interaction between the circadian and homeostasis process can be modelled by the model of arousal dynamics (MAD). The MAD simplified the interaction between between these two processes into a single variable, namely total sleep drive (TSD), only by using light intensity profile as an input. Sleep quality is a detrimental factor that affect the sleepiness level. While there exist many ways to quantify sleep quality, sleep efficiency (SE) is the most commonly being referred to in the literature of sleep studies since it covers the whole stage of sleep (i.e., cycling through NREM and REM). Hence, in this thesis, sleep quality is measured from the perspective of sleep efficiency.

Then, the three-information can be effectively organise using the Bayesian Network (BN) and subsequently utilising the Bayesian Theorem to obtain the final verdict on the sleepiness level. The fusion of multiple sleepiness indicators to improve the SDS performance thus served for the fulfilment of Objective 2.a. Cases of missing KSS data are possible. Selecting a candidate to impute the subjective sleepiness state is a critical issue. The direct and simplest way is to use the TSD information generated from the MAD. Taking advantage of the strong correlation between the TSD dynamics against the self-reported sleepiness that collected during constant routine and force desynchrony protocol, this thesis propose, whenever the subjective sleepiness is not available, the decision from TSD was referred; thus fulfilling the research objective 2.b. This information sharing mechanism open the possibility to extend the prediction horizon. This is because, it is possible to estimate the sleepiness state by only utilizing the TSD and SE, to estimate the sleepiness state at some time ahead of the future; thus fulfilling the research objective 2.c.

In this thesis, a supervised learning is employed whereby the model is given a set of labelled training data and learn to make the classification based on these training data. Then, the prediction from the learned model and known outcomes are compared, while the model's parameters are tune until the two outcomes align. The training data were derived from the $RT_{10\,\mathrm{fast}}$, which obtained from performance of psychomotor vigilance task (PVT).

3.2 Variables

This section explains the derivation of IvSRS, TSD and SE that serves as variables for the proposed BN.

3.2.1 Improvement Of Self-evaluated Assessment

One can consciously express their instantaneous sleepiness level by using subjective self-reporting measure [295]. In this study, the KSS was employed to obtain vSRS of individuals. Nevertheless, the application of subjective sleepiness assessment is susceptible to response biases, which derives from the inability of one to interpret his perception correctly or may define each scales differently [228]. Response biases are manifested as spikes or noise.

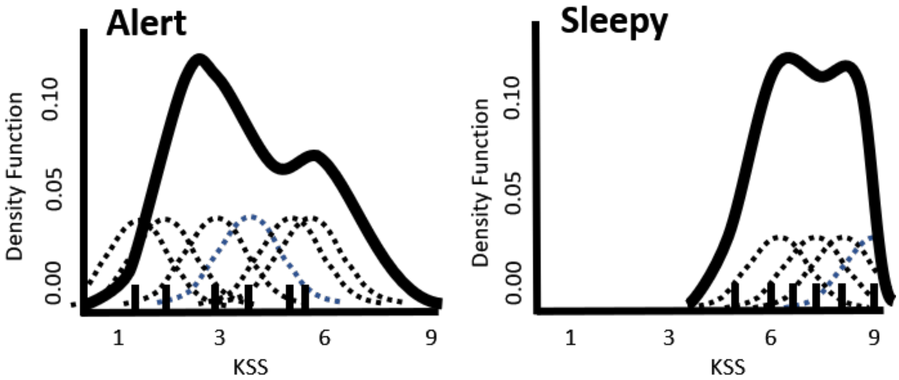
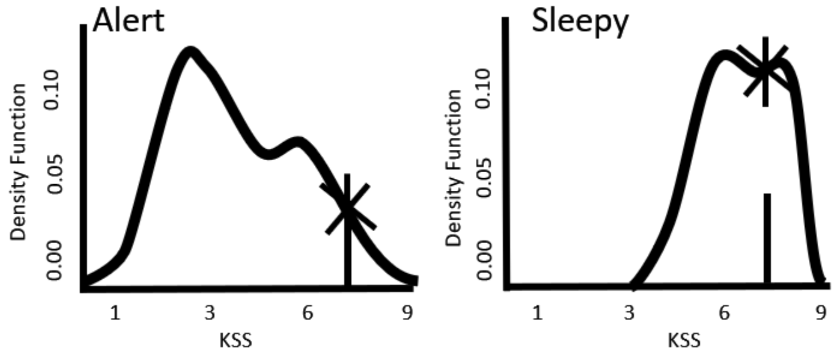


Figure 3.2: KDE smoothing



$$\hat{f}_{\text{Sleepy}}(7) \; \& \; \hat{f}_{\text{Alert}}(7)$$

Figure 3.3: Retrieving the density estimate given an instances of self-reported sleepiness value.

Since the distribution of vSRS scores was uneven, the thesis proposes the application of KDE and LR to improve the subjective sleepiness estimation (IvSRS). The process first transform the self-reported sleepiness value of a group of subjects into a density estimate using the kernel density estimator as shown in Figure 3.2. Indirectly, this process smoothen the noise (response bias) and generate the estimated densities of different classes. In the second module, the vSRS scores were assigned to varying sleepiness levels in accordance to the values retrieved from the LR test as shown in Figure 3.3.

3.2.2 Total Sleep Drive

As sleepiness is greatly affected by the duration of wakefulness and time of the day, a model that accounts for such interaction may be applied to accurately predict the level of sleepiness amongst shift workers. In this study, the MAD proposed by Postnova and colleague [215] had been employed.

As explained in Section 2.2.1.2. the MAD can simplified the interaction between the circadian and homeostatic process into a single variable namely total sleep drive (TSD).

3.2.3 Sleep Efficiency

In this work, SE was applied as an indicator to measure sleep quality. SE is the percentage of total sleep time to the actual time in bed (TIB) [125]. The total sleep time is the addition of the total time spent in Stages N1, N2, and N3, as well as REM sleep, while TIB refers to the period that starts with light-out and ends with light-off. A subject is deemed to have poor SQ when SE $\leq$ 85% [126, 127]. Long sleep onset latency, wake after sleep onset, and early sleep offset may lead to low SE [99], which suggests that one has spent a large proportion of sleep opportunity awake [124].

3.3 Detection Of Marine Shift Worker Sleepiness Based On Bayesian Network

The information retrieved from IvSRS, TSD, and SE can be used to infer one's level of sleepiness. Nonetheless, each listed component has uncertainty and may produce ambiguity if applied as a standalone decision-making tool. For instance, a subjective assessment may be affected by self uncertainty and intentional manipulation [170]. The parameters employed in this study to generate TSD estimate had been based on group-average values.

As sleepiness is an intricate concept, the correlations between sleepiness and sleepiness-related variables may be multi-directional. From the stance of multiple information fusion, there are n hypotheses for n information about subject sleepiness level [48]. Integrating such information decreases the ambiguity of conflicting (if present) hypotheses to a minimum. In this study, the fusion system was represented by BN. The BN refers to the combination of Bayesian theory and graph theory, which can effectively organise the link between dependent and independent variables [250].

Four random variables, total sleep deprivation, sleep efficiency, sleep disorder, and sleepiness level, had been considered and denoted as TSD, SE, SD, and SL, respectively. The joint probability over variables $P(SD, TSD, SE, SL)$ can be factorised as a product of conditional probability, as expressed in Equation 3.1,

$$P(SD, TSD, SE, SL) = P(SD) \cdot P(TSD|SD) \cdot P(SE|SD, TSD) \cdot P(SL|SD, TSD, SE) \tag{3.1}$$

The above factorisation portrays the possible dependency between the variables, such that each variable relies on every other variable. Nevertheless, if the factorisation is allowed to be presented in the following form,

$$P(SD, TSD, SE, SL) = P(SD) \cdot P(TSD) \cdot P(SE|SD) \cdot P(SL|TSD, SE) \tag{3.2}$$

This factorisation describes the conditional independence between some variables. The conditional independence between SD and SL given SE and TSD is displayed in the following:

$$\begin{aligned} P(SL, SD|TSD, SE) &= \frac{P(SD, TSD, SE, SL)}{P(TSD, SE)} \\ &= \frac{P\ (SD)\ P(TSD)\ P(SE|SD)\ P(SL|TSD, SE)}{\int\int P\ (SD).\ P(TSD).\ P(SE|SD).\ P(SL|TSD, SE) dSDdSL} \\ &= \frac{P\ (SD)\ P(SE|SD)\ P(SL|TSD, SE)}{P(SE)} \\ &= P(SD|SE)P(SL|TSD, SE) \end{aligned} \tag{3.3}$$

A directed acyclic graph is used to portray this factorisation, as illustrated in Figure 3.4. A directed acyclic graph contains two elements; nodes and arcs. The nodes refer to random variables (e.g., SE and SD), while the directed arc

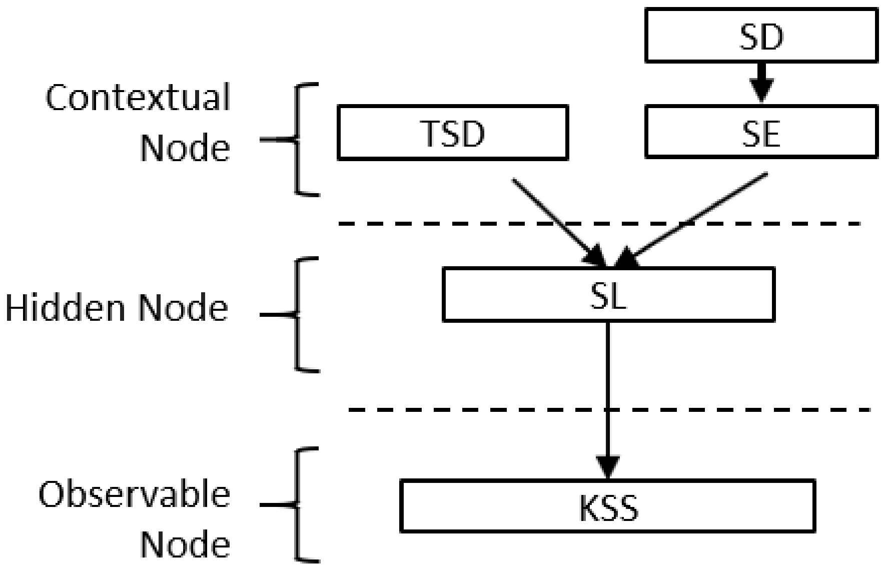


Figure 3.4: Three main groups of nodes: contextual nodes, hidden nodes, and observable nodes. The dotted lines separate the groups. The pointing arrows showed the dependency of SL subnode to TSD and SE toward sleepiness level, originating from the nodes of TSD and SE.

indicates the condition of dependency between the nodes. In the example given above, a directed arc that links SD node to SE node can be established, with the direction pointing from SD to SE nodes. The SD node is the parent node of the SE node, while the SE node is the child node of the SD node. A root node, meanwhile, has no parent node.

These nodes may be grouped into three main nodes, namely contextual, observable, and hidden nodes (see Figure 3.4). Contextual node is the knowledge of some contextual information that may affect sleepiness. In this case, the level of sleepiness is the unobserved event and therefore, represented by hidden node. Finally, the observable node is the changes in specific observations due to varied sleepiness levels. Subjective sleepiness (which can be rated using KSS) is an observable measure of sleepiness. Generally, each node is positioned based on an expert's view or results from multiple empirical studies.

Both parents and children of SL must be treated separately when evaluating the probabilities at the SL node. Figure 3.4 displays that TSD and SE are parents of SL, while subjective sleepiness is the children of SL. The probabilities at the SL nodes derive from all evidence, e, throughout the rest of the network $P(SL|e)$. However, the dependency from both the parents and children can be divided into the following:

$$P(SL|e) \propto P(e^{\mathcal{C}}|SL)P(SL|e^{\mathcal{P}}) \tag{3.4}$$

where e refers to all evidence, $e^{\mathcal{P}}$ is the parent node, and $e^{\mathcal{C}}$ denotes the children node.

Generally, the dependency is found on the evidence of the children nodes.

$$P\left(e^{\mathcal{C}}\middle|\, SL\right) = P(e_{\mathcal{C}_1}, e_{\mathcal{C}_2}, \cdots, e_{\mathcal{C}_{|\mathcal{C}|}}|SL) \tag{3.5}$$

By assuming that unconnected children nodes (i.e., no lines connected between them) are statistically independent of each other, the children nodes become conditionally independent to each other.

$$\begin{aligned} &= P\left(e_{\mathcal{C}_1}\middle|\, SL\right)P\left(e_{\mathcal{C}_2}\middle|\, SL\right), \cdots, P(e_{\mathcal{C}_{|\mathcal{C}|}}|SL) \\ &= \prod_{j=1}^{|\mathcal{C}|}\left(e_{\mathcal{C}_i}\middle|\, SL\right) \end{aligned} \tag{3.6}$$

Where $\mathcal{C}_i$ denotes the jth child node, $e_{\mathcal{C}_i}$ is the probability value of its state, and $\mathcal{C}$ represents the number of elements in the set.

Since the network is composed of a single child node, its evidence is in the following form:

$$P(e_{IvSRS}|SL) = P(e_{IvSRS}|SL) \tag{3.7}$$

Upon integrating evidence from the parent nodes for the purpose of completeness, it is assumed that the probabilities at SL depend on the evidence from the parents, $\mathcal{P}$.

$$P(SL|e^{\mathcal{P}}) = P(SL|e_{\mathcal{P}_1}, e_{\mathcal{P}_2}, \cdots, e_{\mathcal{P}_{|\mathcal{P}|}}) \tag{3.8}$$

Based on the law of total probability, the equation is expanded into the following:

$$= \sum_{all\ i,j,\cdots,k} P(SL|\mathcal{P}_{1i}, \mathcal{P}_{2j}, \cdots, \mathcal{P}_{|\mathcal{P}|k})\ P(\mathcal{P}_{1i}, \mathcal{P}_{2j}, \cdots, \mathcal{P}_{|\mathcal{P}|k}|e_{\mathcal{P}_1}, \cdots, e_{\mathcal{P}_{|\mathcal{P}|}}) \tag{3.9}$$

Assume that the unconnected parent nodes are statistically independent of each other:

$$= \sum_{all\ i,j,\cdots,k} P(SL|\mathcal{P}_{1i}, \mathcal{P}_{2j}, \cdots, \mathcal{P}_{|\mathcal{P}|k})\ P(\mathcal{P}_{1i}|e_{\mathcal{P}_1}), \cdots, P(\mathcal{P}_{|\mathcal{P}|k}|e_{\mathcal{P}_{|\mathcal{P}|}}) \tag{3.10}$$

Based on the local Markov properties (also known as local independencies, topological ordering) for BN, the node SL becomes conditionally independent of its non-descendants (e.g., all variables that affect SE) given its parents [296]. Hence, in order to compute the probabilities at node SL, the dependencies beyond parents and node SL can be neglected. The equation is simplified into the following:

$$P(SL|e^{\mathcal{P}}) = \sum_{all\ i,j,\cdots,k} P(SL|\mathcal{P}_{1i}, \mathcal{P}_{2j}, \cdots, \mathcal{P}_{|\mathcal{P}|k}) \prod_{i=1}^{|\mathcal{P}|} P(\mathcal{P}_i) \tag{3.11}$$

The conditional probability of SL, given the occurrence of the parent's node, is written as

$$P(SL|e^{\mathcal{P}}) = \sum_i \sum_k P(SL|TSD, SE)P(TSD)P(SE) \tag{3.12}$$

According to Bayes theorem, the conditional probability of the state at the hidden node, given the contextual information (prior probability) and observation (posterior probability), can be determined from Equations 3.7 and 3.12. Therefore, based on Bayes Theorem,

$$P\ (SL|e) = \frac{P(SL|e^{\mathcal{P}})P((e_{IvSRS}|SL)}{\sum P(SL|e^{\mathcal{P}})P\,(e_{IvSRS}|SL)} \tag{3.13}$$

For a set of evidence $x(n)$ derived from the n^{th} sampling time, the decision rule of BN is to assign $x(n)$ to the most probable sleepiness level, which is given as follows:

$$d(x(n)) = \underset{i}{\operatorname{argmax}}\{P(\omega_i|x(n)), i = 0, \cdots, m\} \tag{3.14}$$

where $P(\omega_i)$ is the i^{th} sleepiness level, and $P(\omega_i|x(n))$ denotes the posterior probability of assigning the n^{th} set of evidence to the i^{th} sleepiness level. From another stance, the numerical value of posterior probability may be interpreted as the confidence estimate of assigning a set of evidence to that sleepiness level. If a class claims that the test sample belongs to a certain sleepiness level with significantly higher posterior probability, then the confidence estimate is sufficiently high. On the contrary, if multiple classes claim that the test sample belongs to varied sleepiness levels with relatively high posterior probability, then the confidence estimate is low. Hence, in order to reduce the misclassification of assigning $x(n)$ to the wrong class, class with larger posterior probability was selected for this study.

3.4 Workaround For Missing KSS And Extending The Prediction Horizon

This study attempted to minimise the dependency between contextual and observational nodes. Hence, if the observation node becomes unavailable, the overall framework can still be functional. Cases of missing KSS data are indeed possible [185, 275]. This can be due to non compliance by the subject. In order to address this issue, the decision from available information can be referred to impute the subjective sleepiness state at the observation node. Selecting candidate to impute the subjective sleepiness state is a critical issue. If the candidate is meant to measure the subjective sleepiness state, the direct and simplest way is to use the BMM. Nonetheless, the suitability of BMM as a potential candidate highly relies on the model capability to assume the probable subjective sleepiness at the point where KSS may be unavailable. Thus, the suitable scientific validation of the BMM outputs against self-reported sleepiness measure must have a considerable face value.

A validation assessment was carried out to compare prediction made by MAD with KSS data gathered from eight diffrent laboratory studies involving constant routine (CR) and forced desynchrony (FD) protocols [211]. Validation using data derived from CR and FD protocols does have a face value, particularly when judging the suitability of the model for application in shift-work schedule. In the study, Root Mean Square Error (RSME) was employed as a comparative metric to identify if the parameter offered good fitting between model prediction and population-average performance observation. Lower RMSE values signify better goodness of fit between

Table 3.1: RMSE for the KSS model prediction when using default parameter value and customise model fitting to individual studies

			Forced Desynchrony			
	Constant Routine		Time Since waking		Circadian Phase	
Study	Default	Study Fit	Default	Study Fit	Default	Study Fit
[297]	0.14	0.08	—	—	—	—
[298]	0.09	0.08	—	—	—	—
[299]	0.21	0.12	—	—	—	—
[300]	—	—	1.4	0.24	0.14	0.09
[301]	—	—	0.42	0.17	0.53	0.23
[302]	—	—	0.97	0.29	1.77	0.15
[303]	—	—	0.48	0.15	0.62	0.12
[304]	—	—	1.10	0.29	0.13	0.10

predicted and observed performance (e.g., KSS data points). The evaluation outcomes exemplified a strong correlation between model dynamics and KSS experimental data, as verified by the low RMSE tabulated in Table 3.1. A visualisation comparison of the model dynamics with three experimental data is displayed in Figure 3.5, which demonstrates near identical match to most KSS data. These evaluation outputs confirm the ability of the MAD model in replicating the dynamic fluctuation in subjective sleepiness, and subsequently, approximate the subjective sleepiness at the point where KSS reading may not unavailable.

The TSD output served as dual-role information (contextual and observational information), as illustrated in Figure 3.6. In an ideal case, where information pertaining to subjective sleepiness, TSD, and SE are available, the calculation of each binary or ternary output using TSD, SE, and IvSRS can be conducted as depicted in Sections 3.3. Next, the three informative points can be integrated by using the Bayesian Theorem to gain the final verdict on the sleepiness level. Upon unavailability of KSS data, the status based on an observable feature can be referred from the TSD. This information-sharing aspect is shown as a double-headed ant arrow displayed in Figure 3.6.

The information imputation open the possibility to extend the prediction

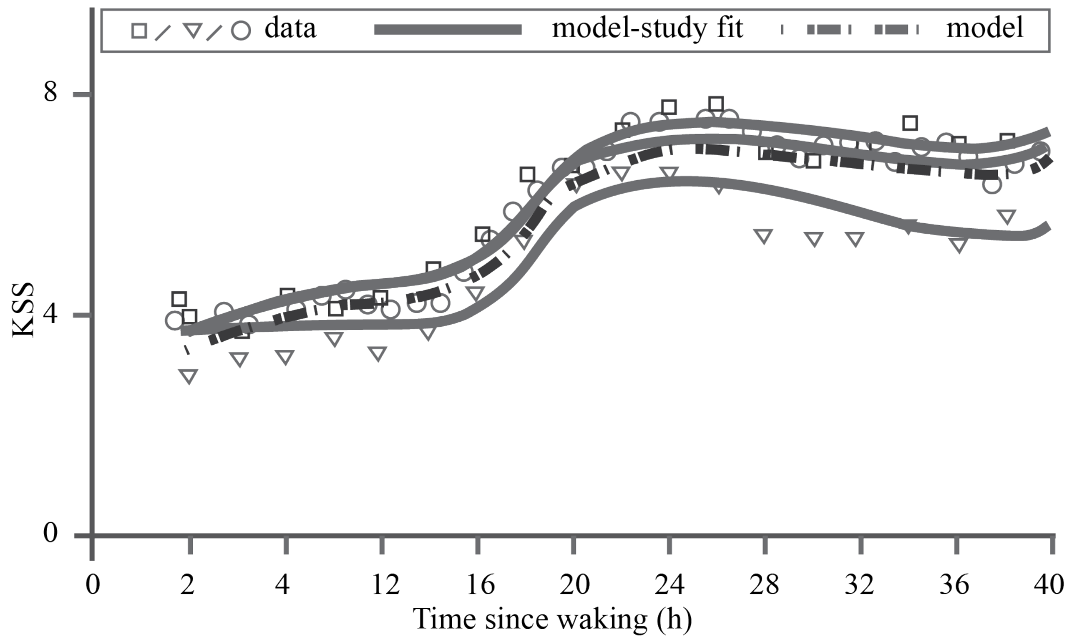


Figure 3.5: Comparison of model projections and information acquired for the KSS experimental data obtained during during the protocols of CR. The experimental data are from [298], [299], and [297] and each are represented by the open symbols □, ▽, and ○, respectively. Model projections with the default values are displayed with ant line. Customise model fitting to individual studies are shown with solid grey line. Figure reproduced from [211]

horizon. This is because, it is possible to estimate the sleepiness state at some time in the future by utilising the TSD and SE decisions under the proposed framework as illustrated in Figure 3.7.

3.5 Preparation For Comparative Study

3.5.1 Criteria Of Sleepiness Level Evaluation

The datasets used for training and testing purposes in this study were retrieved from three shift patterns (see Appendix B). The PVT was applied as the validation metric (ground truth), mainly because operating and maintaining a tool would demand continuous vigilance and immediate reaction from the SW [151], otherwise, accidents could take place. Based on some PVT outputs, the fastest 10% of response time ($RT_{10\,\text{fast}}$), which had been defined as the average RT of the fastest 10% of all responses, excluding false starts and lapses [160], was chosen as the indicator to assess the

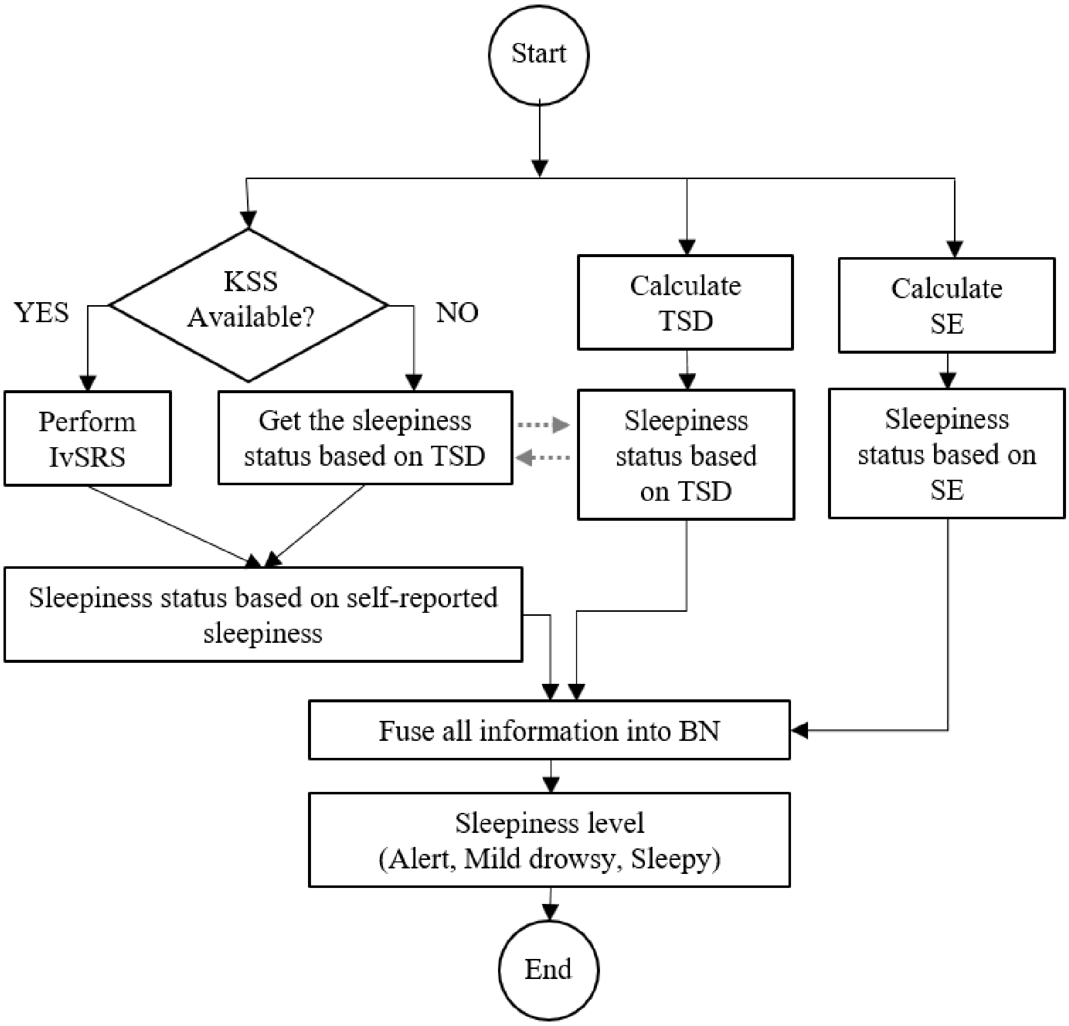


Figure 3.6: Workflow of the proposed method.

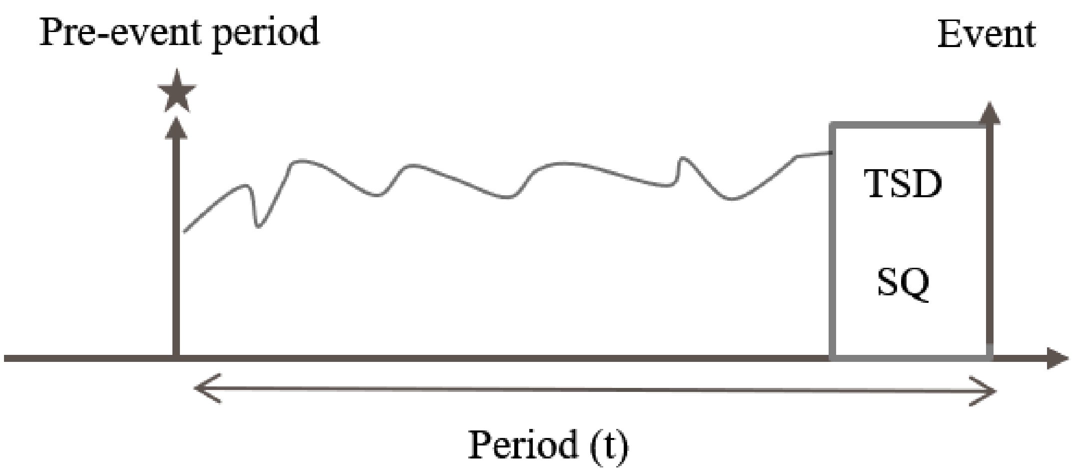


Figure 3.7: Extending the prediction horizon using only SE and TSD.

performance exerted by the subjects. False start and lapses are trials with RT below 100 ms and above 500 ms, respectively.

For each subject, the $\mathrm{RT}_{10\,\mathrm{fast}}$ performance throughout the nine-day experiment is given in the following form: $\{Z(n), Y(n)\}_{n=1}^{p} \in \mathbb{R} \times \{1, \cdots, m\}$, where $Z(n)$ denotes the $\mathrm{RT}_{10\,\mathrm{fast}}$ and $Y(n)$ reflects the sleepiness level determined via manual classification. Let n_{max} and n_{min} denote the highest and the lowest values of $Z(n)$. In case of a binary issue, max $Y(n)$=k, the separation threshold can be retrieved by using the $\frac{100}{k}$% splitting ratio of $\mathrm{RT}_{10\,\mathrm{fast}}$ performance [289, 305]. Let's assume that the fastest and slowest $\mathrm{RT}_{10\,\mathrm{fast}}$ scores obtained by a subject during the nine-day study was 10 ms and 110 ms, respectively. Hence, for a 50% splitting ratio (where k=2 for binary issue), the separation threshold is 60 ms. As such, the $\mathrm{RT}_{10\,\mathrm{fast}}$ score that are below and above 60 ms are attributed to alert (m=1) and sleepy (m=2) states, respectively. As for the ternary issue, the $\mathrm{RT}_{10\,\mathrm{fast}}$ performance (the highest to the lowest of $\mathrm{RT}_{10\,\mathrm{fast}}$ score) for a subject throughout the nine-day experiment is divided into three segments of alert, mild drowsy, and sleepy states, with the following distribution of data points; 736, 271, and 88, respectively.

3.5.2 Subject-wise Cross Validation for Performance Evaluation

The objective of the proposed SDS is addressed by the empirical outcomes. In order to assess the generalisation performance of the proposed SDS, the "leave-one-proband-out" [306] scheme was implemented. For each N subject, the data can be represented in the following form: $\{x(n), y(n)\}_{n=1}^{z} \in \mathbb{R}^{G} \times \{1, \cdots, m\}$, where $x(n)$ refers to the G feature vector (e.g., KSS, TSD) derived from the n^{th} sampling point (i.e., sampled at times during a cognitive test), while $y(n)$ denotes the corresponding sleepiness level, and m is determined based on validation metric. In this particular scheme, for the total data subsets, $\{\mathbb{D}_i\}_{i=1}^{N}$ was gathered from N subjects, wherein the data from $N-1$ subjects were concatenated into training set $\mathbb{D}_{tra}$, while the samples from the left-out subjects were employed as testing set $\mathbb{D}_{tes}$. As this procedure was repeated N times, it resulted in N pairs of $\mathbb{D}_{tra}$ and $\mathbb{D}_{tes}$. In each iteration for the $\mathbb{D}_{tra}$ and $\mathbb{D}_{tes}$ pair, the $\mathbb{D}_{tra}$ is held out for tuning the ML algorithm. Next, assessment can be performed by feeding $\mathbb{D}_{tes}$ into the trained classifier and comparing the classification outcomes with the corresponding label given by the $\mathbb{D}_{tes}$. Evaluation that was carried out on other unseen subject minimised the impacts of individual variances, thus ascertaining strict evaluation of the subject-independent performance using the proposed SDS [175, 289].

3.5.3 Evaluation Metric

3.5.3.1 F_1-measure

In this study, the representation of varied arousal levels was unequal. Within the ML domain, such an event is called 'class imbalance problem' [307]. It is vital to select the most suitable evaluation metric to assess the performance of a classifier on imbalanced dataset [307]. The correlation between recall and precision is a suitable candidate to address such an issue. Both precision and recall can be determined from Equation 3.15 and 3.16, respectively [307],

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \tag{3.15}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \tag{3.16}$$

True positive (TP) is data points that reflect the sleepy state correctly, thus classified as sleepy. False positive (FP) denotes the data points that refer to the alert state being classified as sleepy state. True negative (TN) is the data points that reflect the alert state correctly classified as an alert. False negative (FN) is the data inputs that refer to sleepy being classified as alert state.

The trades off between precision and recall can be minimised to a single scalar value, namely F_β-measure, which is computed using Equation 3.17. The F_β-measure is a commonly applied evaluation metric in the development of SDS meant to quantify the performance of the proposed system [308–310]. Higher F_β-measure value signifies better performance of a model than that of the reference model. For comparison purpose, β=1 (hence the name F_1-measure), which refers to the harmonic mean between precision and recall, had been applied in this thesis to assess the performance of the proposed system.

$$F_\beta = (1+\beta) . \frac{\text{Precision+Recall}}{(\beta \text{ . Precision}) + \text{Recall}} \tag{3.17}$$

3.5.3.2 *Box-Whisker Plot*

Box-whiskers plot and average F_1-measure of the 24 selected subjects are reported in this study. Each box plot displays the median, the first and third quartile, as well as the minimum and maximum values. The notches around each median give a rough idea on the significantly varied medians: if the notches do not overlap, the medians differ at 5% significance level.

3.5.3.3 *Statistical Analyses*

The paired t-test was applied to ascertain the significant variance between the average F_1-measure of the proposed system and other techniques [311]. The presence of outlier was determined by inspecting the boxplot for a value exceeding 1.5 box-length from the edge of the box in a boxplot. For each classifier performance, the normal distribution of the F_1-measure score was examined using Shapiro-Wilk's test. As for F_1-measure performance that is not normally distributed, two t-tests were conducted by including and excluding the value of the outlier. The statistical significance was fixed at $p \leq 0.05$.

3.6 Chapter Summary

This chapter is composed up of five primary sections. The first section discussed the approach taken in this thesis in developing the SDS. The second section shows that instead of using direct KSS, this study proposes to improve the subjective sleepiness estimation by using the combination of KDE and LR technique. Under the same section, the derivation of total sleep drive and sleep efficiency are described. In the third section, the formulation and construction of SDS based on BN are described in detail. As for the fourth section, the issue of missing KSS data is addressed by employing TSD, which could be applied to approximate the state at the observation node.

The final section presents the application of the retrieved empirical data to train the proposed method. Next, the trained system was evaluated on unseen testing sets. Both training and testing had been based on the leave-one-proband-out scheme, while the concordance with the testing sets was quantified via F_1-measure.

CHAPTER 4

RESULTS AND DISCUSSION

This chapter presents the empirical outcomes and discusses in detail their significance to the field of study, as well as to the research community.

4.1 Improved-vSRS

4.1.1 Improved-vSRS Against Simple Thresholding Technique Results

The value of KSS $\geq$ 8 has been linked with higher risk of car accidents [169, 188], long duration of eye closure, as well as increased brain activity in both theta and alpha bands [171]. High levels of theta energy signify the onset of sleep [190–192], while increment in alpha activities reflect the increased mental effort in maintaining vigilance [193]. This particular observation has motivated the researcher to group the driver performance into two or three classification based on their subjective sleepiness (i.e., obtained from KSS). In precise, the driving performances noted during the driving period with $\leq$ 6 and KSS $\geq$ 8 were designated as alert and sleepy, respectively. As for the ternary problem, the driving performances recorded for driving durations with $\leq$ 5, 6 $\leq$ KSS $\leq$ 7, and KSS $\geq$ 8 were designated as alert, mild drowsy, and sleepy, respectively [175, 176, 194–197]. The performances were classified into two arousal levels, namely simple threshold technique binary (STT_b) and simple-threshold-technique ternary (STT_t).

The performance of IvSRS against STT_b and STT_t had been compared. The classification outcomes for the two techniques in two varied scenarios are

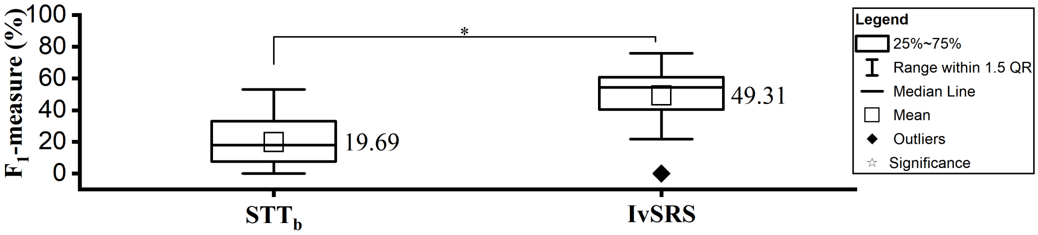


Figure 4.1: Box-Whiskers plots comparing, as evaluated by F_1-measure, classification of binary problem with STT_b and IvSRS.

illustrated as box plot (see Figure 4.1 and 4.2). As for the binary problem, the average F_1-measure for IvSRS and STT_b were 49.31% and 19.69%, respectively, for the average of the 24 selected subjects. Meanwhile, for the ternary problem, the average F_1-measure for IvSRS and STT_t were 46.52% and 41.41%, respectively, for the average of the 24 subjects. The IvSRS increased the F_1-measure by averages of +29.62% and +5.11%, when compared to STT_b and STT_t, respectively. While the improvement in the binary problem appeared significant (p <0.05), otherwise was noted for the ternary problem, based on the t-test.

4.1.2 Discussion Of The Results

So far, the study outcomes demonstrate that the prediction using subjective sleepiness assessment may be enhanced by transforming it into density estimate and examined using the LR test. Clearly, response biases were present, which resulted in the inability of the subjects to interpret their own perception or ended up defining each scale differently [228]. Such response biases are denoted as spikes or noise. In this study, the KDE approach was employed to smoothen the vSRS scores. Despite the easy computation via KSS thresholding, the IvSRS method gave better performance in distinguishing the varying arousal states. This notion verifies the hypothesis formulated in this study that post-processing the subjective assessment does enhance the application of KSS in separating the varied levels of sleepiness.

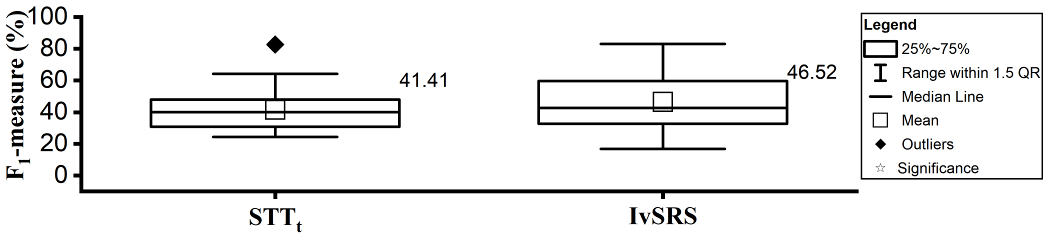


Figure 4.2: Box-Whiskers plots comparing, as evaluated by F_1-measure, classification of ternary problem with STT_t and IvSRS.

4.2 Bayesian Network

4.2.1 Bayesian Network Against Standalone Information Results

The benefits of combining multiple types of sleepiness information in separating the ternary sleepiness class problems had been determined. The classification results for individual sleepiness indicator and after the fusion are illustrated as boxplot in Figure 4.3. The average F_1-measure of sole IvSRS, sole TSD, and sole SQ were 46.52%, 62.76%, 66.83%, respectively. The numerical values signified that amongst the standalone information, the use of SQ gave better performance in terms of F_1-measure, in comparison to IvSRS and TSD. Nevertheless, the variance between the performances was statistically insignificant (p >0.05). Meanwhile, combining all the sleepiness indicators resulted in an F_1-measure of 82.09%. As predicted, significant improvements (p <0.05) were noted in the F_1-measure, especially when compared to the standalone approaches of SQ, TSD, and IvSRS with average improvements of +15.26%, +19.33%, and +35.57%, respectively.

4.2.2 Discussion Of The Results

The empirical outcomes revealed that the proposed system significantly enhanced the classification performance, when compared with the three standalone approaches (p <0.05). This confirms the hypothesis formulated in this study, which denotes that employing BN to fuse multiple sleepiness information can improve sleepiness detection. The BN was adopted in this study to integrate multiple sleepiness information, inclusive of the decisions

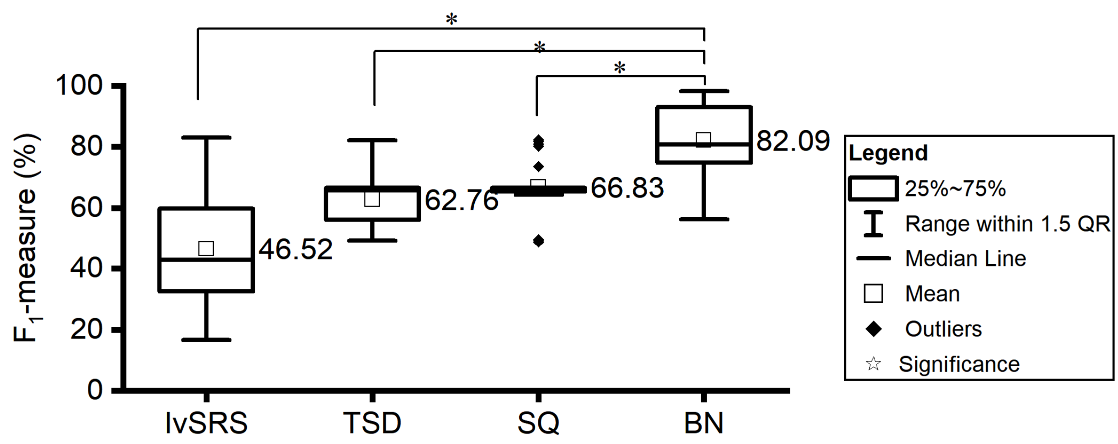


Figure 4.3: Box-whiskers plots comparing, as evaluated by F_1-measure, ternary classification with SQ, TSD, IvSRS, and BN.

obtained based on IvSRS, TSD, and SQ. One plausible reason for these improvements is that combining multiple sleepiness indicators can effectively reduce the prediction of uncertainties posed by each component. Hence, it is promising to embed additional contextual information, such as work environment [256, 312].

Other SDS studies have also employed the BN as the framework. The reported findings showed that improvement was attained by systematically fusing the multiple sleepiness indicators under the BN. Nonetheless, these studies were constricted to the binary sleepiness problem. As depicted in Chapter 2, such design poses limited practicality for the "real world" setting. Thus, this thesis, for the first time, successfully demonstrated the capability of BN in extending to ternary SDS.

4.3 Handling Missing KSS Data

The efficacy of the proposed SDS was examined under adverse setting (unavailable or missing observable information). The validation experimental process was repeated for 500 cycles, each with *M* percent (%) of missing KSS for each subject, where *M* ranged between 0 and 100% with a step size of 1%. In each cycle (every iteration of the 500 cycles), a new validation dataset was generated by randomly substituting *M*% of the KSS data from the complete dataset with Not-a-Number (NaN). Each NaN represented a missing KSS data point in the validation dataset.

The proposed SDS was compared with a similar SDS, but with single standalone input feature. With IvSRS being the standalone feature, the missing data points were assigned manually as either sleepy ($\text{IvSRS}_{\text{sleepy}}$), mild drowsy ($\text{IvSRS}_{\text{mild}}$) or alert ($\text{IvSRS}_{\text{alert}}$) in the dataset. The total average for the averages of 500-repetitive performances per subject for all the 24 subjects is reported in this study.

4.3.1 Ability Of The Proposed System In Handling Missing KSS Data Results

Figure 4.4 illustrates a decrease in performances for both $\text{IvSRS}_{\text{sleepy}}$ and $\text{IvSRS}_{\text{mild}}$ with increment in the number of missing or lost KSS data points. On the contrary, the performance of $\text{IvSRS}_{\text{alert}}$ enhanced when the adverse setting was set to its maximum. Figure 4.4 also displays that while the proposed system outperformed the other standalone methods, a steady drop was observed as a function of missing data. It is noteworthy to highlight that the network performance converged to that of TSD at 100% point. The performances of SQ and TSD exerted a linear horizontal line (see Figure 4.4). It is also interesting to observe that SQ outperformed the other standalone methods for all missing data cases. The classification ability displayed by SQ exceeded that of the proposed approach, especially when the percentage of missing KSS data exceeded 85%.

4.3.2 Discussion Of The Results

Figure 4.4 portrays that the performances of $\text{IvSRS}_{\text{sleepy}}$ and $\text{IvSRS}_{\text{mild}}$, as well as $\text{IvSRS}_{\text{alert}}$, decreased and increased, respectively, along with increasing number of missing KSS data points. The contradiction noted in the performances between $\text{IvSRS}_{\text{sleepy}}$, $\text{IvSRS}_{\text{mild}}$, and $\text{IvSRS}_{\text{alert}}$ is attributable to

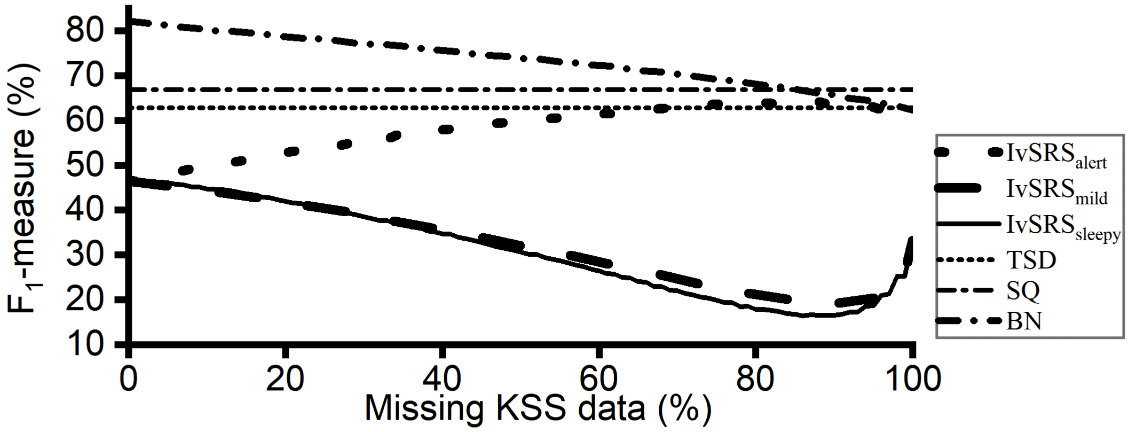


Figure 4.4: Comparing proposed technique BN with $\text{IvSRS}_{\text{sleepy}}$, $\text{IvSRS}_{\text{mild}}$, $\text{IvSRS}_{\text{alert}}$, TSD and SQ in handling missing subjective data. Zero value on the X-axis represents an ideal case (i.e. no missing KSS data).

dataset imbalance, whereby a higher number of alert data was obtained when compared to mild and sleepy data. Thus, the performance of $\text{IvSRS}_{\text{alert}}$, in which it assumes missing KSS data point as alert state, enhanced with increment in the percentage of missing KSS data points. The outputs displayed in Figure 4.4 suggest that the IvSRS method should not be allowed to grossly assume the state of subjective sleepiness assessment upon unavailable information.

Addressing missing data is an objective of this study. The merit of the proposed model is best displayed when more than 1% of the subjective sleepiness data are lost. With dependency between variables presented graphically, BN can still function despite missing data entries [313]. The level of sleepiness was deduced based on contextual (TSD and SQ) and observable (KSS) features (see Section 3.4). With nil KSS data, the ternary sleepiness state may still be inferred solely using contextual information. Despite its adverse impact on the average performance across the validation set, the system as a whole can still function with only TSD and SQ.

The performances of SQ and TSD were expected to exhibit a linear horizontal line, as both variables are unaffected by missing KSS. The outcomes revealed an interesting pattern when SQ outperformed all the other standalone approaches and even exceeded the proposed approach when the missing KSS data exceeded 85%. This signifies that the proposed system of F_1-measure should, after the 85% missing KSS data, cap around to that of SQ empirical value (F_1-measure of SQ = 66.83%). On the contrary, the performance of the proposed system exerted a monotonic downward trend and converged to that of TSD towards the 100% point of the X-axis. This behaviour is partly attributed to the system design, whereby the ternary sleepiness state was inferred solely from the TSD information with unavailable KSS data. This raises a question if SQ information should be weighed in upon addressing missing observational information? This is excluded from this study due to time constraint, but remains a piece of critical information that demands investigation.

4.4 Ability To Increase Prediction Horizon

In the previous experiment (Section 4.3), due to the nature that the dataset being generated randomly, we were unable to pinpoint at what prediction horizon the proposed system behaves poorly. To further investigate, new training and validation datasets were generated. The validation process had been repeated for six cycles for each sampling hour (2 h, 4 h, 6 h, 8 h, 10 h, and 12 h) for each subject. For instance, for the 2 h group, a new validation dataset was generated by substituting the KSS samples with NaN. This similar procedure was repeated for samples from the other remaining groups. Likewise, the IvSRS method, as a standalone indicator, existed as $\text{IvSRS}_{\text{sleepy}}$, $\text{IvSRS}_{\text{mild}}$ and $\text{IvSRS}_{\text{alert}}$. The total average of averages of all six-prediction horizon per subject for all the 24 subjects is reported in this study.

4.4.1 Results From The Assessment Of Prediction Horizon

Figure 4.5 shows that $\text{IvSRS}_{\text{sleepy}}$ and $\text{IvSRS}_{\text{mild}}$ had a small fluctuation in their performances with increment in prediction horizon. As expected, due to the independence to the event of missing KSS, SQ and TSD performances displayed a linear horizontal line, as illustrated in Figure 4.4. Figure 4.5 showcases that the proposed system could perform considerably well and even outperformed other approaches for all the studied prediction horizons. Massive improvement was noted in the classification due to the fusion of multiple sleepiness indicators at varying prediction horizons. Table 4.1 presents that the fusion increased the average F_1-measure, when compared to $IvSRS_{sleepy}$, $IvSRS_{mild}$, $IvSRS_{alert}$, TSD, and SQ for the following prediction horizons; 2 h, 4 h, 6 h, 8 h, 10 h, and 12 h. Amongst all the prediction horizons, the BN obtained an average F_1-measure of 78.49%.

4.4.2 Discussion Of The Results

Another objective of this study is to extend the prediction horizon. Taking the missing KSS event from another stance (prediction horizon), the benefit of using both TSD and SQ has become more apparent. It was indeed possible to extend the time (prediction horizon) by using only TSD and SQ decisions under the proposed framework. The findings exhibited that the proposed SDS achieved a prediction horizon of up to 12 h with F_1-measure greater than 78%. This is especially beneficial when the proposed SDS was implemented for those working in the 12 h shift. It is worthy to highlight that the prediction horizon of the proposed model was higher than those reported in [279, 280, 308, 314], which ranged in a minute scale. Nonetheless, as previously mentioned, the use of PRM and SQ in the proposed framework met some trades off in the classification performance.

Initially, it had been predicted that the capability of the proposed model at a longer prediction horizon was attributable to mostly the TSD feature, but the outcomes displayed otherwise. Two studies have fused the three-process-model of alertness (TPM), which is a variant of BMM, with extra features, such as driving behaviour [175] and eye movement [194]. Upon employing the TPM in isolation, [175] and [194] had managed to obtain as

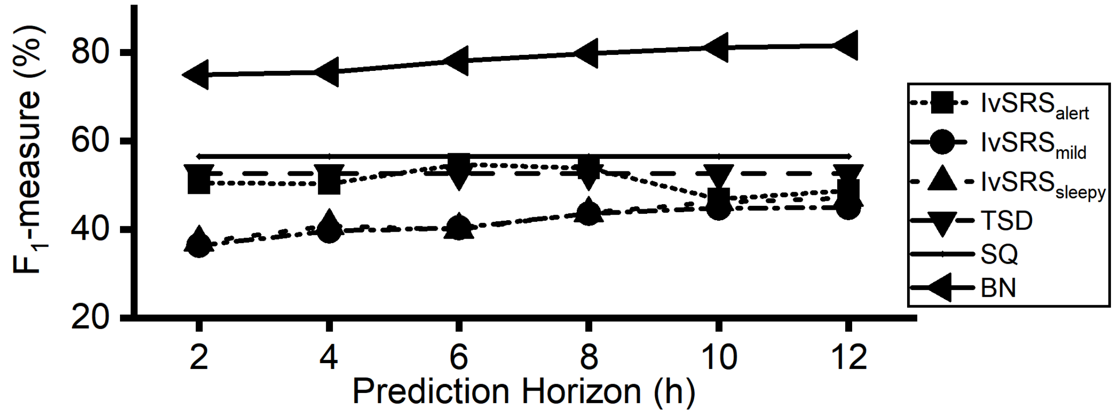


Figure 4.5: Comparing proposed technique BN with IvSRS$_{sleepy}$, IvSRS$_{mild}$, IvSRS$_{alert}$, TSD and SQ in term of different prediction horizon values.

Table 4.1: Table showed the F$_1$-measure improvement achieved by the proposed technique compared to IvSRS$_{sleepy}$, IvSRS$_{mild}$, IvSRS$_{alert}$, TSD, and SQ at prediction horizon of 2h, 4h, 6h, 8h, 10h and 12h.

	Prediction Horizon (h)					
Sleepiness Indicator	2	4	6	8	10	12
IvSRS$_{sleepy}$	+24.38	+25.08	+23.28	+25.80	+34.19	+32.74
IvSRS$_{mild}$	+38.53	+36.00	+37.69	+36.13	+36.30	+36.68
IvSRS$_{alert}$	+37.86	+34.74	+38.06	+36.09	+34.90	+34.34
TSD	+22.17	+22.80	+25.29	+26.97	+28.36	+28.79
SQ	+18.38	+19.01	+21.50	+23.18	+24.57	+25.00

much as 78% (accuracy) and 0.69 (r-value). It was reported that the TPM achieved the best result, in comparison to other standalone approaches. This study found that the MAD underperformed the SDS with standalone features, such as SQ. Besides applying a range of varied sleepiness BMM (TPM vs MAD), the methodological variances were likely to decipher the difference in the performance. Other studies validated their proposed SDSs against the ground truth obtained from driving performances. Therefore, comparison with these studies should be carried out mindfully [175, 194].

Table 4.2: Comparison of the proposed algorithm with other ternary studies.

Study	Task	benchmark method	Parameter(s)	Classifier	Acc (%)
[197]	Driving	KSS	EEG, EOG, CF	SVM	52.00
[292]	Driving	KSS	SWA	SVM	63.86
[294]	Driving	KSS	PERCLOS, SWA	MLO	64.15
[290]	Driving	FE	FE	NB	69.90
[291]	Driving	FE	FE	NB	74.50
[293]	Driving	KSS	PERCLOS, EM	ANN	77.36
[256]	Driving	EEG, EOG	EEG, EOG	CDR	80.60
Proposed method	Cognitive Task	PVT	KSS, TSD, SQ	BN	67.20

4.5 Qualitative Comparison With Other Studies

This study probed into enhancing the classification resolution. Capturing the intermediate states during the alert-to-sleep transition appears valuable to activate a warning signal at optimal time. Comparing the performance of the proposed ternary SDS with outcomes of other related studies unfolds interesting findings. Table 4.2 presents a summary of the empirical setup for each study. It was found that almost each study gathered data from the driving task. This is mainly because the SDSs were developed particularly to detect drowsiness among drivers, thus being more realistic to induce drowsiness from the driving activity itself. Varying indicators were employed as the benchmark to build SDSs, including KSS, FE, EEG, and EOG. Selection of validation metric, at least within the context of driver SDS, is attributed by flexibility, quick assessment, and convenient (e.g., KSS), while with intrusive nature by using camera (e.g., FE), and motivated by the need for using the most predictive and reliable sleepiness indicator (e.g., EEG). Additionally, a number of parameters have been proposed as input signals for the varied classifiers. Table 4.2 summarises the accuracies achieved by several related studies based on their evaluation metric.

At the time of this study, no similar public simulated shift work dataset has been applied to benchmark the efficacy of the proposed method. Table 4.2 presents a numerical comparison between state-of-the-art ternary sleepiness detection methods and the proposed method. Nevertheless, the comparison seemed somewhat difficult due to the variances in the experimental setup. The accuracy of the proposed method was calculated to perform a fair comparison with the existing ternary SDSs depicted in the literature. The comparison outcomes exerted that the average accuracy (67.20%) of the proposed method was indeed comparable with the method proposed by [197, 290, 292, 294]. However, the proposed framework underperformed, when compared to three reported approaches found in the literature [256, 291, 293]. These three proposed methods employed features retrieved from camera or electrode-based technique, which constricted their generalisation to other applications.

4.6 Chapter Summary

This chapter presents the results and related discussions based on five comparative studies. As for the initial comparative study, the experimental findings showed that the IvSRS, when compared to STT_b, was significantly better in assigning data point to the alert and sleepy states. Nevertheless, a similar advantage of post-processing the KSS was not reflected in the ternary problem, whereby the variance in performances appeared to be insignificant despite IvSRS having greater average F_1-measure, in comparison to STT_t.

The second part of the comparative study presents the outcomes after integrating multiple sleepiness information using BN, including its performance that was compared with three standalone indicators based on F_1-measure. As a result, it appeared that integrating multiple sleepiness information using BN had significantly enhanced the classification of SW ternary sleepiness level, in comparison to individual SQ, TSD, and IvSRS approaches.

As for the third part of the comparative study, outcomes derived from the proposed SDS under adverse conditions are revealed, along with a comparison of its performance with $IvSRS_{sleepy}$, $IvSRS_{mild}$, $IvSRS_{alert}$, SQ, and TSD. As a result, the proposed system outperformed the other standalone approaches. Nevertheless, its performance slumped steadily with increment in missing data points. This finding proves that the proposed system is robust in handling missing KSS data.

The fourth part of the comparative study presents the capability of the proposed SDS in extending the prediction horizon, along with a comparison of its performance with $IvSRS_{sleepy}$, $IvSRS_{mild}$, $IvSRS_{alert}$, SQ, and TSD. The empirical findings exemplified that the proposed system had the ability to predict the ternary state up to 12 h prediction horizon. Apart from that, the numerical outcomes signified that the proposed system displayed better performance, when compared with the three standalone approaches.

Finally, seven studies associated with the subject matter at hand, which had weighed in ternary SDS, had been compared with the proposed system. The outcomes exhibited that the proposed system performed considerably better than most of the reported studies. Nevertheless, it is emphasised here that the seven selected studies did not share similar experimental setup as that employed in this thesis, thus suggesting somewhat subjective benchmarking.

CHAPTER 5

CONCLUSION AND FUTURE DIRECTION

5.1 Overview

Sleepiness is a major safety concern within the maritime industry. Despite the active investigation made in light of SDS development, the gap to enhance its measurement performance is still present. Hence, in the attempt of bridging this knowledge gap, this thesis constructed and assessed a novel SDS in the context of mariners. This chapter is divided into three parts, namely (1) conclusion derived from the empirical outcomes, (2) highlights of the main contributions of this study to the vast research field, and (3) several suggestions and future work to further enhance the proposed SDS in this study.

5.2 Conclusion Of The Study

Sleepiness is a major safety concern in shift-related industry. To address this issue, a new framework based on BN has been proposed, which systematically fused multiple information to estimate the SW sleepiness level. By using the data obtained from an experiment that involved six-hour and eight-hour shift schedules, the proposed framework successfully demonstrated its feasibility and effectiveness in making objective inferences. The experimental findings exemplified that the IvSRS, in comparison to STT_b, was better in assigning data points into the binary problem. However, the same benefit of post-processing the KSS was not reflected for the ternary problem, whereby the difference in performance was not significant despite the IvSRS having greater average F_1-measure compared to STT_t.

The comparative study presents that integrating multiple sleepiness information using BN had significantly enhanced the classification of SW ternary sleepiness level, in comparison to individual SQ, TSD, and IvSRS approaches. The robustness of the proposed SDS was clearly demonstrated, especially in handling missing data, and the prediction horizon was extended to 12 h. In short, the proposed SDS is effective and robust, despite only contextual information was applied for the operation. Finally, the outcomes retrieved from this study were qualitatively compared with those reported by other researchers. Due to the simplicity of this proposed SDS, practical deployment of SW sleepiness detection can be easily proliferated without installing any intricate equipment or making major modification to the occupation setting.

5.3 Contributions Of The Study

The contributions of this study are detailed in the following:

1. Improvement of the self-evaluated sleepiness assessment

   This present study demonstrated for the first time, the benefits of post-process subjective sleepiness assessment. In particular, the KDE and LR framework was employed to enhance the subjective sleepiness prediction. The proposed method was implemented by first translating the subjective sleepiness level into a probability density function via KDE, which was later assessed using the LR test.

2. Information fusion using Bayesian Network for ternary SDS

   This thesis advances the application of efficient knowledge fusion, particularly when dealing with uncertain information in light of sleepiness detection implementation. The BN refers to a probabilistic graphical model that has been established since past decades now. This thesis has systematically portray how the existing BN framework may be implemented to address uncertainties posed by a range of modalities, apart from enhancing the performance of information fusion-based system. Although SDSs studies have embedded BN into their framework, those that had looked into ternary sleepiness issues are in scarcity. Therefore, this thesis has proven that BN can also be employed for ternary SDS.

3. Ability to function in the event of missing features and to increase prediction horizon

   This study has initiated the use of contextual information as the surrogate for subjective sleepiness input upon unavailability of subjective sleepiness input. This ascertains continued functionality of the system as a whole. By taking advantage of the system's ability to function in the event of missing subjective, prediction horizon may be increased up to 12-hour prediction horizon. The approximate inference made by Bayesian theorem solely relies on contextual information.

4. Impact to the industry

   The thesis respond to PETRONAS initiative to curb fatigue at work place. The advantage of the proposed method is that it can be apply to those in low risk and high risk worker in the platform. Another highlight of the proposed SDS was that the data collection (i.e., subjective sleepiness assessment) during the working period did not interrupt the ongoing activities and requires merely 10-15 s. Moreover, the sampling interval for KSS information could be conducted every 2 h. This, in turn, reduces the inconvenience to the SW. This allow the management to perform a periodic sleepiness assessment warn if one's sleepiness level is not safe for the them to resume operation.

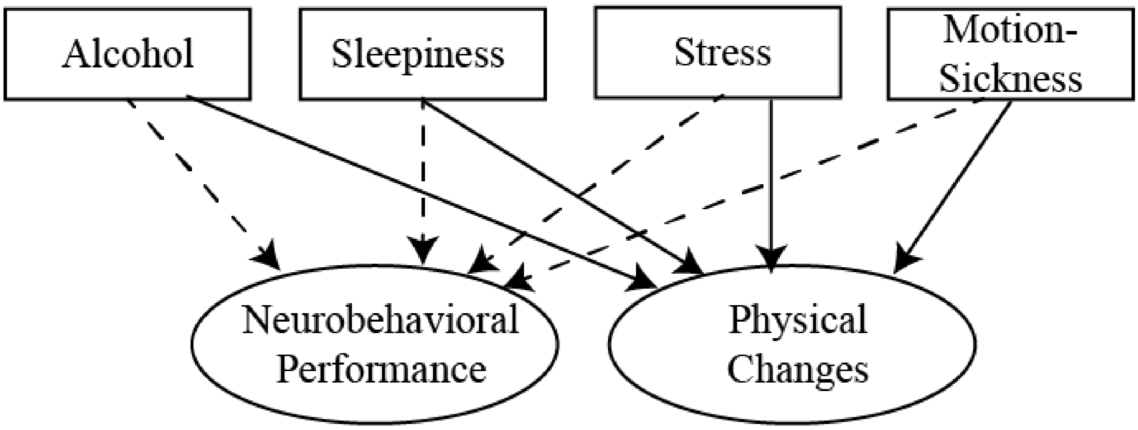


Figure 5.1: A BN based approach to detect general cognitive impairment

5.4 Future Work

This section presents some limitations of this study and a guideline to extend this work.

In this study, SQ was monitored with standard polysomnography (PSG), where the assessment was made from multiple physiological sensor readings, including EEG, EMG, and ECG. While such a setup is impractical to be used for a routine application [315], it is crucial at the current stage of our experiment to validate the applicability of SQ as one of the contextual information's main inputs. For this reason, we deliberately used data derived from PSG, which is known as the gold standard in assessing SQ [316]. The assessment of SQ can also be obtained using other methods, such as wrist actigraphy, which has been shown to have good correlation with PSG [316]. We left this for future study.

The proposed methodology of sleepiness detection can be extended to detect general cognitive impairment. For example, to detect deteriorated cognitive performance under the impact of alcohol or distraction. The related nodes in the BN framework can simply be updated through the understanding about the causal and effect under each impairment of the SW states. Figure 5.1 shows the extended version of the Figure 3.4 with other possible cognitive impairment stressors.

LIST OF PUBLICATIONS

1. R. P. Balandong, T. T. Bong, M. Short, and M. N. Saad, ”Maritime Shift Workers Sleepiness Detection System With Multi-Modality Cues,” *IEEE Access*, vol. 7, December. 2019. **(Impact Factor: 4.098)**

2. R. P. Balandong, R. F. Ahmad, M. N. Saad, and A. S. Malik, ”A Review on EEG-Based Automatic Sleepiness Detection Systems for Driver,” *IEEE Access*, vol. 6, May. 2017. **(Impact Factor: 3.557)**

[1] S. K. Lal and A. Craig, "Reproducibility of the spectral components of the electroencephalogram during driver fatigue," *International Journal of Psychophysiology*, vol. 55, no. 2, pp. 137–143, 2005.

[2] A. Vassalli and D.-J. Dijk, "Sleep function: current questions and new approaches," *European Journal of Neuroscience*, vol. 29, no. 9, pp. 1830–1841, 2009.

[3] A. Avidan and P. Zee, *Handbook of Sleep Medicine*, ser. Lippincott Williams & Wilkins Handbook Series, Y. A. Avidan and C. Z. Phyllis, Eds. Wolters Kluwer Health, 2011.

[4] G. Kerkhof and H. P. V. Dongen, *Human Sleep and Cognition: Basic Research*, 1st ed. Elsevier Science, 2010, vol. 185.

[5] A. Shahid, J. Shen, and C. M. Shapiro, "Measurements of sleepiness and fatigue," *Journal of Psychosomatic Research*, vol. 69, no. 1, pp. 81–89, 2010.

[6] D. Neu, P. Linkowski, and O. Le Bon, "Clinical complaints of daytime sleepiness and fatigue: How to distinguish and treat them, especially when they become 'excessive' or 'chronic'?" *Acta Neurologica Belgicapresents*, vol. 110, no. 1, pp. 15–25, 2010.

[7] D. Dawson, D. Darwent, and G. D. Roach, "How should a bio-mathematical model be used within a fatigue risk management system to determine whether or not a working time arrangement is safe?" *Accident Analysis & Prevention*, vol. 99, pp. 469–473, 2017.

[8] International Labour Organization, *Night Work Convention (No.171)*, C171 Std., 1990. [Online]. Available: https://www.ilo.org/dyn/normlex/en/f?p=NORMLEXPUB:12100:0::NO::P12100_INSTRUMENT_ID:312316

[9] R. Phillips, "Sleep, watchkeeping and accidents: a content analysis of incident at sea reports," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 3, no. 4, pp. 229–240, 2000.

[10] M. A. Short, A. Agostini, K. Lushington, and J. Dorrian, "A systematic review of the sleep, sleepiness, and performance implications of limited wake shift work schedules," *Scandinavian Journal of Work, Environment & Health*, vol. 41, no. 5, pp. 425–440, 2015.

[11] M. Akhatar, "Fatigue at sea - a manning problem," *Journal of Maritime Research*, vol. 11, no. 3, pp. 27–42, 2014.

[12] M. Lützhöft, M. R. Grech, and T. Porathe, "Information environment,

fatigue, and culture in the maritime domain," *Reviews of Human Factors and Ergonomics*, vol. 7, no. 1, pp. 280–322, 2011.

[13] N. Lamond, D. Darwent, and D. Dawson, "Train drivers' sleep and alertness during short relay operations," *Applied Ergonomics*, vol. 36, no. 3, pp. 313–318, 2005.

[14] N. Lamond and D. Darwent, "How well do train driver's sleep in relay vans?" *Industrial Health*, vol. 43, no. 1, pp. 98–104, 2005.

[15] W. M. V. Leeuwen, M. Barnett, and C. Pekcan, "Investigation of the 8-hours on/8-hours off seafarer watch keeping system," Maritime and Coastguard Agency, Report, 2016.

[16] W. M. V. Leeuwen, C. Pekcan, M. Barnett, M. Williams, and G. Kecklund, "Modelling watch keeper sleep and fatigue in the maritime industry," *Sleep Medicine*, vol. 40, p. e334, 2017.

[17] C. Hetherington, R. Flin, and K. Mearns, "Safety in shipping: The human element," *Journal of Safety Research*, vol. 37, no. 4, pp. 401–411, 2006.

[18] M. Celik, S. M. Lavasani, and J. Wang, "A risk-based modelling approach to enhance shipping accident investigation," *Safety Science*, vol. 48, no. 1, pp. 18–27, 2010.

[19] M. J. Akhtar and I. B. Utne, "Human fatigue's effect on the risk of maritime groundings – A Bayesian network modeling approach," *Safety Science*, vol. 62, pp. 427–440, 2014.

[20] O. Ulurlu, U. Yildirim, and E. Basar, "Analysis of grounding accidents caused by human error," *Journal of Marine Science and Technology*, vol. 23, no. 5, pp. 748–760, 2015.

[21] Marine Accident Investigation Branch, "Bridge watchkeeping safety study," Southampton: Marine Accident Investigation Branch, Report, 2004. [Online]. Available: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/377400/Bridge_watchkeeping_safety_study.pdf

[22] T. Eisenberg, M. T. Wells, and M. Heise, "Variability in punitive damages: Empirically assessing exxon shipping co. v. baker," *Journal of Institutional and Theoretical Economics*, vol. 166, no. 1, p. 5, 2018.

[23] M. Dorsett, "Exxon Valdez oil spill continued effects on the Alaskan economy," *Colonial Academic Alliance Undergraduate Research Journal*, vol. 1, pp. 1–17, 2010.

[24] M. Bloor, M. Thomas, and T. Lane, "Health risks in the global shipping industry: An overview," *Health, Risk & Society*, vol. 2, no. 3, pp. 329–340, 2000.

[25] K. S. Gould and V. F. Koefoed, "Facing the facts on fatigue on sea," *Haugesund: Norwegian Maritime Directorate*, 2007.

[26] Accident Investigation Board of Finland, "Factors contributing to fatigue and its frequency in bridge work," Report S3/2004M, 2008.

[27] A. Smith, T. Lane, M. Bloor, P. Allen, A. Burke, and N. Ellis, "Fatigue offshore: Phase 2 the short sea and coastal shipping industry:," Seafarers International Research Centre, Tech. Rep. 1-900174-21-9, 2003.

[28] Australian Transport Safety Bureau, "Marine safety investigations & reports," ATSB, Tech. Rep., 2019. [Online]. Available: https://www.atsb.gov.au/marine/

[29] Marine Accident Investigation Branch, 2018. [Online]. Available: https://www.gov.uk/government/organisations/marine-accident-investigation-branch

[30] P. Gander and M. U. S. R. Centre, *A Review of Fatigue Management in the Maritime Sector*. Sleep Wake Research Centre, Massey University, 2005.

[31] L. Torsvall, T. Åkerstedt, K. Gillander, and A. Knutsson, "Sleep on the night shift: 24-hour EEG monitoring of spontaneous sleep/wake behavior," *Psychophysiology*, vol. 26, no. 3, pp. 352–358, 1989.

[32] G. Kecklund and T. Åkerstedt, "Effects of timing of shifts on sleepiness and sleep duration," *Journal of Sleep Research*, vol. 4, pp. 47–50, 1995.

[33] P. Boudreau, S. Lafrance, and D. B. Boivin, "Alertness and psychomotor performance levels of marine pilots on an irregular work roster," *Chronobiology International*, vol. 35, no. 6, pp. 773–784, 2018.

[34] J. Horne and L. Reyner, "Vehicle accidents related to sleep: A review," *Occupational and Environmental Medicine*, vol. 56, no. 5, pp. 289–294, 1999.

[35] S. Bioulac, J.-A. Micoulaud-Franchi, M. Arnaud, P. Sagaspe, N. Moore, F. Salvo, and P. Philip, "Risk of motor vehicle accidents related to sleepiness at the wheel: A systematic review and meta-analysis," *Sleep*, vol. 41, no. 7, 2018.

[36] National Highway Traffic Safety Administration, "Drowsy driving," Tech. Rep., 2018. [Online]. Available: https://www.nhtsa.gov/risky-driving/drowsy-driving

[37] W. Vanlaar, H. Simpson, D. Mayhew, and R. Robertson, "Fatigued and drowsy driving: A survey of attitudes, opinions and behaviors," *Journal of Safety Research*, vol. 39, no. 3, pp. 303–309, 2008.

[38] K.-Q. Shen, "Signal processing methods for mental fatigue measurement and monitoring using EEG," Ph.D. dissertation, Department of Mechanical Engineering, National University of Singapore, Singapore, 2008, accessed on: Mar. 28, 2017. [Online]. Available: https://scholarbank.nus.edu.sg/handle/10635/28288.

[39] J. A. Caldwell, "Fatigue in aviation," *Travel Medicine and Infectious*

Disease, vol. 3, no. 2, pp. 85–96, 2005.

[40] J. Dorrian, G. D. Roach, A. Fletcher, and D. Dawson, "Simulated train driving: Fatigue, self-awareness and cognitive disengagement," *Applied Ergonomics*, vol. 38, no. 2, pp. 155–166, 2007.

[41] C. A. Czeisler, C. A. Pellegrini, and R. M. Sade, "Should sleep-deprived surgeons be prohibited from operating without patients' consent?" *The Annals of Thoracic Surgery*, vol. 95, no. 2, pp. 757–766, 2013.

[42] J. Sonati, M. D. Martino, R. Vilarta, É. Maciel, E. Moreira, F. Sanchez, G. D. Martino, and R. Sonati, "Quality of life, health, and sleep of air traffic controllers with different shift systems," *Aerospace Medicine and Human Performance*, vol. 86, no. 10, pp. 895–900, 2015.

[43] G. Legault, A. Clement, G. P. Kenny, S. Hardcastle, and N. Keller, "Cognitive consequences of sleep deprivation, shiftwork, and heat exposure for underground miners," *Applied Ergonomics*, vol. 58, pp. 144–150, 2017.

[44] D. Dawson and K. Reid, "Fatigue, alcohol and performance impairment," *Nature*, vol. 388, no. 6639, pp. 235–235, 1997.

[45] M. Lútzhóft, A. Dahlgren, A. Kircher, B. Thorslund, and M. Gillberg, "Fatigue at sea in Swedish shipping- A field study," *American Journal of Industrial Medicine*, vol. 53, no. 7, pp. 733–40, 2010.

[46] A. M. Williamson, "Moderate sleep deprivation produces impairments in cognitive and motor performance equivalent to legally prescribed levels of alcohol intoxication," *Occupational and Environmental Medicine*, vol. 57, no. 10, pp. 649–655, 2000.

[47] D. Lee and R. Pan, "Predictive maintenance of complex system with multi-level reliability structure," *International Journal of Production Research*, vol. 55, no. 16, pp. 4785–4801, 2017.

[48] Q. Ji, P. Lan, and C. Looney, "A probabilistic framework for modeling and real-time monitoring human fatigue," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 36, no. 5, pp. 862–875, 2006.

[49] M. Neil, N. Fenton, and L. Nielson, "Building large-scale Bayesian networks," *The Knowledge Engineering Review*, vol. 15, no. 3, pp. 257–284, 2000.

[50] A. C. Constantinou, N. Fenton, and M. Neil, "Integrating expert knowledge with data in Bayesian networks: Preserving data-driven expectations when the expert variables remain unobserved," *Expert Systems with Applications*, vol. 56, pp. 197–208, 2016.

[51] Q. Ji, Z. Zhu, and P. Lan, "Real-time nonintrusive monitoring and prediction of driver fatigue," *IEEE Transactions on Vehicular Technology*, vol. 53, no. 4, pp. 1052–1068, 2004.

[52] C. Ning and F. You, "Data-driven decision making under uncertainty integrating robust optimization with principal component analysis and kernel smoothing methods," *Computers & Chemical Engineering*, vol. 112, pp. 190–210, apr 2018.

[53] R. Zafar, S. C. Dass, and A. S. Malik, "Electroencephalogram-based decoding cognitive states using convolutional neural network and likelihood ratio based score fusion," *PLOS ONE*, vol. 12, no. 5, 2017.

[54] M. A. Short, S. Centofanti, C. Hilditch, S. Banks, K. Lushington, and J. Dorrian, "The effect of split sleep schedules (6h-on/6h-off) on neurobehavioural performance, sleep and sleepiness," *Applied Ergonomics*, vol. 54, pp. 72–82, 2016.

[55] A. Borbély, "A two process model of sleep regulation," *Human Neurobiology*, vol. 1, no. 3, pp. 195–204, 1982.

[56] R. Y. Moore, "Suprachiasmatic nucleus in sleep-wake regulation," *Sleep Medicine*, vol. 8, no. SUPPL. 3, pp. 27–33, 2007.

[57] T. Roenneberg, S. Daan, and M. Merrow, "The art of entrainment," *Journal of Biological Rhythms*, vol. 18, no. 3, pp. 183–194, 2003.

[58] D. A. Golombek and R. E. Rosenstein, "Physiology of circadian entrainment," *Physiological Reviews*, vol. 90, no. 3, pp. 1063–1102, 2010.

[59] V. K. Sharma and M. K. Chandrashekaran, "Zeitgebers (time cues) for biological clocks," *Current Science*, vol. 89, no. 7, pp. 1136–1146, 2005.

[60] C. Cajochen, K. Kräuchi, and A. Wirz-Justice, "Role of melatonin in the regulation of human circadian rhythms and sleep," *Journal of Neuroendocrinology*, vol. 15, no. 4, pp. 432–7, 2003.

[61] H. P. V. Dongen and D. F. Dinges, "Sleep, circadian rhythms, and psychomotor vigilance," *Clinics in Sports Medicine*, vol. 24, no. 2, pp. 237–249, 2005.

[62] H. P. V. Dongen, "Shift work and inter-individual differences in sleep and sleepiness," *Chronobiology International*, vol. 23, no. 6, pp. 1139–1147, 2006.

[63] A. Kosmadopoulos, C. Sargent, D. Darwent, X. Zhou, D. Dawson, and G. D. Roach, "The effects of a split sleep–wake schedule on neurobehavioural performance and predictions of performance under conditions of forced desynchrony," *Chronobiology International*, vol. 31, no. 10, pp. 1209–1217, 2014.

[64] T. G. Raslear, S. R. Hursh, and H. P. V. Dongen, "Predicting cognitive impairment and accident risk," *Progress in Brain Research*, pp. 155–167, 2011.

[65] C. A. Czeisler, E. Weitzman, M. Moore-Ede, J. Zimmerman, and R. Knauer, "Human sleep: its duration and organization depend on its circadian phase," *Science*, vol. 210, no. 4475, pp. 1264–1267, 1980.

[66] M. Gibbs, S. Hampton, L. Morgan, and J. Arendt, "Adaptation of the circadian rhythm of 6-sulphatoxymelatonin to a shift schedule of seven nights followed by seven days in offshore oil installation workers," *Neuroscience Letters*, vol. 325, no. 2, 2002.

[67] M. J. Midwinter and J. Arendt, "Adaptation of the melatonin rhythm in human subjects following night-shift work in antarctica," *Neuroscience Letters*, vol. 122, no. 2, pp. 195–8, 1991.

[68] R. Barnes, M. Forbes, and J. Arendt, "Shift type and season affect adaptation of the 6-sulphatoxymelatonin rhythm in offshore oil rig workers," *Neuroscience Letters*, vol. 252, no. 3, pp. 179–82, 1998.

[69] K. R. Parkes, "Sleep patterns, shiftwork, and individual differences: a comparison of onshore and offshore control-room operators," *Ergonomics*, vol. 37, no. 5, pp. 827–44, 1994.

[70] A. Furnham and K. Hughes, "Individual difference correlates of nightwork and shift-work rotation," *Personality and Individual Differences*, vol. 26, no. 5, pp. 941–959, 1999.

[71] M. Ingre, G. Kecklund, T. Åkerstedt, M. Soderstrom, and L. Kecklund, "Sleep length as a function of morning shift-start time in irregular shift schedules for train drivers: self-rated health and individual differences," *Chronobiol International*, vol. 25, no. 2, pp. 349–58, 2008.

[72] M.-H. Chung, F.-M. Chang, C. C. Yang, T. B. Kuo, and N. Hsu, "Sleep quality and morningness-eveningness of shift nurses," *Journal of Clinical Nursing*, vol. 18, no. 2, pp. 279–84, 2009.

[73] R. Barnes, S. Deacon, M. Forbes, and J. Arendt, "Adaptation of the 6-sulphatoxymelatonin rhythm in shiftworkers on offshore oil installations during a 2-week 12-h night shift," *Neuroscience Letters*, vol. 241, no. 1, pp. 9–12, 1998.

[74] J. H. Hansen, I. H. Geving, and R. E. Reinertsen, "Adaptation rate of 6-sulfatoxymelatonin and cognitive performance in offshore fleet shift workers: a field study," *International Archives of Occupational and Environmental Health*, vol. 83, no. 6, pp. 607–15, 2010.

[75] D. J. Mollicone, H. P. V. Dongen, N. L. Rogers, S. Banks, and D. F. Dinges, "Time of day effects on neurobehavioral performance during chronic sleep restriction," *Aviation, Space, and Environmental Medicine*, vol. 81, no. 8, pp. 735–744, 2010.

[76] W. P. Colquhoun, J. Rutenfranz, H. Goethe, B. Neidhart, R. Condone, R. Plett, and P. Knauth, "Work at sea: a study of sleep, and of circadian rhythms in physiological and psychological functions, in watchkeepers on merchant vessels," *International Archives of Occupational and Environmental Health*, vol. 60, no. 5, pp. 321–329, 1988.

[77] R. Condon, P. Colquhoun, R. Plett, D. D. Vol, and N. Fletcher, "Work at sea: a study of sleep, and of circadian rhythms in physiological and psychological functions, in watchkeepers on merchant vessels,"

International Archives of Occupational and Environmental Health, vol. 60, no. 6, pp. 405–411, 1988.

[78] J. Rutenfranz, R. Plett, P. Knauth, R. Condon, D. DeVol, N. Fletcher, S. Eickhoff, K. H. Schmidt, R. Donis, and W. P. Colquhoun, "Work at sea: a study of sleep, and of circadian rhythms in physiological and psychological functions, in watchkeepers on merchant vessels," *International Archives of Occupational and Environmental Health*, vol. 60, no. 5, pp. 331–9, 1988.

[79] T. F. Sanquist, M. Raby, A. Forsythe, and A. B. Carvalhais, "Work hours, sleep patterns and fatigue among merchant marine personnel," *Journal of Sleep Research*, vol. 6, no. 4, pp. 245–251, 1997.

[80] M. Härmä, M. Partinen, R. Repo, M. Sorsa, and P. Siivonen, "Effects of 6/6 and 4/8 watch systems on sleepiness among bridge officers," *Chronobiology International*, vol. 25, no. 2-3, pp. 413–423, 2008.

[81] W. M. V. Leeuwen, A. Kircher, A. Dahlgren, M. Lützhöft, M. Barnett, G. Kecklund, and T. Åkerstedt, "Sleep, sleepiness, and neurobehavioral performance while on watch in a simulated 4 hours on/8 hours off maritime watch system," *Chronobiology International*, vol. 30, no. 9, pp. 1108–15, 2013.

[82] C. A. Eriksen, M. Gillberg, and P. Vestergren, "Sleepiness and sleep in a simulated "six hours on/six hours off" sea watch system," *Chronobiology International*, vol. 23, no. 6, pp. 1193–202, 2006.

[83] J. R. Jepsen, Z. Zhao, and W. M. V. Leeuwen, "Seafarer fatigue: A review of risk factors, consequences for seafarers' health and safety and options for mitigation," *International Maritime Health*, vol. 66, no. 2, pp. 106–117, 2015.

[84] S. M. Jay, D. Dawson, S. A. Ferguson, and N. Lamond, "Driver fatigue during extended rail operations," *Applied Ergonomics*, vol. 39, no. 5, pp. 623–629, 2008.

[85] D. Darwent, N. Lamond, and D. Dawson, "The sleep and performance of train drivers during an extended freight-haul operation," *Applied Ergonomics*, vol. 39, no. 5, pp. 614–622, 2008.

[86] C.-S. Huang, C.-L. Lin, L.-W. Ko, S.-Y. Liu, T.-P. Su, and C.-T. Lin, "Knowledge-based identification of sleep stages based on two forehead electroencephalogram channels," *Frontiers in Neuroscience*, vol. 8, 2014.

[87] Y.-H. Yu, S.-F. Chen, C.-S. Huang, L.-W. Ko, and C.-T. Lin, "Portable and wireless EEG device used in sleep quality tracking," in *2015 IEEE International Conference on Consumer Electronics - Taiwan*. IEEE, 2015.

[88] D. J. Buysse, C. F. Reynolds, T. H. Monk, S. R. Berman, and D. J. Kupfer, "The pittsburgh sleep quality index: A new instrument for psychiatric practice and research," *Psychiatry Research*, vol. 28, no. 2, pp. 193–213, 1989.

[89] N. Watson and B. Vaughn, *Clinician's Guide to Sleep Disorders*, ser. Neurological Disease and Therapy. CRC Press, 2006.

[90] S. Centofanti, "The impact of split shifts and naps on sleep, performance and mood," Ph.D. dissertation, Division of Education, Arts and Social Sciences, The University of South Australia, 1234, accessed on: Mar. 28, 2017. [Online]. Available: http://researchoutputs.unisa.edu.au/11541.2/123160.

[91] A. Rechtschaffen and A. Kales, *A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects*. Washington, DC: United States Government Printing Office, 1968.

[92] C. Iber, S. Ancoli-Israel, A. Chesson Jr., and S. Quan, "The new sleep scoring manual - The evidence behind the rules," *Journal of Clinical Sleep Medicine*, vol. 3, no. 2, p. 107, 2007.

[93] Z.-J. Cai, "The cholinergic-regulated thalamocortical imagining as mind differentiated in two phases of sleep: A view from the Guangdong people," *Sleep and Hypnosis*, vol. 04, pp. 1–6, 2017.

[94] B. Solheim, K. Langsrud, H. Kallestad, A. Olsen, B. Bjorvatn, and T. Sand, "Difficult morning awakening from rapid eye movement sleep and impaired cognitive function in delayed sleep phase disorder patients," *Sleep Medicine*, vol. 15, no. 10, pp. 1264–1268, 2014.

[95] R. F. Schmidt, "Integrative functions of the central nervous system," in *Human Physiology*. Springer Berlin Heidelberg, 1989, pp. 124–165.

[96] S. Keenan and M. Hirshkowitz, "Monitoring and staging human sleep," in *Principles and Practice of Sleep Medicine*. Elsevier, 2011, pp. 1602–1609.

[97] J. R. Deatherage, R. D. Roden, and K. Zouhary, "Normal sleep architecture," *Seminars in Orthodontics*, vol. 15, no. 2, pp. 86–87, 2009.

[98] D. V. Raju and R. A. Radtke, "Sleep/wake electroencephalography across the lifespan," *Sleep Medicine Clinics*, vol. 7, no. 1, pp. 13–22, 2012.

[99] D. Shrivastava, S. Jung, M. Saadat, R. Sirohi, and K. Crewson, "How to interpret the results of a sleep study," *Journal of Community Hospital Internal Medicine Perspectives*, vol. 4, no. 5, p. 24983, 2014.

[100] R. Allada and J. M. Siegel, "Unearthing the phylogenetic roots of sleep," *Current Biology*, vol. 18, no. 15, pp. R670–R679, 2008.

[101] M. H. Schmidt, "The energy allocation function of sleep: A unifying theory of sleep, torpor, and continuous wakefulness," *Neuroscience & Biobehavioral Reviews*, vol. 47, pp. 122–153, 2014.

[102] M. H. Schmidt, T. W. Swang, I. M. Hamilton, and J. A. Best, "State-dependent metabolic partitioning and energy conservation: A theoretical framework for understanding the function of sleep," *PLOS ONE*, vol. 12, no. 10, p. e0185746, 2017.

[103] J. H. Benington and H. C. Heller, "Restoration of brain energy metabolism as the function of sleep," *Progress in Neurobiology*, vol. 45, no. 4, pp. 347–360, 1995.

[104] M. T. Scharf, N. Naidoo, J. E. Zimmerman, and A. I. Pack, "The energy hypothesis of sleep revisited," *Progress in Neurobiology*, vol. 86, no. 3, pp. 264–280, 2008.

[105] J.-M. Petit, S. Burlet-Godinot, P. J. Magistretti, and I. Allaman, "Glycogen metabolism and the homeostatic regulation of sleep," *Metabolic Brain Disease*, vol. 30, no. 1, pp. 263–279, 2014.

[106] D. Kaufman, "Sleep disorders," in *Clinical Neurology for Psychiatrists*. Elsevier, 2007, pp. 371–399.

[107] A. Tarasiuk, N. Berdugo-Boura, A. Troib, and Y. Segev, "Role of growth hormone-releasing hormone in sleep and growth impairments induced by upper airway obstruction in rats," *European Respiratory Journal*, vol. 38, no. 4, pp. 870–877, 2011.

[108] E. V. Cauter and L. Plat, "Physiology of growth hormone secretion during sleep," *The Journal of Pediatrics*, vol. 128, no. 5, pp. S32–S37, 1996.

[109] S. Bohlhalter, H. Murck, F. Holsboer, and A. Steiger, "Cortisol enhances non-REM sleep and growth hormone secretion in elderly subjects," *Neurobiology of Aging*, vol. 18, no. 4, pp. 423–429, 1997.

[110] S. Sharma and M. Kavuru, "Sleep and metabolism: An overview," *International Journal of Endocrinology*, vol. 2010, pp. 1–12, 2010.

[111] J. Born and I. Wilhelm, "System consolidation of memory during sleep," *Psychological Research*, vol. 76, no. 2, pp. 192–203, 2011.

[112] B. Rasch and J. Born, "About sleep's role in memory," *Physiological Reviews*, vol. 93, no. 2, pp. 681–766, 2013.

[113] J. M. Siegel, "REM sleep," in *Principles and Practice of Sleep Medicine*. Elsevier, 2011, pp. 92–111.

[114] U. Lundberg, "Stress hormones in health and illness: The roles of work and gender," *Psychoneuroendocrinology*, vol. 30, no. 10, pp. 1017–1021, 2005.

[115] A. N. Goldstein and M. P. Walker, "The role of sleep in emotional brain function," *Annual Review of Clinical Psychology*, vol. 10, no. 1, pp. 679–708, 2014.

[116] S. A. Cairney, S. J. Durrant, R. Power, and P. A. Lewis, "Complementary roles of slow-wave sleep and rapid eye movement sleep in emotional memory consolidation," *Cerebral Cortex*, vol. 25, no. 6, pp. 1565–1575, 2014.

[117] H. S. Ephron and P. Carrington, "Rapid eye movement sleep and cortical homeostasis." *Psychological Review*, vol. 73, no. 6, pp. 500–526, 1966.

[118] J. Horne, “REM sleep — by default?” *Neuroscience & Biobehavioral Reviews*, vol. 24, no. 8, pp. 777–797, 2000.

[119] C. A. Czeisler, “Duration, timing and quality of sleep are each vital for health, performance and safety,” *Sleep Health*, vol. 1, no. 1, pp. 5–8, 2015.

[120] M. Hirshkowitz, K. Whiton, S. M. Albert, and et al, “National sleep foundation’s updated sleep duration recommendations: final report,” *Sleep Health*, vol. 1, no. 4, pp. 233–243, 2015.

[121] A. Vallières, A. Azaiez, V. Moreau, M. LeBlanc, and C. M. Morin, “Insomnia in shift work,” *Sleep Medicine*, vol. 15, no. 12, pp. 1440–1448, 2014.

[122] S. Fattinger, T. T. de Beukelaar, K. L. Ruddy, C. Volk, N. C. Heyse, J. A. Herbst, R. H. R. Hahnloser, N. Wenderoth, and R. Huber, “Deep sleep maintains learning efficiency of the human brain,” *Nature Communications*, vol. 8, p. 15405, 2017.

[123] M. Ohayon, E. M. Wickwire, M. Hirshkowitz, and et al, “National sleep foundation's sleep quality recommendations: First report,” *Sleep Health*, vol. 3, no. 1, pp. 6–19, 2017.

[124] D. L. Reed and W. P. Sacco, “Measuring sleep efficiency: What should the denominator be?” *Journal of Clinical Sleep Medicine*, vol. 12, no. 2, pp. 263–266, 2016.

[125] R. W. Matthews, S. A. Ferguson, X. Zhou, A. Kosmadopoulos, D. J. Kennaway, and G. D. Roach, “Simulated driving under the influence of extended wake, time of day and sleep restriction,” *Accident Analysis & Prevention*, vol. 45, no. SUPPL., pp. 55–61, 2012.

[126] G. J. Landry, J. R. Best, and T. Liu-Ambrose, “Measuring sleep quality in older adults: A comparison using subjective and objective methods,” *Frontiers in Aging Neuroscience*, vol. 7, no. SEP, 2015.

[127] M. Kim, “Association between objectively measured sleep quality and obesity in community-dwelling adults aged 80 years or older: A cross-sectional study,” *Journal of Korean Medical Science*, vol. 30, no. 2, pp. 199–206, 2015.

[128] T. Kongsvik, K. Størkersen, and J. Hansen, “The possible impact of different watch keeping regimes at sea on sleep, fatigue, and safety,” in *Advances in Safety, Reliability and Risk Management*, 2012, pp. 2910–2918.

[129] J. Arendt, B. Middleton, P. Williams, G. Francis, and C. Luke, “Sleep and circadian phase in a ship’s crew,” *Journal of Biological Rhythms*, vol. 21, no. 3, pp. 214–221, 2006.

[130] T. Åkerstedt, “Is there an optimal sleep-wake pattern in shift work?” *Scandinavian Journal of Work, Environment and Health*, vol. 24, no. SUPPL. 3, pp. 18–27, 1998.

[131] S. Postnova, S. W. Lockley, and P. A. Robinson, “Sleep propensity under

forced desynchrony in a model of arousal state dynamics," *Journal of Biological Rhythms*, vol. 31, no. 5, pp. 498–508, 2016.

[132] T. Åkerstedt and K. P. Wright, "Sleep loss and fatigue in shift work and shift work disorder," *Sleep Medicine Clinics*, vol. 4, no. 2, pp. 257–271, 2009.

[133] M. L. Jackson, S. Banks, and G. Belenky, "Investigation of the effectiveness of a split sleep schedule in sustaining sleep and maintaining performance," *Chronobiology International*, vol. 31, no. 10, pp. 1218–1230, 2014.

[134] R. Plett, W. P. Colquhoun, R. Condon, P. Knauth, J. Rutenfranz, and S. Eickhoff, "Work at sea: a study of sleep, and of circadian rhythms in physiological and psychological functions, in watchkeepers on merchant vessels," *International Archives of Occupational and Environmental Health*, vol. 60, no. 6, pp. 395–403, 1988.

[135] S. M. Jay, B. Aisbett, M. Sprajcer, and S. A. Ferguson, "Sleeping at work: Not all about location, location, location," *Sleep Medicine Reviews*, vol. 19, pp. 59–66, 2015.

[136] M. Grech, "Fatigue risk management: A maritime framework," *International Journal of Environmental Research and Public Health*, vol. 13, no. 2, p. 175, 2016.

[137] A. Smith and R. McNamara, "Noise and disturbed sleep aboard ships and on oil installation," in *Paper presented at the 31st International Congress on Noise Control Engineering, Dearborn, Michigan*, 2002.

[138] Y. Tamura, T. Kawada, and Y. Sasazawa, "Effect of ship noise on sleep," *Journal of Sound and Vibration*, vol. 205, no. 4, pp. 417–425, 1997.

[139] Y. Tamura, T. Horiyasu, Y. Sano, K. Chonan, T. Kawada, Y. Sasazawa, M. Kuroiwa, and S. Suzuki, "Habituation of sleep to a ship's noise as determined by actigraphy and a sleep questionnaire," *Journal of Sound and Vibration*, vol. 250, no. 1, pp. 107–113, 2002.

[140] T. Dobie and T. May, "Effects of fatigue and motion sickness on naval task performance," National Biodynamics Laboratory. University of New Orleans, Tech. Rep. UK Contract Report Number CU0000000553, 2002.

[141] P. Pisula, C. Lewis, and R. Bridger, "Vessel motion thresholds for maintaining physical and cognitive performance: a study of naval personnel at sea," *Ergonomics*, vol. 55, no. 6, pp. 636–649, 2012.

[142] S. Lerman, D. Mollicone, and S. Coats, "Use of the psychomotor vigilance test in fitness for work assessments," *Journal of Occupational and Environmental Medicine*, vol. 59, no. 8, pp. 716–720, 2017.

[143] D. Dinges and M. Mallis, "Managing fatigue by drowsiness detection: Can technological promises be realized?" *Managing Fatigue in Transportation*, pp. 209–229, 1998.

[144] S. Kerick, J. Metcalfe, T. Feng, A. Ries, and K. McDowell, "Review of fatigue management technologies for enhanced military vehicle safety and performance," US Army Research Laboratory, Aberdeen Proving Ground, MD, techreport ARL-TR-6571, 2013.

[145] A. Vadeby, Å. Forsman, G. Kecklund, T. Åkerstedt, D. Sandberg, and A. Anund, "Sleepiness and prediction of driver impairment in simulator studies using a cox proportional hazard approach," *Accident Analysis & Prevention*, vol. 42, no. 3, pp. 835–841, 2010.

[146] R. Feng, G. Zhang, and B. Cheng, "An on-board system for detecting driver drowsiness based on multi-sensor data fusion using Dempster-Shafer theory," in *2009 International Conference on Networking, Sensing and Control*. IEEE, 2009.

[147] A. Sahayadhas, K. Sundaraj, and M. Murugappan, "Detecting driver drowsiness based on sensors: A review," *Sensors*, vol. 12, no. 12, pp. 16 937–16 953, 2012.

[148] D. Amso and G. Scerif, "The attentive brain: insights from developmental cognitive neuroscience," *Nature Reviews Neuroscience*, vol. 16, no. 10, pp. 606–619, 2015.

[149] M. Catherine and H. Thomas, *Neuropsychiatry*. Lippincott Williams & Wilkins, 2003, book section 17, pp. 405–425.

[150] N. Goel, M. Basner, and D. F. Dinges, "Phenotyping of neurobehavioral vulnerability to circadian phase during sleep loss," in *Methods in Enzymology*. Elsevier, 2015, pp. 285–308.

[151] B. Mathias and D. F. Dinges, "Maximizing sensitivity of the psychomotor vigilance test (PVT) to sleep loss," *Sleep*, vol. 34, no. 5, pp. 581–591, 2011.

[152] D. F. Dinges and J. W. Powell, "Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations," *Behavior Research Methods, Instruments, & Computers*, vol. 17, no. 6, pp. 652–655, 1985.

[153] M. Basner and D. F. Dinges, "An adaptive-duration version of the PVT accurately tracks changes in psychomotor vigilance induced by sleep restriction," *Sleep*, vol. 35, no. 2, pp. 193–202, 2012.

[154] Ambulatory Monitoring, "PVT-192 user's guide and reference manual," 2019, accessed on: Mar. 28, 2017. [Online]. Available: https://www.artisantg.com/Scientific/69923-1/AMI_PVT_192_Psychomotor_Vigilance_Task_Monitor.

[155] M. Basner, E. Hermosillo, J. Nasrini, S. McGuire, S. Saxena, T. M. Moore, R. C. Gur, and D. F. Dinges, "Repeated administration effects on psychomotor vigilance test performance," *Sleep*, vol. 41, no. 1, 2017.

[156] Y. Zhu, Y. Xi, N. Fei, Y. Liu, X. Zhang, L. Liu, Z. Xu, J. Sun, X. Yang, H. Yin, J. Tian, and W. Qin, "Dynamics of cerebral responses to sustained attention performance during one night of sleep deprivation," *Journal of*

Sleep Research, vol. 27, no. 2, pp. 184–196, 2017.

[157] S. A. Drummond, A. Bischoff-Grethe, D. F. Dinges, L. Ayalon, S. C. Mednick, and M. Meloy, "The neural basis of the psychomotor vigilance task." *Sleep*, vol. 28 9, pp. 1059–68, 2005.

[158] J. Lim and D. F. Dinges, "Sleep deprivation and vigilant attention," *Annals of the New York Academy of Sciences*, vol. 1129, no. 1, pp. 305–322, 2008.

[159] K. A. Cote, C. E. Milner, S. L. Osip, M. L. Baker, and B. P. Cuthbert, "Physiological arousal and attention during a week of continuous sleep restriction," *Physiology & Behavior*, vol. 95, no. 3, pp. 353–364, 2008.

[160] H. Q. Yu, Y. Tian, C. H. Wang, S. G. Chen, and J. H. Guo, "Effects of sleep restriction and circadian rhythm on human vigilance," in *Bioinformatics and Biomedical Engineering: New Advances*, H. Zhou and J. J. Chou, Eds. CRC Press/Balkema, 2016, Conference Proceedings, pp. 367–374.

[161] Y. Li, S. Wang, Y. Zhao, and Q. Ji, "Simultaneous facial feature tracking and facial expression recognition," *IEEE Transactions on Image Processing*, vol. 22, no. 7, pp. 2559–2573, 2013.

[162] R. J. Apparies, T. C. Riniolo, and S. W. Porges, "A psychophysiological investigation of the effects of driving longer-combination vehicles," *Ergonomics*, vol. 41, no. 5, pp. 581–592, 1998.

[163] C. Zhao, M. Zhao, J. Liu, and C. Zheng, "Electroencephalogram and electrocardiograph assessment of mental fatigue in a driving simulator," *Accident Analysis & Prevention*, vol. 45, pp. 83–90, 2012.

[164] M. J. Khan and K.-S. Hong, "Passive BCI based on drowsiness detection: an fNIRS study," *Biomedical Optics Express*, vol. 6, no. 10, p. 4063, 2015.

[165] J. Horne and C. Burley, "We know when we are sleepy: Subjective versus objective measurements of moderate sleepiness in healthy adults," *Biological Psychology*, vol. 83, no. 3, pp. 266–268, 2010.

[166] A. Williamson, R. Friswell, J. Olivier, and R. Grzebieta, "Are drivers aware of sleepiness and increasing crash risk while driving?" *Accident Analysis & Prevention*, vol. 70, pp. 225–234, 2014.

[167] O. Mairesse, E. D. Valck, S. Quanten, D. Neu, A. Cortoos, N. Pattyn, P. Theuns, R. Cluydts, and J. Hofmans, "Sleepiness phenomics: Modeling individual differences in subjective sleepiness profiles," *International Journal of Psychophysiology*, vol. 93, no. 1, pp. 150–161, 2014.

[168] J. Dorrian, N. Lamond, A. L. Holmes, H. J. Burgess, G. D. Roach, A. Fletcher, and D. Dawson, "The ability to self-monitor performance during a week of simulated night shifts," *Sleep*, vol. 26, no. 7, pp. 871–877, 2003.

[169] M. Ingre, T. Åkerstedt, B. Peters, A. Anund, G. Kecklund, and A. Pickles, "Subjective sleepiness and accident risk avoiding the ecological fallacy,"

Journal of Sleep Research, vol. 15, no. 2, pp. 142–148, 2006.

[170] T. Åkerstedt, A. Anund, J. Axelsson, and G. Kecklund, "Subjective sleepiness is a sensitive indicator of insufficient sleep and impaired waking function," *Journal of Sleep Research*, vol. 23, no. 3, pp. 242–254, 2014.

[171] T. Åkerstedt and M. Gillberg, "Subjective and objective sleepiness in the active individual," *International Journal of Neuroscience*, vol. 52, no. 1-2, pp. 29–37, 1990.

[172] M. Gillberg, G. Kecklund, and T. Åkerstedt, "Relations between performance and subjective ratings of sleepiness during a night awake," *Sleep*, vol. 17, no. 3, pp. 236–241, 1994.

[173] P. Sagaspe, J. Taillard, T. Åkerstedt, V. Bayon, S. Espié, G. Chaumet, B. Bioulac, and P. Philip, "Extended driving impairs nocturnal driving performances," *PLOS ONE*, vol. 3, no. 10, 2008.

[174] A. J. Filtness, L. A. Reyner, and J. A. Horne, "Driver sleepiness-comparisons between young and older men during a monotonous afternoon simulated drive," *Biological Psychology*, vol. 89, no. 3, pp. 580–583, 2012.

[175] D. Sandberg, T. Åkerstedt, A. Anund, G. Kecklund, and M. Wahde, "Detecting driver sleepiness using optimized nonlinear combinations of sleepiness indicators," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 1, pp. 97–108, 2011.

[176] D. Hallvig, A. Anund, C. Fors, G. Kecklund, and T. Åkerstedt, "Real driving at night – predicting lane departures from physiological and subjective sleepiness," *Biological Psychology*, vol. 101, no. 1, pp. 18–23, 2014.

[177] T. Åkerstedt, D. Hallvig, A. Anund, C. Fors, J. Schwarz, and G. Kecklund, "Having to stop driving at night because of dangerous sleepiness - awareness, physiology and behaviour," *Journal of Sleep Research*, vol. 22, no. 4, pp. 380–388, 2013.

[178] K. Kaida, M. Takahashi, T. Åkerstedt, A. Nakata, Y. Otsuka, T. Haratani, and K. Fukasawa, "Validation of the karolinska sleepiness scale against performance and EEG variables," *Clinical Neurophysiology*, vol. 117, no. 7, pp. 1574–1581, 2006.

[179] K. Kaida, T. Åkerstedt, G. Kecklund, J. P. Nilsson, and J. Axelsson, "Use of subjective and physiological indicators of sleepiness to predict performance during a vigilance task," *Industrial Health*, vol. 45, no. 4, pp. 520–526, 2007.

[180] J. A. Horne and S. D. Baulk, "Awareness of sleepiness when driving," *Psychophysiology*, vol. 41, no. 1, pp. 161–165, 2004.

[181] J. V. D. Berg, G. Neely, L. Nilsson, A. Knutsson, and U. Landström,

"Electroencephalography and subjective ratings of sleep deprivation," *Sleep Medicine*, vol. 6, no. 3, pp. 231–240, 2005.

[182] M. G. Lenné and E. E. Jacobs, "Predicting drowsiness-related driving events: A review of recent research methods and future opportunities," *Theoretical Issues in Ergonomics Science*, vol. 17, no. 5-6, pp. 533–553, 2016.

[183] A. Lowden, G. Kecklund, J. Axelsson, and T. Åkerstedt, "Change from an 8-hour shift to a 12-hour shift, attitudes, sleep, sleepiness and performance," *Scandinavian Journal of Work, Environment and Health*, vol. 24, no. SUPPL. 3, pp. 69–75, 1998.

[184] M. Ingre, G. Kecklund, T. Åkerstedt, and L. Kecklund, "Variation in sleepiness during early morning shifts: A mixed model approach to an experimental field study of train drivers," *Chronobiology International*, vol. 21, no. 6, pp. 973–990, 2004.

[185] S. Waage, A. Harris, S. Pallesen, I. B. Saksvik, B. E. Moen, and B. Bjorvatn, "Subjective and objective sleepiness among oil rig workers during three different shift schedules," *Sleep Medicine*, vol. 13, no. 1, pp. 64–72, 2012.

[186] A. Lowden and T. Åkerstedt, "Assessment of a new dynamic light regimen in a nuclear power control room without windows on quickly rotating shiftworkers—effects on health, wakefulness, and circadian alignment: A pilot study," *Chronobiology International*, vol. 29, no. 5, pp. 641–649, 2012.

[187] D. Sandberg, "Detecting driver sleepiness," Ph.D. dissertation, Department of Applied Mechanic, Chalmers University of Technology, Göteborg, Sweden, 2011.

[188] L. A. Reyner and J. A. Horne, "Falling asleep whilst driving: Are drivers aware of prior sleepiness?" *International Journal of Legal Medicine*, vol. 111, no. 3, pp. 120–123, 1998.

[189] J. A. Horne and L. A. Reyner, "Counteracting driver sleepiness: Effects of napping, caffeine, and placebo," *Psychophysiology*, vol. 33, no. 3, pp. 306–309, 1996.

[190] M. Corsi-Cabrera, "Effect of 38 h of total sleep deprivation on the waking eeg in women: Sex differences," *International Journal of Psychophysiology*, vol. 50, no. 3, pp. 213–224, 2003.

[191] C. Papadelis, Z. Chen, C. Kourtidou-Papadeli, P. D. Bamidis, I. Chouvarda, E. Bekiaris, and N. Maglaveras, "Monitoring sleepiness with on-board electrophysiological recordings for preventing sleep-deprived traffic accidents," *Clinical Neurophysiology*, vol. 118, no. 9, pp. 1906–1922, 2007.

[192] C.-T. Lin, C.-J. Chang, B.-S. Lin, S.-H. Hung, C.-F. Chao, and I.-J. Wang, "A real-time wireless brain–computer interface system for drowsiness detection," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 4, no. 4, pp. 214–222, 2010.

[193] T. Brown, R. Johnson, and G. Milavetz, "Identifying periods of drowsy driving using EEG," *Annals of Advances in Automotive Medicine*, vol. 57, pp. 99–108, 2013.

[194] C. Ahlstróm, M. Nystrólym, K. Holmqvist, C. Fors, D. Sandberg, A. Anund, G. Kecklund, and T. Åkerstedt, "Fit-for-duty test for estimation of drivers' sleepiness level: Eye movements improve the sleep/wake predictor," *Transportation Research Part C: Emerging Technologies*, vol. 26, pp. 20–32, 2013.

[195] I. Daza, L. Bergasa, S. Bronte, J. Yebes, J. Almazán, and R. Arroyo, "Fusion of optimized indicators from advanced driver assistance systems (ADAS) for driver drowsiness detection," *Sensors*, vol. 14, no. 1, pp. 1106–1131, 2014.

[196] H. Martensson, O. Keelan, and C. Ahlstróm, "Driver sleepiness classification based on physiological data and driving performance from real road driving," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–10, 2018.

[197] S. Barua, M. U. Ahmed, C. Ahlström, and S. Begum, "Automatic driver sleepiness detection using EEG, EOG and contextual information," *Expert Systems with Applications*, vol. 115, pp. 121–135, 2019.

[198] H. Peng and F. Bouak, "Development of bio-mathematical models for human performance under fatigue," no. DRDC Scientific Report, DRDC-RDDC-2015-R280, 2015.

[199] Dédale Asia Pacific, "Biomathematical fatigue models guidance document," Civil Aviation Safety Authority, Technical Report, 2014, accessed on: Mar. 28, 2017. [Online]. Available: https://www.casa.gov.au/files/biomathematical-fatigue-models-guidance.

[200] A. Gundel, K. Marsalek, and C. ten Thoren, "A critical review of existing mathematical models for alertness," *Somnologie - Schlafforschung und Schlafmedizin*, vol. 11, no. 3, pp. 148–156, 2017.

[201] F. O. James, L. B. Waggoner, P. M. Weiss, P. D. Patterson, J. S. Higgins, E. S. Lang, and H. P. V. Dongen, "Does implementation of biomathematical models mitigate fatigue and fatigue-related risks in emergency medical services operations? A systematic review," *Prehospital Emergency Care*, vol. 22, no. sup1, pp. 69–80, 2018.

[202] D. Dawson, Y. I. Noy, M. Härmä, T. Åkerstedt, and G. Belenky, "Modelling fatigue and the use of fatigue models in work settings," *Accident Analysis & Prevention*, vol. 43, no. 2, pp. 549–564, 2011.

[203] J. Reifman and P. Gander, "Commentary on the three-process model of alertness and broader modeling issues," *Aviation, Space, and Environmental Medicine*, vol. 75, no. SUPPL.1, pp. A84–A88, 2004.

[204] S. Folkard, T. Åkerstedt, I. Macdonald, P. Tucker, and M. B. Spencer,

"Beyond the three-process model of alertness: Estimating phase, time on shift, and successive night effects," *Journal of Biological Rhythms*, vol. 14, no. 6, pp. 579–589, 1999.

[205] M. E. Jewett and R. E. Kronauer, "Interactive mathematical models of subjective alertness and cognitive throughput in humans," *Journal of Biological Rhythms*, vol. 14, no. 6, pp. 588–597, 1999.

[206] T. Åkerstedt, S. Folkard, and C. Portin, "Predictions from the three-process model of alertness," *Aviation, Space, and Environmental Medicine*, vol. 75, pp. 75–83, 2004.

[207] A. Belyavin and M. Spencer, "Modeling performance and alertness: The QinetiQ approach," *Aviation, Space, and Environmental Medicine*, vol. 75, pp. 93–103, 2004.

[208] B. Fulcher, A. Phillips, and P. Robinson, "Quantitative physiologically based modeling of subjective fatigue during sleep deprivation," *Journal of Theoretical Biology*, vol. 264, no. 2, pp. 407–19, 2010.

[209] M. Puckeridge, B. Fulcher, A. Phillips, and P. Robinson, "Incorporation of caffeine into a quantitative model of fatigue and sleep," *Journal of Theoretical Biology*, vol. 273, no. 1, pp. 44–54, 2011.

[210] A. Phillips, C. A. Czeisler, and E. B. Klerman, "Revisiting spontaneous internal desynchrony using a quantitative model of sleep physiology," *Journal of Biological Rhythms*, vol. 26, no. 5, pp. 441–53, 2011.

[211] S. Postnova, S. W. Lockley, and P. A. Robinson, "Prediction of cognitive performance and subjective sleepiness using a model of arousal dynamics," *Journal of Biological Rhythms*, vol. 33, no. 2, pp. 203–218, 2018.

[212] A. Phillips and P. Robinson, "Sleep deprivation in a quantitative physiologically based model of the ascending arousal system," *Journal of Theoretical Biology*, vol. 255, no. 4, pp. 413–23, 2008.

[213] M. A. S. Hilaire, E. B. Klerman, S. B. S. Khalsa, K. P. Wright, C. A. Czeisler, and R. E. Kronauer, "Addition of a non-photic component to a light-based mathematical model of the human circadian pacemaker," *Journal of Theoretical Biology*, vol. 247, no. 4, pp. 583–99, 2007.

[214] A. Phillips and P. Robinson, "A quantitative model of sleep-wake dynamics based on the physiology of the brainstem ascending arousal system," *Journal of Biological Rhythms*, vol. 22, no. 2, pp. 167–79, 2007.

[215] S. Postnova, A. Layden, P. A. Robinson, A. J. Phillips, and R. G. Abeysuriya, "Exploring sleepiness and entrainment on permanent shift schedules in a physiologically based model," *Journal of Biological Rhythms*, vol. 27, no. 1, pp. 91–102, 2012.

[216] S. Postnova, P. A. Robinson, and D. D. Postnov, "Adaptation to shift work: Physiologically based modeling of the effects of lighting and shifts' start

time," *PLOS ONE*, vol. 8, no. 1, 2013.

[217] S. Postnova, D. D. Postnov, M. Seneviratne, and P. A. Robinson, "Effects of rotation interval on sleepiness and circadian dynamics on forward rotating 3-shift systems," *Journal of Biological Rhythms*, vol. 29, no. 1, pp. 60–70, 2014.

[218] G. S. Larue, A. Rakotonirainy, and A. N. Pettitt, "Driving performance impairments due to hypovigilance on monotonous roads," *Accident Analysis & Prevention*, vol. 43, no. 6, pp. 2037–2046, 2011.

[219] J. Gale, T. L. Signal, A. L. Garden, and P. H. Gander, "Electroencephalography artifacts in workplace alertness monitoring," *Scandinavian Journal of Work, Environment & Health*, vol. 33, no. 2, pp. 148–152, 2007.

[220] C.-H. Chuang, C.-S. Huang, L.-W. Ko, and C.-T. Lin, "An EEG-based perceptual function integration network for application to drowsy driving," *Knowledge-based Systems*, vol. 80, pp. 143–152, 2015.

[221] Y.-T. Wang, K.-C. Huang, C.-S. Wei, T.-Y. Huang, L.-W. Ko, C.-T. Lin, C.-K. Cheng, and T.-P. Jung, "Developing an EEG-based on-line closed-loop lapse detection and mitigation system," *Frontiers in Neuroscience*, vol. 8, 2014.

[222] G. Li, B.-L. Lee, and W.-Y. Chung, "Smartwatch-based wearable EEG system for driver drowsiness detection," *IEEE Sensors Journal*, vol. 15, no. 12, pp. 7169–7180, 2015.

[223] Y. M. Chi, T.-P. Jung, and G. Cauwenberghs, "Dry-contact and noncontact biopotential electrodes: Methodological review," *IEEE Reviews in Biomedical Engineering*, vol. 3, pp. 106–119, 2010.

[224] Q. He, Z. Fei, X. Fan, and W. Li, "Driver fatigue evaluation model with integration of multi-indicators based on dynamic Bayesian network," *IET Intelligent Transport Systems*, vol. 9, no. 5, pp. 547–554, 2015.

[225] X. Li, E. Seignez, A. Lambert, and P. Loonis, "Real-time driver drowsiness estimation by multi-source information fusion with Dempster–Shafer theory," *Transactions of the Institute of Measurement and Control*, vol. 36, no. 7, pp. 906–915, 2013.

[226] J. Dorrian, N. Lamond, and D. Dawson, "The ability to self-monitor performance when fatigued," *Journal of Sleep Research*, vol. 9, no. 2, pp. 137–144, 2000.

[227] N. Rogers and D. Dinges, "Subjective surrogates of performance during night work." *Sleep*, vol. 26, no. 7, pp. 790–791, 2003.

[228] A. Čolić, O. Marques, and B. Furht, "Driver drowsiness detection and measurement methods," in *Driver Drowsiness Detection*. Springer International Publishing, 2014, pp. 7–18.

[229] A. K. Hyder, *Multisensor Fusion*, 1st ed., ser. Nato Science Series II, A. K.

Hyder, E. Shahbazian, and E. Waltz, Eds. Springer Netherlands, 2002, vol. 1.

[230] A. D. McDonald, J. D. Lee, C. Schwarz, and T. L. Brown, "A contextual and temporal algorithm for driver drowsiness detection," *Accident Analysis & Prevention*, vol. 113, pp. 25–37, 2018.

[231] Z. Li-Hong, D. Ning, and Y. Cai-Kun, "Fatigue-driving recognition based on the state of the human eye and electroencephalographic signal," in *2016 Chinese Control and Decision Conference*. IEEE, 2016, Conference Proceedings, pp. 1567–1572.

[232] V. J. Kartsch, S. Benatti, P. D. Schiavone, D. Rossi, and L. Benini, "A sensor fusion approach for drowsiness detection in wearable ultra-low-power systems," *Information Fusion*, vol. 43, pp. 66–76, 2018.

[233] R. N. Khushaba, S. Kodagoda, S. Lal, and G. Dissanayake, "Driver drowsiness classification using fuzzy wavelet-packet-based feature-extraction algorithm," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 1, pp. 121–131, 2011.

[234] R. N. Khushaba, S. Kodagoda, and S. Lal, "Uncorrelated fuzzy neighborhood preserving analysis based feature projection for driver drowsiness recognition," *Fuzzy Sets and Systems*, vol. 221, pp. 90–111, 2013.

[235] C. Zhang, C.-X. Zheng, and X.-L. Yu, "Automatic recognition of cognitive fatigue from physiological indices by using wavelet packet transform and kernel learning algorithms," *Expert Systems with Applications*, vol. 36, no. 3, pp. 4664–4671, 2009.

[236] B.-G. Lee, B.-L. Lee, and W.-Y. Chung, "Mobile healthcare for automatic driving sleep-onset detection using wavelet-based EEG and respiration signals," *Sensors*, vol. 14, no. 10, pp. 17 915–17 936, 2014.

[237] X.-Q. Huo, W.-L. Zheng, and B.-L. Lu, "Driving fatigue detection with fusion of EEG and forehead EOG," in *2016 International Joint Conference on Neural Networks*, vol. 2016-October. IEEE, 2016, Conference Proceedings, pp. 897–904.

[238] Z. Guo, Y. Pan, G. Zhao, S. Cao, and J. Zhang, "Detection of driver vigilance level using EEG signals and driving contexts," *IEEE Transactions on Reliability*, vol. 67, no. 1, pp. 370–380, 2018.

[239] W. Sun, X. Zhang, S. Peeta, X. He, and Y. Li, "A real-time fatigue driving recognition method incorporating contextual features and two fusion levels," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 12, pp. 3408–3420, 2017.

[240] T. Nguyen, S. Ahn, H. Jang, S. C. Jun, and J. G. Kim, "Utilization of a combined EEG/NIRS system to predict driver drowsiness," *Scientific Reports*, vol. 7, p. 43933, 2017.

[241] P. Boyraz, M. Acar, and D. Kerr, "Multi-sensor driver drowsiness

monitoring," *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, vol. 222, no. 11, pp. 2041–2062, 2008.

[242] S. Samiee, S. Azadi, R. Kazemi, A. Nahvi, and A. Eichberger, "Data fusion to develop a driver drowsiness detection system with robustness to signal loss," *Sensors*, vol. 14, no. 9, pp. 17 832–17 847, 2014.

[243] C. Xu, S.-J. Pei, and X.-S. Wang, "Driver drowsiness detection based on non-intrusive metrics considering individual difference," *Zhongguo Gonglu Xuebao/China Journal of Highway and Transport*, vol. 29, no. 10, pp. 118–125, 2016.

[244] Z. Li, L. Chen, J. Peng, and Y. Wu, "Automatic detection of driver fatigue using driving operation information for transportation safety," *Sensors*, vol. 17, no. 6, p. 1212, 2017.

[245] M. B. Kurt, N. Sezgin, M. Akin, G. Kirbas, and M. Bayram, "The ANN-based computing of drowsy level," *Expert Systems with Applications*, vol. 36, no. 2, pp. 2534–2542, 2009.

[246] C. Zhang, H. Wang, and R. Fu, "Automated detection of driver fatigue based on entropy and complexity measures," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 1, pp. 168–177, 2014.

[247] Q. He, X. Fan, Z. Fei, and W. Li, "Evaluation of driver fatigue with multi-indicators based on artificial neural network," *IET Intelligent Transport Systems*, vol. 10, no. 8, pp. 555–561, 2016.

[248] W. Sun, X. Zhang, S. Peeta, X. He, Y. Li, and S. Zhu, "A self-adaptive dynamic recognition model for fatigue driving based on multi-source information and two levels of fusion," *Sensors*, vol. 15, no. 9, pp. 24 191–24 213, 2015.

[249] G. Yang, Y. Lin, and P. Bhattacharya, "Multimodality inferring of human cognitive states based on integration of neuro-fuzzy network and information fusion techniques," *EURASIP Journal on Advances in Signal Processing*, vol. 2008, no. 1, 2007.

[250] L.-L. Chen, Y. Zhao, J. Zhang, and J.-Z. Zou, "Automatic detection of alertness/drowsiness from physiological signals using wavelet-based nonlinear features and machine learning," *Expert Systems with Applications*, vol. 42, no. 21, pp. 7344–7355, 2015.

[251] P. Ranganathan, C. S. Pramesh, and R. Aggarwal, "Common pitfalls in statistical analysis: Logistic regression," *Perspectives in Clinical Research*, vol. 8, no. 3, pp. 148–51, 2017.

[252] A. Murata, "Proposal of a method to predict subjective rating on drowsiness using physiological and behavioral measures," *IIE Transactions on Occupational Ergonomics and Human Factors*, vol. 7323, pp. 1–13, 2016.

[253] S. Alayón, R. Robertson, S. K. Warfield, and J. Ruiz-Alzola, "A fuzzy

system for helping medical diagnosis of malformations of cortical development," *Journal of Biomedical Informatics*, vol. 40, no. 3, pp. 221–235, 2007.

[254] P. Harliana and R. Rahim, "Comparative analysis of membership function on mamdani fuzzy inference system for decision making," *Journal of Physics: Conference Series*, vol. 930, no. 1, p. 012029, 2017.

[255] R. Senaratne, D. Hardy, B. Vanderaa, and S. Halgamuge, "Driver fatigue detection by fusing multiple cues," in *Advances in Neural Networks*, D. Liu, S. Fei, Z. Hou, H. Zhang, and C. Sun, Eds. Springer Berlin Heidelberg, 2007, pp. 801–809.

[256] A. Picot, S. Charbonnier, and A. Caplier, "On-line detection of drowsiness using brain and visual information," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 42, no. 3, pp. 764–775, 2012.

[257] G. Yang, Y. Lin, and P. Bhattacharya, "A driver fatigue recognition model based on information fusion and dynamic Bayesian network," *Information Sciences*, vol. 180, no. 10, pp. 1942–1954, 2010.

[258] R. Fu, H. Wang, and W. Zhao, "Dynamic driver fatigue detection using hidden markov model in real driving condition," *Expert Systems with Applications*, vol. 63, pp. 397–411, 2016.

[259] C. Craye, A. Rashwan, M. S. Kamel, and F. Karray, "A multi-modal driver fatigue and distraction assessment system," *International Journal of Intelligent Transportation Systems Research*, vol. 14, no. 3, pp. 173–194, 2015.

[260] L. Gang, X. Song, M. Zhang, B. Yao, and L. Zhou, "Analysis of driver fatigue causations based on the Bayesian metwork model," *Simulation*, vol. 93, no. 7, pp. 553–565, 2016.

[261] C. Xu, X. Wang, X. Chen, and H. Zhang, "Driver drowsiness level analysis and predication based on decision tree," *Journal of Tongji University*, vol. 43, pp. 75–81, 2015.

[262] N. Zhang, W.-L. Zheng, W. Liu, and B.-L. Lu, "Continuous vigilance estimation using LSTM neural networks," in *Neural Information Processing*. Springer International Publishing, 2016, pp. 530–537.

[263] L. Hagenmeyer, "Development of a multimodal, universal human-machine-interface for hypovigilance-management-systems," Ph.D. dissertation, Institut für Arbeitswissenschaft und Technologiemanagement, Universität Stuttgart, Stuttgart, 2007, accessed on: Mar. 28, 2017. [Online]. Available: https://elib.uni-stuttgart.de/handle/11682/6670.

[264] R. P. Balandong, R. F. Ahmad, M. N. M. Saad, and A. S. Malik, "A review on EEG-based automatic sleepiness detection systems for driver," *IEEE Access*, vol. 6, pp. 22 908–22 919, 2018.

[265] T. Azim, M. A. Jaffar, and A. M. Mirza, "Fully automated real time fatigue detection of drivers through fuzzy expert systems," *Applied Soft Computing*, vol. 18, pp. 25–38, 2014.

[266] A. G. Correa, L. Orosco, and E. Laciar, "Automatic detection of drowsiness in EEG records based on multimodal analysis," *Medical Engineering & Physics*, vol. 36, no. 2, pp. 244–249, 2014.

[267] R. Chai, S. H. Ling, P. P. San, G. R. Naik, T. N. Nguyen, Y. Tran, A. Craig, and H. T. Nguyen, "Improving EEG-based driver fatigue classification using sparse-deep belief networks," *Frontiers in Neuroscience*, vol. 11, 2017.

[268] A.-M. Chang, D. Aeschbach, J. F. Duffy, and C. A. Czeisler, "Evening use of light-emitting eReaders negatively affects sleep, circadian timing, and next-morning alertness," *Proceedings of the National Academy of Sciences*, vol. 112, no. 4, pp. 1232–1237, 2014.

[269] J. Burt, G. Barber, and D. Rigby, *Elementary Statistics for Geographers*, 3rd ed. New York, London: Guilford Press, 2009.

[270] H. Cevallos Valdiviezo and S. Van Aelst, "Tree-based prediction on incomplete data using imputation or surrogate decisions," *Information Sciences*, vol. 311, pp. 163–181, 2015.

[271] M. Saar-Tsechansky and F. Provost, "Handling missing values when applying classification models," *Journal of Machine Learning Research*, vol. 8, pp. 1625–1657, 2007.

[272] L. Tran, X. Liu, J. Zhou, and R. Jin, "Missing modalities imputation via cascaded residual autoencoder," in *2017 IEEE Conference on Computer Vision and Pattern Recognition*, vol. 2017-January, 2017, pp. 4971–4980.

[273] P. García-Laencina, J.-L. Sancho-Gómez, and A. Figueiras-Vidal, "Pattern classification with missing data: A review," *Neural Computing and Applications*, vol. 19, no. 2, pp. 263–282, 2010.

[274] M. Flores, J. Armingol, and A. de la Escalera, "Driver drowsiness warning system using visual information for both diurnal and nocturnal illumination conditions," *Journal on Advances in Signal Processing*, vol. 2010, no. 1, p. 438205, 2010.

[275] B. Bjorvatn, K. Stangenes, N. Øyane, K. Forberg, A. Lowden, F. Holsten, and T. Åkerstedt, "Subjective and objective measures of adaptation and readaptation to night work on an oil rig in the north sea," *Sleep*, vol. 29, no. 6, pp. 821–829, 2006.

[276] S. Dinerstein, J. Dinerstein, and D. Ventura, "Robust multi-modal biometric fusion via multiple SVMs," in *2007 IEEE International Conference on Systems, Man and Cybernetics*, 2007, pp. 1530–1535.

[277] B. Zhu, C. He, and P. Liatsis, "A robust missing value imputation method for noisy data," *Applied Intelligence*, vol. 36, no. 1, pp. 61–74, 2012.

[278] L. Belanche, V. Kobayashi, and T. Aluja, "Handling missing values in kernel methods with application to microbiology data," *Neurocomputing*, vol. 141, pp. 110–116, 2014.

[279] A. Ali, H. Ying, D. Filev, F. Syed, K. O. Prakah-Asante, F. Tseng, and H.-H. Yang, "A two-stage-training support vector machine approach to predicting unintentional vehicle lane departure," *Journal of Intelligent Transportation Systems*, vol. 21, no. 1, pp. 41–51, 2016.

[280] J. M. Ambarak, H. Ying, F. Syed, and D. Filev, "A neural network for predicting unintentional lane departures," in *2017 IEEE International Conference on Industrial Technology (ICIT)*. IEEE, 2017, pp. 492–497.

[281] A. D. McDonald, J. D. Lee, C. Schwarz, and T. L. Brown, "Steering in a random forest," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 56, no. 5, pp. 986–998, 2013.

[282] Y. Liang, W. J. Horrey, M. E. Howard, M. L. Lee, C. Anderson, M. S. Shreeve, C. S. O'Brien, and C. A. Czeisler, "Prediction of drowsiness events in night shift workers during morning driving," *Accident Analysis & Prevention*, vol. 126, pp. 105–114, 2017.

[283] G. S. Larue, A. Rakotonirainy, and A. N. Pettitt, "Predicting reduced driver alertness on monotonous highways," *IEEE Pervasive Computing*, vol. 14, no. 2, pp. 78–85, 2015.

[284] K. R. Baykaner, M. Huckvale, I. Whiteley, S. Andreeva, and O. Ryumin, "Predicting fatigue and psychophysiological test performance from speech for safety-critical environments," *Frontiers in Bioengineering and Biotechnology*, vol. 3, 2015.

[285] S. Rajaraman, A. V. Gribok, N. J. Wesensten, T. J. Balkin, and J. Reifman, "Individualized performance prediction of sleep-deprived individuals with the two-process model," *Journal of Applied Physiology*, vol. 104, no. 2, pp. 459–468, 2008.

[286] S. S. H. Nazari, A. Moradi, and K. Rahmani, "A systematic review of the effect of various interventions on reducing fatigue and sleepiness while driving," *Chinese Journal of Traumatology*, vol. 20, no. 5, pp. 249–258, 2017.

[287] S. Postnova and P. A. Robinson, "Forced wakefulness for entrainment to permanent shift work: A computational study," in *Advances in Cognitive Neurodynamics (III)*. Springer Netherlands, 2013, pp. 105–111.

[288] S. Lawoyin, D. Y. Fei, O. Bai, and X. Liu, "Evaluating the efficacy of an accelerometer-based method for drowsy driving detection," *International Journal of Vehicle Safety*, vol. 8, no. 2, pp. 165–179, 2015.

[289] K.-Q. Shen, X.-P. Li, C.-J. Ong, S.-Y. Shao, and E. P. Wilder-Smith, "EEG-based mental fatigue measurement using multi-class support vector machines with confidence estimate," *Clinical Neurophysiology*, vol. 119, no. 7, pp. 1524–33, 2008.

[290] C. Zhao, Y. Zhang, X. Zhang, and J. He, “Recognition of driver’s fatigue expression using local multiresolution derivative pattern,” *Journal of Intelligent & Fuzzy Systems*, vol. 30, no. 1, pp. 547–560, 2015.

[291] L. Zhao, Z. Wang, X. Wang, and Q. Liu, “Driver drowsiness detection using facial dynamic fusion information and a DBN,” *IET Intelligent Transport Systems*, vol. 12, no. 2, pp. 127–133, 2018.

[292] M. Chai, S.-W. Li, W.-C. Sun, M.-Z. Guo, and M.-Y. Huang, “Drowsiness monitoring based on steering wheel status,” *Transportation Research Part D: Transport and Environment*, vol. 66, pp. 95–103, 2019.

[293] L. Zilin, G. Tan, Y. Pang, Y. Tang, and K. Qian, “Driving fatigue detection based on blink frequency and eyes movement,” in *SAE Technical Paper Series*. SAE International, 2017.

[294] X. Wang and C. Xu, “Driver drowsiness detection based on non-intrusive metrics considering individual specifics,” *Accident Analysis & Prevention*, vol. 95, pp. 350–357, 2016.

[295] S. Baiardi, C. L. Morgia, L. Sciamanna, A. Gerosa, F. Cirignotta, and S. Mondini, “Is the epworth sleepiness scale a useful tool for screening excessive daytime sleepiness in commercial drivers?” *Accident Analysis & Prevention*, vol. 110, pp. 187–189, 2018.

[296] K. Murphy, “Dynamic Bayesian networks: Representation, inference and learning,” Ph.D. dissertation, Computer Science, University of California, Berkeley, 2002, accessed on: Mar. 28, 2017. [Online]. Available: https://ibug.doc.ic.ac.uk/media/uploads/documents/courses/DBN-PhDthesis-LongTutorail-Murphy.pdf.

[297] S. Ftouni, S. A. Rahman, K. E. Crowley, C. Anderson, S. M. Rajaratnam, and S. W. Lockley, “Temporal dynamics of ocular indicators of sleepiness across sleep restriction,” *Journal of Biological Rhythms*, vol. 28, no. 6, pp. 412–424, 2013.

[298] J. Shekleton, S. Rajaratnam, J. Gooley, E. Van Reen, C. Czeisler, and S. Lockley, “Improved neurobehavioral performance during the wake maintenance zone,” *Journal of Clinical Sleep Medicine*, vol. 9, no. 4, pp. 353–362, 2013.

[299] E.-P. Chua, W.-Q. Tan, S.-C. Yeo, P. Lau, I. Lee, I. Mien, K. Puvanendran, and J. Gooley, “Heart rate variability can be used to estimate sleepiness-related decrements in psychomotor vigilance during total sleep deprivation,” *Sleep*, vol. 35, no. 3, pp. 325–334, 2012.

[300] E. J. Silva, W. Wang, J. M. Ronda, J. K. Wyatt, and J. F. Duffy, “Circadian and wake-dependent influences on subjective sleepiness, cognitive throughput, and reaction time performance in older and young adults,” *Sleep*, vol. 33, no. 4, pp. 481–490, apr 2010.

[301] T. M. Burke, F. A. J. L. Scheer, J. M. Ronda, C. A. Czeisler, and K. P. Wright, “Sleep inertia, sleep homeostatic and circadian influences on

higher-order cognitive functions," *Journal of Sleep Research*, vol. 24, no. 4, pp. 364–371, mar 2015.

[302] J. K. Wyatt, C. Cajochen, A. R.-D. Cecco, C. A. Czeisler, and D.-J. Dijk, "Low-dose repeated caffeine administration for circadian-phase-dependent performance degradation during extended wakefulness," *Sleep*, vol. 27, no. 3, pp. 374–381, may 2004.

[303] S. Grady, D. Aeschbach, K. P. Wright, and C. A. Czeisler, "Effect of modafinil on impairments in neurobehavioral performance and learning associated with extended wakefulness and circadian misalignment," *Neuropsychopharmacology*, vol. 35, no. 9, pp. 1910–1920, may 2010.

[304] J. K. Wyatt, A. R.-D. Cecco, C. A. Czeisler, and D.-J. Dijk, "Circadian temperature and melatonin rhythms, sleep, and neurobehavioral function in humans living on a 20-h day," *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, vol. 277, no. 4, pp. R1152–R1163, oct 1999.

[305] K.-Q. Shen, C.-J. Ong, X.-P. Li, Z. Hui, and E. Wilder-Smith, "A feature selection method for multilevel mental fatigue EEG classification," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 7, pp. 1231–1237, 2007.

[306] S. N. Lahiri, *Resampling Methods for Dependent Data*, 1st ed., ser. Springer Series in Statistics. Springer-Verlag New York, 2003.

[307] T. Saito and M. Rehmsmeier, "The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets," *PLOS ONE*, vol. 10, no. 3, 2015.

[308] M. Stikic, R. R. Johnson, D. J. Levendowski, D. P. Popovic, R. E. Olmstead, and C. Berka, "Eeg-derived estimators of present and future cognitive performance," *Frontiers in Human Neuroscience*, vol. 5, 2011.

[309] D. Qian, B. Wang, X. Qing, T. Zhang, Y. Zhang, X. Wang, and M. Nakamura, "Drowsiness detection by Bayesian-Copula discriminant classifier based on EEG signals during daytime short nap," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 4, pp. 743–754, 2017.

[310] J. Yu, S. Park, S. Lee, and M. Jeon, "Representation learning, scene understanding, and feature fusion for drowsiness detection," in *Asian Conference on Computer Vision*, ser. Lecture Notes in Computer Science, vol. 10118. Springer, Cham, 2017, pp. 165–177.

[311] J. Demšar, "Statistical comparisons of classifiers over multiple data sets," *Journal of Machine Learning Research*, vol. 7, pp. 1–30, 2006.

[312] M. R. Stevenson, J. Elkington, L. Sharwood, L. Meuleners, R. Ivers, S. Boufous, A. Williamson, N. Haworth, M. Quinlan, R. Grunstein, R. Norton, and K. Wong, "The role of sleepiness, sleep disorders, and the work environment on heavy-vehicle crashes in 2 Australian states," *American Journal of Epidemiology*, vol. 179, no. 5, pp. 594–601, 2014.

[313] X. Li and Q. Ji, "Active affective state detection and user assistance with dynamic Bayesian networks," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 35, no. 1, pp. 93–105, 2005.

[314] A. A. Albousefi, H. Ying, D. Filev, F. Syed, K. O. Prakah-Asante, F. Tseng, and H.-H. Yang, "A support vector machine approach to unintentional vehicle lane departure prediction," in *2014 IEEE Intelligent Vehicles Symposium Proceedings*. IEEE, 2014, pp. 299–303.

[315] A. T. M. Van De Water, A. Holmes, and D. A. Hurley, "Objective measurements of sleep for non-laboratory settings as alternatives to polysomnography - A systematic review," *Journal of Sleep Research*, vol. 20, no. 1 PART II, pp. 183–200, 2011.

[316] A. Kosmadopoulos, C. Sargent, D. Darwent, X. Zhou, and G. D. Roach, "Alternatives to polysomnography (PSG): A validation of wrist actigraphy and a partial-PSG system," *Behavior Research Methods*, vol. 46, no. 4, pp. 1032–1041, 2014.

APPENDICES

Appendix A: Model of the Arousal State Dynamics

The Model of the Arousal State Dynamics (MAD) is based on the well-substantiated neuroanatomical and neurophysiological underlying the regulatory mechanisms responsible for the sleep-wake cycle, and its illustration and example of dynamic are depicted in Figure 5.2. The model was developed based on the mutually inhibitory interaction between the wake-promoting monoaminergic neurons (MA) in the brain stem and sleep-active neurons which located predominantly in the ventrolateral preoptic (VLPO) areas of the anterior hypothalamus, The mutual inhibition between the MA and VLPO resulting in only one group being active at a time and is analogous to the 'flip-flop' switch in electronic [208, 214]. Such that, VLPO and MA become dominant during sleep and wake, respectively. The VLPO is driven by inputs from the homeostasis and circadian processes while the MA receive its input from the cholinergic and orexinergic neurons. In this model, the interaction between MA and VLPO group is treated as the average properties over all neuronal population in each group.

The time evolutions of the mean neuronal potential for each of the MA and VLPO population, in response of each of the input are given by the equation:

$$\tau_m \frac{dV_{\mathrm{m}}}{dt} = -V_{\mathrm{m}} + v_{mv} Q_v + D_m \tag{5.1}$$

$$\tau_v \frac{dV_v}{dt} = -V_v + v_{vm} Q_m + D_v \tag{5.2}$$

The variable τ_v and τ_m are the time constants involved in *charging* and discharging the MA and VLPO neuronal population. Here v_{vm} and v_{mv} are the coupling strength between VLPO-to-MA and MA-to-VLPO, respectively. The drive from other neuronal population to the MA network including the responses from the orexinergic, acetylcholine-related can be simplified by averaging the combined inputs in a simplified constant drive, D_m. The mean firing rates, Q_m and Q_v , across a population of neurons expressed as a sigmoidal function

$$Q_i = \frac{Q_{i\,\max}}{1 + \exp\left[\frac{\theta - V_i}{\sigma}\right]} \tag{5.3}$$

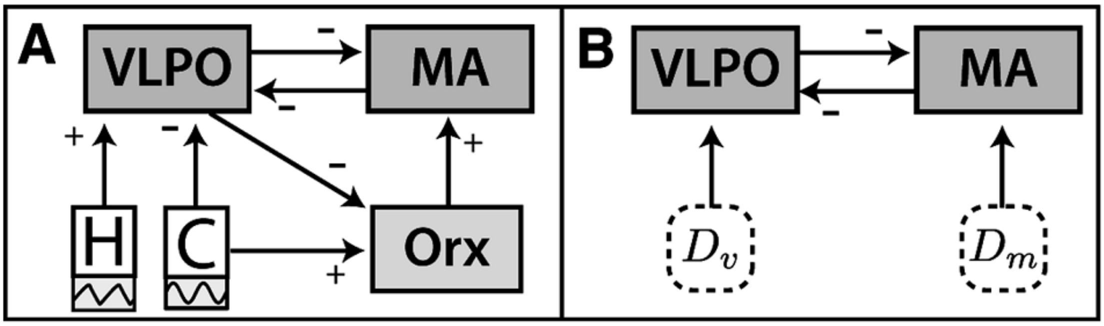


Figure 5.2: Illustration of dynamic interaction between sleep-wake networks regulated by interaction of circadian(C) process and homeostatic (H) processes. The ventrolateral preoptic nucleus (VLPO) nuclei and monoaminergic (MA) are activated by D_v and D_m, respectively, and consequently influence the sleep–wake transition. D_v is a combination of homeostasis and circadian processes, Figure adapted from [208].

Where $i = v, m$, Q_{max} is the maximum mean firing rate for any given neuronal population, θ is a mean firing threshold, and σ denote the standard deviation of the firing threshold that is responsible in controlling the steepness of the response curve (slope).

The net drive into the VLPO resulting from the homeostasis (H), circadian (C) and initial level for sleep drive level D_0 is group under the variable D_v as shown in below.

$$D_V = V_{vh}H + V_{vc}C + D_{\mathbf{0}} \tag{5.4}$$

The weight parameter,V_{vh} *and* V_{vc} govern the distinct contribution of each of the homeostasis and circadian processes to the VLPO.

The D_v also referred as total sleep drive as it primarily responsible in generating the overall drive to sleep.

The differential equation that govern the accumulation and dissipation homeostatic H component during wakefulness and sleep, respectively can be represented according to

$$\frac{dH}{dt} = \frac{-H + \mu Q_m}{\chi} \tag{5.5}$$

The χ represent the rise and decay time constant of the homeostatic sleep pressure and, μ is a constant

Visual transduction begins when a photon of light reach to each photoreceptor cell that in a “ready” state in the retina. Subsequently, the light energy is converted to an electrical impulse, a process known as photo-transduction and, the photoreceptor is now in the condition referred to as ”used” state

The transition rate α of the photoreceptor from the “dark” to “activated” state with the dependency to light intensity can be represented as

$$\alpha(I) = \alpha_0(\frac{I}{I_0})^p \frac{I}{I + I_1} \tag{5.6}$$

Where I, I_0 and p are tweaking constant that usually varied to fit the model’s output against certain experimental data. I represent intensity of the illuminating light in the surrounding region.

The number of available photoreceptor in the “dark” state at time instant t, can be expressed in the form

$$\frac{dn}{dt} = [\alpha\,(1-n) - \beta n] \tag{5.7}$$

β reflects the rate of recovery process by which the activated photoreceptor cell returns from the “activated” to “dark” state.

The net photic drive B to the circadian pacemaker is directly proportional to the transition rate α of the photoreceptor from the ”dark” to ”activated” state and, the number of photoreceptor (1-n) that is ”dark” to be converted to the ”activated” state during the next cycle of light activation.

$$B = G\alpha\,(1-n)(1-\varepsilon\mathcal{X})(1-\varepsilon\mathcal{X}_C) \tag{5.8}$$

Where the value for constant G and ε were specifically tailored to fit the model’s predictions to the observed experimental data [96].

Then, the influence of light input into the circadian pacemaker activity can be represented by

$$\frac{d\mathcal{X}}{dt} = \Omega\,[\mathcal{X}_C + \gamma\left(\frac{1}{3}\mathcal{X} + \frac{4}{3}\mathcal{X}^3 - \frac{256}{105}\mathcal{X}^7\right) + B] \tag{5.9}$$

The parameter Ω is used to scale the biological period to 24 circadian hours per cycle, and γ determine the oscillator stiffness

The circadian sensitivity modulator is represented by the following differential equation

$$\frac{d\mathcal{X}_c}{dt} = \Omega\,[qB\mathcal{X}_C - \mathcal{X} < \left(\frac{\delta}{\tau_c}\right)^2 kB >] \tag{5.10}$$

The parameters k and q are responsible in controlling the photic drive strength and τ_c denote the intrinsic period.

It has been shown that non-photic stimuli, such as meals and locomotor timing, may influence the phase of the circadian pacemaker. The non-photic drive N_s can be represented by

$$N_s = \rho\left(\frac{1}{3} - s\right)[1 - \tanh(rx)] \tag{5.11}$$

Where responsible for the sensitivity of the non-photic stimulation, and r govern the timing of the non-photic effect relative to the core body temperature minimum. The variable s is set equal to 1 and 0 during awake and sleep, respectively.

Thus, the contribution by both the light-dark cycle and non-photic inputs to the circadian pacemaker activity x can be expressed as

$$\frac{d\mathcal{X}}{dt} = \Omega[\mathcal{X}_C + \gamma\left(\frac{1}{3}\mathcal{X} + \frac{4}{3}\mathcal{X}^3 - \frac{256}{105}\mathcal{X}^7\right) + B + N_s] \tag{5.12}$$

In the human circadian pacemaker model of St. Hilaire, the oscillations of the circadian variable $\mathcal{X}$ follow an approximately sinusoidal pattern with respect to time and, the maximum and minimum amplitude of displacement was in between -1 and 1. Since the SCN firing rate can only be a magnitude with positive real number, [210] introduced the following linear function to normalize the magnitude of the circadian variable into the VLPO in the range of 0 to 1.

$$C = 0.1C_{old} + C_{adj}^2 \tag{5.13}$$

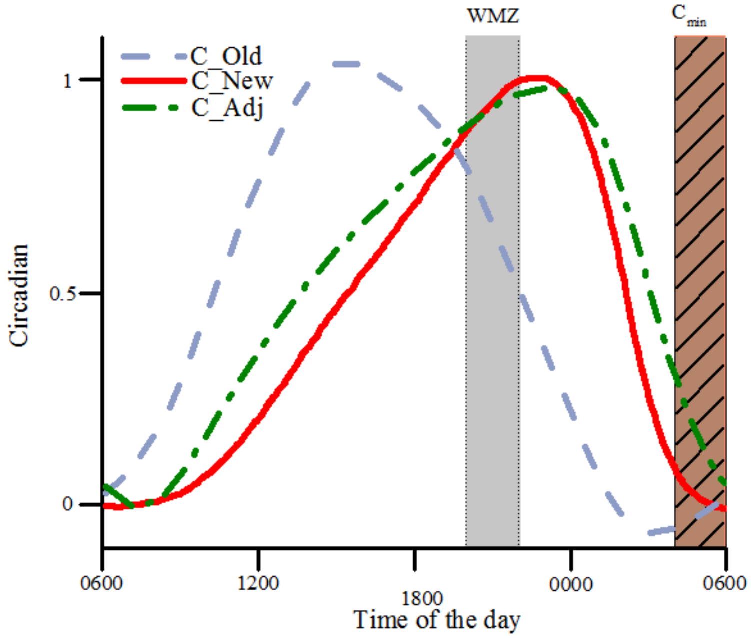


Figure 5.3: Improvement in the circadian oscillation.

Where

$$C_{old} = \frac{X+1}{2} \tag{5.14}$$

In earlier version, by using the following equation, the circadian variable maximizes at 1530 h under entrained condition. However, this contradict with the finding showing wake maintenance zone, a ∼3-h window of minimum sleep propensity, appear in the early evening around 2000 to 2200 h. C_{adj} was introduced to reflect the minimum sleep propensity occurred at 2220 h and minimum at 0540 h under entrain condition. Because of C_{adj} the sleep propensity grow steeply after WMZ . C_{adj} can be expressed as

$$C_{adj} = \frac{3.1X - 2.5Y + 4.2}{3.7\,(X+2)} \tag{5.15}$$

Appendix B. Dataset

The significance of a controlled laboratory experiment is beyond doubt to establish reliable and valid measures of the related constructs. The data that had been employed to develop the proposed system was retrieved from the controlled laboratory experiment carried out by Short *et al* [54]. This particular dataset was chosen mainly because the experiment mimicked the schedule applied in the maritime industry. The experiment looked into the impact of naps and varying shift systems implemented in the maritime industry on the aspects of neurobehavioural performance, sleep, and sleepiness.

5.4.1 Ethics Approval

Ethical approval for this study was given by the University of South Australia Human Research Ethics Committee. All subjects gave their written informed consent after they were given a comprehensive explanation regarding the nature, the probable consequences, and several privacy issues related to this study. All the subjects were given an honorarium for their participation.

5.4.2 Participants

A total of 24 healthy adults were selected to participate in the behavioural experiment (mean age ± standard deviation: 24.5 ± 4.2 years old). The subjects were recruited via public and online advertisements. All the subjects were physically sound, non-smokers, did not take drugs (urinary and blood screenings were performed prior to the study) or on medication, as well as free from medical, psychiatric, and sleep disorders. None of the subjects were habitual napper, consumed excessive alcohol and caffeine-containing beverages, nor engaged in shift work or travelled trans meridian within two months' preceding the study. In order to discard a considerable inter-individual variabilities in the circadian phase; extreme morning and

Figure 5.4: Example of activity monitor used for screening prior to laboratory studies. Figure reprinted from [90]

evening types based on the Scale of Morningness-Eveningness had been omitted. Additionally, only those who scored below five for the Pittsburgh's Sleep Quality Index participated in this study. The subjects maintained a bedtime routine with sleep onset between 22:00 and 24:00 h, as well as sleep offset between 07:00 and 09:00 h for a week as outpatients prior to the study. This sleep-wake activity was recorded and verified via sleep diaries and wrist actigraphy (see Figure 5.4). The subjects were prohibited from taking habitual nap, consuming alcohol, and performing intense exercise throughout the week.

5.4.3 Experimental Design

Each experimental session involved a group of eight subjects who were randomly assigned to either the 6h-on/6h-off: Early ($\text{shift}_{6/6_{\text{Early}}}$), or Late 6h-on/6h-off ($\text{shift}_{6/6_{\text{Late}}}$), or the 8h-on/8h-off ($\text{shift}_{8/8}$) split-sleep-schedules (see Figure 5.5). The subjects spent a total of nine days and eight nights in the sleep laboratory that was segregated into two baseline days, four days for one of the two types of split-sleep-schedules, one post-shift day (except for $\text{shift}_{6/6_{\text{Late}}}$), and two recovery nights. The working period for the $\text{shift}_{6/6_{\text{Late}}}$ shift was between 08:30 and 14:30 h, as well as between 20:30 and 02:30 h, which later commenced at 08:30 h on day three until day seven at 02:30 h.

Subjects selected for under shift$_{6/6_{\text{Early}}}$ condition carried out their simulated shift work that began from 08:30 to 13:30 h, and from 20:30 to 01:30 h on day two until day six. Meanwhile, subjects under the shift$_{8/8}$ condition began their shift work from 08:00 to 17:00 h on the second and fifth day, from 01:00 to 09:00 h on the third and sixth day, and from 16:00 to 00:00 h on the fourth and seventh day.

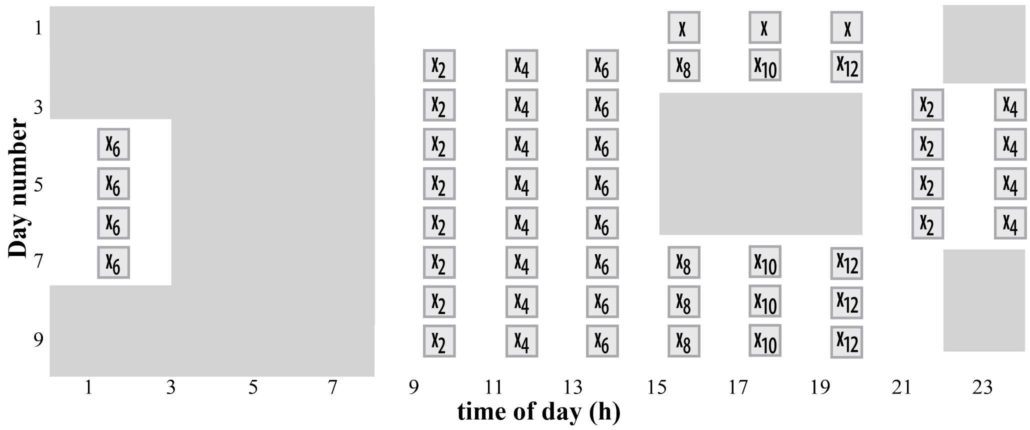
The experiment was conducted in a controlled environment of a sleep laboratory situated at the Centre for Sleep Research in the University of South Australia. The sleep laboratory environment was free of natural light, but had artificial lighting that was kept below 50 lux (to prevent direct alerting effect and circadian phase-shifting effect of brightness) during scheduled wakefulness and below 1 lux (darkness) during scheduled sleep. The temperature was set at 22 °C (±1 °C). Showering and calorie-controlled meals were provided to the subjects at scheduled times. When not engaged in assessment, the subjects spent their time in the laboratory socialising, reading, or watching movies. The subjects were under continuous human supervision in the laboratory to ascertain maintenance of wakefulness.

5.4.4 Measured Data

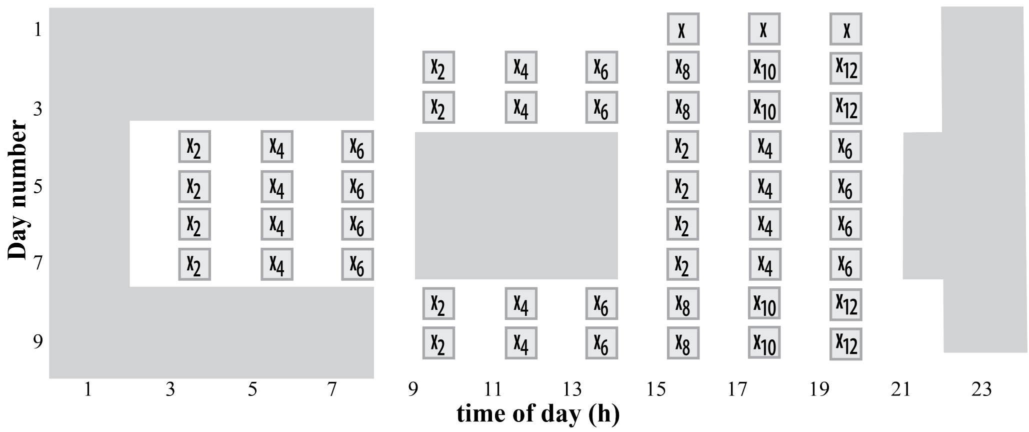
This section explains the collection of sleep measures, PVT and KSS.

5.4.4.1 Polysomnography

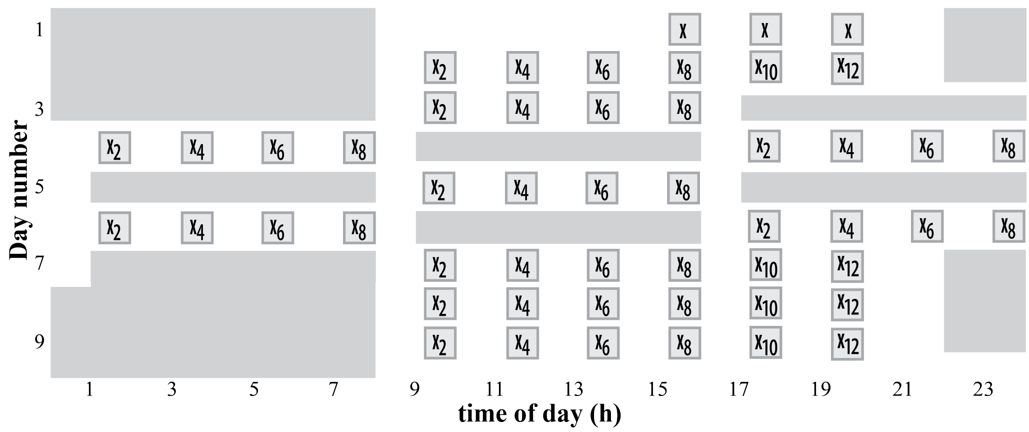
The PSG signals were recorded during sleep, particularly from sleep initiation to sleep termination, by using the Compumedics Grael Sleep System and Compumedics Profusion PSG 3 Software (Melbourne, Australia). The electrodes were placed in accordance to the international 10-20 system of electrode placement. The standard PSG montage was applied, inclusive of EEG (frontal: F_3, F_4; central: C_3, C_4; as well as occipital: O_1 and referenced against contralateral A_1/A_2), EOG, EMG, and ECG. The retrieved sleep data



(a) shift$_{6/6_{\text{Late}}}$



(b) shift$_{6/6_{\text{Late}}}$



(c) shift$_{8/8}$

Figure 5.5: Protocol schematics of the shift$_{6/6_{\text{Early}}}$ (a), shift$_{6/6_{\text{Late}}}$ (b) and shift$_{8/8}$ (c) schedules. Time (24 h) is on the horizontal axis, and the vertical axis displays the nine days of the study. The background shade of grey indicates the sleep times, whereas the scheduled shift periods are indicated by dotted boxes filled in grey X_i. Black crosses ($\mathbf{X}_i$) shows the timing of the PVT tests and subjective sleepiness assessment. Here, i indicate the hour at which the two activities were conducted, i =[2, 4, 6, 8, 10, 12] .

Table 5.1: Measures derived from PSG and definitions according to American Academy of Sleep Medicine guidelines.

Sleep variable	Definition
Time in bed (TIB)	Period that begins with light-out and end with light-on
Sleep stages	The summation of all epochs corresponding to each sleep stage throughout the TIB
Total sleep time (TST)	The total time spent in Stage N1, Stage N2, Stage N3, and REM sleep throughout the TIB
Sleep efficiency (SE)	The percentage of time in bed spent asleep and quantifying as a scalar value of TST divided by actual TIB

were analysed and scored visually in continuous 30 s epochs in adherence to the American Academy of Sleep Medicine by a senior medical scientist, who was blind to the experimental setting, by using commercial software (ProFusion PSG 3; Compume discs Limited).

Some sleep measures were derived from the PSG analysis, such as TIB, total sleep time, SE, and time spent at each sleep stage (i.e., N1, N2, N3, REM). The definition of each sleep measure is presented in Table 5.1.

5.4.4.2 *Psychomotor Vigilance Test*

A 10-min PVT was performed to evaluate the vigilant and sustained attention. Since PVT has the advantage of virtually no learning curve, several practice trials were carried out on the adaptation day (day 1) for each session. The PVT was performed on a hand-held purposely-built response box (see Figure 5.6) with an extended display to the computer screen (see Figure 5.7) using the parallel port of the computer. Within the 10-min PVT test, the subjects were asked to respond by pressing the button with their dominant thumb rapidly every time a visual stimulus appears on the computer screen. Each stimulus was presented at random intervals and was varied from two to ten seconds, in which the stimulus ceased once the button was pressed. Right away, the time taken to react to the display of stimulus on the screen was

Figure 5.6: Psychomotor vigilance task response box. Figure reprinted from [90]

recorded for a second as feedback. The outcome measure obtained from the 10-min PVT applied in this study referred to $RT_{10\,fast}$. This particular outcome measure was selected as it reflected the optimal alertness level [156], and it was not skewed by the occurrence of lapses.

5.4.4.3 Karolinska Sleepiness Scale

The subjective sleepiness of KSS was employed to assess subjective sleepiness. A range of studies have validated the correlation of the scale with polysomnographic, behavioural, and subjective indicators of sleepiness [178]. The subjects were requested to choose one of the nine statements, which ranged from one ('extremely alert') to nine ('very sleepy, great effort to stay awake, fighting sleep') that most accurately reflected their subjective sleepiness at that particular time (see Figure 5.8).

Figure 5.7: Computer-based psychomotor vigilance task (PVT). Figure reprinted from [90].

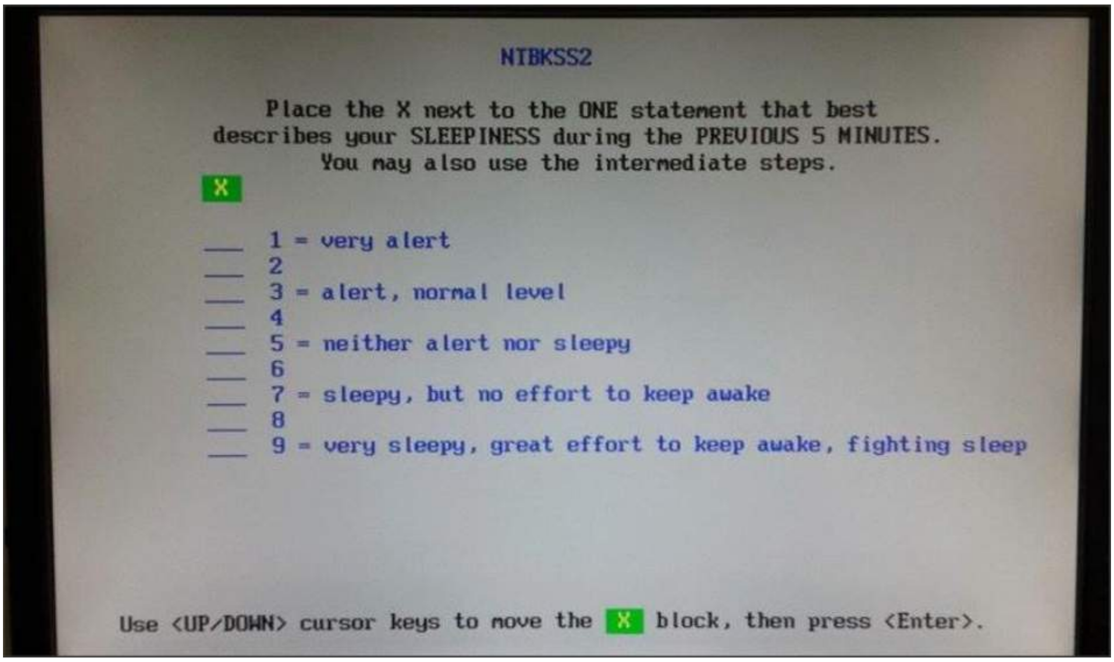


Figure 5.8: Computer-based Karolinska Sleepiness Scale. Figure reprinted from [90].

Appendix C: Proving Independence

Assume

$$P(X,Y,W,Z) = P(W)P(X)P(Y|W)P(Z|X,Y) \tag{5.16}$$

Proof the following

$$P(Z,W|X,Y) = \frac{P(X,Y,W,Z)}{P(X,Y)} \tag{5.17}$$

Rewrite P(X,Y) in the double-integral form using the concept of marginalization.

$$= \frac{P(W)P(X)P(Y|W)P(Z|X,Y)}{\int\int P(X,Y,W,Z)\,dW\,dZ} \tag{5.18}$$

$$= \frac{P(W)P(X)P(Y|W)P(Z|X,Y)}{\int\int P(W)P(X)P(Y|W)P(Z|X,Y)\,dW\,dZ} \tag{5.19}$$

Let's work on the denominator Using conditional probability's rule $P(Y|W)P(W) = P(Y,W)$, we get:

$$\int\int P(W)P(X)P(Y|W)P(Z|X,Y)\,dW\,dZ \tag{5.20}$$

- $\int\int P(X)P(Y,W)P(Z|X,Y)\,dW\,dZ$
- $P(X)\int\int P(Y,W)P(Z|X,Y)\,dW\,dZ$ Note[1]
- $P(X)\int(\int P(Y,W)\,dW)P(Z|X,Y)\,dZ$
- $P(X)\int P(Y)P(Z|X,Y)\,dZ$ Note[2]
- $P(X)P(Y)\int P(Z|X,Y)\,dZ$
- $P(X)P(Y)$ Note[3]

[1] Since P(X) is constant w.r.t both W and Z
[2] Since $\int P(Y,W)\,dW = P(Y)$ (by marginalizing W)
[3] since $\int P(Z|X,Y)\,dZ = 1$ (by definition of probability)

Thus,

$$= \frac{P(W)P(X)P(Y|W)P(Z|X,Y)}{P(X)P(Y)}$$

$$= \frac{P(W)P(Y|W)P(Z|X,Y)}{P(Y)}$$

$$= \frac{P(W)P(Y,W)P(Z|X,Y)}{P(W)P(Y)} \qquad \text{Note}^{4}$$

$$= \frac{P(Y,W)P(Z|X,Y)}{P(Y)}$$

$$= P(W|Y)P(Z|X,Y) \qquad \text{Note}^{5}$$

[4]from the relationship $P(Y|W) = \frac{P(Y,W)}{P(W)}$

[5]from the relationship $P(W|Y) = \frac{P(W,Y)}{P(Y)}$