1-1-2019

Physiological Approach To Characterize Drowsiness In Simulated Flight Operations During Window Of Circadian Low

Bijay Guragain

PHYSIOLOGICAL APPROACH TO CHARACTERIZE DROWSINESS IN SIMULATED FLIGHT OPERATIONS DURING WINDOW OF CIRCADIAN LOW

by

Bijay Guragain

Bachelor of Engineering, Anna University, 2014

A Thesis

Submitted to the Graduate Faculty

of the

University of North Dakota

in partial fulfillment of the requirements

for the degree of

Master of Science

Grand Forks, North Dakota

December

2019

This thesis, submitted by Bijay Guragain in partial fulfillment of the requirements for the Degree of Master of Science from the University of North Dakota, has been read by the Faculty Advisory Committee under whom the work has been done and is hereby approved.

Kouh

Kouhyar Tavakolian, Chairperson

Naima Kaabouch, Committee Member

Nick Wilson, Committee Member

Annie Tangpong

Digitally signed by Annie Tangpong
DN: cn=Annie Tangpong, o=North Dakota State University, ou=Mechanical Engineering, email=annie.tangpong@ndsu.edu, c=US
Date: 2019.12.03 11:47:03 -06'00'

Annie Tangpong, Committee Member

F. Richard Ferraro, Committee Member

Rashid Ahmed, Committee Member

This thesis is being submitted by the appointed advisory committee as having met all of the requirements of the School of Graduate Studies at the University of North Dakota and is hereby approved.

Dr. Chris Nelson, Associate Dean

School of Graduate Studies

12/10/19

Date

PERMISSION

Title	Physiological Approach to Characterize Drowsiness in Simulated Flight Operations during Window of Circadian Low
Department	Biomedical Engineering
Degree	Master of Science

In presenting this thesis in partial fulfillment of the requirements for a graduate degree from the University of North Dakota, I agree that the library of this University shall make it freely available for inspection. I further agree that permission for extensive copying for scholarly purposes may be granted by the professor who supervised my thesis work or, in this absence, by the Chairperson of the department or the dean of the School of Graduate Studies. It is understood that any copying or publication or other use of this dissertation or part thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of North Dakota in any scholarly use which may be made of any material in my thesis.

Bijay Guragain

November 2019

Abstract

Drowsiness is a psycho-physiological transition from awake towards falling sleep and its detection is crucial in aviation industries. It is a common cause for pilot's error due to unpredictable work hours, longer flight periods, circadian disruption, and insufficient sleep. The pilots' are prone towards higher level of drowsiness during window of circadian low (2:00 am- 6:00 am). Airplanes require complex operations and lack of alertness increases accidents. Aviation accidents are much disastrous and early drowsiness detection helps to reduce such accidents. This thesis studied physiological signals during drowsiness from 18 commercially-rated pilots in flight simulator. The major aim of the study was to observe the feasibility of physiological signals to predict drowsiness. In chapter 3, the spectral behavior of electroencephalogram (EEG) was studied via power spectral density and coherence. The delta power reduced and alpha power increased significantly ($p<0.05$) during microsleep (MS) whereas, alpha coherence was significantly high ($p<0.05$) during MS compared to alert state at many electrode interactions with the aviation headset. In chapter 4, heart rate variability was analyzed from electrocardiogram (ECG) signals and observed significant changes in both time and frequency domains during drowsiness. Moreover, photoplethysmogram (PPG) signal was found to be surrogate of ECG and suggested simplistic way to monitor pilot's physiological state via PPG sensor embedded within the aviation headset. Further, in chapter 5, the classification performance of MS was obtained as 86.23% in terms of F1 score with Quadratic Discriminant classifier and the performance for predicting drowsiness 4 sec before eye-closure was 83.80% using Random Forest classifier from electrodes within the aviation headset. In chapter 6, the overall performance in terms of unweighted mean of F1 score was 83.70% among four classes of drowsiness providing approach for early prediction. These results strongly corroborated the plausibility of automatic drowsiness detection in aviation from physiological signals.

Dedication

To my mom Saraswati and dad Agni,

I am indebted to their sweat and toil,

Lots of Love!

Acknowledgements

I would like to thank my advisor Dr. Kouhyar Tavakolian for accepting me as a graduate student and providing me the opportunity to perform research in his Biomedical Engineering Research Complex- BERC laboratory. The skills that I harnessed from him helped me to drive my study during two years and inspired me to succeed in future endeavors. I am very much thankful to him for investing his valuable time in me. I am thankful to Nicholas Wilson and Lewis Archer for providing the pilot dataset along with constructive feedback in my work. The efforts on data collection during late nights is highly praised and without their exertions, such valuable datasets would not have obtained. Special thanks to Dr. Ali Bahrami Rad for guiding me in Machine Learning. This helped me to develop a foundation and interest in the field of Artificial Intelligence. I am thankful to Dr. Ajay Verma and Dr. Chunwu Wang for initial description of the task along with dataset. I am thankful to my thesis committee members Dr. Naima Kaabouch, Nick Wilson, Dr. Annie Tangpong, Dr. F. Richard Ferraro, and Dr. Rashid Ahmed for their constructive feedback. Thanks to the department chair Dr. Ryan Adams, and graduate program director Dr. Naima Kaabouch for their key advice. Thanks to the School of Electrical Engineering & Computer Science- Biomedical Engineering Program for providing continuous funding. Thanks to Shubha Majumder and Abdiaziz Mohamud for assisting in data recording. Thanks to everyone who missed their sleeps during data acquisition time that jeopardized their day-time activities. Thanks to Department of Aviation for providing flight simulator and North Dakota Department of Commerce for financial assistance. Thanks to industrial partners Seeing Machines and Collins Aerospace. Thanks to my parents, sisters, and entire family who always inspired me to move ahead. Thanks to my wife Sabina for her unconditional love, care, and support. Last but not the least, thanks to my friends and colleagues from BERC lab and other labs for their usual support and care.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

- Acc: Accuracy
- ANS: Autonomic Nervous System
- BPF: Band Pass Filter
- BS: Baseline
- CV: Cross-validation
- DMS: Driver Monitoring System
- ECG: Electrocardiogram
- EEG: Electroencephalogram
- EOG: Electrooculogram
- FAA: Federal Aviation Administration
- HACM: Human Annotation Classification Method
- HF: High Frequency Power
- HR: Heart Rate
- HRV: Heart Rate Variability
- ICA: Independent Component Analysis
- ICAO: International Civil Aviation Organization
- KSS: Karolinska Sleepiness Scale
- LDA: Linear Discriminant Analysis
- LEC: Long Eyelid Closure
- LF: Low Frequency Power
- LS: Long Sleep
- MS: Microsleep

- NASA: National Aeronautics and Space Administration
- NTSB: National Transportation Safety Board
- n.u.: Normalized Unit
- PERCLOS: Percentage of Eyelid Closure
- PPG: Photoplethysmogram
- PPI: Pulse-to-Pulse Interval
- PPV: Positive Predictive Value
- PRV: Pulse Rate Variability
- PSD: Power Spectral Density
- PSG: Polysomnography
- PSN: Parasympathetic Activity
- PVT: Psychomotor Vigilance Test
- QDA: Quadratic Discriminant Analysis
- RF: Random Forest
- RRI: R-to-R Interval
- SE: Standard Error
- Sen: Sensitivity
- SM: Seeing Machines
- SN: Sympathetic Activity
- TP: Total Power
- WOCL: Window of Circadian Low

Chapter 1.

Introduction

1.1. Motivation

Every individual requires sleep and it's a daily routine. Sleep is a transitional process with changes in psycho-physiological behavior. Failure to maintain proper sleep makes an individual drowsy and fatigued. Drowsiness is common during monotonous task and is identified as a contributing factor for both road and airline accidents. The 4-7% of civil aviation mishaps are caused due to drowsy pilots and the cost of a single major civil aviation accident exceeds $500 million financial losses, while costs in terms of personal sufferings are inestimable [1]. Aviation accidents are disastrous and early identification of drowsiness will reduce such accidents.

Drowsiness is characterized by behavioral signs such as nodding head, facial expression, and yawning [2]. From such behaviors, drowsiness can be predicted only after falling towards sleep which is too late to prevent an accident. However, physiological signals change at the earlier stage of drowsiness [3], [4]. Therefore, an efficient reliable system to characterize drowsiness based on physiological changes among pilots is essential. To address this concern, this thesis proposes to study drowsiness based on the analysis of three physiological signals: electroencephalogram (EEG), electrocardiogram (ECG), and photoplethysmogram (PPG) during the window of circadian low (WOCL). During this time of 2:00 am- 6: 00 am, the participants are highly prone towards sleep [5]. Upon quantification of above-mentioned three signals, drowsiness can be predicted before eye-closure that highlights the potential of automatic drowsiness detection system.

1.2. Literature Review

1.2.1. Need for Identifying Drowsiness in Aviation

Airplanes require complex human interactions, complete attention, and proficient skills to fly. The professional pilots performing such critical functions are impaired by drowsiness or fatigue during long-haul flights due to circadian disruption [6]. Drowsiness is a propensity to fall asleep. It is a transition from awake to onset of sleep [7]. It is a leading cause of aviation accidents mainly due to unpredictable work hours, long duty periods, circadian disruption, and insufficient sleep [8]. For example, the scheduled flight time of the longest non-stop flight from Singapore to New York is approximately 17 hours. Such long-haul flights are painful and pilots in the cockpit may suffer from fatigue due to disruption in circadian rhythm due to which they may undergo multiple drowsy states. Flying in a drowsy state can lead to substandard performance, spatial disorientation, and even serious aircraft accidents [9]. Aviation accidents are disastrous and early identification of drowsiness will reduce such accidents.

Drowsiness is marked by several factors such as increase in reaction time, reduction in performance, changes in brain rhythm, frequency of eye blinking, changes in respiration and heart rate [10]. Federal Aviation Administration (FAA) classified pilot fatigue or drowsiness as one of the four common aviation hazards [11]. National Aeronautics and Space Administration (NASA) flight controller fatigue assessment report showed that 70% of survey respondents working midnight shifts caught themselves “about to doze off” while working actively, and 18% reported being involved in an operational error among which about 56% cited fatigue as a contributing factor [12]. Another report by National Transportation Safety Board (NTSB) investigated from January 1, 2001 to December 31, 2012 identified fatigue as a contributing factor of accidents. In

2013, a business plane crashed short of runway at Birmingham, Alabama, and the captain and the first officer were fatally injured. The findings indicated they were fatigued due to improper off-duty time management [6]. According to the National Sleep Foundation, about 20% of U.S. adult drivers admit to fall sleep behind the wheel during driving [13]. Further, the 20% of medium-haul and 40% of long-haul flights are affected by drowsiness [14]. Drowsiness is involved in 4-7% of civil aviation mishaps and the cost of a single major civil aviation accident can be millions, while costs in terms of personal sufferings are inestimable [1]. Flight operations require constant vigilance and even a minor mistake due to drowsiness can cause high devastation, thereby requiring a need for early detection. Therefore, the development of supplemental technology to accurately identify transitional state from alertness towards drowsiness can be of great importance in the aviation industry.

1.2.2. Approaches for Drowsiness Detection

There are mainly three approaches for drowsiness detection: vehicular, behavioral, and physiological [15]–[17]. The vehicular approach detects drowsiness based on the vehicle speed and position. Whereas, the behavioral approach monitors various signs such as eye and head movements, facial expressions, eye closure, and yawning based on vision [2]. This approach has been commonly used for the development of drowsiness detection technologies in the market to alert operators. Some of the vision based technologies are *CoPilot*® (Attention Technologies, Pittsburgh, PA), *Optalert*® (Optalert Pty Ltd, Australia), and *Driver State Monitor*® (Delphi Electronics & Safety, Inc., Kokomo, IN) that provide drowsiness feedback based on percentage of eyelid closure (PERCLOS). They use infrared illumination to identify PERCLOS [18]–[20]. Both these approaches are system-dependent that have different evaluation criteria and restricted to certain environmental conditions such as proper lighting and inclusion of objects within their

focus. They detect only after falling towards sleep which is too late to prevent an accident. However, physiological signals change at the earlier stage of drowsiness [3], [4]. Thus, there needs an efficient reliable technology to characterize drowsiness in real time, before falling sleep, based on physiological performance of the pilots.

1.2.3. Subjective Measures for Analyzing Drowsiness

Along with the physiological performances as a primary measure to analyze drowsiness, subjective measures such as Karolinska Sleepiness Scale (KSS) and Psychomotor Vigilance Task (PVT) based on participant's feelings before and after the experiment are also used to assess drowsiness levels [21]–[23]. KSS is a subjective measure of sleepiness that indicates level of psychological state experienced by a participant in last 10 min. It helps to assess changes in the state due to environmental factors, circadian rhythm, and effect of drugs. This is a self-report measure based on 9-point scale (1=extremely alert, 3=alert, 5=neither alert nor sleepy, 7=sleepy, and 9=extremely sleepy) [24]. PVT is a measure of behavioral alertness where sustained or vigilant attention is obtained by recording response times to visual stimuli at random inter-stimulus intervals. The most reliable effect of sleep deprivation is degradation of attention, especially vigilant attention which is measured by 10 min PVT quantifying the number of lapses in attention of the tested participant [25]. However, these methods have limited usability to quantify drowsiness. They are not suitable for continuous sleepiness evaluation and self-assessment of drowsiness is often wrong [2], [26]–[29]. Further, drowsiness is a transitional recurring state and its estimation from the subjective analysis is not applicable in the instrumentation design for real time detection purpose [30]–[32]. To address this limitation, there is a need of physiological approach for drowsiness analysis.

1.2.4. Physiological Approach for Drowsiness Detection

Various physiological signals such as electroencephalogram (EEG), electrocardiogram (ECG), electrooculogram (EOG), and photoplethysmogram (PPG) have been studied for drowsiness detection. The changes during drowsiness are directly reflected in these physiological signals that are used to estimate drowsiness.

EEG is a record of electrical signals generated by brain cells and physiological changes in the brain due to drowsiness are directly reflected from EEG due to which it is commonly considered as a drowsiness marker [2], [17], [33], [34]. EEG is non-invasive, easy to acquire from limited electrodes and drowsiness effects are vividly reflected in the spectral behavior [4], [35], [36]. The following rhythms have been identified in EEG: delta (0.5-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), beta (13-30 Hz), and gamma (above 30 Hz) [37]. The slow waves (delta and theta) increases during transition from awake to drowsiness and deep sleep. The alpha rhythm is a marker of relaxed wakefulness and increases when a participant prolongs towards drowsiness, whereas decreases during deep sleep. The beta rhythm is associated with alertness, arousal, and excitement [17]. Hence, EEG behavior has been extensively studied during simulated and real driving scenarios to gain inference pertaining to drowsiness [38]–[40].

ECG is a record of electrical activity of heart and can be a potential marker to predict drowsiness or fatigue based on heart rate variability (HRV) [41]–[45]. HRV is a marker of autonomic nervous system (ANS) that influences modulations of cardiac activities such that homeostasis of cardiovascular system is properly maintained [46]. ANS activity changes during stressed, fatigued, and drowsy periods. Wakefulness is marked by increase in sympathetic activity (SN) and/or decrease in parasympathetic activity (PSN), whereas drowsiness or fatigue is marked by increase in PSN or decrease in SN [44], [47], [48]. Thus, ECG can be utilized to detect drowsiness or fatigue in pilots. ANS can be measured noninvasively by HRV from surface ECG. Various time and

frequency domain methods have been applied to analyze HRV from RR intervals (RRI) which is the duration of two adjacent R peaks in QRS complexes of ECG [49]. The insight towards drowsiness can be obtained from the changes in these time and frequency domain features from ECG.

PPG is another noninvasive signal obtained via optical technique containing an infrared emitter and a detector inside a probe that reflects beat to beat relative blood volume changes in peripheral tissues usually on the forefinger or earlobe [50]. The amount of infrared light reflected back to the detector determines the amount of blood flowing to the tissue at any time. So, the pulsating PPG signal depicts blood volume fluctuations observed from a finger or earlobe. The mean time delay between ECG R-wave and PPG foot at rest typically ranges from 180 to 260 ms (finger), 125 to 155 ms (ear), and 180 to 330 ms (toe) [51]. However, the median time delay between ECG peak and PPG peak, that is, pulse arrival time (PAT) peak at rest is about 436 ms (thumb), 397 ms (ear), and 515 ms (toe) [52]. In literature, HRV features obtained from ECG and PPG are found to have high correlation [50], [53]–[55]. Thus, PPG may be used as a substitute of ECG for HRV analysis.

EOG is a corneo-retinal standing potential between front and back of the eye which measures eye movements to detect drowsiness. The blink rate is slower as well as eye closure duration is longer during drowsiness [56]–[58]. The two main features extracted from EOG are blink duration and PERCLOS. The blink duration is the time duration between two eye blinks and PERCLOS is the percentage of time that the eyelids are closed more than 80% [59]. These features of ocular behavior can be used to analyze drowsiness.

This thesis investigates the potential of three physiological signals: EEG, ECG, and PPG to characterize drowsiness. The major portion of the work is carried on analyzing EEG and EOG is used to further validate the eye closure events during drowsy state of pilots in a flying simulator

during window of circadian low (WOCL). This thesis underscores the potential of non-invasive approach for drowsiness detection from physiological signals and also proposes the idea of modifying the signal acquisition system to the region of aviation headset.

1.3. Thesis Outline

This thesis is outlined as follows. Chapter 2 describes the methods used for data analysis. Chapter 3 illustrates the spectral behavior of EEG to analyze drowsiness. This chapter also discusses the interactions of different EEG electrodes within the region of aviation headset to characterize drowsiness. Chapter 4 discusses the approach of heart rate variability to estimate drowsiness and fatigue. This chapter also compares ECG and PPG and demonstrates the feasibility of estimating drowsiness from PPG alone using aviation headset. Chapter 5 explains the classification and prediction model designed for drowsiness analysis. The issues and challenges of automatic drowsiness detection using standalone EEG signal is also discussed in this chapter. Chapter 6 further extends to four-way classification of drowsiness and proposes an idea for early drowsiness prediction. Chapter 7 concludes the paper identifying limitations and proposing future work to address the limitations.

1.4. Thesis Contribution

This thesis illustrated the potential of non-invasive physiological signals to characterize drowsiness in aviation operations. The key contributions are as follows.

i. Depicted the spectral changes in different brain regions during drowsiness (Chapter 3).
ii. Illustrated the significant interactions between different EEG electrodes within the region of aviation headset (chapter 3).

iii. Calculated HRV to estimate drowsiness and fatigue (chapter 4).

iv. Demonstrated the plausibility of estimating drowsiness using PPG obtaining its surrogacy with ECG (chapter 4).

v. Designed a machine learning model to classify drowsiness from alert state using EEG features (chapter 5).

vi. Discussed the potential of automatic drowsiness annotation using standalone EEG signal (chapter 5).

vii. Classified drowsiness among four classes and depicted the plausibility for early drowsiness prediction (chapter 6).

viii. Discussed spectral behavior among four drowsiness classes (chapter 6).

ix. Highlighted the limitations of the work proposing ideas to address the limitations (chapter 7).

1.5. Publications

Some of the publications of this work along with the related other works are as follows.

Journal Articles:

1) Chunwu Wang*, **Bijay Guragain***, Ajay K. Verma, Lewis Archer, Shubha Majumder, Abdiaziz Mohamud, Emily Flasherty-Woods, Geoffrey Shapiro, Mahathir Almashor, Mike Lenne, Rama Myers, Jonny Kuo, Shiyan Yang, Nicholas Wilson†, and Kouhyar Tavakolian†, "Spectral Analysis of EEG during Microsleep Events Annotated via Driver Monitoring System to Characterize Drowsiness", IEEE Transactions on Aerospace and Electronic Systems, 2019 (* and † contributed equally).

2) Nicholas D. Wilson, **Bijay Guragain**, Ajay Verma, Lewis Archer, and Kouhyar Tavakolian, "Blending Human and Machine: Feasibility of Measuring Fatigue through the Aviation Headset", The Journal of Human Factors and Ergonomics Society, 2019.
3) **Bijay Guragain**, Ali Haider, and Reza Fazel-Rezai, "Hybrid Brain-Computer Interface Systems: Approaches, Features, and Trends", Evolving BCI Therapy- Engaging Brain State Dynamics, IntechOpen, 2018.

Conference Proceedings:

1) **Bijay Guragain**, Ali Bahrami Rad, Chunwu Wang, Ajay K. Verma, Lewis Archer, Nicholas Wilson, and Kouhyar Tavakolian, "EEG-based Classification of Microsleep by Means of Feature Selection: An Application in Aviation", Engineering in Medicine and Biology Conference (EMBC), Berlin, 2019.
2) **Bijay Guragain**, Ali Haider, and Reza Fazel-Rezai, "A New Region Based SSVEP BCI Speller", Seventh International BCI Meeting, May 21-25, Pacific Grove, California, 2018.

Chapter 2.

Methodology

2.1. Power Spectral Analysis of EEG

EEG has been considered as a fundamental tool in neurophysiology for humans since its discovery in 1924 by German Psychiatrist Hans Berger. The changes in neural activities can be used to study physiology of humans via analysis of EEG rhythms [60], [61]. The common four EEG rhythms used for analysis are: delta (0.5- 4 Hz), theta (4-8 Hz), alpha (8-13 Hz), and beta (13-30 Hz). The characteristics of EEG depends on demographics, behavioral, and psychophysiological state of a person. The EEG pattern is influenced by neuro-pathological conditions, metabolic disorders, and drug actions [62].

EEG is generally recorded non-invasively from scalp electrodes using common International 10-20 system. This is a standard testing method for EEG based on the relationship between location of an electrode and underlying area of cerebral cortex. The 10 and 20 refers the actual distance between adjacent electrodes are either 10% or 20% of the total front-back or right-left distance of the skull. This protocol standardized physical placements and designations of 21 electrodes on the scalp. The name of each electrode has a letter and a number. The letter refers to brain region (F: frontal, C: central, T: temporal, P: parietal, and O: occipital), and the number is related to cerebral hemisphere (even in right and odd in left hemisphere) [63]. The EEG montage for recording is depicted in Figure 2.1. EEG can be used for various physiological monitoring.

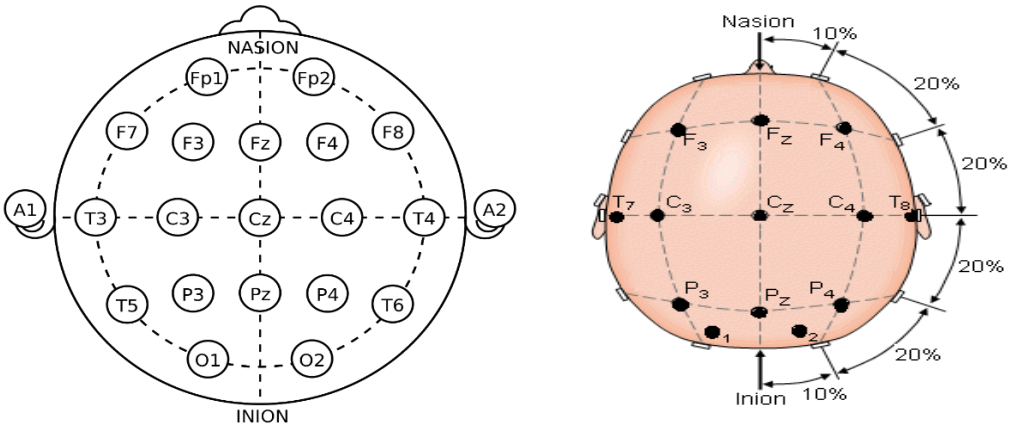


Figure 2.1. The 10-20 International electroencephalogram (EEG) electrode placement system along with EEG montage [37].

The power spectral density (PSD) is a measure of signal's power with respect to frequency. The PSD from EEG is usually calculated within the frequency range of 0.5- 30 Hz that contains useful information. The major four frequency bands in EEG as mentioned before are: delta (0.5- 4 Hz), theta (4-8 Hz), alpha (8-13 Hz), and beta (13-30 Hz). The four EEG rhythms are briefly described here [37].

1) *Delta activity:* A low frequency and high amplitude slow delta waves are characteristic EEG feature recorded during deep sleep.
2) *Theta activity:* The theta wave is active during emotional or some cognitive states.
3) *Alpha activity:* The alpha wave is predominant during wakefulness and mostly observed during eyes closed and relaxation state.
4) *Beta activity:* This wave is a characteristic EEG for the states of increased alertness and focused attention.

EEGLAB toolbox (Ver. 14.1.1b) is used to process the raw EEG signals [64]. The EEG signals of 500 Hz sampling frequency from all 13 channels are resampled to 125 Hz. Then, high amplitude

artifacts due to body movements or muscle activities are removed through visual inspection and confirmed by two researchers, and 0.5-30 Hz band pass FIR filter (BPF) is used to remove potential high or low frequency artifacts. After that, INFOMAX algorithm for Independent Component Analysis (ICA) is applied to decompose the filtered data, and ADJUST algorithm is used to automatically detect and remove ocular artifacts [39], [65]. Next, PSD is calculated using Welch's power spectral density estimate having Hamming window size of 128 samples with 50% overlap and 256 discrete Fourier transform (DFT) points. This is obtained using command 'pwelch' in MATLAB (Version R2017A, Mathworks Inc., MA, USA) with appropriate parameters. After finding the absolute power at respective frequency bands, the relative power is obtained by dividing each band's absolute power with the total power. The processing steps for calculation of EEG relative PSD is shown in Figure 2.2.

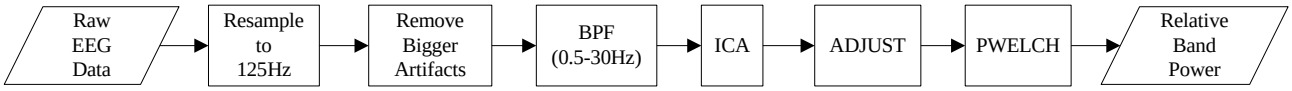


Figure 2.2. Step-by-step signal processing block diagram to compute relative spectral power distributed in the respective EEG frequency bands [66].

The median frequency (MDF) is defined as a frequency at which the EEG power spectrum is divided into two regions of equal power. MDF is also defined as the half of the total power and is given as follows [67], [68].

$$\sum_{j=1}^{MDF} P_j = \sum_{j=MDF}^{M} P_j = \frac{1}{2}\sum_{j=1}^{M} P_j \tag{2.1}$$

where, MDF is the median frequency, M is the total frequency and P_j is the j^{th} power that sums to total power over entire M.

2.2. Magnitude Squared and Wavelet Transform Coherence of EEG

EEG coherence is a measure of functional connectivity between brain regions recorded from spatially separated scalp electrodes to estimate the similarity of waveform components generated by the mass action of neurons in underlying cortical regions [69], [70]**.** Functional connectivity is an undirected relationship and a measure of correlations between two or more nodes in frequency domain can be quantified through coherence [71]. The two coherence methods are briefly described. Magnitude Squared Coherence (MSC) is used to obtain correlations between two different signals x and y in frequency domain. It is the ratio of cross-spectrum of x and y to the product of auto-spectrum of x and y and is calculated from the expression as follows [70].

$$C_{xy}(f) = \frac{|P_{xy}(f)|^2}{|P_{xx}(f)|.|P_{yy}(f)|} \tag{2.2}$$

where, $C_{xy}(f)$ is MSC, $P_{xy}(f)$ is cross-spectral power of x and y, $P_{xx}(f)$ is auto-spectral power of x, and $P_{yy}(f)$ is auto-spectral power of y. MSC is computed with a Hamming window size of 256 samples at 50% overlap having 256 DFT points.

Neural signals are non-linear and non-stationary. Wavelet Transform Coherence (WTC) can detect short episodes of coupling between neural signals measuring coherence as a function of time and frequency [72]. The time-frequency coherence analysis from WTC solves the problem of non-stationarity than the classical Fourier based MSC which is only a function of frequency [73]. In WTC, x and y signals are correlated with Morlet wavelet function. The function consists of product of a sinusoidal wave at frequency f and a Gaussian function centered at time τ with standard deviation σ and is given as [74]:

$$\psi_{\tau,f}(u) = \sqrt{f}e^{i2\Pi f(u-\tau)}e^{\frac{-(u-\tau)^2}{\sigma^2}} \tag{2.3}$$

WTC is computed using the Morlet wavelt with 12 equal number of voices per octave and scales to smooth. MSC and WTC are compared in four frequency spectra: delta (δ, 0.5-4Hz), theta (θ, 4-8Hz), alpha (α, 8-13Hz), and beta (β, 13-30Hz) and the mean coherence is used for statistical comparison.

2.3. Heart and Pulse Rate Variability

HRV is a measure of time variation between each heartbeat in ECG signal. The changes in heartbeats are controlled by ANS. HRV is a marker of homeostasis and analyzed in time and frequency domains [75]. The results from the various studies depicted that PPG is a surrogate of ECG and an analogous approach of HRV in PPG signal is Pulse Rate Variability (PRV) [76]–[79]. HRV from ECG or PRV from PPG first relies on detection of the 'R' peaks in 'QRS' complex of ECG. That means the PPG analysis is referenced on ECG. The normal ECG waveform is shown in Figure 2.3.

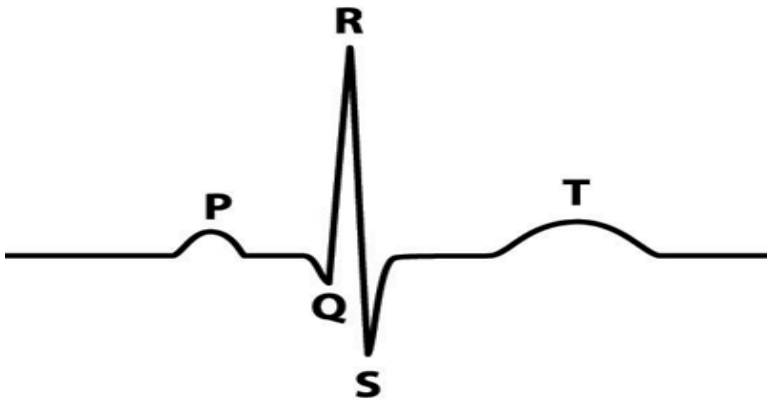


Figure 2.3. The normal electrocardiogram (ECG) signal depicting P, Q, R, S, and T (source: EKG Academy).

The ECG 'R' peaks are generally detected from widely used Pan-Tompkins algorithm [49]. After detecting 'R' peaks, the locations of 'R' peaks are obtained and successive RR intervals (RRI) are

calculated. Heart rate (HR) can be calculated from RRI, and mean HR is obtained as the reciprocal of mean RRI within a certain duration and gives as follows.

$$Avg.HR = \frac{60}{Avg.RRI\ (sec)}\ beats/\ min \tag{2.4}$$

Similar to ECG, PPG has also 'P' peaks after a duration of PAT and HR can also be calculated from pulse-to-pulse intervals (PPI). The detection of 'R' and 'P' peaks in ECG and PPG as well as RRI and PPI waveforms is illustrated in Figure 2.4.

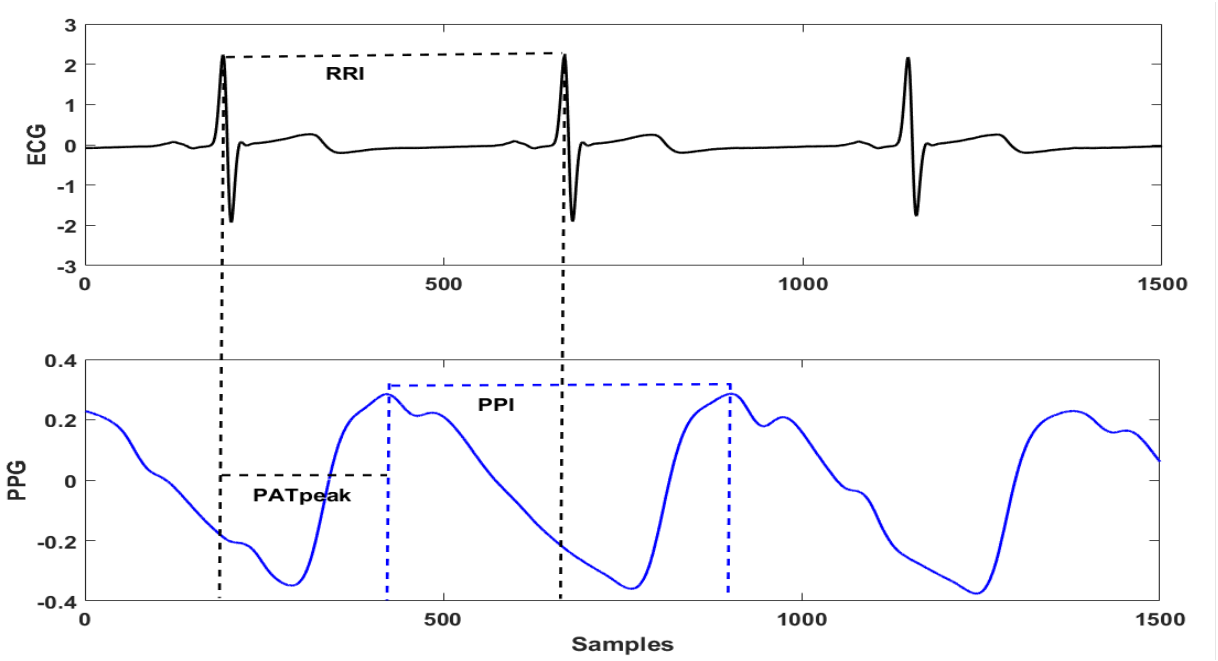


Figure 2.4. R-to-R interval (RRI), pulse-to-pulse interval (PPI), and pulse arrival time (PAT) peak from simultaneously recorded electrocardiogram (ECG) and photoplethysmogram (PPG) for three heart beats from a representative participant [80]. The different time and frequency domain features extracted from ECG/PPG are described in appendix A.

2.4. Machine Learning

Machine learning is one of the most useful statistical method to perform a specific task when the structure of the task is not well understood but can be characterized by dataset with strong statistical regularity. It has numerous applications such as search and recommendation, automatic speaker identification, face identification, objects tracking, medical diagnosis, and drug delivery. It is divided into supervised and unsupervised learning. In supervised learning, the data has labels whereas the data has no labels in unsupervised learning. The problem of predicting a discrete random variable y from another random variable x is called classification. Consider data $(x_1, y_1), (x_2, y_2) \dots \dots \dots . (x_n, y_n)$ where

$$x_i = \begin{bmatrix} x_{i1} \\ . \\ . \\ . \\ x_{id} \end{bmatrix} \epsilon \ X \ \mathrm{c} \ \mathbb{R}^d$$

is a d-dimensional vector and y_i takes values in some finite set Y. A classification rule is a function $h: X \to Y$ and y is predicted as $h(x)$ from a new x [81].

2.4.1. Linear and Quadratic Discriminant Analysis

The Bayes classifier is based on the conditional probability rule to perform classification. The conditional probability of an event B is the probability that the event will occur given the knowledge that an event A has already occurred. This probability is written as P(B|A). Bayes classifier is a fundamental statistical approach towards the problem of classification. Let us consider two classes y={0, 1}. Then, the probability distribution of an unknown quantity in class 1 can be obtained as:

$$\mathrm{P}(y = 1 | x = x_0) = \frac{P(x = x_0 | y = 1)\, P(y=1)}{P(x=x_0)} \qquad (2.5)$$

where, $P(y = 1|x = x_0)$ is called posterior probability distribution of an unknown quantity x_0, $P(x = x_0|y = 1)$ is called class conditional density or likelihood, and $P(x = x_0)$ is called marginal. The decision boundary is a set of points such that the probability of being in class 1 is same as the probability of being in class 0. These points are used to set a boundary such that any test points can be classified. If a test point lies to one specific side of a boundary, then it belongs to that class. The linear discriminant analysis (LDA) and quadratic discriminant analysis (QDA) are two classifiers based on Bayes approach.

In LDA, we assume class conditional as a multivariate Gaussian distribution and both the classes share the same co-variance matrix. The decision boundary is linear and any point in this line has an equal probability of lying in class 0 or 1. In real, the distribution of data and co-variance matrix of each class may be different, however, we assume to have Gaussian distribution and same class co-variance in case of LDA. If we relax on the second assumption of LDA and consider the co-variance matrix of both classes to be different, then shape of the decision boundary will change to quadratic that makes QDA classifier. That means, in QDA, we assume multivariate Gaussian distribution with unequal class co-variance matrices [81].

2.4.2. Random Forest Classifier

A random forest (RF) classifier is an ensemble of decision trees. An ensemble is a group or combination of multiple machine learning algorithms to obtain better predictive performance than any of the individual decision trees. During training, multiple decision trees are constructed. These individual decision trees turn out to be different because only a subset of features are given to each of them, using a technique called bagging. During prediction, mode of the classes of individual decision trees constitute the output of the forest. Bootstrap aggregation or bagging averages

prediction over a collection of bootstrap samples, thereby reducing its variance. Each bootstrap tree will involve different features than the original and might have different number of terminal nodes. Thus, the problem of low bias and high variance causing instability of decision trees is reduced due to bagging. The bagged estimate is the modal prediction from these trees. The random forest is a substantial modification of bagging that builds a large collection of de-correlated trees. The depth of tree i.e. leaf node size and feature subset size are the two important parameters for tuning the RF classifier. For classification tasks using RF classifier, the default feature subset size is the square root of total feature size and minimum leaf size is 1. The binary splitting is preferred in RF classifier at each node because the problem with multiway split is that it fragments the data too quickly, leaving insufficient data at the next level down [81].

2.5. Nested Cross-Validation

Cross-validation (CV) is a model validation technique to assess the results of a statistical model to an independent dataset. It is a popular strategy for model selection. The idea behind CV is to split the data into training and test sets. The training dataset is used to design a predictive machine learning model and the designed model is validated on the test datasets, also called validation datasets. CV helps to select the algorithm with the smallest error avoiding over and under fitting [82]. The CV are beneficial in use of all available data considering less availability of medical data, reliable performance with random data selection for validation and parameters tuning to select the best parameters. In this work, a nested CV is used for features selection as well as model assessment. The word nesting simply means placing one inside the other. That means, inner CV is used for feature selection and the outer CV is used for model evaluation. The architecture of nested CV is depicted in Figure 2.5.

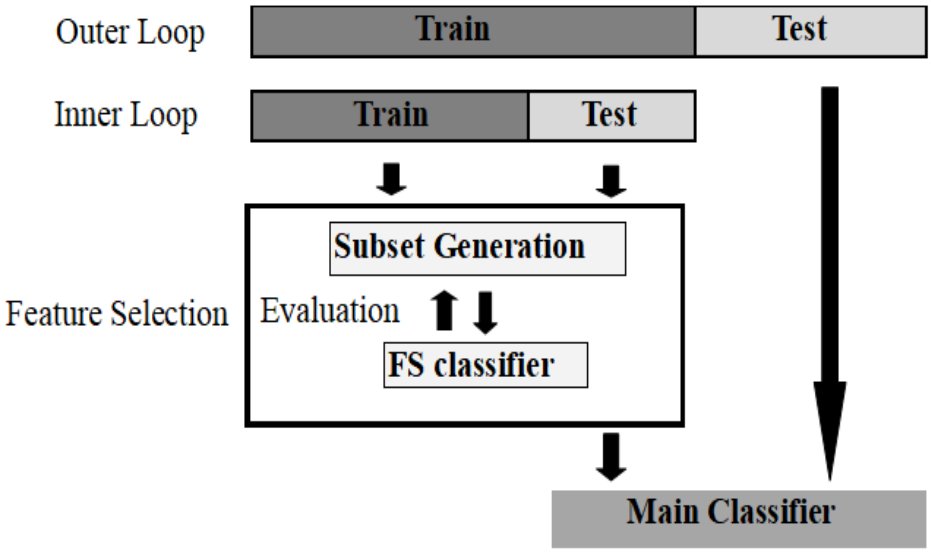


Figure 2.5. Nested cross-validation (CV) architecture for feature selection and model assessment [83]. The features are selected using feature selection (FS) classifier in the inner loop and the classifier model is evaluated in the outer loop.

In the first step, the whole dataset is divided into training and testing using leave-one-participant-out CV. That means the entire samples from a particular participant is left for testing in the outer loop. Then, the training samples are again divided randomly into further training and test samples using 5-fold CV in the inner loop. After that, feature subsets are generated from feature selection classifier in the inner loop using a feature selection procedure. The minimal error can be used as a feature selection criterion. Finally, the selected features are used to assess the performance of the classifier at the outer loop in leave-one-participant-out procedure. Nested CV ensures that the parameter selection and model validation are independent, thus preventing the over fit of tuning the parameters to the train set [84].

2.6. Data Acquisition and Processing

2.6.1. Experimental Protocol

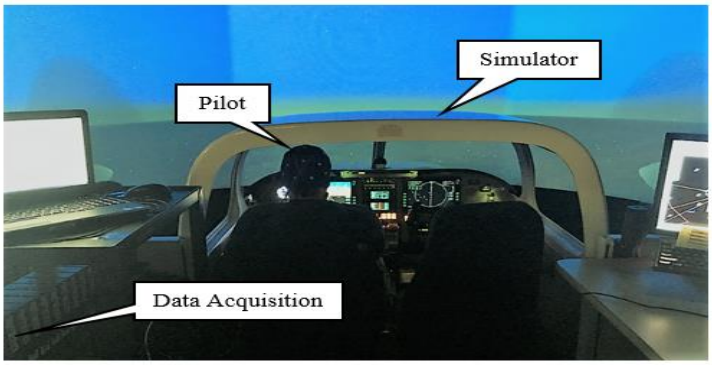
The simulated flight scenario and data collection took place in a Piper Seminole fixed-base aviation training simulator at John D. Odegard School of Aerospace Sciences, University of North Dakota (UND). FAA certified commercially rated pilots participated in data acquisition. Participants were asked to report for a routine flight scenario which would take place during the window of circadian low (2:00 am- 6:00 am) defined by the FAA. Each participants were requested to avoid naps, sleep, and caffeine or similar energy beverages from the time they woke up the previous day in advance of showing for the study. All participants were requested to remain awake for 18-20 hours before starting the experiment.

Each participant's data acquisition process was divided into two phases: a takeoff phase of approximately six minutes and a cruise phase using autopilot of about 2 hours. From the design of two phases of flight, no participants reached their destination that facilitated consistent low-workload cruise conditions for data acquisition. This was done so to maintain constant/ monotonous work environment during the second phase of flight to observe drowsiness. If participants had to fly till destination by modifying the experimental protocol, then they would have different workload during landing phase near to the destination. The data recording sessions were concluded upon meeting one of the three conditions: (1) the session reached 2.5 hours of flight time, (2) the participant had been awoken twice from a period of sleep more than five minutes or (3) the participant elected to discontinue the session. This experiment was approved by the Institutional Review Board at UND.

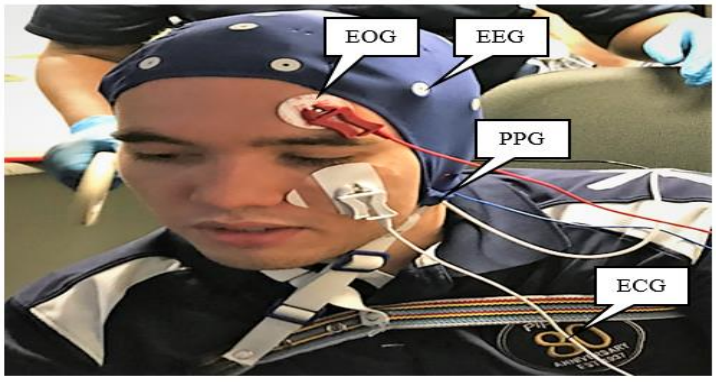
2.6.2. Data Recording

The four major recorded signals are EEG, ECG, PPG, and EOG. EEG was recorded using 13 electrodes from five brain regions: frontal (F3, Fz, F4), central (C3, Cz, C4), parietal (P3, Pz, P4),

temporal (T7, T8), and occipital (O1, O2) using international 10-20 EEG electrode placement system, and the unilateral ear of the participant was used as a reference in that hemisphere. ECG was recorded from a single channel using Lead II Einthoven's Triangle and PPG was recorded from the earlobe. All the signals were recorded at a sampling rate of 500Hz using BIOPAC data acquisition system (MP160, BIOPAC Systems, CA). The facial information were simultaneously acquired from the DMS camera sensor, which included an infrared light and an infrared camera. The camera was glare shield-mounted and rearward facing towards the participant. Vertical EOG signal was also recorded to verify the DMS MS events. In addition, the drowsiness level through the KSS was also self-reported by each participant before and after the experiment. The overall experimental setup for data recording is shown in Figure 2.6.



(a)



(b)

Figure 2.6. The experimental setup for data acquisition using Piper Seminole fixed-base aviation training simulator is shown on the top (a) and a pilot participating in the experiment with embedded electrodes is shown on the bottom (b). The acquisition of four physiological signals: electroencephalogram (EEG), electrocardiogram (ECG), photoplethysmogram (PPG), and electrooculogram (EOG) are depicted in the figure [66].

Among 18 pilots recruited for data collection, one participants recording was interrupted due to technical problems and other participant had no MS data, hence these two participants were discarded from data analysis. Therefore, signals acquired from the remaining 16 participants (age: 21.5±2.4 year, 2 females) were used for further analysis. All participants had previous mean flight time experience of 282 hours, and had mean simulated flight time of about two hours between takeoff and end time during the experimental session. The participants were also requested to self-report KSS scale before and after experiment and a significant difference (p<0.05) was observed between them using t-test as shown in Table 2.1.

Table 2.1. The basic information of valid participants. A significant difference was observed in the KSS value at the end of the experiment compared to the beginning. '*' represents alpha level at 0.05.

Participant ID	Age (years)	Gender	Flight Time (hours)	Takeoff time (am)	End Time (am)	KSS (start)	KSS (end)
105	20	Male	250	2:04	4:26	6.0	9.0
122	27	Male	400	2:14	4:21	6.0	8.0
126	21	Male	280	2:11	4:11	7.0	8.0

133	21	Male	315	2:11	4:18	4.0	9.0
135	28	Male	419	1:57	4:04	6.0	9.0
136	21	Female	185	2:08	4:08	7.0	9.0
140	21	Male	200	2:06	3:25	7.0	9.0
143	20	Male	400	2:06	4:18	7.0	9.0
145	20	Male	430	2:10	4:15	6.0	8.0
149	21	Male	200	2:09	4:23	3.0	8.0
165	21	Male	215	2:00	4:26	3.0	9.0
172	20	Male	280	2:05	4:20	7.0	9.0
179	21	Male	250	2:09	4:33	4.0	8.0
183	21	Male	300	2:13	3:42	7.0	9.0
198	21	Female	195	2:20	4:06	5.0	9.0
199	20	Male	200	2:08	4:15	1.0	9.0
mean	**21.5±2.4**	**14 / 2**	**282±87**	**2:08**	**4:11**	**5.4±1.9**	**8.7±0.5***

2.6.3. Data Preparation

Seeing Machines DMS was used to detect MS from long eyelid closure (LEC) events by non-invasively measuring eyelid behavior. Once LEC events had been identified by the DMS, a manual HACM review process was followed to determine if an instance was indeed a microsleep. HACM identified MS by classifying these LEC events, specifically the events between 1.5s and 15s duration were classified as MS, and greater than 15s as long-sleep (LS). MS and LS classifications

were used to investigate potential differences between shorter and longer drowsy periods. A detailed description is provided in Appendix B.

Once LEC events were identified by DMS, manual HACM was used to determine if the LEC were indeed MS events (MS and LS periods) via three strict rules: (1) During the eye closure, the pilot was not engaged in any secondary behavior. (2) During the eye closure, the pilot appeared to exhibit signs of drowsiness or reduced cognitive capacity (e.g. slumping head position/ posture). (3) The eye closure event exhibited reduced velocity of eyelid closure and opening [85]. Finally, the MS events annotated via DMS technology were further confirmed using EOG signal by manual verification that no blinks were observed during those periods. These steps were followed to ensure the reliability of annotated drowsy periods. A flowchart for the annotation process is illustrated in Figure 2.7.

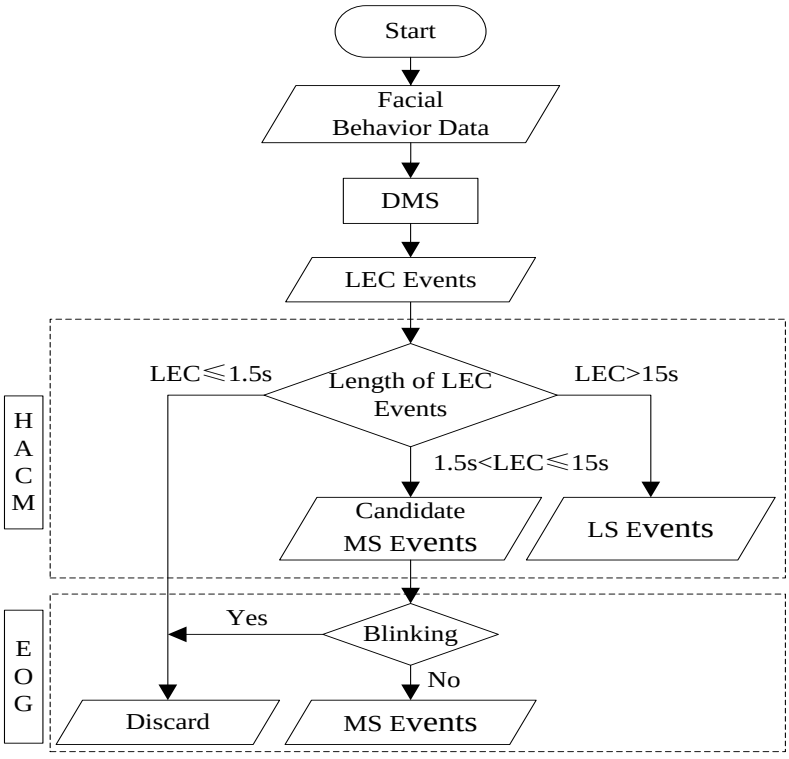


Figure 2.7. The flowchart for annotation of drowsy periods. The Seeing Machines Driver Monitoring System (DMS) was used to detect long eyelid closure (LEC) events and a manual Human Annotation Classification Method (HACM) is followed to ensure the microsleep candidate (microsleep (MS) and longsleep (LS)). Finally, all candidate MS events were further verified by electrooculogram (EOG) [66].

2.6.4. EEG Data Processing

For PSD and coherence analysis, we compared spectral power from three physiological states: BS, MS, and LS periods. EEG signals from 13 channels were resampled to 125 Hz. Then, high amplitude artifacts were removed and band pass filtered to 0.5-30 Hz and then ICA and ADJUST algorithm were used as mentioned in power spectral analysis section above.

PSD was calculated using Welch's power spectral density estimate having Hamming window size of 128 samples with 50% overlap and 256 discrete Fourier transform (DFT) points. Finally, the relative spectral power including delta (δ, 0.5-4Hz), theta (θ, 4-8Hz), alpha (α, 8-13Hz), and beta (β, 13-30Hz) were calculated [86]. Furthermore, four indices (i) (θ+α)/β, (ii) α/β, (iii) (θ+α)/(α+β), and (iv) θ/β were computed as the additional indicators of drowsiness [87]. The mean of each indicator (four frequency band indicators and four derived indices) during different psychological states (BS, MS, and LS) from the channels corresponding to the same region of brain was calculated, then all those mean values of each indicator from the five regions were used for statistical analysis. For example, the mean of O1 and O2 was considered as the occipital signal.

For coherence analysis, the resampled and filtered signals from five EEG electrodes: T7, C3, Cz, C4, and T8 within the aviation headset region are used for analysis. The coherence from each channel is obtained using a 4 sec window. This particular window length is selected at which the

coherence value is found to be higher than the other lengths. Non-drowsy EEG is obtained from the baseline starting from 2 min and the signal is cut with a 4 sec segment sliding with 3 sec overlap for 20 iterations in each subjects. For MS EEG, MS events greater than 5 sec is chosen and cut with 4 sec length from the onset of MS. Similarly for LS EEG, LS events greater than 20 sec is chosen and cut with 4 sec length after 15 sec of LS. This is done because LEC events up to 15 sec constitute MS. Then, the mean of MSC and WTC on the three physiological states: BS, MS, and LS are computed and used for statistical analysis. A simple aviation headset with the electrodes location within the headset region is depicted in Figure 2.8.

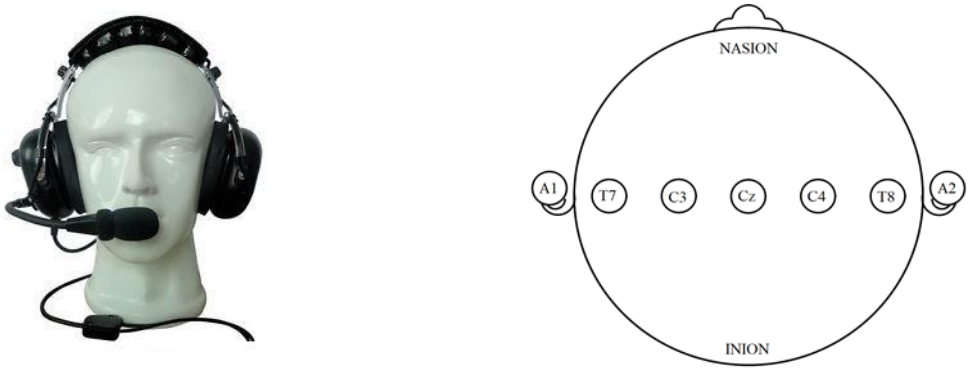


Figure 2.8. A sample of aviation headset [88] and the electrodes location within the headset region from left to right respectively.

2.6.5. ECG Data Processing

The acquired ECG signal from a single channel using Lead II Einthoven's Triangle is segmented into four experimental time periods (T1, T2, T3, and T4) as shown in Figure 2.9.

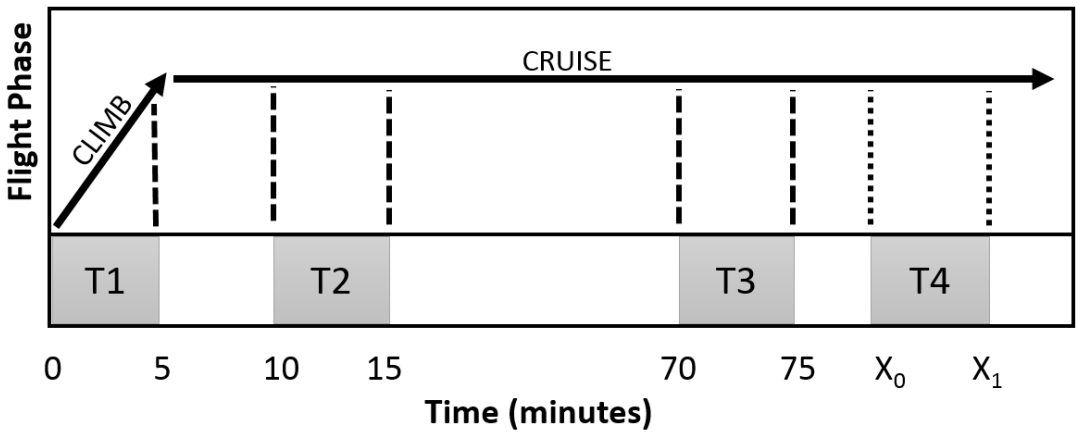


Figure 2.9. The experimental protocol illustrating four time periods (T1, T2, T3, and T4). The three time periods are pre-defined as: baseline (T1, 0-5 min), relax (T2, 10-15 min), and fatigued (T3, 70-75 min), whereas drowsy period (T4) is obtained from longest eye-closure based on vertical electrooculogram (EOG) and vary with participants [80].

During the baseline period (T1, 0-5 min), the pilots were busy conducting flight takeoff and climb operations and can be considered as an alert time. During relax period (T2, 10-15 min), the flight is in autopilot mode and assumed that the pilots are in relaxation state. The third period (T3, 70-75 min) is hypothesized as fatigued period assuming that the pilots are stressed and fatigued due to longer flight duration and sleep deprivation from the previous day. The final 5 min period (T4) is calculated as a drowsy period where pilots are falling asleep and the period is obtained from long sleep events based on the analysis of EOG or eye blink periods selecting the periods of no eye blink. The complete segment of T4 is obtained from additional recorded data prior to the long sleep event to make the periods of 5 min for each participant.

The R peaks are detected from the segmented data in each experimental time periods using Pan-Tompkins method [49] and RR intervals are calculated. Ten time domain features: heart rate (HR), standard deviation (SD), standard deviation of successive difference (SDSSD), root mean square of successive difference (RMSSSD), triangular index (TI), sample entropy (SA), NN20, PNN20,

NN50, and PNN50 based on HRV were calculated from RR intervals. To find frequency domain features, RR intervals are interpolated using spline interpolation with 500 sample points. This maintains the continuity of RR intervals. Then, the interpolated RR interval is resampled to 1 Hz since the frequency spectrum of HRV lies from 0 to 0.5 Hz. Finally, power is calculated using Welch's power spectral density method having Hamming window size of 64 samples with 50% overlap. Power in three frequency bands were obtained: very low frequency (VLF, 0-0.04 Hz), low frequency (LF, 0.04-0.15 Hz), and high frequency (HF, 0.15- 0.5 Hz). LF represents both sympathetic and parasympathetic modulations of heart, whereas HF represents mainly parasympathetic influences and LF/HF ratio represents symapthovagal balance. LF and HF were measured in absolute values of power (ms^2) as well as in normalized units (n.u.), which represent relative value of each power component in proportion to the total power minus the VLF component [75]. The description of time and frequency domain features are provided in appendix A. The mean±SE from each features (ten time domain and six frequency domain) during four psychophysiological states (T1, T2, T3, and T4) were calculated and the mean values of each states were used for statistical analysis. The steps for calculation of time and frequency domain features are depicted in Figure 2.10.

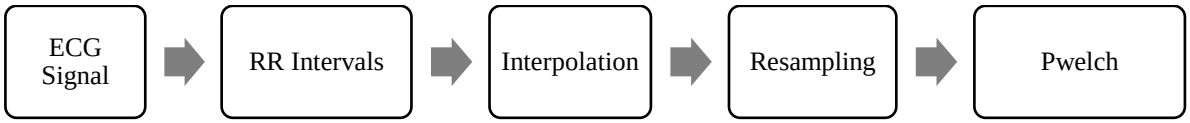


Figure 2.10. Steps for calculation of time and frequency domain features from electrocardiogram (ECG). Time domain features are obtained from RR intervals and frequency domain features from Pwelch spectrum.

2.6.6. Statistical Analysis

The normality test using Shapiro-Wilk method among three physiological states (BS, MS, and LS) of EEG spectrum showed mixed behavior due to limited sample size and uneven distribution in three states. Therefore, Kruskal-Wallis test followed by Bonferroni correction was used to calculate the significant difference at $\alpha=0.05$ between the three states in the four EEG spectrum. Similarly for ECG analysis, a Shapiro-Wilk test is used to examine the normality of data among four defined periods (T1, T2, T3, and T4) to determine the use of a parametric or nonparametric method based on the data distribution. The normality test shows mixed behavior in case of ECG analysis too due to limited sample size; therefore, a Kruskal-Wallis test followed by a Bonferroni correction as a post hoc analysis is used to test the significant difference at $\alpha=0.05$ between the four periods. All the tests are conducted using a statistical toolbox of MATLAB (The MathWorks, Inc., Natick, MA, USA). All the graphs and tables are presented as mean±SE and mean±SD respectively.

Chapter 3.

Spectral Analysis of Electroencephalogram during Drowsiness

3.1. Summary

Background: Detecting drowsiness plays an important role in aviation industry. The use of feasible approach to quantify drowsiness gives an added advantage. EEG is non-invasive and simpler to acquire so, it is widely used to characterize drowsiness. However, its behavior at the exactly annotated drowsiness periods remains to be investigated. The subjective measures such as KSS and PVT are often used references to analyze drowsiness that has high inter participant's variability and numerous limitations. **Methods:** To address the limitations, Seeing Machines (SM) Driver Monitoring System (DMS) is used to annotate microsleep (MS) events which are further verified using vertical EOG. EEG, EOG, and facial behavior data are recorded simultaneously from sixteen commercially-rated pilots in simulated flights during WOCL. The PSD within the five brain regions: frontal, central, parietal, occipital, and temporal and coherences (MSC and WTC) from EEG electrodes within the region of aviation headset: T7, C3, Cz, C4, and T8 using 10-20 international system are computed. The analysis is performed within the four EEG spectra: delta (0.5-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), and beta (13-30 Hz). **Results:** On analyzing PSD, delta power reduces during MS (p<0.05 in all regions), alpha power increases during MS (p<0.01 in all regions), while theta and beta powers did not change (p>0.05). The coherence analysis among aviation headset depicts that WTC provides more profound results than MSC with significant synchronization (p<0.05) among most of the electrodes at alpha band which is the main indicator for drowsiness. **Conclusion:** The PSD analysis highlights the capability of EEG delta and alpha spectrum towards characterizing MS events. Further, the coherence results illustrate functional

neural connectivity highlighting the effectiveness of WTC to detect drowsiness using EEG from suitable electrodes placement within the aviation headset. The results depict the potential of EEG spectra to characterize drowsiness and the plausibility for detection from the typical aviation headset.

3.2. Background

Polysomnography (PSG), also known as sleep study, involves the monitoring of various aspects of body activity during sleep. It records brain waves, oxygen level in blood, heart rate, breathing rate, and eye as well as leg movements during the study to diagnose sleep disorders [89]. Sleep cycle is broadly divided into five stages: rapid eye movement (REM), non-REM1 (NREM-1), NREM-2, NREM-3, and NREM-4 where REM sleep and NREM sleep alternate across the night in a 90 min cycle. The description of various sleep stages are given in Table 3.1 [90]. The table includes all the five sleep stages, patterns of EEG during each stages, and the response of participants.

Table 3.1. Description of five sleep stages (awake, NREM-1, NREM-2, NREM-3, NREM-4, and REM) in correspondence to EEG patterns and human physiological states.

Sleep Stage	Status	EEG Patterns	Comments
Zero	Awake	Rapid varying, low voltage EEG and prominent beta waves	Patient's eyes open
Stage 1	NREM-1, drowsiness	Alpha waves predominant and EEG frequency around 6-8 Hz	Very slow eye movements

Stage 2	NREM-2, light sleep	K-complexes and sleep spindles begin to appear, EEG medium amplitude with frequency 4-7 Hz	Eye movements stop with slower brain waves
Stage 3	NREM-3, deep sleep	Delta waves begin to appear, EEG high amplitude with frequency around 1-3 Hz	Awakening is rare
Stage 4	NREM-4, deeper sleep	Slow delta waves prominent, Very high EEG amplitude with frequency less than 2 Hz	Very difficult to awake person
Stage 5	REM, dreaming	Brain waves more active than stage 3 and 4, theta wave is more predominant	Rapid eye movements with occasional muscular twitches, arousal is easy

Drowsiness is defined as the propensity to fall asleep. It is NREM-1 sleep stage, a transition from awake to onset of sleep. Driver drowsiness has been considered as one of the major causes of road accidents [7]. A report of US National Highway Traffic Safety Administration found that drivers' drowsiness results in 1,550 deaths, 71,000 injuries, and $ 12.5 billion losses in revenue every year [7], [91]. According to the National Sleep Foundation, about half of U.S. adult drivers admit to consistently getting behind the wheel while feeling drowsy. The drivers may not realize how much drowsy driving puts themselves and others at risk. The Governors Highway Safety Association report states that an estimated 5,000 people died in 2015 due to crashes involving drowsy driving. Driving after going more than 20 hours without sleep is equivalent to driving with a blood-alcohol concentration of 0.08%, the US legal limit. When a driver experiences 4 to 5 seconds of MS, then

the vehicle will travel the length of a football field [92]. Drowsiness has also detrimental effects in pilots during flight operations. The long and monotonous flights make pilots in the cockpit drowsy and fatigue. FAA maintains a strict protocol of pilot's fitness for flight. The standards for medical certification are contained in 14 CFR part 67. These medical conditions also include drowsiness and fatigue [11]. Self-assessment of drowsiness and fatigue state by pilots may not be the effective approach to evaluate their psycho-physiological conditions that may result aviation hazards. There are various aviation accidents due to drowsiness and fatigue and the accidents occur in space, the loss is tremendously high [6]. Thus, there is a requirement of efficient and reliable drowsiness detection system in aviation to mitigate such life-taking hazards.

There are various approaches for drowsiness detection. Among them, the three main approaches are: vehicular, behavioral, and physiological as mentioned earlier [15]–[17]. The first two approaches are system dependent and affected by external environmental conditions. They detect drowsiness only after eye-closure which is too late to prevent an accident. The changes in physiological signals due to drowsiness occur before eye-closure and the analysis of these signals can characterize drowsiness before falling sleep that is practical in real time to detect drowsiness [3], [4].

Among various physiological signals such as EEG, ECG, EOG, and PPG; EEG is a commonly used drowsiness indicator [2], [17], [33], [34]. EEG is a reflectance of brain dynamics and drowsiness effects are vivid in the brain, it is an effective indicator. EEG is non-invasive, easy to acquire from limited electrodes, and drowsiness effects are clearly reflected in the spectral behavior [4], [36], [93]. To study the EEG behavior, various features are extracted in time, frequency, and non-linear domains. Time domain features include statistical (mean, standard deviation, root mean square, skewness, and kurtosis) and Hjorth parameters (activity, mobility,

and complexity) [94], whereas frequency domain features include power spectral analysis, median frequency, and wavelet transform. The commonly used non-linear features are entropies, largest Lyapunov exponent, and Hurst exponent [95]. Out of these, spectral power changes in EEG frequency bands: delta, theta, alpha, and beta are common in literature to estimate shift of psychological state from alertness towards drowsiness [86], [96]. The alpha has been observed as the most relevant drowsiness indicator within the spectrum [65], [97], [98]. However, there are contradicting views regarding the spectral changes in different brain regions due to traditional approaches for drowsiness annotations [65], [97]–[100].

Apart from PSD, coherence analysis is also common in literature to estimate the interactions between different electrode pairs. EEG coherence is a measure of functional connectivity between brain regions recorded from spatially separated scalp electrodes to estimate the similarity of waveform components generated by the mass action of neurons in underlying cortical regions [69], [70]. Estimating coherence across the physiological states help to understand the neural mechanisms towards drowsiness and may provide early assessment to prevent aviation accidents. In this chapter, the MSC and WTC are also compared to illustrate significant interactions among five EEG electrodes: T7, C3, Cz, C4, and T8 within the headset region.

The alterations in electrophysiological organization between brain regions were first observed by [69] through inter and intra-hemispheric MSC during light drowsiness. In contrast to the MSC based on Fourier transform, WTC was first used in neural signals by [72] to estimate coherence in course of time between single trial brain signals. Neural signals are non-linear and non-stationary. WTC can detect short episodes of coupling between neural signals measuring time- frequency coherence [72]. This coherence analysis from WTC solves the problem of non-stationarity than the classical Fourier based coherence which is a only a function of frequency [73].

A number of EEG electrodes have been used for drowsiness estimation ranging from 64 to 1 [101]–[105]. Even, drowsiness has been detected from non-hair bearing areas using six electrodes to avoid interference of hair and the skin [106]. Embedding EEG electrode cap with even a single sensor at unfeasible brain region may be cumbersome to pilots during flight operations. Therefore, we propose a novel idea to place electrodes along an aviation headset around the central sulcus of the hemisphere. Embedding electrodes through the headset will eliminate the extra sensor placement due to which drowsiness can be estimated from the coherence analysis among the electrode pairs.

As stated earlier, there are varying results in literature due to different approaches adopted for drowsiness annotation. Annotating drowsy EEG periods is critical to evaluate the spectral behavior and failure to do so produces misleading results. Manual scanning and interpretation by highly trained experts is time consuming and expensive [107]. Currently, the subjective measures such as KSS and PVT based on participants' feelings are used to assess drowsiness levels [21]–[23] and these measures have limitations. To overcome the limitation of annotating drowsy EEG periods, we integrated Seeing Machines DMS in our experiment that performed computer vision analysis of facial behavior i.e. face, eye, and eyelid tracking. Then, eyelid closure data was used to obtain PERCLOS value and candidate sleep events. Finally, Human Annotation Classification Method (HACM) was used to obtain microsleep (MS) events [85] and simultaneously recorded EOG signal was used to verify the annotated MS events to make sure no eye blink exists in them. The existence of no eye blink was essential for long eyelid closure, eye closure greater than 0.5s [85], to annotate MS events to ensure that the participant was in drowsy state. In this research, we comprehensively investigated the characteristics of EEG during SM annotated drowsy periods via DMS technology. The brain region suitable for drowsiness detection was also explored.

In this chapter, we combine two approaches: PSD and coherence to characterize drowsiness. The PSD analysis in EEG spectra among the different physiological states helps to understand the significant changes in EEG power that exist in specific brain regions during drowsiness. The coherence analysis provides the significant interactions within a particular EEG spectra at specific interactions within the region of aviation headset. The both approaches aid in characterizing drowsiness using EEG spectral changes.

3.3. Results

The relative PSD changes (mean±SD) in the δ, θ, α, and β and indices (i) (θ+α)/β, (ii) α/β, (iii) (θ+α)/(α+β), and (iv) θ/β in BS, MS, and LS states for five specific brain regions are listed in Tables 3.2 and 3.3. In addition, their trends (mean±SE) are further illustrated in Figures 3.1 and 3.2 respectively. The post hoc *p*-value for significant difference between each two states are listed in Table 3.4. The analyzed regions include frontal (F3, Fz, F4), central (C3, Cz, C4), parietal (P3, Pz, P4), temporal (T7, T8), and occipital (O1, O2). Statistical analysis on comparing MS and baseline showed that delta power had significant decrease ($p<0.05$ for temporal, and $p<0.001$ for others), while alpha had significant increase ($p<0.001$) in all regions. Index (i) had significant increase ($p<0.05$ for frontal and central, and $p<0.001$ for parietal and occipital) in all regions except temporal. For index (ii), significant increase ($p<0.05$ for temporal, and $p<0.001$ for others) was found in all regions. In addition, comparison of LS with baseline showed that delta power had significant decrease in parietal whereas, alpha had significant increase in all regions except temporal ($p<0.05$). Furthermore, index (i) had significant increase ($p<0.05$) in central and parietal, and index (ii) had significant increase ($p<0.05$ for central and occipital, and $p<0.001$ for parietal) in all regions except frontal and temporal. However, there were no significant ($p>0.05$) changes in all the indicators between MS and LS.

Table 3.2. Relative spectral power distribution (mean±SD) for four frequency bands in five brain regions ((Frontal, F), (Central, C), (Parietal, P), (Temporal, T), and (Occipital, O)) under three conditions. '*' and '†' represent the alpha level: p=0.05, and p=0.001 respectively for significant difference.

Indicator	Region	BS (%)	MS (%)	LS (%)
δ	**F**	68.08±10.91	**49.24±12.57†**	**57.10±9.56**
	C	64.92±9.87	**42.83±12.23†**	52.24±11.03
	P	63.28±11.72	**36.33±13.43†**	**47.25±14.85***
	T	66.15±11.89	**50.30±11.99***	56.97±11.49
	O	60.43±13.33	**38.59±12.79†**	47.19±14.16
θ	**F**	13.58±4.21	15.78±3.43	16.18±2.45
	C	13.41±3.31	15.81±5.49	15.69±2.90
	P	13.75±3.70	13.03±4.65	15.27±3.41
	T	9.83±3.18	12.50±3.96	12.40±2.39
	O	11.71±3.44	12.60±6.59	13.99±3.32
α	**F**	8.06±3.80	**23.65±12.83†**	**15.86±9.20***
	C	9.56±3.55	**28.68±14.00†**	**20.28±11.70***
	P	10.52±3.92	**39.58±16.34†**	**27.01±17.05***
	T	8.04±3.89	**20.57±11.34†**	14.42±8.43
	O	10.44±3.54	**33.71±16.61†**	**22.19±13.37***
β	**F**	10.28±4.57	11.32±3.44	10.86±5.13
	C	12.10±4.91	12.68±4.77	11.80±6.40

	P	12.44±5.54	11.06±3.72	10.48±5.16
	T	15.98±8.73	16.63±6.84	16.20±7.67
	O	17.42±9.32	15.10±6.17	16.63±8.12

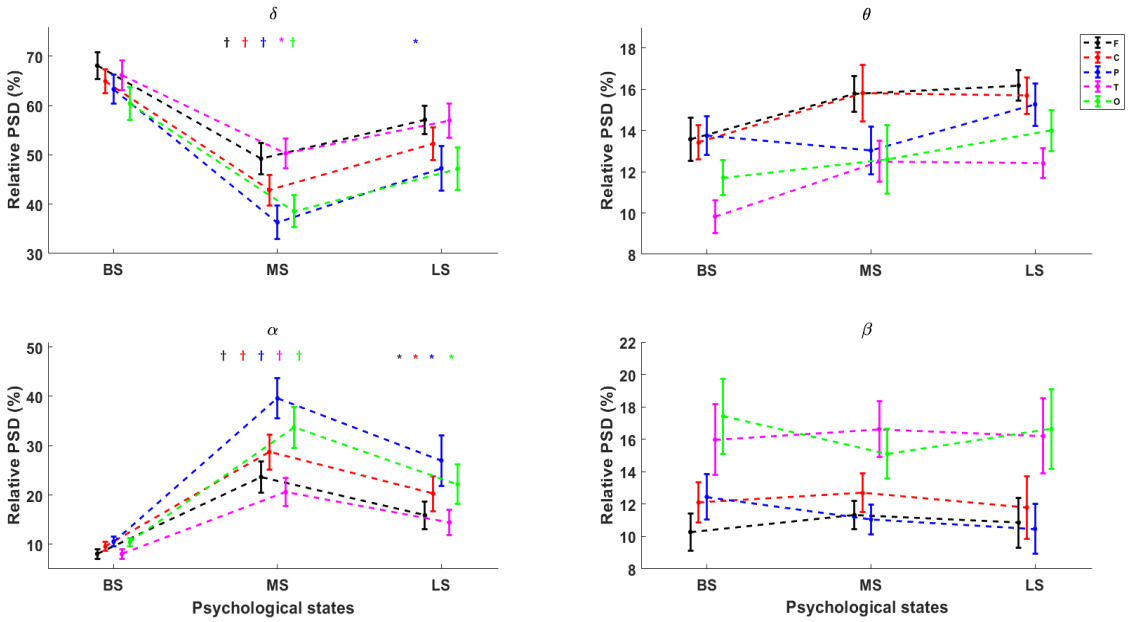


Figure 3.1. Relative spectral power distributed (mean±SE) in four EEG frequency bands: δ, θ, α, and β under three psychological states (BS, MS, and LS) in five brain regions ((Frontal, F), (Central, C), (Parietal, P), (Temporal, T), and (Occipital, O)). The x-axis represents three psychological states and y-axis represents relative power spectral density (PSD). The '*' and '†' represent significant difference at p = 0.05 and p = 0.001 respectively in MS and LS compared to BS. BS=Baseline, MS=Microsleep, LS=Long sleep [66].

Table 3.3. Value (mean±SD) for four indices in five brain regions ((Frontal, F), (Central, C), (Parietal, P), (Temporal, T), and (Occipital, O)) under three conditions. '*' and '†' represent the alpha level: p=0.05, and p=0.001 respectively for significant difference.

Indicator	Region	BS (%)	MS (%)	LS (%)
(i) (θ+α)/β	F	2.43±0.89	**3.92±2.20***	**3.49±1.63**
	C	2.20±0.86	**4.09±2.25***	**3.98±2.81***
	P	2.22±0.69	**5.49±2.86†**	**5.12±3.66***
	T	1.64±0.87	2.49±1.33	2.19±1.29
	O	1.57±0.62	**3.85±2.45†**	2.70±1.41
(ii) α/β	F	0.82±0.20	**2.41±1.76***	1.78±1.35
	C	0.84±0.22	**2.72±2.04†**	**2.35±2.51***
	P	0.89±0.23	**4.25±2.79†**	**3.44±3.54†**
	T	0.63±0.23	**1.59±1.13***	1.21±1.08
	O	0.69±0.22	**2.93±2.34†**	**1.67±1.28***
(iii) (θ+α)/(α+β)	F	1.32±0.41	1.14±0.14	1.25±0.22
	C	1.17±0.40	1.10±0.18	1.18±0.23
	P	1.16±0.33	1.05±0.10	1.18±0.18
	T	0.96±0.46	0.91±0.17	0.93±0.17
	O	0.90±0.27	0.97±0.18	0.98±0.20
(iv) θ/β	F	1.61±0.79	1.51±0.57	1.70±0.51
	C	1.36±0.76	1.37±0.47	1.63±0.63
	P	1.33±0.61	1.24±0.38	1.68±0.53
	T	1.01±0.74	0.90±0.34	0.98±0.31
	O	0.88±0.43	0.92±0.39	1.02±0.37

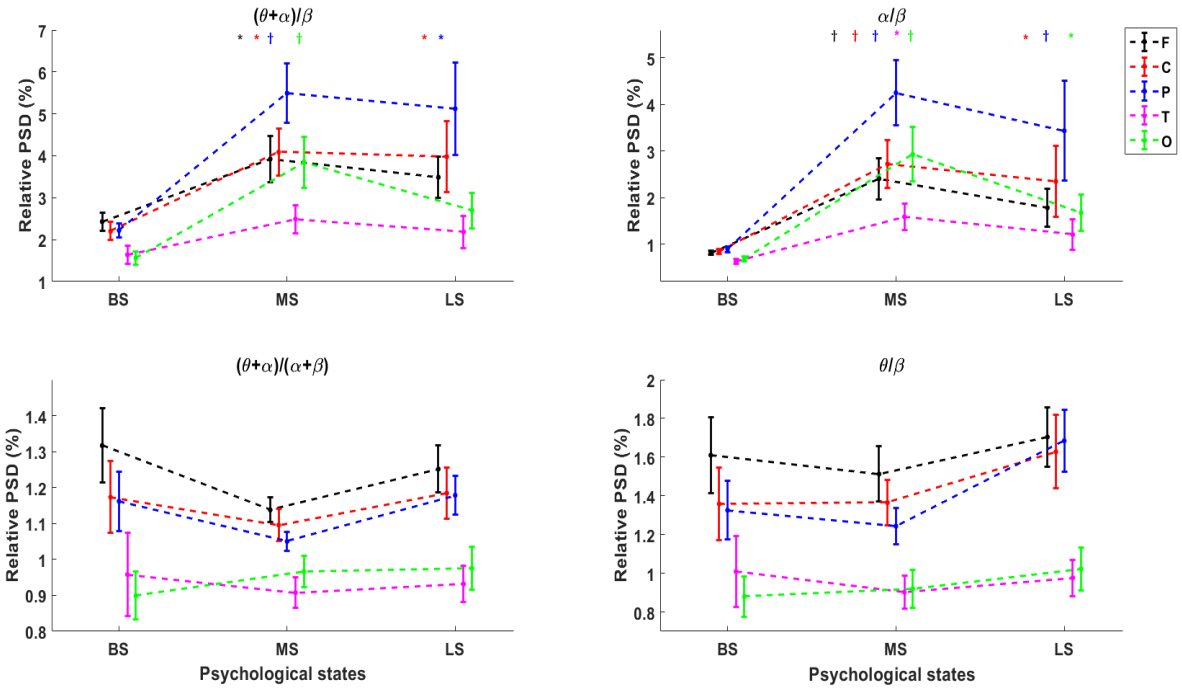


Figure 3.2. The behavior (mean±SE) of indices (i)–(iv) under three psychological states (BS, MS, and LS) in five brain regions ((Frontal, F), (Central, C), (Parietal, P), (Temporal, T), and (Occipital, O)). The x-axis represents three psychological states and y-axis represents relative PSD. The '*' and '†' represent significant difference at p = 0.05 and p = 0.001 respectively in MS and LS compared to baseline. BS=Baseline, MS=Microsleep, LS=Long sleep [66].

Table 3.4. The table lists multiple comparison p-values for five brain regions under three psychological states (BS, MS, and LS). '*' and '†' represent the alpha level: p=0.05, and p=0.001 respectively for significant difference.

Indicators	**Conditions to Compare**		**Frontal**	**Central**	**Parietal**	**Temporal**	**Occipital**
δ	MS	BS	**<0.001†**	**<0.001†**	**<0.001†**	**0.003***	**<0.001†**

	LS	BS	0.092	0.068	**0.046***	0.251	0.134
	LS	MS	0.614	0.276	0.343	0.599	0.432
θ	MS	BS	0.235	0.429	0.816	0.178	1.000
	LS	BS	0.149	0.214	0.889	0.075	0.625
	LS	MS	1.000	1.000	0.125	1.000	0.421
α	MS	BS	**<0.001†**	**<0.001†**	**<0.001†**	**<0.001†**	**<0.001†**
	LS	BS	**0.042***	**0.006***	**0.005***	0.092	**0.007***
	LS	MS	0.290	0.317	0.243	0.259	0.237
β	MS	BS	1.000	1.000	1.000	1.000	1.000
	LS	BS	1.000	1.000	0.346	1.000	1.000
	LS	MS	1.000	1.000	1.000	1.000	1.000
(i) (θ+α)/β	MS	BS	**0.037***	**0.002***	**<0.001†**	0.124	**0.001†**
	LS	BS	0.145	**0.017***	**0.006***	0.697	0.058
	LS	MS	1.000	1.000	1.000	1.000	0.951
(ii) α/β	MS	BS	**0.001†**	**<0.001†**	**<0.001†**	**0.005***	**<0.001†**
	LS	BS	0.086	**0.005***	**0.001†**	0.184	**0.018***
	LS	MS	0.823	1.000	0.795	0.987	0.440
(iii) (θ+α)/(α+β)	MS	BS	0.854	1.000	0.912	1.000	1.000
	LS	BS	1.000	0.811	1.000	1.000	0.834
	LS	MS	0.658	0.501	0.202	1.000	1.000
(iv) θ/β	MS	BS	1.000	1.000	1.000	1.000	1.000
	LS	BS	1.000	0.267	0.125	1.000	0.685
	LS	MS	0.733	0.656	0.098	1.000	0.715

During coherence analysis, ten interactions are possible from five electrodes and the possible interactions include: T7-C3, T7-Cz, T7-C4, T7-T8, C3-Cz, C3-C4, C3-T8, Cz-C4, Cz-T8, and C4-T8. The trends of MSC and WTC in terms of mean±SE in four frequency bands: δ, θ, α, and β in ten interactions among five electrodes in three psychological states: BS, MS, and LS within the EEG spectrum (0.5-30 Hz) are depicted in Figures 3.3 and 3.4.

The results illustrated increase in coherence in most of the EEG spectra during MS from WTC method, in addition to this, four interactions: T7-Cz, T7-C4, T7-T8, and C3-T8 exhibited a significant increase (p<0.05) in alpha band illustrated in Figure 3.5. The WTC in alpha interaction T7-C3 is just near to significant p-value (p=0.06). However, alpha coherence from WTC decreased significantly at C3-T8 during LS compared to MS. Apart from these, there are no significant changes at all in any frequency bands from the both methods. WTC is only significant (p<0.05) at alpha interactions which is well recognized indicator of drowsiness and fatigue [108]. The highest coherence is observed at C3-Cz, and Cz-C4 interactions at all the three physiological states from both the coherence methods. This implies high waveform similarity around the central brain region.

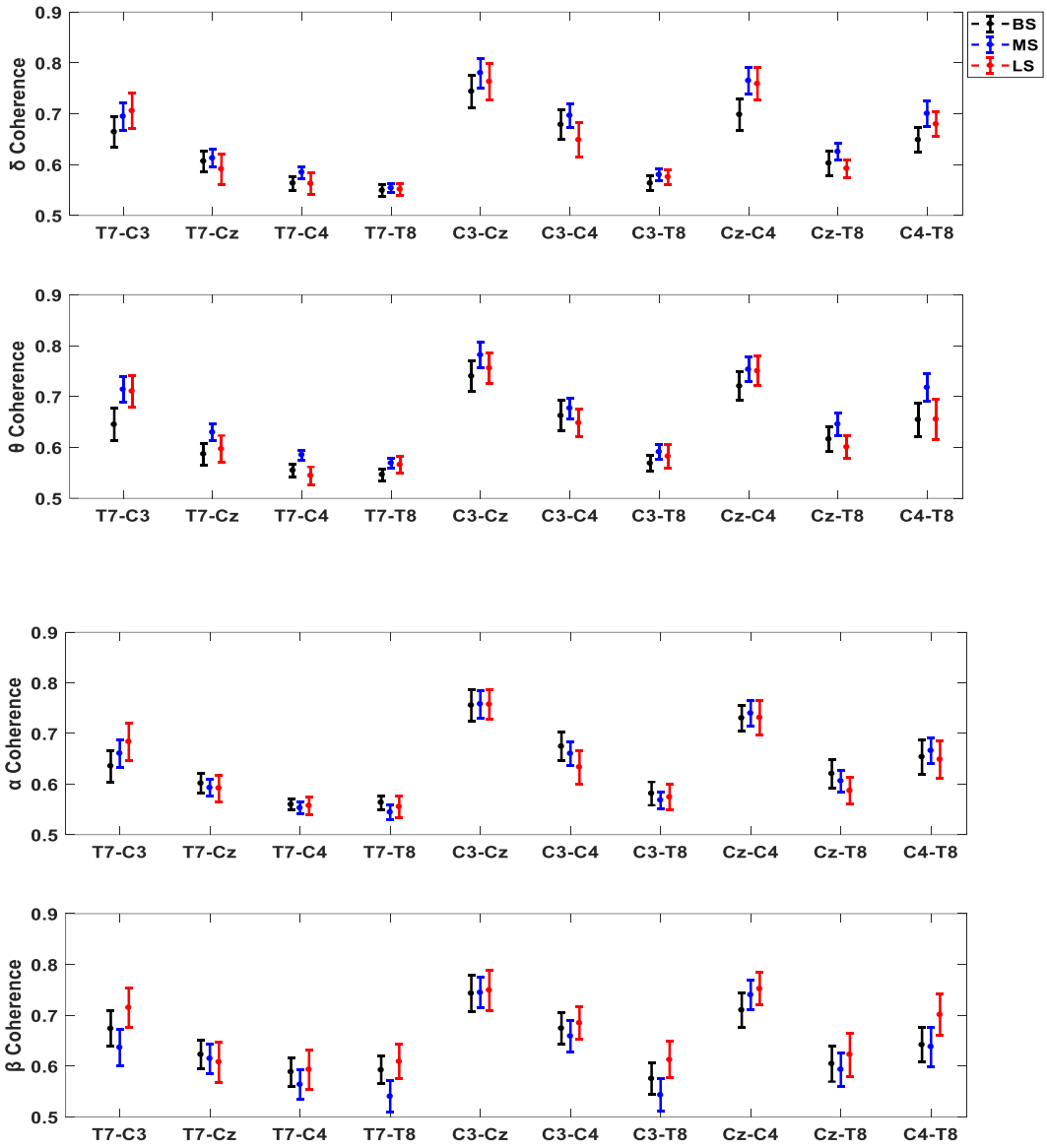


Figure 3.3. The trend (mean±SE) of Magnitude Squared Coherence (MSC) of five EEG electrodes within headset region in four EEG spectra. Black, blue, and red lines represent baseline (BS), microsleep (MS), and longsleep (LS) respectively. The x-axis represents ten interactions among five electrodes and y-axis represents mean MSC in four frequency spectra. There exists no significant difference at all.

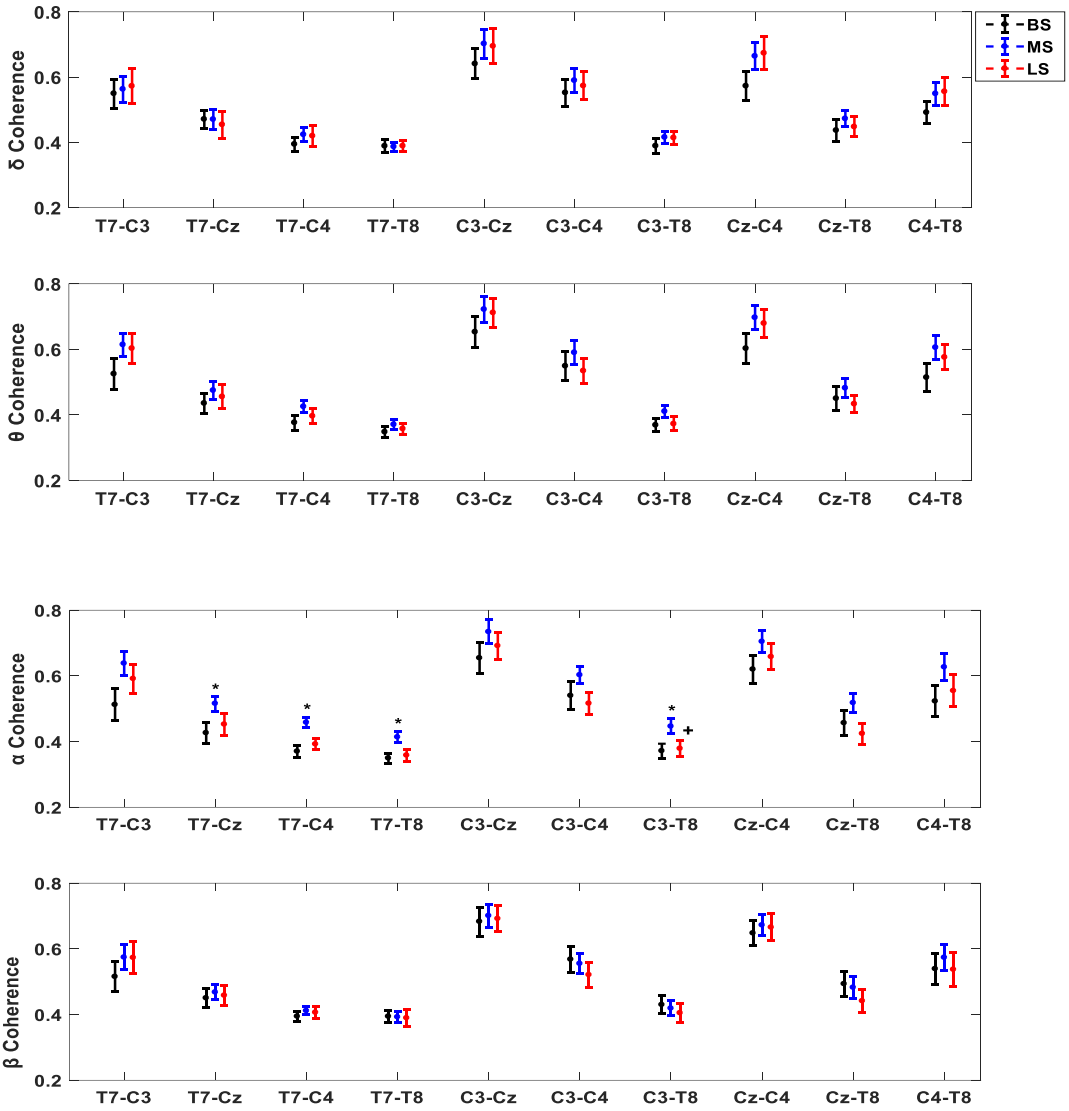


Figure 3.4. The trend (mean±SE) of Wavelet Transform Coherence (WTC) of five EEG electrodes within headset region in four EEG spectra. Black, blue, and red lines represent baseline (BS), microsleep (MS), and longsleep (LS) respectively. The x-axis represents ten interactions among five electrodes and y-axis represents mean WTC in four frequency spectra. '*' and '+' represent significant difference (p=0.05) between BS and MS, and MS and LS respectively.

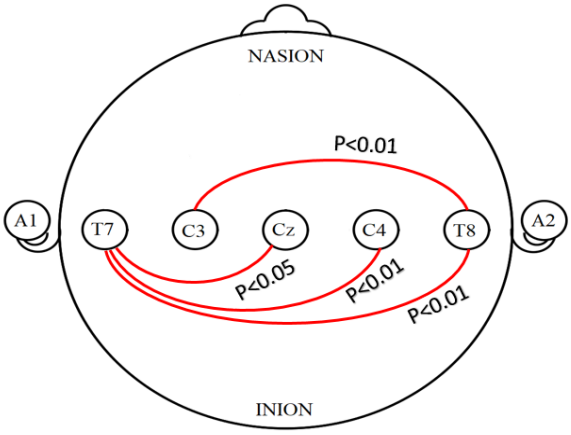


Figure 3.5. The significant alpha interactions from Wavelet Transform Coherence (WTC) during baseline (BS) and microsleep (MS). These significant interactions can be a marker for drowsiness detection using aviation headset.

3.4. Discussion

In this chapter, we comprehensively investigate the potential changes in relative spectral powers and coherences in respective EEG frequency bands as an indicator of transition to drowsiness during WOCL. We performed a study during simulated flight of the FAA certified commercially rated pilots. In order to accurately obtain the characteristics of EEG, we utilized Seeing Machines DMS technology along with EOG, for the first time, to annotate MS events.

During PSD analysis, the relative power spectra of the four typical bands and their four derived indices were calculated, and all the indicators in the BS, MS, and LS states were statistically analyzed. In addition, we also examined drowsiness self-reported (KSS) by each participant before and after the experiment. Our observation of statistically significant ($p=0.05$) differences in the EEG spectral power during MS events compared to baseline underscored the feasibility of early stage drowsiness detection via EEG (Table 3.2). The major novelty of this study was in the experimental protocol, where multimodal data, including EEG, EOG, and DMS, were acquired

during the window of circadian low which is known to disrupt the circadian activity [5]. Therefore, data acquisition at this period of circadian low could offer a good potential for observing drowsiness within a typical 24 hours window. In addition, the simulated physiological conditions in our experiment are similar to the extended flight time in some flight operations scenarios. Referring to KSS scale [109], the statistical result indicated that the degree of drowsiness was significantly increased ($p<0.05$) during the course of the experiment (Table 2.1). Thus, spectral characteristics obtained from EEG during the window of circadian low was beneficial to assess drowsiness for pilots.

The subjective analysis via self-assessments such as KSS and PVT have limited usability to quantify drowsiness and implementation of such methods always exhibits a challenge. Compared with these methods, DMS is used to perform visual analysis of facial expression and acquire drowsiness information based on PERCLOS. To assist DMS, manual evaluation process by HACM is further carried out and EOG is used to double check the annotated drowsy periods, due to which the reliability of annotation for drowsy periods is enhanced. The reliable annotation approach and data recording during the window of circadian low from the professional participants justify the effectiveness of our experiment. Therefore, our experimental results characterizing spectral behavior provide evidence towards the development of future technology for monitoring psychological states and warn pilots based on different degrees of drowsiness.

In the existing literature, some researchers have found a significant change in delta and theta power, more specifically in theta during drowsy state [86], [96], [110]. In a study of simulated driving, it was found that theta power increased significantly during night time compared to baseline taken during daytime [111]. Another study on prolonged awakening revealed that theta power showed significant increase only at 9th, 21th and 24th hours [112]. For the study of brain

regions, EEG from the first and last 20 minutes of driving sessions were analyzed and found lower alpha power in frontal and higher in parieto-occipital regions [108]. However, the initial study by [17] had shown increase in alpha towards frontal and decrease towards occipital regions. Another experiment on simulated driving had shown higher alpha power in parietal and lower in frontal regions. Also, the beta power had shown dramatic decrease towards frontal site [34]. The low performance driving due to drowsiness was evaluated by [113] and observed to have delta, theta, and alpha spectral powers dominance in parietal and occipital regions. The trend of EEG spectrum within brain regions was varied due to difference in experimental protocols. As most of these experiments were carried on simulated driving, there still lags the characterization of EEG spectrum within the brain regions in aviation using a favorable experimental protocol incorporating efficient drowsiness annotations. Our experimental analysis accomplished to fulfill this gap in the literature.

In our experiment of about two hours of simulated flight, delta power ($p<0.05$) had significant decrease in all regions during MS as well as in parietal region for LS, while theta and beta power had no significant change ($p>0.05$). Similar result for delta activity has been reported in a monotonous simulated driving where total driving time was segmented into 10 equal sections to identify changes in EEG spectra. Their results showed significant decrease in sections 6, 8, and 10 compared to section 4 in terms of delta activity [40]. In the literature, alpha activity has been shown to be associated with microsleep [65], [97], [98]. Gharagozlou et al. found significant increase in absolute alpha power between first and the last 10 minutes of driving simulation [114]. Papadelis et al. found a significant increase in relative alpha power on-road experiment for assessment of drowsiness [96].

Additionally, other researchers revealed alpha burst can be related to drowsiness [97], [115]. These studies have comprehensively revealed that alpha is an important drowsiness indicator. In our current study, based on the DMS, we found that relative alpha power showed significant increase in MS (p<0.001, in all regions), and in LS (p<0.05, in all regions except temporal) compared to the BS. Furthermore, the alpha power has a decreasing but insignificant trend in LS compared to MS (Figure 3.1). Our results further state the significance of alpha in drowsiness assessment. In this study, the behavior of delta, theta, alpha, and beta power was primarily related to baseline selection, experimental protocol, and duration of MS event. To investigate the pilot's EEG behavior within the window of circadian low, our experiment was conducted in the early morning and the initial 3-5 minutes of simulated flight was chosen as baseline. The participants in our experiment had been continuously awake for 18-20 hours, therefore they were already prolonged into the early stage of drowsiness before starting the experiment. This is in a good agreement with the results of the KSS assessment (Table 2.1). The average pre-flight KSS value of 5.4 clearly corroborates our statement that the pilots seem to have some signs of sleepiness [109]. Additionally, in the literature, there is no set criteria for "microsleep" and the authors use a variety of definitions and measurements. Therefore, we set sleep duration from 1.5s-15s as MS and greater than 15s as LS in this experiment [85].

The change in a single EEG indicator may not adequately indicate the participant's level of drowsiness. The combination of multiple indicators ensures the reliability to characterize drowsiness. Therefore, in this research we also discussed the behavior of four derived indices which showed the ratio of slow and fast waves [87]. In our study, the alpha power increased significantly during MS events, while theta and beta power had no significant changes, thus

resulting in a significant increase in indices (i) and (ii) (Tables 3.3 and 3.4). This observation corroborates with existing finding in the literature regarding such behavior [40], [116].

To study the interactions between different brain regions, we analyzed coherence via MSC and WTC with the region of aviation headset. We illustrated the potential of neural interactions via WTC between five electrodes within the aviation headset to mark drowsiness from EEG signal. The mean coherences are computed from each electrode pairs during ten interactions in four frequency bands and found to have significant performances from WTC at alpha band. The significant four alpha interactions: T7-Cz, T7-C4, T7-T8, and C3-T8 can be used for drowsiness estimation since alpha is considered as a predictive marker of drowsiness [97]. There exist no significant interactions from MSC. The low frequency bands are more prone to skull muscle artifacts, and the reliability cannot be trusted from the significance of low frequency interactions. Therefore, WTC at alpha can be a suitable coherence estimate to detect drowsiness from the aviation headset.

Changes in the inter-hemispheric coherence with the onset of drowsiness were first reported by [117] and further stated that inter-hemispheric coherence was significantly lower and intra-hemispheric coherence was significantly higher at some frequency spectra during light drowsiness compared to wakefulness [69]. The effect of monotonous driving on inter-hemispheric coherence from five homologous EEG electrode pairs was investigated by [70] and found that frontal and occipital coherence were significantly higher than central, parietal, and temporal sites for all four frequency bands. Increase or decrease of coherence during drowsiness is still unclear. However, in our result, coherence increases mostly during MS with a more significant rise in alpha band from WTC. Alpha spindles are likely to relate attention tasks [108] whose coherence increases during MS state. Higher coherence might increase brain synchronization overloading the brain

structures involved in consciousness processing and preventing them from treating new incoming information [118] that may led to drowsiness. Thus, drowsiness can alter the cerebral functional connectivity between brain regions. However, change in coherence value is found to be bi-directional during LS from the two methods.

In summary, this study not only distinguished MS and LS from baseline, but also investigated the EEG's ability to differentiate LS from MS. The results showed no significant change between LS and MS during both PSD and coherence analysis. A possible cause is that LS duration is not long enough to induce deep sleep. In our experimental protocol, participants will be woken up if they sleep more than 5 minutes and the experiment will be terminated if awakened twice. Due to the short LS duration (15s-300s) and limited sample size, we did not observe significant changes between MS and LS. Our result further showed significant changes between LS and baseline in some of the regions that were also varied significantly during MS compared to baseline.

3.5. Limitations and Future Work

The characterization of EEG in terms of PSD and coherence have posed interesting results useful for drowsiness detection, however, this study has certain limitations which needs to be discussed. Whenever simulated events, machinery or scenarios are used to proxy a real-life scenario, there can be outcomes which are not exact replicas of the real-world. The pilots in this study may be less alert in the simulator than they would be in a real aircraft and may feel less stressed due to lack of life-threatening consequences that are naturally absent in a simulated environment compared to an actual flight. Conversely, the simulator experience and minor control differences between the real aircraft may cause some pilots to show an increase in anxiety, however, this was not measured as a part of this study. In addition, our sample size was limited to final 16

participants. A larger sample size with even demographic distribution would allow greater confidence in generalizing the results to a larger population of commercial pilots. The choice of baseline is also very crucial. In future experiments, baseline should be considered during day to further analyze the behavior of EEG in MS events.

MS is a cognitive state and inferring such state is subjective. It can occur even when a participant's eyes are fully opened [7], but the DMS technology only intends to identify those MS events during eye closure. Moreover, there is a limitation in tracking availability when a participant's face moves outside the field-of-view of the single DMS camera sensor. In this study the mean eyelid tracking availability was 91.99% across the participant population, so it is plausible that any temporary loss of tracking could lead to some MS and LS events going undetected by DMS. To address this problem, the number of cameras can be increased in future experiments to increase the tracking availability further.

In future research, other physiological parameters should be combined to extract additional drowsiness-related features and use effective machine learning methods to study the feasibility of real-time drowsiness monitoring systems during flight. The analysis of all brain regions is narrowed down to the region of central sulcus over the scalp, the region for embedding aviation headset and coherence is analyzed. Then, the feasibility of drowsiness detection and prediction from EEG within the headset region will be studied along with the instrumental design of the headset. The information on PSD analysis during drowsiness can also be found from [66]. The installment of EEG electrodes within the headset will prevent from unnecessary burden during EEG recording via montage. In future, we will expand the work towards directed interactions to estimate the flow of information around the electrode pairs within the aviation headset. Furthermore, the data acquisition will be made from the headset with the specified electrodes and

the feasibility of the method will be studied in real time. In addition to this, interactions of other physiological signals along with EEG will also be analyzed.

3.6. Conclusion

The incorporation of Seeing Machines DMS technology as a reference for drowsiness highlighted the efficacy of EEG spectral behavior in terms of PSD and coherence to characterize drowsiness. This technology overcomes the limitations of subjective drowsiness evaluation and thus, be able to reveal the detailed features of EEG spectral behavior during MS events. Our findings showed changes in EEG spectral behavior during drowsiness. Our proposed method is effective in identifying significant EEG spectra at specific brain regions to discern drowsiness that eventually reduces the number of embedded EEG electrodes. Further, this study also shows the potential use of conventional pilot headsets embedded with limited EEG electrodes pertaining to specific brain regions that can monitor drowsiness in aviation. This will reduce the extra burden of wearing EEG acquisition system by pilots reducing the system's complexity and improving their comfort. Those significant spectra can also be used to develop a drowsiness prediction model.

Chapter 4.

Estimation of Drowsiness and Fatigue Using Heart Rate Variability

4.1. Summary

Background: Drowsiness and fatigue are common workplace problems especially in aviation. Our aim is to determine their association with heart rate variability (HRV) features that can be used for early prediction in aviation pilots. The dynamics of HRV contains pertinent information regarding the sympathetic and parasympathetic arm of the autonomic nervous system (ANS). The psychological states of drowsiness and fatigue are shown to affect the balance of the ANS. Thus, quantification of the behavior of respective arm of ANS can assist early detection. **Methods:** HRV features are derived using electrocardiogram (ECG) recorded non-invasively from 16 commercially rated pilots during a simulated flight. We have computed ten temporal and six frequency domain HRV features during four time periods: T1 (baseline), T2 (relax state), T3 (fatigue state), and T4 (drowsy state) to assess drowsiness and fatigue. **Results:** Three time domain features: heart rate, NN50, and PNN50 exhibited significant changes (p<0.05) during drowsy state whereas, two frequency domain features had significant changes (p<0.05). Low frequency power increased significantly during fatigue state compared to relax state and total power increased significantly during drowsy state compared to baseline. However, LF/HF ratio, a marker of sympathovagal balance did not change significantly during drowsiness and fatigue. **Conclusions:** The results depict that drowsiness and fatigue can be estimated from the HRV features that vary significantly and further helps to prevent aviation accidents from changes in these cardiovascular parameters. The bigger sample size can aid higher confidence in justifying the results.

4.2. Background

International Civil Aviation Organization (ICAO) defines fatigue as a physiological state of reduced mental or physical performance resulting from sleep loss or extended wakefulness, circadian phase disruption, or workload (mental and/or physical activity) that can impair a crew member's alertness and ability to safely operate an aircraft or perform safety related duties [119], [120]. Pilot fatigue is a significant problem in modern aviation operations, largely because of the unpredictable work hours, long duty periods, circadian disruptions, and insufficient sleep that are commonplace in both civilian and military flight operations. Fatigued pilots think and move more slowly, and have memory difficulties which can lead to aviation errors and accidents [8]. FAA classifies pilot fatigue as one of the four common aviation hazards [11] and three out of four aviation accidents result from human errors [121]. Human poorly monitors own mental condition and level of drowsiness or fatigue, therefore often unidentified by pilots about their mental health condition that impairs judgement during flights causing accidents. FAA reported about 247 people died in 209 general aviation accidents from October 2016 through September 2017 and loss of control was one of the cause of these accidents [11].

Drowsiness is a tendency to fall asleep affected by circadian rhythm, workload, monotonous work environment, and liquor or some medicine consumption [122]. Monotonous work environments rapidly induces drowsiness [123] and common in pilots during long duration flights. Drowsiness is different from fatigue. Drowsiness is a feeling of sleep requirement whereas fatigue is a feeling of tiredness or lack of energy. When a person becomes tired, they may prone towards falling sleep. Drowsiness and fatigue can cause many problems such as decreased short term memory, slowed reaction time, increased errors, worsened mood, difficulties in memory, making mistakes, health problems, impaired work performance, and decreased work-private life balance [124]. Drowsiness

or fatigue has been investigated from three approaches: vehicular approach measuring deviations from lane positions, movement of steering wheel, and pressure on the acceleration pedal; behavioral approach measuring driver's behavior, eye closure, eye blinking, and head position; and physiological approach measuring physiological signals non-invasively [3]. Drowsiness and fatigue are closely associated with ANS that can be investigated using HRV. Drowsiness and fatigue can be predicted analyzing physiological signals such as ECG, EEG, EOG, and PPG [125], [126]. Among these, ECG can be a potential marker to predict drowsiness or fatigue based on HRV [3], [43].

HRV is a marker of autonomic nervous system that influences modulations of cardiac activities such that homeostasis of cardiovascular system is properly maintained. Various time and frequency domain methods have been applied to analyze HRV from RR intervals [75]. A few time and frequency domain features have been analyzed for early detection of drowsiness or fatigue [44], [47]. Further, [127] extended time domain features and developed a model for prediction of drowsy accident using multivariate statistical process control.

In this chapter, we analyzed time and frequency domain features in detail assessing fatigue and drowsy state useful for early prediction among aviation pilots from four psycho-physiological states (baseline, relax, fatigue, and drowsy states). The objective of this work is to early assess fatigue or drowsiness from HRV through statistical analysis that helps prevent aviation hazards based on physiological measurements.

4.3. Results

The data is recorded from eighteen commercially rated pilots during WOCL. One pilot's data was interrupted due to technical problems and other three pilot's ECG data suffered from high noise, so they are not considered for analysis. ECG data recorded from fourteen pilots (age: 22±2.77

years, 13 males, 1 female, previous flight experience: 293.14±88.26 hours) are used for our analysis. Time and frequency domain features are assessed in terms of mean±SE in four psychophysiological states illustrated in Figures 4.1 and 4.2.

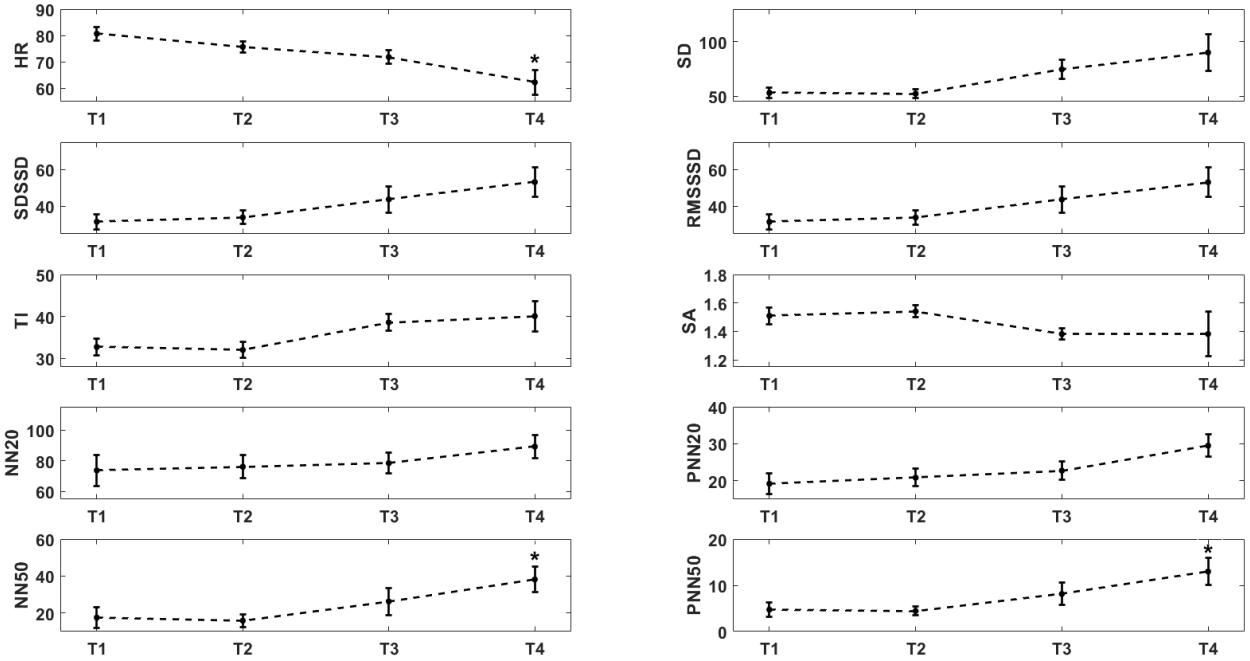


Figure 4.1. The trend (mean±SE) of time domain features based on HRV in four experimental time periods: T1, T2, T3, and T4. The x-axis represents four experimental time periods and y-axis represents each time domain feature values. '*' represents significant difference (p=0.05) from T1.

The values of mean±SD during four experimental time periods are shown in Table 4.1. The results depicted that three time and two frequency domain features changed significantly (p=0.05) during fatigued and drowsy states as compared to baseline and relax states. More specifically, HR decreased significantly (p<0.05) at T4 compared to T1, NN50 and PNN50 increased significantly (p<0.05) during T4 compared to T1 in time domain. Other seven time domain features changes noticeably but are not found significant (p<0.05). In frequency domain, absolute LF and TP increased significantly at T3 and T4 compared to T2 and T1 respectively. Other frequency domain

features did not change significantly (p=0.05). The LF/HF ratio, a marker of sympathovagal balance corresponding to ANS, increases at T3 and T4 compared to T1 but not significant (p=0.05).

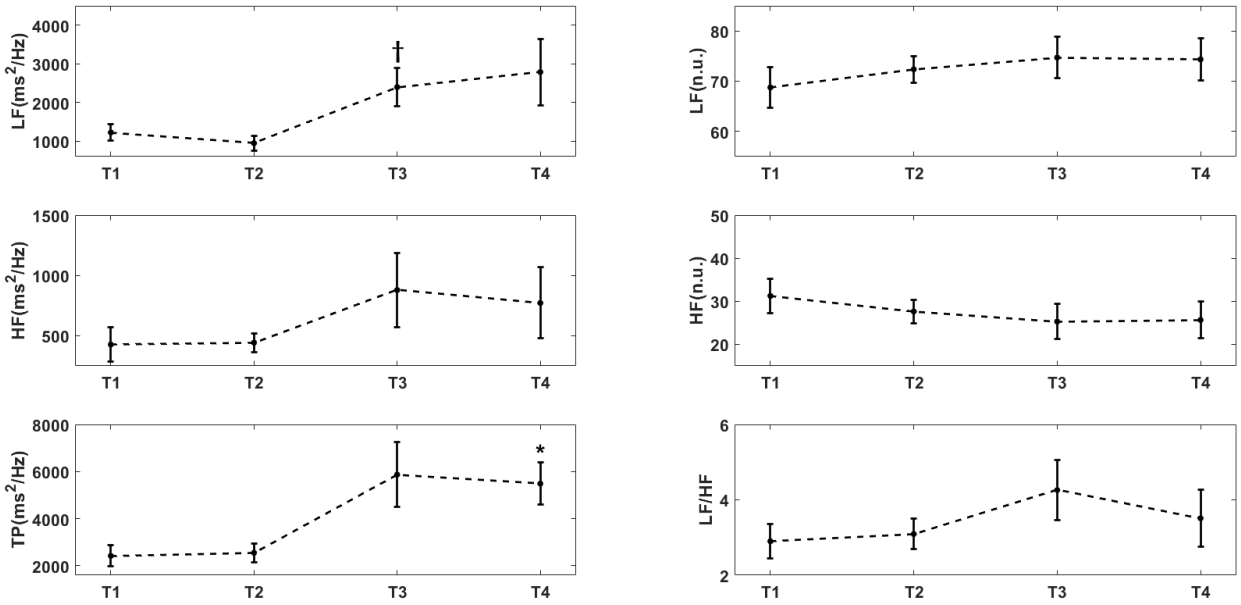


Figure 4.2. The trend (mean±SE) of frequency domain features based on HRV in four experimental time periods: T1, T2, T3, and T4. The x-axis represents four experimental conditions and y-axis represents each frequency domain feature values. '*' and '†' represent significant difference (p=0.05) from T1 and T2 respectively.

Table 4.1. Mean±SD values in four experimental time periods using HRV features. '*' and '†' represent significant difference (p=0.05) from T1 and T2 respectively.

Experimental Time Periods				
Features	**T1**	**T2**	**T3**	**T4**
HR	80.7±9.64	75.71±8.1	71.85±9.43	**62.26±11.28***
SD	53.09±17.92	52.28±15.48	74.64±32.23	90.07±40.69

SDSSD	31.6±15.61	33.96±13.88	43.82±26.65	53.26±19.93
RMSSSD	31.56±15.59	33.91±13.86	43.75±26.6	53.18±19.89
TI	32.75±7.55	32.06±7.02	38.59±7.25	40.1±8.74
SA	1.51±0.22	1.54±0.16	1.38±0.15	1.38±0.38
NN20	73.71±37.26	76.07±28.15	78.43±25.27	89.17±18.4
PNN20	19.21±10.6	20.85±8.88	22.75±9.2	29.49±7.42
NN50	17.5±21.38	15.86±12.51	26.29±27.08	**38.17±17.1***
PNN50	4.72±5.97	4.42±3.62	8.12±9.05	**12.97±7.23***
LF (abs)	1223.39±810.21	945.15±720.22	**2396.79±1874.14†**	2786.28±2098.23
LF (nu)	68.75±14.9	72.37±9.96	74.71±15.39	74.34±10.4
HF (abs)	427.05±525.43	443.09±295.25	879.09±1157.9	771.6±722.14
HF (nu)	31.25±14.9	27.63±9.96	25.29±15.39	25.66±10.4
TP (abs)	2420.81±1677.17	2549.67±1493.33	5870.54±5151.46	**5486.65±2182.86***
LF/HF	2.9±1.73	3.09±1.51	4.26±2.95	3.51±1.84

4.4. Discussion

Drowsiness and fatigue are assessed among pilots from HRV based on the variation of time and frequency domain features obtained from RR intervals. The window of 5 minute duration was taken at each four time periods to obtain an appropriate HRV results [75]. The mean value of time and frequency domain features are compared statistically at four experimental conditions: T1, T2, T3, and T4. The significant changes in those features can be used as a marker for drowsiness or fatigue prediction.

Mental stress, fatigue, and drowsiness are closely associated with ANS and their relationship with HRV has been studied from the earlier date [128]. Changes in sympathetic and parasympathetic activity induces alternations in heart rate. The mean heart rate decreases during drowsiness or fatigue [47] due to longer intervals of heart beat which ultimately increases standard deviation among the RR intervals. Slower heart rate increases triangular index which is obtained from the total number of RR intervals with respect to the number of most frequently occurring RR intervals. Sample entropy is a measure of rate of new information production [129] and low entropy is observed at drowsy and fatigued states reflecting high degree of similarity among RR intervals due to slower heart rate causing longer RR intervals.

Wakefulness is marked by an increase of sympathetic and/or a decrease of parasympathetic activity, while extreme relaxation is characterized by an increase of parasympathetic and/or decrease of sympathetic activity [44]. Drowsiness and fatigue levels are associated with sympathovagal balance corresponding to ANS, however, the relationship found so far is inconsistent. Some authors claimed that LF/HF ratio increases during drowsiness, however, others claimed that the ratio decreases [43]. In our result, increased LF (n.u.) relating mainly to sympathetic activity may be associated with stressed arousal as pilots are compelling themselves to stay awake that ultimately induces fatigue. As pilots are awake from the previous day of experiment, longer awake time may increase the level of fatigue due to which LF/HF ratio is increased at T3 though the change is insignificant. Then, the ratio decreases slightly during drowsy state. The reason might be the pilots fell less stressed during the drowsy state. The significance level might be increased using larger sample sizes. From the results, we observed that time and frequency domain indicators can be a reliable measure to assess drowsiness and fatigue from ECG.

The HRV features can also be acquired from PPG similar to ECG. More particularly, this phenomenon in PPG is called pulse rate variability (PRV). PPG is another non-invasive but optical technique containing an infrared emitter and a detector inside a probe that reflects beat to beat relative blood volume changes in peripheral tissues usually on the forefinger or earlobe [50]. The amount of infrared light reflected back to the detector determines the amount of blood flowing to the tissue at any time. So, the pulsating PPG signal depicts blood volume fluctuations observed from a finger or earlobe. The literature shows that HRV features obtained from ECG and PPG are found to have high correlation [50], [53]–[55]. Thus, PPG can be a substitute of ECG for HRV analysis. In our result, we obtained significant decrease of heart rate during drowsy state compared to baseline. This motivated to calculate heart rate from PPG. The PPG was acquired from earlobe with the same experimental setting as ECG. The selection of earlobe for PPG data acquisition was to determine feasibility of sensor placement within the confines of aviation headset. Similar to RR intervals, PP intervals are also calculated and heart rate is also obtained from PPG. All the analysis of PPG are similar to ECG. The heart rate at all the four experimental time periods using ECG and PPG is found to have high correlation depicted in Figures 4.3 and 4.4. The results show high correlation between ECG and PPG and also heart rate show a consistent decline from T1 to T4 with statistical significance. The results also show a collection of PPG data from earlobe may be a feasible alternative to other more complex forms of physiological signals. Thus, PPG is simpler to acquire and can be an alternative of ECG for HRV analysis.

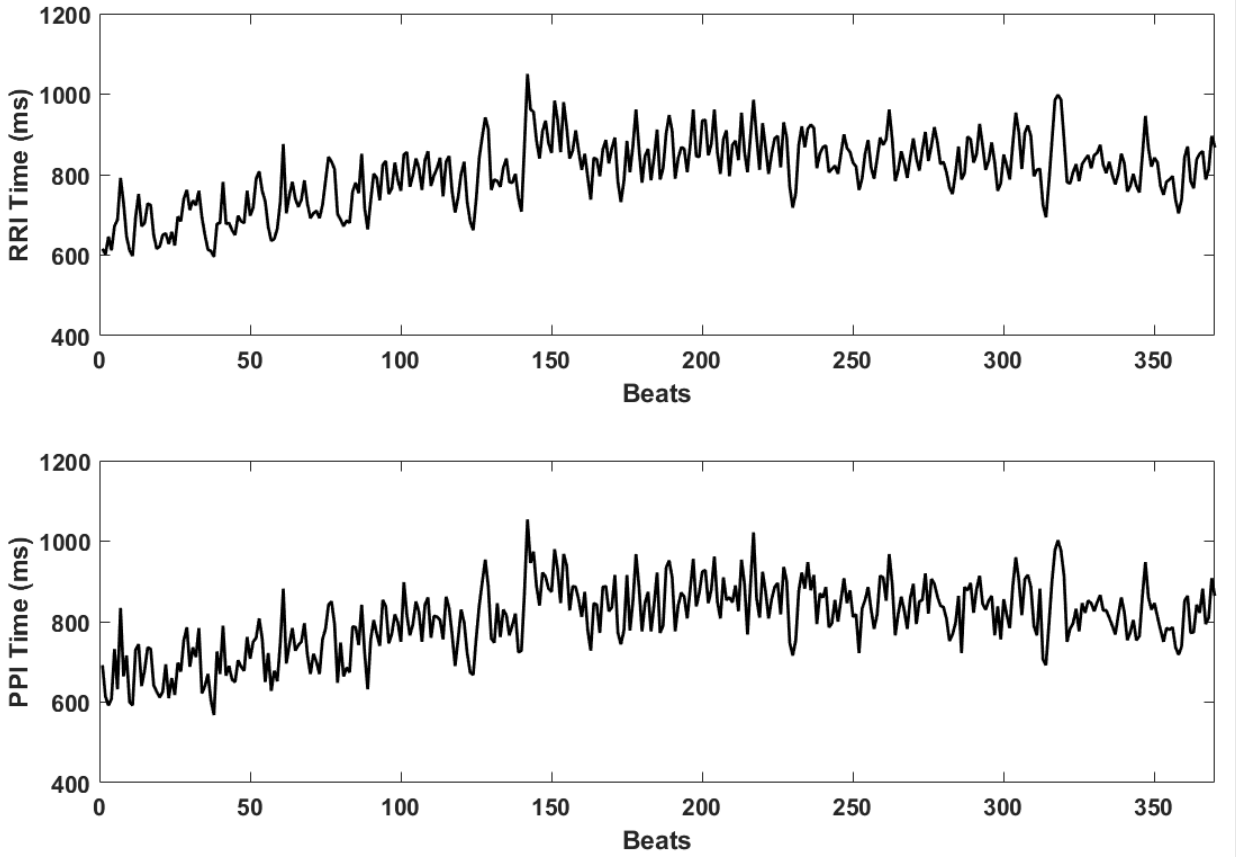


Figure 4.3. Illustration of RR intervals (RRI) and pulse-to-pulse intervals (PPI) per beats for the duration of 5 min baseline period (T1) from a representative participant. Graphs are placed vertically for comparison depicting a high correlation between RRI and PPI [80].

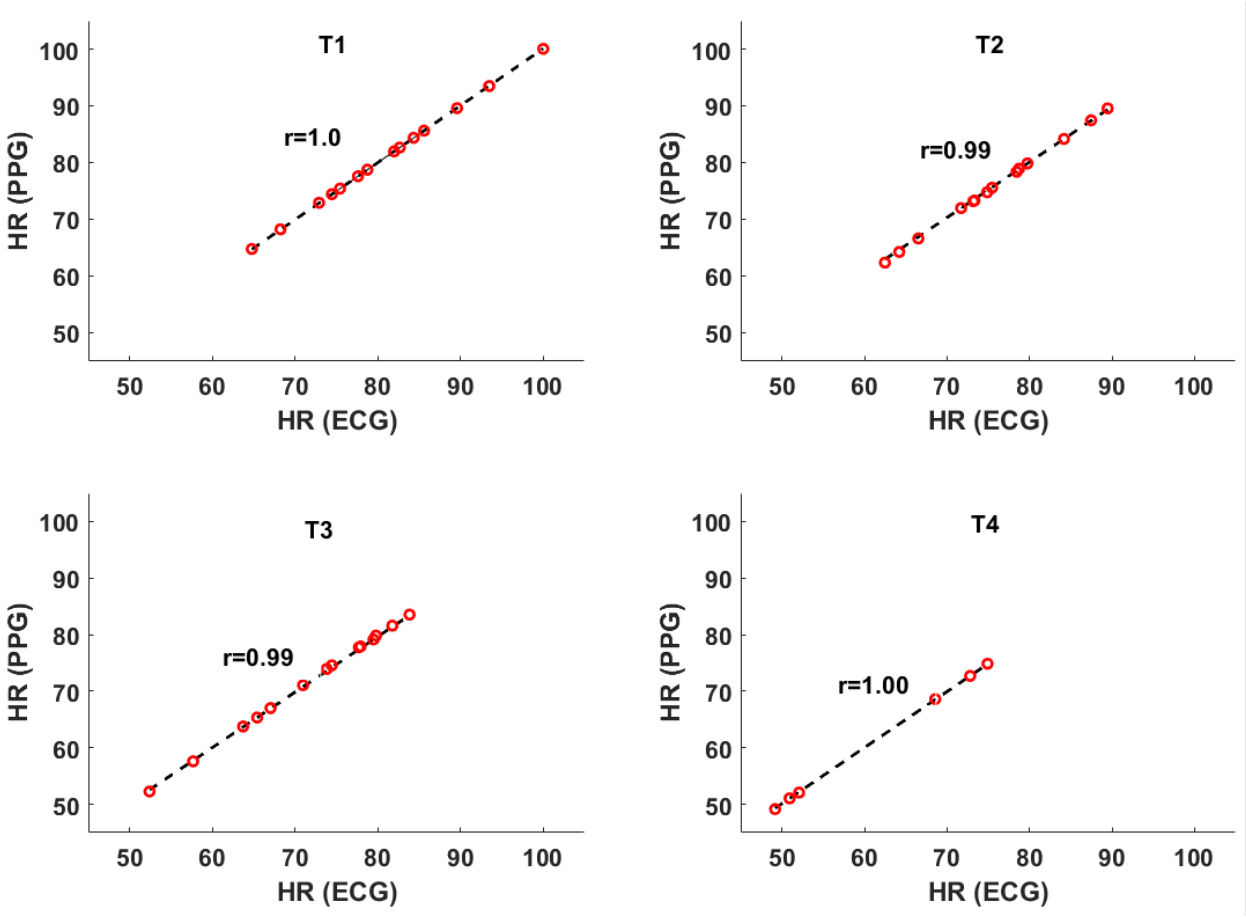


Figure 4.4. Pearson correlation coefficient of heart rate (HR) between electrocardiogram (ECG) and photoplethysmogram (PPG) for a representative participant [80].

4.5. Limitations and Future Scope

This research provides a basis to characterize drowsiness from cardiac signals and there are several limitations that need to be considered. The availability of higher sample size may provide greater confidence to validate the results. The results are based on flying simulator and greater practicability of the results could have been obtained if the data was acquired from the headset within the actual aircraft considering noises and other physical disturbances. HRV can vary on different environmental conditions apart from drowsiness and fatigue. Those conditions may include extensive exercise, stressful condition, sickness condition, and smoking or alcohol

consumption. Apart from these, age and gender differences, athletes and non-athletes, geographical locations, and personal temperament are some other factors that need to be considered on to estimate the results from these cardiovascular parameters over the wide range of populations.

The use of physiological signals such as ECG and PPG are common in medical fields to monitor the physiological states of a patient. Despite their prevalence in medicine, there exists opportunity to expand their usage in aviation operations. More importantly, the use of ECG can be supplemented by a simpler PPG signal and the PPG sensor can be placed on either earlobe or similar soft tissue embedded within an aviation headset. In future, we will recruit more participants to acquire data in real time from the cockpit in real flights using the headset. The results obtained from so could strengthen the claim for feasibility to monitor drowsiness and fatigue based on changes in cardiovascular parameters.

4.6. Conclusion

HRV features in time and frequency domain show significant changes during drowsiness and fatigue. Our flight simulating experiment corroborates a possibility of non-invasive method in real time for early prediction of drowsiness and fatigue based on HRV. This further directs towards the development of smart technologies using sensors acquiring physiological signals to monitor drowsiness or fatigue so that aviation hazards resulting from drowsiness or fatigue can be mitigated.

Chapter 5.

Drowsiness classification using Electroencephalogram: the feasibility within aviation headset

5.1. Summary

Background: Drowsiness is a psycho-physiological transition from awake towards sleep. During this transition, there are short episodes of eye-closure events called microsleeps (MS). The detection and early prediction of such MS events are crucial in aviation industry. This chapter presents a method for classification of MS from baseline utilizing linear and non-linear features derived from electroencephalography (EEG) and an approach towards early prediction. **Methods:** The EEG data is acquired from sixteen commercially-rated pilots during the window of circadian low (2:00 am- 6:00 am). MS events are annotated using Driver Monitoring System and further verified using vertical electrooculogram (EOG). A total of 55 features are extracted from EEG. A subset of these features is then selected using a wrapper-based method. The selected features are fed into a linear or quadratic discriminant analysis (LDA or QDA) classifier to classify baseline from MS states. Further, the signals from the region of aviation headset are used to predict drowsiness using Random Forest (RF) Classifier. In RF classifier, features are selected based on feature ranking procedure. **Results:** The overall classification performance of the best- proposed discriminant classifier is 86.23% in terms of F1 score and the prediction performance using RF classifier within headset electrodes is 83.80% in terms of F1 score. **Conclusion:** The results obtained highlight the potential of proposed method towards automatic drowsiness detection which could assist mitigating aviation accidents in future and also the plausibility of developing automatic prediction model within the region of aviation headset.

5.2. Background

As stated in earlier chapters, drowsiness is a serious issue in airlines and requires early prediction. Among the three main approaches (vehicular, behavioral, and physiological) for drowsiness detection, physiological approach is only dependent on participants' physiological condition [34]. The physiological signals used in drowsiness analysis are EEG, ECG, EOG, and PPG. Among them, EEG has been widely used due to easiness in signal recording and the brain dynamics during drowsiness can be revealed from EEG [33], [34]. Thus, EEG has been considered as the most effective predictive marker [17], [130].

Efficient drowsiness detection and prediction has been a major goal in the transportation industry. Various drowsiness detection techniques have been developed based on driver's behaviors and performances incorporating physiological signals as well. The use of Random Forest for detecting drowsiness related lane departures [131], deep learning for drowsiness detection from EEG [132], a long short-term memory (LSTM) recurrent neural networks using EEG [133], convolutional two-stream network using multi-facial feature fusion [134], neural network classifier with discrete wavelet EEG features [135], and Fischer's discriminant classifier coupling EEG and NIRs (near-infrared spectroscopy) are a few machine learning techniques implemented for drowsiness analysis. Table 5.1 briefly lists the recent EEG based drowsiness detection systems. The drowsiness detection system has also been recorded from a single individual [136] although the reliability remains questionable. The observation from Table 5.1 depicts that drowsiness can be estimated from non-hair bearing areas [106], even from a single channel [137], [138] and PSD is found to be the most useful features with Support Vector Machine (SVM) classifier.

Table 5.1. The list of drowsiness detection systems with various classifiers varying different number of electrodes.

Sub. #	Electrodes	Features	Classifier	Accuracy	Ref.
1	One channel (PO6)	Signal peak and variance	SVM	93.52%	[136]
10	Non-hair bearing areas (F7, F8, A1, & A2)	Logarithmic band power	SVM	80%	[106]
10	Two channels (O1 and C3)	Power spectral density (PSD)	ANN	89%	[139]
20	Seven channels (P1-7)	PSD	Least-squares SVM	84.1%	[101]
10	Eight channels (FP1-2, C3-4, P7-8, & O1-2)	PSD	Multilayer Neural Network	96.7%	[140]
23	Scalp-EEG (C3, C4) & Ear-EEG (Ch1, Ch2)	Multi-scale fuzzy entropy	SVM	88.8% (Scalp EEG), 82.9% (Ear-EEG)	[141]

43	Thirty-two channels	PSD	Sparse-deep belief networks	93.1%	[142]
25	Four channels (C3, C4 & O1, O2)	Expert based (ratio, amplitude, symmetry, and extension)	Bayesian-copula discriminant classifier	94.3%	[143]
22	One channel (TP7)	Entropy	Decision tree	94%	[137]
12	One channel (CP4)	Entropy	Random Forest	96.6%	[138]
20	Two channels (O1 and O2)	EEG Power Percentage	SVMPPM	91.92%	[144]
30	Eight channels	PSD of DWT	ANN	95%	[135]

Note: SVM=Support Vector Machine, ANN=Artificial Neural Network, SVMPPM=SVM Posterior Probabilistic Model

Most of the drowsiness experiments are conducted using simulated cars [125], [145] and more opportunity exists to research drowsiness during flight operations. In this chapter, we implement two tasks: (i) algorithms based on Bayesian decision theory and feature selection methods for drowsiness classification approach using five brain regions, and (ii) algorithms based on Random Forest classifier and feature ranking procedure for drowsiness prediction within the region of aviation headset. For the first task, linear and non-linear features from EEG are extracted from five brain regions: frontal (F), central (C), parietal (P), occipital (O), and temporal (T) during baseline

(BS) and microsleep (MS) period. A set of high discriminative features are selected using different wrapper-based feature selection algorithms. The selected features are then fed into linear or quadratic discriminant analysis (LDA or QDA) classifiers. The Bayesian classifier with optimal performance is proposed. For the second task, the same linear and non-linear features from EEG as the first task are extracted from five EEG electrodes: T7, C3, Cz, C4, and T8 within the region of aviation headset during baseline (BS) and before microsleep (BMS) periods. The features are selected using Random Forest classifier in a feature ranking procedure. Then, the subset of highest significant features are used for prediction.

5.3. Feature Engineering and Model Assessment

5.3.1. Evaluation Criteria

The algorithm performance is assessed using the following metrics: sensitivity, specificity, positive predictive value, accuracy, and F1 score. Sensitivity (also called recall) measures the proportion of correctly classified positive samples from the true positive samples. Specificity measures the proportion of correctly classified negative samples from true negative samples. Positive predictive value (also known as precision) measures the proportion of correctly classified positive samples from positive classified samples. The accuracy measures the proportion of correctly classified samples from all samples. The F1 score is the harmonic mean of precision and recall and is a useful criterion in class imbalance problem. Thus, it is adopted as the overall performance criteria. These matrices are defined based on the contingency table shown in Table 5.2. and derived from [83]. Here, $N_{i,j}$ represents the total number of samples of class i (true label) classified as class j (predicted label). In this case $i, j \in \{\mathrm{BS}, \mathrm{MS}\} \equiv \{1, 2\}$. The total number of true samples R_i in class i and the total number of predicted samples C_j in class j can be expressed as:

$$R_i = \sum_{j=1}^{2} N_{i,j} \text{ and } C_j = \sum_{i=1}^{2} N_{i,j} \quad (5.1)$$

The total number of samples is:

$$N_{tot} = \sum_{i=1}^{2} \sum_{j=1}^{2} N_{i,j} = \sum_{i=1}^{2} R_i = \sum_{j=1}^{2} C_j \quad (5.2)$$

Table 5.2. The contingency table for two-way classification of drowsiness among baseline (BS) and microsleep (MS).

		Predicted Label		
		BS	**MS**	
True Label	**BS**	$N_{1,1}$	$N_{1,2}$	R_1
	MS	$N_{2,1}$	$N_{2,2}$	R_2
	$\sum$	C_1	C_2	N_{tot}

Note: Numerical class equivalencies are: 1≡BS (baseline), 2≡MS (microsleep)

The true positive (TP), false positive (FP), false negative (FN), and true negative (TN) for class *i* can be generalized as follows:

$$TP_i = N_{i,i}$$

$$FP_i = C_i - N_{i,i}$$

$$FN_i = R_i - N_{i,i}$$

$$TN_i = N_{tot} - TP_i - FP_i - FN_i = N_{tot} - R_i - C_i + N_{i,i} \quad (5.3)$$

Thus, the performance metrics for class *i* are obtained as:

$$Sen_i = \frac{TP_i}{R_i} = \frac{N_{i,i}}{R_i}, \quad PPV_i = \frac{TP_i}{C_i} \text{ and } F1_i = \frac{2 * Sen_i * PPV_i}{Sen_i + PPV_i} \quad (5.4)$$

5.3.2. Feature Extraction

For linear and non-linear analysis of EEG, features such as power spectral density (linear), Shannon entropy, sample entropy, and Renyi entropy (non-linear) are extracted from 13 channels and grouped into five brain regions as well into the region of aviation headset. All the features from each channel are obtained using a 4 sec window. Non-drowsy EEG is obtained from the baseline starting from 2 min and the signal is cut with a 4 sec segment sliding with 3 sec overlap for 20 iterations in each participant. The baseline of 3 min duration is not completely used to maintain the similar size of baseline and drowsy segments. For drowsy EEG segments, MS events greater than 5 sec is chosen and 4 sec signal is taken after 1 sec of MS onset. For the second task of prediction within the headset region, the 4 sec segments before every MS events are selected.

The PSD is obtained from Welch's method having a Hamming window size of 128 samples with 50% overlap and 256 discrete Fourier transform points. Then, the relative spectral power in four bands is calculated, i.e., delta (δ, 0.5-4 Hz), theta (θ, 4-8 Hz), alpha (α, 8-13 Hz), and beta (β, 13-30 Hz). The next set of features are the following indices derived from EEG spectra: (θ+α)/β, α/β, (θ+α)/(α+β), and θ/β. Shannon entropy is a measure of probability distribution reflecting the predictability of signal behavior. Sample entropy describes the rate of new information production and is used to quantify the complexity of EEG. Sample entropy can be represented as SampEn (s,m,r) where the symbols are represented as (s: input signal), (m: embedding dimension or maximum template length), and (r: tolerance window) [146], [147]. In our experiment, we choose m=2, and r=0.1 from the literature [148]. Renyi entropy is considered as a superclass of ShanEn and is given by RenEn (s, q) where the symbols are represented as (s: input signal), and (q: order of entropy). When q approaches 1, RenEn converges to ShanEn [146].

5.3.3. Feature Selection and Model Assessment

We use two approaches for feature selection and model assessment with Bayesian classifiers LDA and QDA and with Random Forest (RF) classifier. A nested cross-validation (CV) is used for feature selection in Bayesian classifiers and in RF classifiers, features are selected based on the reduction of entropy.

For Bayesian classifiers, a nested CV is used with leave-one-out participant-wise external CV for model assessment (i.e. estimation of the performance of the classification algorithms) and 5-fold internal CV for feature selection. The wrapper method is designed for feature selection based on the performance of an inner classifier. In this work, the inner classifier is the same type as the main classifier in the external CV loop. In the inner CV loop, a set of features with minimum misclassification rate (1-F1) is selected and pass to the external CV loop for model assessment. In addition, the sequential forward and backward selection (SFS and SBS) search algorithms are used to find the best features subset in the internal CV loop. SFS starts from an empty feature set and sequentially add features until the minimum criterion is met, whereas SBS starts from full features set deleting each feature sequentially [149].

For prediction purpose within the aviation headset, a RF classifier is used to rank all the handcrafted 55 features based on the decreasing order of entropy. The out of bag (OOB) feature importance profile is depicted in Figure 5.1. The OOB error is a way of predicting error in RF utilizing bootstrap aggregation (bagging) to sub-sample the training data. The more description of the classifier is in classification section below. A subset of ranked features with varying sizes are selected to analyze the classifier performances.

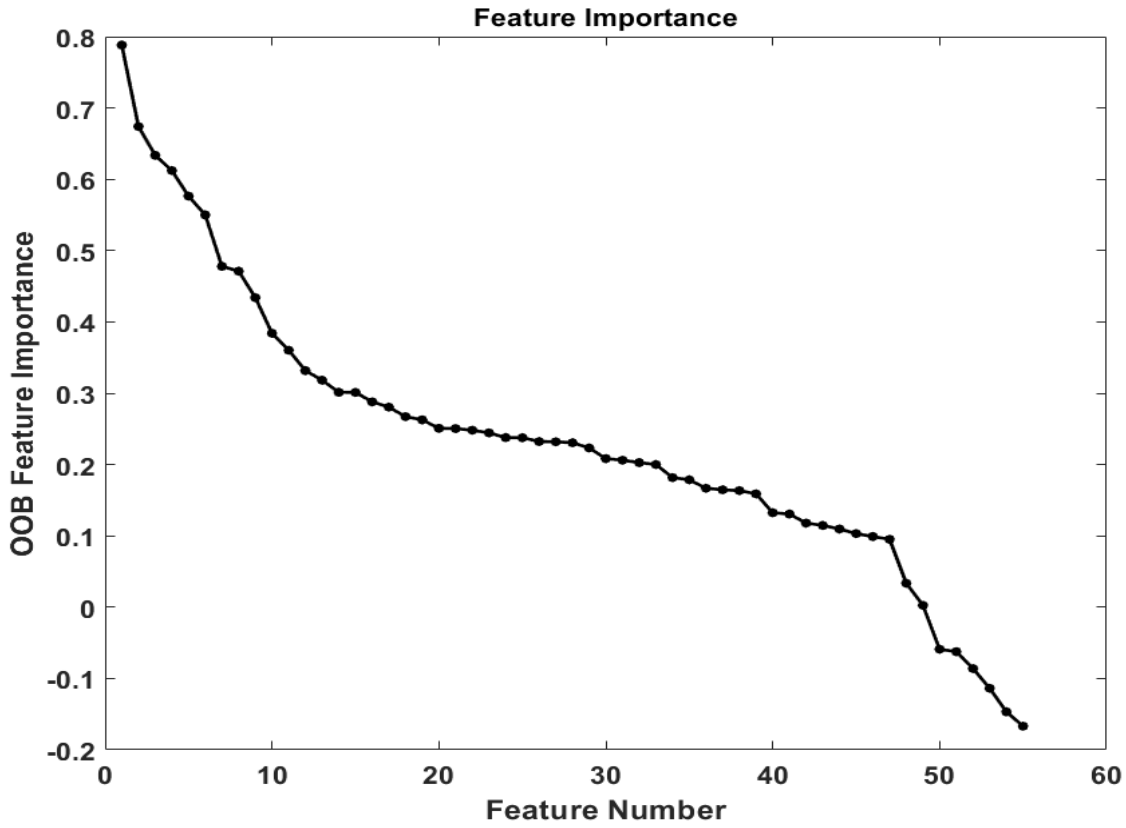


Figure 5.1. Feature ranking based on the decreasing order of entropy from Random Forest classifier with 300 decision trees and unit leaf size using uniform prior.

5.3.4. Classification

For the purpose of Bayesian classification, we use the same type of classifier for both the inner and outer CV loops. In our work, we use LDA and QDA classifiers whose class conditional densities are modeled with multivariate normal distributions with equal covariance matrices in LDA and unequal covariance matrices in QDA. Moreover, we assume to have a uniform class prior. The combination of these two classifiers (LDA and QDA) and two feature selection methods (SFS and SBS) results in four algorithms: LDA+SFS, LDA+SBS, QDA+SFS, and QDA+SBS that we use to evaluate the performance. The method implemented here is obtained from [149].

For the purpose of prediction within the region of aviation headset, RF classifier is used to classify two psyho-physiological states: baseline (BS) and before-microsleep (BMS) from the selected top 15 features in a feature selection procedure. A random forest is an ensemble of decision trees. An ensemble is a group or combination of multiple machine learning algorithms to obtain better predictive performance than any of the individual algorithms. During training, multiple decision trees are constructed. These individual decision trees turn out to be different because only a subset of features are given to each of them, using a technique called bootstrap aggregation or bagging. This is a technique for reducing variance of an estimated prediction function. During prediction, mode of the classes of individual decision trees constitute the output of the RF [81]. To obtain suitable number of trees and leaf size in RF, OOB error is computed with 500 trees and varying leaf sizes 1, 5, 10, 20, 50, and 100 that is shown in Figure 5.2.

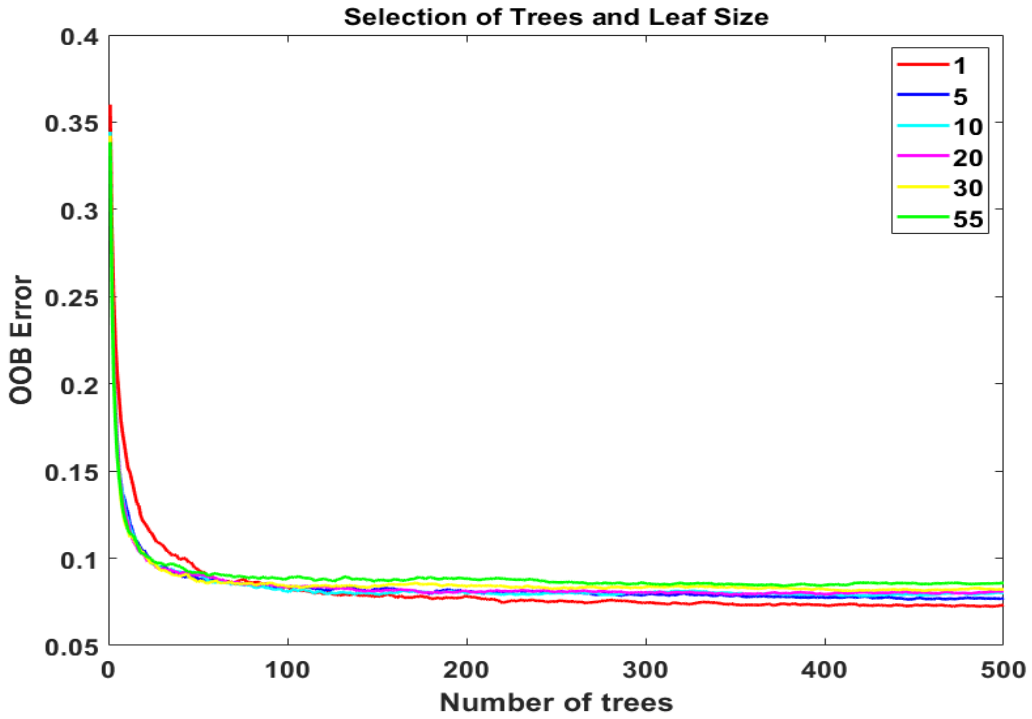


Figure 5.2. Out-of-bag (OOB) error with 500 decision trees and varying number of leaf sizes: 1, 5, 10, 20, 30, and 55 in a Random Forest classifier.

From Figure 5.2, the error in the classifier is least with a unit leaf size and is consistent after 300 number of trees. Thus, a RF classifier with 300 decision trees and 1 leaf size is selected. Each decision trees in the RF are trained with bootstrapped samples and 5 (square root of 15 and up rounding results 5) features are selected at random for each node. Then, the entropy measure decides the specific feature to split at each node.

5.4. Results

Data from sixteen participants (age: 21.5±2.4 year, 2 females, 14 males, previous average flight experience: 282 hours) were used for both classification within five brain regions and prediction from the region of aviation headset. Out of eighteen participants, two were not considered for data analysis because data recording was interrupted for one participant due to technical problems and no MS events in another. A total of 830 EEG segments (320 non-drowsy and 510 drowsy or before drowsy observations) are extracted from the aviation dataset for both classification and prediction purposes.

The performances of four different algorithms for classification task using Bayesian approach are depicted in Table 5.3. The confusion matrix of the algorithm with the highest overall F1 score (QDA+SBS) is illustrated in Table 5.4 where rows represent the true class and columns represent the classifier prediction. The list of 13 selected features from the highest performance algorithm is provided in Table 5.5 which are annotated as ShanEn (brain region), RenEn (brain region), and PSD (EEG band or derived indices, brain region) to illustrate a feature from a specific brain region for ShanEn, RenEn and also at specific frequency band in case of PSD.

Table 5.3. The performances of four algorithms (LDA+SFS), (LDA+SBS), (QDA+SFS), and (QDA+SBS) used for classification.

Used Algorithm	Sen (%)	Spe (%)	PPV (%)	Acc (%)	F1 (%)	# Features
LDA+SFS	75.50 (25.16)	88.43 (27.24)	89.61 (23.41)	81.25 (20.27)	77.05 (22.99)	10
LDA+SBS	76.49 (21.43)	82.18 (28.28)	84.86 (24.62)	76.62 (14.26)	74.88 (16.65)	10
QDA+SFS	83.62 (18.04)	86.56 (27.36)	89.20 (18.93)	84.63 (15.02)	83.62 (14.95)	12
QDA+SBS	81.54 (16.49)	92.81 (15.80)	94.58 (9.52)	87.22 (12.16)	**86.23 (9.84)**	13

Note: Here, sensitivity (Sen), specificity (Spe), positive predictive value (PPV), accuracy (Acc), and F1 score (F1) are calculated from the selected number of features (# Features) in each algorithm. Two sequential feature selection methods: sequential forward selection (SFS), and sequential backward selection (SBS) are adopted. The numbers are the unweighted average of the results in the external CV loop together with their standard deviations in parentheses.

Table 5.4. The confusion matrix of QDA+SBS classification algorithm. Here, 'ND' represents Non-drowsy and 'D' represents drowsy states.

		Predicted class	
		ND	**D**
True Class	**ND**	**297**	**23**
	D	**95**	**415**

Table 5.5. The list of 13 selected features from the algorithm (QDA+SBS).

PSD (θ, C)	**PSD (α/β, C)**	**PSD ((θ+α)/β, P)**
PSD ((θ+α)/(α+β), P)	**PSD (δ, O)**	**PSD (θ, O)**
PSD (β, O)	**PSD ((θ+α)/(α+β), O)**	**PSD (θ, T)**
PSD (α, T)	**ShanEn (F)**	**ShanEn (C)**
RenEn (T)	---	---

For prediction task, the classification performances of RF classifier with 300 decision trees and 1 leaf size in leave-one-participant-out cross-validation architecture varying the number of selected features from 10 to 55 among two classes baseline (BS) and before-microsleep (BMS) is shown in Table 5.6. The average best performance in terms of F1 is 84.01% in classifying positive samples obtained from the top selected 15 features. Further, the classifier show the potential in detecting both positive (before- drowsiness) and negative (baseline) with 84.23% and 83.79% respectively which is fairly an equal performance. The top 15 selected features comprising the best performance results are listed in Table 5.7. The features in the table are illustrated in terms of feature name (brain region) or feature name (EEG spectra, brain region).

Table 5.6. The performance of Random Forest classifier with varying number of selected features.

#Features	**BS**			**BMS**		
	Sen (%)	**PPV (%)**	**F1 (%)**	**Sen (%)**	**PPV (%)**	**F1 (%)**
10	73.44	70.57	71.98	79.88	82.07	80.96
15	**76.25**	**75.08**	**75.66**	**83.37**	**84.23**	**83.80**

20	75.31	75.08	75.20	83.57	83.74	83.66
25	74.69	74.69	74.69	83.37	83.37	83.37
30	74.06	74.29	74.18	83.16	82.99	83.08
35	75.94	73.64	74.77	82.14	83.86	82.99
40	72.19	71.74	71.96	81.31	81.65	81.48
45	74.06	73.15	73.60	82.14	82.82	82.47
50	74.38	72.78	73.57	81.72	82.92	82.32
55	72.50	70.95	71.72	80.49	81.67	81.08

Table 5.7. The list of 15 top ranked features obtained from Random Forest classifier.

Rank	**Features**	**Rank**	**Features**
1	PSD (δ, F)	9	PSD (β, O)
2	PSD (β, F)	10	PSD (α, T)
3	PSD (α/β, F)	11	PSD (β, T)
4	PSD (δ, C)	12	PSD ((θ+α)/(α+β), T)
5	PSD (α/β, C)	13	PSD (θ/β, T)
6	PSD (δ, P)	14	RenEn (F)
7	PSD (θ, P)	15	SampEn (T)
8	PSD (δ, O)	---	---

5.5. Discussion

This chapter describes classification of drowsiness using two approaches: classification of microsleep (short episodes of eye-closure) from baseline (alert state) among five brain regions and classification of before-microsleep (4 sec before eye-closure) from baseline for prediction task within the region of aviation headset. Two main tasks of classification are detecting the microsleep sates from the whole brain region and more importantly studying the feasibility of EEG electrodes within the region of aviation headset for prediction purposes. We obtain overall F1 score of 86.23% in classifying microsleep from baseline using QDA+SBS and a similar performance of 83.80% in predicting samples before eye-closure using RF classifier. The selected features used are also similar, 13 in case of QDA+SBS and 15 in case of RF. One problem with the former classifier i.e. QDA with SBS is that it takes longer time to select the features in backward selection approach where features in RF are selected based on the reduction of entropy and takes comparatively lesser time for execution. It is wise to avoid sequential feature selection techniques (SFS or SBS) during feature selection from large features set and fairly better to use procedures from RF.

The results demonstrate the possibility of automatic drowsiness detection among pilots using the hand-engineered feature and machine learning algorithms. This method could be potentially applied to intelligent monitoring and warning systems to prevent drowsiness or fatigue. The classification performances of four algorithms using Bayesian approach are compared using a nested CV architecture. The best performance in terms of F1 score (86.23%) belongs to QDA+SBS possibly because QDA has greater flexibility than LDA due to its quadratic decision surface. The classes in QDA have different co-variance matrix and assuming the data distribution as Gaussian, the decision boundary between two classes is quadratic in nature incorporating various sample attributes thereby increasing the classification performance.

The analysis is extended constricting the electrodes within the region of aviation headset and to speed of the execution time RF classifier is used in a feature ranking architecture based on reduction of entropy. Further, the analysis of EEG segments are done 4 sec ahead of eye-closure to explore the potential of EEG features to predict drowsiness using RF classifier. The predictive performance is obtained as 83.80% in terms of F1 score using top ranked 15 features. The result depicts the possibility towards early drowsiness prediction before eye-closure from the electrodes within the aviation headset. The performance of headset electrodes on classifying microsleep from baseline is also observed using RF and observed as 83.50% in terms of F1 score. The performance on detecting and predicting mircosleep from cut datasets within the aviation headset is found to be similar. The prediction performance has more significant usage in aviation industry and this approach motivates towards embedding EEG electrodes underneath the aviation headset useful for predicting drowsy state during flights.

5.6. Limitations and Future Work

The classification of before-microsleep with baseline with an acceptable performance has posed an important notion for early prediction of microsleep events 4 sec ahead of eye-closure from the cut datasets which is crucial for timely warning to the pilots during flight operations. Although, this study drives toward the early prediction approach, there are a few limitations. The experiment is conducted in a simulated environment and the results may vary during real flights due to different environmental conditions such as noises and levels of concentration. The cockpit noise and multiple motion artifacts can hinder the signal quality that affects the drowsiness detection performances. The results shown in this chapter are based on the cut datasets and the performance may unlikely follow the similar trend with the continuous EEG signals in real time scenarios.

In future, we will increase the sample size, compare with other classifiers, and expand the analysis from cut datasets to drowsiness annotation from continuous EEG signal developing a drowsiness annotation engine to predict drowsiness in real time. Eventually, we will record the signals from the aviation headset in real flights and annotate the drowsy segments from the developed automatic drowsiness annotation engine. The other physiological signals such as ECG, PPG, and EOG will also be incorporated into the machine learning model for continuous drowsiness detection. Lastly, the baseline (non-drowsy) data collection would happen outside of participant's window of circadian low, ideally during daylight hours when it is expected that the drowsiness levels would be lower.

5.7. Conclusion

The classification of microsleep from baseline using Bayesian classifier and classification of before-microsleep events using Random Forest classifier examine the feasibility of machine learning algorithms to detect drowsiness from EEG. The use of Random Forest classifier classifying before-microsleep events depict the plausibility of predicting microsleep events 4 sec ahead of eye-closure with 83.8% in terms of F1 score from the cut dataset using the selected 15 hand-engineered features. The results demonstrate the feasibility for early prediction using EEG data and leaves the room for further improvement. The concept of automatic drowsiness detection from continuous signals is discussed and the development of the system with reliable performance is yet to achieve. The proposed method could be potentially useful for developing warning systems to prevent drowsiness or fatigue in aviation operations and further be extended to other driving or flying related tasks.

Chapter 6.

Four-way Classification of Drowsiness: An Approach towards Early Prediction

6.1. Summary

Background: Drowsiness is a psycho-physiological transition from awake towards sleepiness. During this transition, there are short episodes of eye-closure events called microsleeps (MS). The early prediction of such MS events are crucial in aviation industry. This chapter presents a four-way classification of drowsiness among baseline (BS), before microsleep (BMS), during microsleep (MS), and longsleep (LS) states utilizing linear and non-linear features derived from electroencephalogram (EEG) and provides an approach towards early prediction. **Methods:** The EEG data is acquired from sixteen commercially-rated pilots during the window of circadian low (2:00 am- 6:00 am). MS events are annotated using Driver Monitoring System and further verified using vertical electrooculogram (EOG). A total of 555 features from time, frequency, and time-frequency domains as well as non-linear features are extracted from EEG separated into five brain regions: frontal, central, parietal, occipital, and temporal. The features are selected in a feature ranking procedure. Random Forest classifier with 300 decision trees and unit leaf size is used to classify the test data from the top selected 40 hand-engineered features into the mentioned four classes using a leave-one-out cross-validation architecture. **Results:** The unweighted mean of F1 (UMF1) score is used as an overall performance criterion and the classifier obtains UMF1 to be 83.70%. The multiway accuracy and unweighted mean of sensitivities of the classifier are achieved as 82.55% and 83.75%. Further, F1 score of before microsleep class is obtained as 81.45%. **Conclusion:** The results show the possibility of predicting drowsiness before eye-closure that helps to design an early drowsiness prediction system to reduce aviation accidents.

6.2. Background

As stated in earlier chapters, efficient drowsiness detection and prediction has been a major goal in the transportation industry. Various drowsiness detection techniques have been developed based on driver's behaviors and performances incorporating physiological signals as well. Most of the analysis are binary and [58] has performed both binary and multi-way classification considering alert, somewhat sleepy, and sleepy states with accuracy of 93% for binary class and 79% for multiclass using EEG, EOG, and contextual information during simulated car driving. Most of the drowsiness experiments are conducted using simulated cars [145], [150] and more opportunity exists to research drowsiness during flight operations.

The drowsiness detections and predictions are not interchangeable. There are many published research papers where prediction and detection of microsleep are considered similar [151]. Many researchers have analyzed drowsiness segments and claimed as prediction [152]–[154]. The prediction for drowsiness should be analyzed before the onset of eye-closure. The capability of predicting drowsiness with high reliability aids in preventing transportation related accidents. In this study, the instances of microsleep, a short episode of eye-closure, are extracted from Seeing Machine Driver Monitoring System that tracks participants' eye movements and label drowsiness based on the duration of eye-closure. This chapter addresses the issue of classifying before microsleep with multiple classifications from the cut datasets with acceptable performance and also provides a concept for early drowsiness prediction from physiological signal that is applicable in real flight operations.

6.3. Feature Engineering and Model Assessment

6.3.1. Evaluation Criteria

Sensitivity (Sen), positive predictive value (PPV), and F1 score of each class is used to evaluate the performance of the classifier. Sensitivity (also called recall) measures the proportion of correctly classified positive samples from the true positive samples. Positive predictive value (also known as precision) measures the proportion of correctly classified positive samples from positive classified samples. The accuracy measures the proportion of correctly classified samples from all samples. The F1 score is the harmonic mean of precision and recall and is a useful criterion in class imbalance problem. Thus, it is adopted as the overall performance criteria. These matrices are defined based on the contingency table shown in Table 6.1 and extracted from [83]. Here, N_{ij} represents the total number of samples of class i (true label) classified as class j (predicted label). In this case $i, j \epsilon \{BS, BMS, MS, LS\} \equiv \{1,2,3,4\}$. The total number of true samples R_i in class i and the total number of predicted samples C_i in class j can be expressed as:

$$R_i = \sum_{j=1}^{4} N_{i,j} \text{ and } C_j = \sum_{i=1}^{4} N_{i,j} \tag{6.1}$$

The total number of samples is:

$$N_{tot} = \sum_{i=1}^{4} \sum_{j=1}^{4} N_{i,j} = \sum_{i=1}^{4} R_i = \sum_{j=1}^{4} C_j \tag{6.2}$$

Table 6.1. The contingency table for four-way classification of drowsiness among baseline (BS), before microsleep (BMS), during microsleep (MS), and longsleep (LS).

		Predicted Label				
		BS	BMS	MS	LS	
True Label	BS	$N_{1,1}$	$N_{1,2}$	$N_{1,3}$	$N_{1,4}$	R_1
	BMS	$N_{2,1}$	$N_{2,2}$	$N_{2,3}$	$N_{2,4}$	R_2
	MS	$N_{3,1}$	$N_{3,2}$	$N_{3,3}$	$N_{3,4}$	R_3
	LS	$N_{4,1}$	$N_{4,2}$	$N_{4,3}$	$N_{4,4}$	R_4
	∑	C_1	C_2	C_3	C_4	N_{tot}

Note: Numerical class equivalencies are: 1≡BS (baseline), 2≡BMS (before microsleep), 3≡ MS (microsleep), and 4≡ LS (longsleep)

The true positive (TP), false positive (FP), false negative (FN), and true negative (TN) for class *i* can be generalized as follows:

$$TP_i = N_{i,i}$$
$$FP_i = C_i - N_{i,i}$$
$$FN_i = R_i - N_{i,i}$$
$$TN_i = N_{tot} - TP_i - FP_i - FN_i = N_{tot} - R_i - C_i + N_{i,i} \quad (6.3)$$

Thus, the performance metrics for class *i* are obtained as:

$$Sen_i = \frac{TP_i}{R_i} = \frac{N_{i,i}}{R_i}, \quad PPV_i = \frac{TP_i}{C_i} \text{ and } F1_i = \frac{2*Sen_i*PPV_i}{Sen_i+PPV_i} \quad (6.4)$$

Further, the major performance measures to summarize the whole confusion matrix are multiway accuracy (*MulAcc*), unweighted mean of sensitivities (UMS), and unweighted mean of F1 score (UMF1). These parameters are obtained as:

$$MulAcc = \frac{1}{N_{tot}}\sum_{i=1}^{4} N_{i,i}, \quad UMS = \frac{1}{4}\sum_{i=1}^{4} Sen_i, \text{ and } UMF1 = \frac{1}{4}\sum_{i=1}^{4} F1_i \quad (6.5)$$

Multiway accuracy measures the total accuracy and unweighted mean of sensitivities measures the total sensitivity of the classifier. The multiway accuracy is sensitive to class imbalance and unweighted mean of sensitivities measure only recall. Therefore, unweighted mean of F1 (UMF1) score is used as a global performance measure of the classifier that measure both precision and recall and is unaffected by class imbalance problem.

6.3.2. Feature Extraction

The EEG features are extracted in time, frequency, time-frequency, and non-linear domains from 13 channels grouped into five brain regions during four psycho-physiological states: BS, BMS, MS, and LS. All the features from each channel are obtained using a 4 sec window. The EEG for baseline is obtained starting from 2 min and the signal is cut with a 4 sec segment sliding with no

overlap as the same number of MS events in each participant. The baseline of 3 min is not completely used to maintain the similar size of BS, BMS, and MS segments. For MS segments, MS events greater than 5 sec is chosen and 4 sec signal is cut after 1 sec of MS onset and the 4 sec segment just before the onset of MS is extracted for BMS segments. The LS segments are also obtained similar to MS after 1 sec of onset. Some of the participants do not have LS events, so the sample size of LS events are less than MS. The features in all four domains are described here.

6.3.2.1. Time Domain Features

Temporal features are obtained from wavelet packet decomposition at depth 7 with Daubechies5 (db5) wavelets using Shannon entropy. The four EEG band-related signals: delta (δ, 0.5-4 Hz), theta (θ, 4-8 Hz), alpha (α, 8-13 Hz), and beta (β, 13-30 Hz) are extracted from the decomposed signal. The statistical information such as mean, standard deviation, root mean square, skewness, kurtosis, median, mode, inter-quartile range, first quartile, third quartile, and Hjorth parameters (activity, mobility, and complexity) are also computed. Hjorth parameters are based on variance of the derivatives of a signal and are activity, mobility, and complexity that are defined as follows [155].

$$Activity = var(x(t))$$

$$Mobility = \sqrt{\frac{Activity\ (\frac{dx(t)}{dt})}{Activity\ (x(t))}}$$

$$Complexity = \frac{Mobility\ (\frac{dx(t)}{dt})}{Mobility\ (x(t))} \tag{6.6}$$

where, $x(t)$ is a time series data. Activity is the signal power, mobility is the mean frequency, and complexity is the change in frequency of the signal obtained from time series signal.

6.3.2.2. Frequency Domain Features

The frequency domain features are obtained from power spectral density (PSD) calculated using Welch's method having a Hamming window size of 128 samples with 50% overlap and 256 discrete Fourier transform points. The features include absolute power, relative power, power ratios, and median frequency. Absolute power corresponds to power in each EEG frequency bands (delta, theta, alpha, and beta) and relative or normalized power is the ratio of the power of each frequency bands to the total spectral power whereas power ratio is the ratio of two or more normalized powers. The four power ratios derived from normalized EEG spectra include (θ+α)/β, α/β, (θ+α)/(α+β), and θ/β. The median frequency quantifies the relative strength of low and high frequency oscillations. It is the frequency value that separates the frequency range of PSD into two equal halves [67].

6.3.2.3. Time-Frequency Features

The time-frequency features used for analysis are discrete wavelet transform coefficients (DWT), and energy, and homogeneity of continuous wavelet transform (CWT) coefficients. DWT coefficients are obtained using wavelet decomposition with Daubechies5 (db5) wavelet and 7 levels of decomposition using Shannon entropy. Only four DWT coefficients (d2, d3, d4, and d5) are extracted that corresponds to the frequency range of four basic EEG spectra: δ, θ, α, and β. The CWT coefficients are obtained using continuous wavelet transform with Morlet wavelet and energy and homogeneity are calculated as:

$$Energy = \sum_{i,j} p(i,j)^2$$
$$Homogeneity = \sum_{i,j} \frac{p(i,j)}{1+|i-j|} \quad (6.7)$$

where, $p(i,j)$ is the probability of bin (i,j) in CWT space. The energy measures uniformity of CWT coefficients and homogeneity measures the closeness of the distribution of coefficients in CWT space.

6.3.2.4. Non-Linear Features

The Hurst exponent and three entropies (Shannon, Sample, and Renyi) are used as non-linear features for classification. The Hurst exponent (H) is used to measure the presence or absence of long-range dependence and its degree in time series. It characterizes non-linear behavior of EEG [90]. The H value is obtained as:

$$H = \frac{log(\frac{R}{S})}{log(T)} \tag{6.8}$$

where, T is the duration of sample data, R is the difference between maximum deviation from the mean and minimum deviation from the mean, and S is the standard deviation. Shannon entropy is a measure of probability distribution that reflects the predictability of signal behavior. Sample entropy describes the production rate of new information that quantifies EEG complexity. In our analysis, we choose embedding dimension of 2 with 0.1 tolerance window for sample entropy [148]. Renyi entropy is a superclass of Shannon entropy and when order of the entropy approaches to 1, Renyi entropy converges to Shannon entropy.

All the features from 13 channels are grouped into five brain regions that ultimately results a total of 555 features. All these features are feed to the classifier for training and assessment.

6.3.3. Feature Selection

A random forest classifier is used to rank all the handcrafted 555 features based on the decreasing order of entropy. The out of bag (OOB) feature importance profile is depicted in Figure 6.1. The

OOB error is a way of predicting error in RF utilizing bootstrap aggregation (bagging) to sub-sample the training data. The more description of the classifier is in classification section below. A subset of ranked features with varying sizes are selected to analyze the classifier performances.

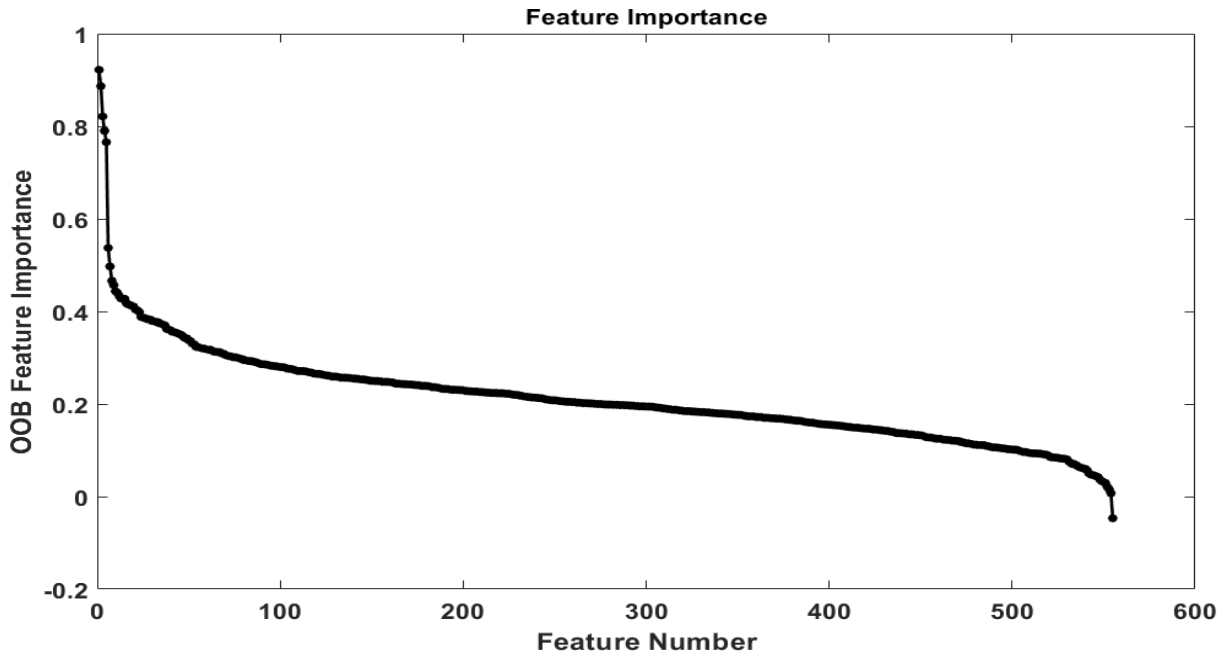


Figure 6.1. Feature ranking based on the decreasing order of entropy from Random Forest classifier with 300 decision trees and unit leaf size among four classes using uniform prior.

6.3.4. Classification

Random forest (RF) classifier is used to classify the four psycho-physiological states: BS, BMS, MS, and LS from the selected top 40 features in the feature selection procedure. A random forest is an ensemble of decision trees. An ensemble is a group or combination of multiple machine learning algorithms to obtain better predictive performance than any of the individual algorithms. During training, multiple decision trees are constructed. These individual decision trees turn out to be different because only a subset of features are given to each of them, using a technique called bootstrap aggregation or bagging. This is a technique for reducing variance of an estimated

prediction function. During prediction, mode of the classes of individual decision trees constitute the output of the RF [81].

6.4. Results

Data from sixteen participants (age: 21.5±2.4 year, 2 females, 14 males, previous average flight experience: 282 hours) were used for four-way classification. Out of eighteen participants, two were not considered for data analysis because data recording was interrupted for one participant due to technical problems and no MS events in another. A total of 1,650 EEG segments (BS=488, BMS=488, MS=488, and LS=186) under four conditions are extracted from the aviation dataset. The performance of the classifier in each class is assessed in terms of Sen, PPV, and F1 score and the performance of the classifier among all the four classes is analyzed from MulAcc, UMS, and UMF1. Moreover, UMF1 is used as a global performance measure.

The classification performances of RF classifier with 300 decision trees and 1 leaf size in leave-one-participant-out cross-validation architecture varying the number of selected features from 10 to 200 is shown in Table 6.2. The best performance in terms of UMF1 is 83.70% obtained from the top selected 40 features. Further, the classifier has the potential of classifying BS, BMS, MS, and LS with performance of 83.79%, 81.45%, 79.46%, and 90.08% respectively in terms of F1 score.

Table 6.2. The performances of Random Forest classifier with varying number of selected features

NumFeat	MulAcc (%)	UMS (%)	UMF1 (%)	BS			BMS			MS			LS		
				Sen (%)	PPV (%)	F1 (%)	Sen (%)	PPV (%)	F1 (%)	Sen (%)	PPV (%)	F1 (%)	Sen (%)	PPV (%)	F1 (%)
10	77.52	79.33	79.06	79.51	77.60	78.54	78.28	77.02	77.64	70.29	74.24	72.21	89.25	86.46	87.83
20	80.55	81.64	81.60	82.79	82.28	82.53	78.07	80.04	79.05	78.07	76.81	77.44	87.63	87.17	87.40

40	**82.55**	**83.75**	**83.70**	**85.25**	**82.38**	**83.79**	**80.53**	**82.39**	**81.45**	**78.89**	**80.04**	**79.46**	**90.32**	**89.84**	**90.08**
50	80.97	82.17	82.16	83.20	80.40	81.77	79.71	82.24	80.96	77.05	77.37	77.21	88.71	88.71	88.71
60	80.55	81.97	81.85	82.58	79.49	81.01	79.10	81.61	80.33	76.43	77.55	76.99	89.78	88.36	89.07
100	79.88	80.58	80.79	82.17	81.17	81.67	75.41	79.65	77.47	80.33	76.26	78.24	84.41	87.22	85.79
150	78.67	79.31	79.56	79.92	80.58	80.25	73.98	78.14	76.00	80.53	74.86	77.59	82.80	86.03	84.38
200	78.79	79.91	79.90	79.71	79.55	79.63	73.77	79.82	76.68	80.12	74.62	77.27	86.02	86.02	86.02

Note: The sensitivity (Sen), positive predictive value (PPV), and F1 score (F1) are computed in each class out of four classes: baseline (BS), before microsleep (BMS), microsleep (MS), and longsleep (LS). Further, multi-way accuracy (MulAcc), unweighted mean of sensitivities (UMS), and unweighted mean of F1 score (UMF1) are calculated from the overall classes.

The confusion matrix of the classifier with best performance results from the ranked 40 features are illustrated in Table 6.3 where the row represents true class and column represents predicted class.

Table 6.3. The confusion matrix of the classifier with best results from the selected features

		Predicted Class			
		BS	**BMS**	**MS**	**LS**
True Class	**BS**	**416**	**20**	**52**	**0**
	BMS	**32**	**393**	**44**	**19**
	MS	**57**	**46**	**385**	**0**
	LS	**0**	**18**	**0**	**168**

The top 40 selected features comprising the best performance results are listed in Table 6.4. The features in the table are illustrated in terms of feature name (brain region) or feature name (EEG spectra, brain region).

Table 6.4. The list of 40 selected features using Random Forest classifier from feature ranking based on entropy.

Rank	Features	Rank	Features	Rank	Features	Rank	Features
1	homogeneity (T)	11	q1 (d2, O)	21	rms (d3, P)	31	PSD ((θ+α)/β, P)
2	homogeneity (P)	12	q1 (β, O)	22	norm_PSD (α, O)	32	PSD ((θ+α)/β, O)
3	homogeneity (F)	13	mobility (δ, F)	23	abs_PSD (α, O)	33	rms (β, O)
4	homogeneity (C)	14	mobility (β, O)	24	abs_PSD (α, P)	34	iqr (β, O)
5	homogeneity (O)	15	std (d2, O)	25	mean (α, T)	35	q1 (α, O)
6	PSD (α/β, P)	16	activity (β, O)	26	mobility (β, C)	36	complexity (θ, T)
7	PSD (α/β, O)	17	std (d3, P)	27	mobility (β, P)	37	mobility (δ, T)
8	abs_PSD (β, O)	18	mobility (β, T)	28	rms (d2, O)	38	complexity (β, T)
9	iqr (d2, O)	19	q3 (β, O)	29	kurtosis (d3, O)	39	kurtosis (α, P)
10	norm_PSD (α, P)	20	kurtosis (d3, P)	30	kurtosis (α, O)	40	rms (d2, P)

6.5. Discussion

This chapter describes the four-way classification of drowsiness into alert (baseline), before falling into drowsiness (before microsleep), short drowsiness episodes (microsleep) and long eye-closure events (longsleep). The two main tasks of classification in this study are detecting microsleep and more importantly predicting before the occurrence of microsleep. From the cut datasets, the classifier achieves F1 scores of 81.45% and 79.46% respectively for classifying before microsleep and during microsleep. The F1 score is observed to be higher in detecting before microsleep due to comparatively higher positive predictive value in the class. The classifier performance for classifying before microsleep events strengthen the claim that drowsiness can be detected with 81.45% of F1 just before 4 sec of the occurrence of microsleep events among all the four drowsiness conditions. This helps to design a drowsiness prediction system. This work is an extension of our previous two works. The binary classification of drowsiness in [156] is extended

to four classes in this work and the comparison of EEG power spectral analysis among three states: baseline, microsleep, and longsleep [66] is extended here in the discussion to four states including beforemicrosleep events.

Out of the features extracted in time, frequency, and non-linear domains, spectral power changes in EEG frequency bands: delta, theta, alpha, and beta are common in literature to estimate the shift of psychological state from alertness towards drowsiness [96], [110]. The alpha has been observed as the most relevant drowsiness indicator within the spectrum. However, there are contradicting views regarding the spectral changes in different brain regions due to traditional approaches for drowsiness annotations [65], [97]–[100]. In our previous work [66], we compared three psychological states mentioned before and observed significant changes in delta and alpha during microsleep and longsleep states. In this work, the power spectral features are extracted in four frequency bands and compared among four psychological conditions: BS, BMS, MS, and LS illustrated in Figure 6.2.

The Kruskal-Wallis test followed by Bonferroni correction was used to calculate the significant difference at α=0.05 between the four states in the four EEG spectrum. On analyzing the power spectral analysis among four psychological states, we observed more significant changes in delta and alpha bands.Even in drowsiness related research, delta and alpha are considered as primary indicators. The relative delta power decreases significantly (p<0.05) during BMS and MS compared to BS among all the five brain regions. There is significant rise (p<0.05) in relative delta power during LS compared to BMS and MS among central, parietal, and occipital brain regions. The decrease in delta power is observed more profound during BMS than MS and then increases during LS. The alpha, another primary drowsiness indicator, also exhibits profound changes during four drowsy states, however, possess antagonistic behavior with respect to delta. The relative alpha

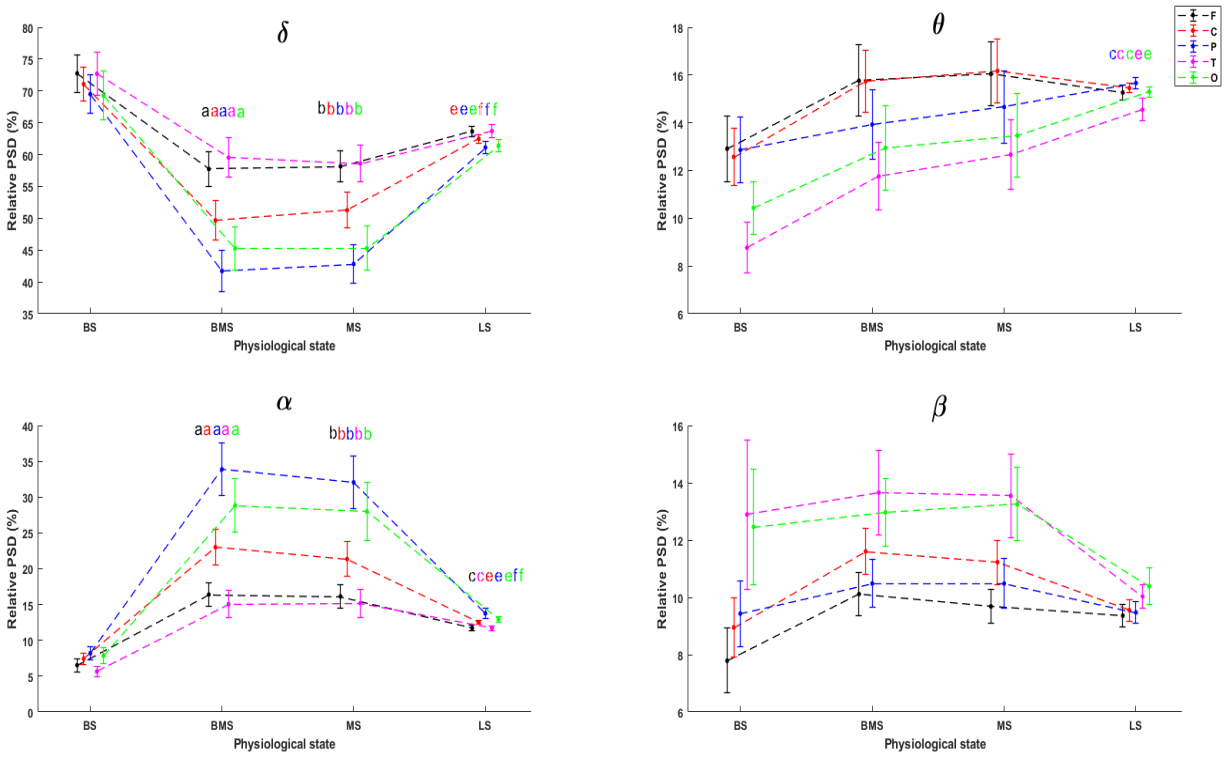


Figure 6.2. Relative spectral power distributed (mean±SE) in four EEG frequency bands: δ, θ, α, and β under four psychological states (BS, BMS, MS, and LS) in five brain regions: F, C, P, T, and O. The x-axis represents four psychological states and y-axis represents relative power spectral density. The significant difference among the four states at p=0.05 is represented by a (BS-BMS), b (BS-MS), c (BS-LS), d (BMS-MS), e (BMS-LS), and f (MS-LS). The color of these symbols represent significant changes (p<0.05) at particular brain region.

power increases significantly (p<0.05) during BMS and MS compared to BS among all the five brain regions. There is also significant rise in alpha power during LS compared to BS at frontal and temporal regions, and majorly in parietal and occipital regions during LS compared to BMS and MS. Theta exhibits a few significant rise (p<0.05) during LS, however, beta does not show any significant changes at all. There are high significant changes (in terms of p value) during BMS than MS compared to BS that justify the claim that drowsiness occurs before eye-closer.

Drowsiness is a progressive psycho-physiological phenomenon that starts before eye-closure and the effect may result to eye-closure due to which it can be detected before eye-closing state [4], [7]. The probable reason for high significant changes before eye-closure could be due to high neural activities just before falling towards drowsiness since the participants are struggling to keep them awake. The capability of EEG in detecting drowsiness before eye-closure helps in the development of physiological signals based early drowsiness prediction system useful for human performance related operations in aviation and road transportations.

The automatic drowsiness prediction system detects drowsiness before eye-closure from the continuously acquired EEG signals after testing the continuous signals of the participants at work from the trained classifier model. This approach can be extended for predicting drowsiness in real time from the modified aviation headset. The 13 channel EEG electrodes used for data acquisition can be modified and EEG electrodes can be embedded underneath the aviation headset used by the pilots in cockpit that incorporates electrodes within the central sulcus of the brain.

The classifier detecting microsleep even though the participants were in baseline state is termed as false positives. The occurrence of more number of false positives is a serious problem for implementation of automatic drowsiness prediction system for aviation operations. One probable reason for such false positives would be that the participants were fatigued on arriving at the experimental session late night at 2 am since they were requested to avoid sleep and even naps at least 18 hours prior to the experiment. The other reason might be the inefficiency of Seeing Machines Driver Monitoring System to capture eye-closure moments when participants are away from the focus of camera. During such situations, the participants were fatigued and drowsy whereas, the true labels were not calibrated as drowsy events. This would result a more number of false positives reducing the performance of automatic drowsiness prediction system.

6.6. Limitations and Future Work

The four-way drowsiness classification has posed an important notion for early prediction of microsleep events 4 sec ahead of eye-closure from the cut datasets which is crucial for timely warning to the pilots during flight operations. Although, this study drives toward the early prediction approach, there are a few limitations. First, the experiment is conducted in a flight simulator although the participants were commercially-rated pilots and the results may differ during real flight scenarios. Second, the pilot's attentiveness may differ between simulated and real flights and may prone towards frequent drowsiness due to safer simulated environment. And, the third limitation is that the results are obtained from a cut dataset and may differ when the training model is implemented in a continuous EEG signals during real flights. The cockpit noise and multiple motion artifacts can hinder the signal quality that affects the drowsiness detection performances. The results are presented from a limited number of 16 participants out of which 2 were females. A larger number of participants with even demographics would allow greater confidence in generalizing results to a population of commercial pilots. The initial time for choosing baseline also affects the results since the participants were prone towards fatigue at the beginning of the experiment as they remained awake at least 18 hours prior.

As discussed above, Seeing Machines Driver Monitoring System has incapability of detecting eye-closure when the participants face is away from the focus of a camera. During this time, the participants can be nodding their head down due to drowsiness, and the system is incapable to assure they are drowsy. Therefore, in future, this problem can be resolved installing more cameras covering 180^0 plane of both horizontal and vertical sights so that the bowed head will be covered. The other physiological signals such as ECG, PPG, and EOG will also be incorporated into the machine learning model for continuous drowsiness detection system. More importantly, the

analysis of brain regions will be constricted around the central sulcus of the brain, an area underlying aviation headset, to study the feasibility of using modified aviation headset for continuous drowsiness detection and prediction. Apart from feature extraction and classification, the interactions and causal influences between brain and heart signals could also be analyzed to understand the effect of cardiac functions during different states of drowsiness.

6.7. Conclusion

The four-way classification results using Random Forest classifier depict the plausibility of predicting microsleep 4 sec ahead with 81.45% of F1 score from the cut dataset using the selected 40 hand-engineered features. Further, the classifier has an overall performance of 83.70% in terms of UMF1. The results demonstrate the feasibility for four-way classification into baseline, before microsleep, during microsleep, and longsleep with a reasonable performance and leaves the room for further improvement. This paper also discusses the power spectral changes during four states providing a notion of more spectral phenomena before eye-closure that aids in design of early prediction systems. The proposed method could be potentially useful for developing warning systems to prevent drowsiness or fatigue in aviation operations and further be extended to other driving or flying related tasks.

Chapter 7.

Conclusions and Future Directions

7.1. Conclusions

Drowsiness is a short momentary period of sleep that lacks consciousness and can cause severe accidents during driving vehicles or flying airplanes. The main challenge is to characterize drowsiness, more importantly predict drowsiness before the eye-closure so that the losses can be minimized. Although there are various techniques available to estimate drowsiness levels, the analysis based on changes in physiological parameters of human system is a reliable one and can be adopted in any environmental situations. The common physiological signals used in this thesis to characterize drowsy events are EEG, ECG, and PPG. EOG is used here as an additional marker of eye-closure events along with Seeing Machines DMS technology. This thesis demonstrates the usefulness of these physiological signals to characterize, detect, as well as predict drowsy events in a simulated flight environment during the WOCL. The major conclusions drawn from the thesis are outlined below.

7.1.1. EEG Spectral Behavior Can Characterize Drowsiness

The study on EEG spectral behavior during drowsiness aids to understand the neural activity of the brain when a participant prone towards sleep. On studying the changes in four spectral bands: delta, theta, alpha, and beta; one can understand the most influential spectra and that helps further to expand the work for the development of reliable automated drowsiness detection system. During our work, we observe significant (p<0.05) reduction in delta power and increment in alpha power in all five brain regions: frontal, central, parietal, temporal, and occipital during microsleep. This

experimental result depicts the plausibility of EEG spectral power to characterize drowsy state. Our long term objective is to develop an automated system incorporating signals within the region of aviation headset. For this, we also tested the interactions between five EEG electrodes: T7, C3, Cz, C4, and T8 underlying the headset region via MSC and WTC and observed no interactions at all via MSC. However, alpha interactions were significantly higher (p<0.05) via WTC during drowsiness computed among four electrode pairs (T7-Cz, T7-C4, T7-T8, and C3-T8). These results on EEG spectra computed via PSD and coherence support the hypothesis to characterize drowsiness via recorded EEG signals among pilots and further the plausibility for extending this work using aviation headset in real flight scenarios. To corroborate our claim, we need to include more number of participants for data recording.

7.1.2. HRV Can Be an Estimator during Drowsiness Detection

HRV is a homeostatic marker for autonomic state of a body obtained from ECG analysis. Mental fatigue and drowsiness are closely associated with ANS and HRV can be used to study the physiological state during drowsiness and fatigue. ANS consists of sympathetic activity (SN) and parasympathetic activity (PSN). Wakefulness is marked by an increase in SN and/ or a decrease in PSN, whereas, drowsiness is characterized by an increase in PSN and/ or decrease in SN. The changes in SN and PSN induces alterations in heart rate that further changes various HRV parameters in time and frequency domain. In our result, we observed significant rise (p<0.05) in HR, NN50, and PNN50 during drowsy state compared to baseline among time domain features. We also obtained significant increase (p<0.05) in LF and TP among frequency domain features during fatigue and drowsiness. The results depict to exhibit changes in cardiovascular parameters during drowsiness that can aid for analyzing drowsy state for the development of drowsiness detection and warning system incorporating other physiological signals such as EEG and PPG.

More importantly, we observed the surrogate behavior of PPG with ECG on analyzing PRV similar to HRV. This helps to claim that the drowsy state can be identified on analyzing PPG signal standalone from a feasible pulse-oximetry locations. On incorporating with the headset, the feasible PPG acquisition location would be earlobe or surrounding locations. On using earlobe, the HR estimation might be affected due to valley formation on accumulation of blood on the lobe and the feasible location could be nearby mastoid process of the ear over the carotid artery that encompasses the aviation headset.

7.1.3. Approach for Early Drowsiness Prediction

Drowsiness detection plays an important role in aviation industry. The airplanes fly in wide space over the ground and have a very little time to rectify the issue, the pilots should remain in alert state. In our work, we classified microsleep from baseline using LDA and QDA classifiers. The performance of best algorithm QDA+SBS on classifying microsleep from baseline is 86.23% in terms of F1 score. This shows the plausibility of classifying microsleep using EEG. Then, we used RF classifier to classify before-microsleep events from baseline using top selected 15 features with the region of aviation headset and obtained an acceptance performance of 83.80% in terms of F1 score. Further, we extended binary classes to four drowsy classes: BS, BMS, MS, and LS and obtained the overall performance of 83.70% in terms of unweighted mean of F1 score. In addition, we also achieved F1 score of 81.45% for classifying before-microsleep among four classes from the EEG signals within five brain regions. Although the result needs improvement, it shows the possibility of predicting drowsiness before eye-closure events which is crucial in drowsiness analysis. In our work, we also used electrodes within the region of aviation headset for prediction task to analyze the feasibility of headset for real time prediction purposes embedding sensors underneath the headset to acquire EEG. The goal of prediction is more important than detection in

drowsiness related works. Identifying drowsiness before eye-closure helps to warn the pilots before they fall into sleep and there will be a time though little for wakefulness that prevents from committing errors in the cockpit. The same scenario applies to drivers in roads. The development of drowsiness prediction classifier from continuous EEG signals collected from bigger sample size would further provide higher confidence on the results and towards the real time hardware implementation.

7.1.4. Drowsiness Detection and Prediction are Different Tasks

The drowsiness detection and prediction terms are not interchangeable. There are many published research papers where prediction and detection of microsleep are considered similar [151]. Many researchers have analyzed drowsy segments after the eye-closure events and claimed as prediction [152]–[154], [157]. The prediction for drowsiness should be made before its onset and the EEG segments should be analyzed before the onset of eye-closure. The capability of predicting drowsiness with high reliability aids in preventing transportation related accidents.

7.2. Future Directions

After performing the simulated flight experiment from commercially rated pilots during WOCL and analyzing the results, we have observed some of the issues that need to be addressed in future that are outlined below.

7.2.1. Experimental Protocol

We conducted the experiment in a flying simulator although the participants were commercially-rated pilots and the results may differ during real flight scenarios. The pilot's attentiveness may differ between simulated and real flights and may prone towards frequent drowsiness due to safer

simulated environment. In our experiment, the baseline is selected within the WOCL at the beginning of experiment and the participants were fatigued due to sleep deprivation from 18-20 hours before arriving to the simulator. In future, we will modify our experimental protocol and select baseline during daytime when the participants remain active without sleep deprivation. We will also acquire the data during real flights from the headset. The cockpit noise and multiple motion artifacts can hinder the signal quality that affects the drowsiness detection performances. This will help us to study the feasibility of headset to record physiological signals and also the potential of physiological signals to characterize drowsiness from the real flights.

7.2.2. Data Size and Demographics

In this work, we managed to collect data from 18 FAA certified commercially-rated pilots using a flight simulator. Among 18 participants, one participant data was interrupted during recording due to technical issues and the other subject had no MS data. So, we had only 16 participants whose data were useful for our analysis out of which only 2 were female. We did EEG analysis from 16 valid participants whereas for ECG analysis we had only 14 valid participants because the two more participants had poor ECG signal quality therefore, we need to discard for our ECG analysis. We have presented our results from a limited sample size with only 2 females. This show uneven distribution of demographics within the population. In future, we will recruit more pilots with varying age and gender that incorporates equal samples distribution. The more sample size (>30) exhibit normal data distribution and also increases confidence level on presenting the results.

7.2.3. Analysis Methods

In this work, we analyze three physiological signals: EEG, ECG, and PPG separately via classical statistical inference as well as machine learning approach using cut datasets. In future, we will

combine all the three signals (EEG, ECG, and PPG) as well as incorporate EOG behavior to develop a machine learning model for prediction in real time using continuous physiological signals. We will also compare the performance with other classifiers and even use deep learning concepts on arranging huge datasets. We will eventually develop a warning system based on physiological parameters to prevent drowsiness or fatigue in aviation operations and further extend to other driving or flying related tasks. In addition to the development of systems, the research can be directed to explore the interactions between cardiovascular and nervous systems during drowsiness. Although, this thesis studied the interactions between different EEG electrodes within the brain via coherence, the work could be extended to estimate cross-interactions between heart and brain signals. As already known from the literature, heart activities are regulated by SN and PSN from the central nervous system, there should be interactions between ECG and EEG signals that needs to be quantified through any of the directed and non-directed interaction techniques.

Appendixes

Appendix A.

A1. Time Domain Features

The time domain features obtained from ECG/PPG are described here.

1) *Heart rate (HR)*: HR can be calculated from RR or PP intervals (RRI or PPI). RRI is the duration between two adjacent R peaks in QRS complexes of ECG, and mean HR is obtained as the reciprocal of mean RR interval within a certain duration. HR is an important indicator to estimate psycho-physiological states.
2) *Standard deviation (SD)*: SD of RRI provides variation in HR within a certain period, and increase in this variation is due to longer RRI, which is indirectly related to the reduction in HR. The comparison of SD should be done in equal duration of experimental conditions.
3) *Standard Deviation of Successive Difference (SDSSD)*: SDSSD is the standard deviation of differences between adjacent RR intervals. It provides the variation of second order difference in R peaks which further helps to estimate variation in heart rate similar to SD but at secondary level.
4) *Root Mean Square of Successive Difference (RMSSSD)*: RMSSSD is the square root of the mean of the sum of the squares of differences between adjacent RR intervals. It gives power associated with variation in RR intervals and slower heart rate has higher RMSSSD.
5) *Triangular Index (TI)*: HRV triangular index (TI) is calculated as the total number of RR intervals divided by number of most frequently occurring RR interval. Faster heart rate has higher density of frequently occurring RR intervals thus, lowering TI.

6) *Sample Entropy (SA)*: Entropy, a non-linear feature, is the rate of generation of new information. Reduction in sample entropy characterizes abnormal heart rate characteristics due to reduced variability caused by vagal dominance.
7) *NN20*: NN20 represents number of RR or NN intervals greater than 20 ms, and there are less NN20 counts if heart beats faster and vice versa.
8) *PNN20*: PNN20 represents percentage of NN20 within the RR interval. It is calculated as the NN20 counts divided by total number of RR intervals.
9) *NN50*: NN20 represents number of RR or NN intervals greater than 50 ms, and more NN50 counts represent slower heart rate.
10) *PNN50*: PNN50 represents percentage of NN50 within the RR interval. It is calculated as the NN50 counts divided by total number of RR intervals.

A2. Frequency Domain Features

The following features are extracted from HRV in frequency domain to estimate drowsiness. PSD provides the variation of power as a function of frequency and the power distribution in respective frequency band is related to the changes in autonomic response of heart beats.

1) *Low-frequency (LF) Power:* LF band (0.04-0.15 Hz) power is influenced by both SN and PSN as well as some other mechanisms. LF is generally considered as a sympathetic dominance when expressed in normalized units and its dominance corresponds to alert state.
2) *High-frequency (HF) Power*: HF band (0.15-0.4 Hz) power is considered due to PSN and the dominance of HF corresponds to relaxed state.
3) *Total Power (TP)*: TP is the total power within the spectrum of 0.4 Hz.

4) *LF/HF ratio*: Balance between SN and PSN is measured by LF/HF ratio, also called as sympathovagal balance. Drowsy states correspond to higher HF from increased PSN while fatigue states correspond to demand for sleep causing higher LF resulting from increased SN. Therefore, sympathovagal balance is decreased at the onset of drowsiness whereas it will increase at fatigue state due to increase in SN. However, there are contradictions with the increase or decrease in the LF/HF ratio during drowsiness and fatigue among researchers. This might be due to the fact that these researches did not separate drowsiness with fatigue.

Appendix B.

- **Driver Monitoring System (DMS):** DMS is a Seeing Machines technology that performed computer vision analysis to track facial behavior i.e. face, eye, and eyelid to annotate drowsy periods. It consists of camera sensor, which included an infrared light and an infrared camera. The camera was glare shield-mounted and facing rearward. DMS is designed to detect microsleep candidates from prolonged long eyelid closure events.
- **Human Annotation Classification Method (HACM):** Once a prolonged eye closure event is identified by DMS, a manual review process is followed known as HACM to determine if that instance is indeed a microsleep candidate using specific rules. These rules include: 1) the instance of prolonged eyelid closure is greater than 1.5s in duration, but less than 15s; 2) the participant is not engaged in an intentional secondary behavior; and 3) the participant appears to exhibit signs of drowsiness or reduced cognitive capacity like slumping head position or posture [85].

- **PERCLOS:** It measures the proportion of eyes closed over a specific time interval and is strongly correlated with fatigue. The Seeing Machines PERCLOS signal is calculated as described by Wierwille et al. [59] as the proportion of time that the eyes of participant are closed 80% or more. For instance, if 5 frames out of 100 are marked as closed, the PERCLOS value becomes 5/100=0.05. A PERCLOS value higher than 0.15 is reported as drowsy [85].

References

[1] J. A. Caldwell, "Fatigue in aviation," *Travel Med. Infect. Dis.*, vol. 3, no. 2, pp. 85–96, May 2005.

[2] R. P. Balandong, R. F. Ahmad, M. N. Mohamad Saad, and A. S. Malik, "A Review on EEG-Based Automatic Sleepiness Detection Systems for Driver," *IEEE Access*, vol. 6, pp. 22908–22919, 2018.

[3] A. Sahayadhas, K. Sundaraj, M. Murugappan, A. Sahayadhas, K. Sundaraj, and M. Murugappan, "Detecting Driver Drowsiness Based on Sensors: A Review," *Sensors*, vol. 12, no. 12, pp. 16937–16953, Dec. 2012.

[4] D. Kim, H. Han, S. Cho, and U. Chong, "Detection of drowsiness with eyes open using EEG-based power spectrum analysis," in *2012 7th International Forum on Strategic Technology (IFOST)*, 2012, pp. 1–4.

[5] D. Powell, M. B. Spencer, D. Holland, and K. J. Petrie, "Fatigue in Two-Pilot Operations: Implications for Flight and Duty Time Limitations," *Aviat. Space. Environ. Med.*, vol. 79, no. 11, pp. 1047–1050, Nov. 2008.

[6] National Transportation Safety Board (NTSB 2017-2018), "Most Wanted List of Transportation Safety Improvements: Reduce Fatigue-Related Accidents," Washington DC, 2018.

[7] A. Chowdhury, R. Shankaran, M. Kavakli, and M. Haque, "Sensor Applications and Physiological Features in Drivers' Drowsiness Detection : A Review," *IEEE Sens. J.*, vol. 1748, no. c, pp. 1–12, 2018.

[8] J. A. Caldwell *et al.*, "Fatigue countermeasures in aviation.," *Aviat. Space. Environ. Med.*, vol. 80, no. 1, pp. 29–59, Jan. 2009.

[9] M. R. Rosekind, K. B. Gregory, E. L. Co, D. L. Miller, and D. F. Dinges, "Crew Factors in Flight Operations XII: A Survey of Sleep Quantity and Quality in On-Board Crew Rest Facilities," *NASA Ames Res. Center, Moffett Field, CA, USA, Tech. Rep. NASA/TM-2000-209611*, pp. 1–84, Sep. 2000.

[10] B. S. Oken, M. C. Salinsky, and S. M. Elsas, "Vigilance, alertness, or sustained attention: physiological basis and measurement.," *Clin. Neurophysiol.*, vol. 117, no. 9, pp. 1885–901, Sep. 2006.

[11] Federal Aviation Administration, "Risk Management Handbook: U.S. Department of Transportation," 2016.

[12] J. Orasanu *et al.*, "Evaluating the Effectiveness of Schedule Changes for Air Traffic Service (ATS) Providers: Controller Alertness and Fatigue Monitoring Study," Dec. 2012.

[13] K. L. Knutson *et al.*, "The National Sleep Foundation's Sleep Health Index," *Sleep Heal.*, vol. 3, no. 4, pp. 234–240, Aug. 2017.

[14] S. Bourgeois-Bougrine, P. Carbon, C. Gounelle, R. Mollard, and A. Coblentz, "Perceived fatigue for short- and long-haul flights: a survey of 739 airline pilots.," *Aviat. Space. Environ. Med.*, vol. 74, no. 10, pp. 1072–7, Oct. 2003.

[15] C. C. Liu, S. G. Hosking, and M. G. Lenné, "Predicting driver drowsiness using vehicle measures: Recent insights and future challenges," *J. Safety Res.*, vol. 40, no. 4, pp. 239–245, 2009.

[16] Z. Zhu and Q. Ji, "Real Time Non-intrusive Monitoring and Prediction of Driver Fatigue Monitoring," in *IEEE Intelligent Transportation Systems*, 2004, pp. 1–6.

[17] S. K. L. Lal and A. Craig, "A critical review of the psychophysiology of driver fatigue," *Biol. Psychol.*, vol. 55, no. 3, pp. 173–194, Feb. 2001.

[18] D. Dinges, G. Maislin, R. Brewster, G. Krueger, and R. Carroll, "Pilot Test of Fatigue Management Technologies," *Transp. Res. Rec. J. Transp. Res. Board*, vol. 1922, no. 1922, pp. 175–182, 2007.

[19] N. Edenborough *et al.*, "Driver state monitor from DELPHI," *Proc. - 2005 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognition, CVPR 2005*, vol. II, no. July, pp. 1206–1207, 2005.

[20] L. Barr, H. Howarth, and S. Popkin, "An Evaluation of Emerging Driver Fatigue Detection Measures and Technologies," *U.S. Dep. Transp. Fed. Mot. Carr. Saf. Adm.*, no. June, pp. 1–27, 2009.

[21] A. Kosmadopoulos *et al.*, "The efficacy of objective and subjective predictors of driving performance during sleep restriction and circadian misalignment," *Accid. Anal. Prev.*, vol. 99, pp. 445–451, Feb. 2017.

[22] M. Poursadeghiyan, A. Mazloumi, G. N. Saraji, A. Niknezhad, A. Akbarzadeh, and M. H. Ebrahimi, "Determination the Levels of Subjective and Observer Rating of Drowsiness and Their Associations with Facial Dynamic Changes.," *Iran. J. Public Health*, vol. 46, no. 1, pp. 93–102, Jan. 2017.

[23] S. Baiardi, C. La 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?," *Accid. Anal. Prev.*, vol. 110, pp. 187–189, Jan. 2018.

[24] A. Shahid, K. Wilkinson, S. Marcu, and C. M. Shapiro, "Karolinska Sleepiness Scale (KSS)," in *STOP, THAT and One Hundred Other Sleep Scales*, New York, NY: Springer New York, 2011, pp. 209–210.

[25] M. Basner and D. F. Dinges, "Maximizing Sensitivity of the Psychomotor Vigilance Test

(PVT) to Sleep Loss," *Sleep*, vol. 34, no. 5, pp. 581–591, May 2011.

[26] S. M. Belz, G. S. Robinson, and J. G. Casali, "Temporal Separation and Self-Rating of Alertness as Indicators of Driver Fatigue in Commercial Motor Vehicle Operators," *Hum. Factors J. Hum. Factors Ergon. Soc.*, vol. 46, no. 1, pp. 154–169, Mar. 2004.

[27] H. J. Moller, L. Kayumov, E. L. Bulmash, J. Nhan, and C. M. Shapiro, "Simulator performance, microsleep episodes, and subjective sleepiness: normative data using convergent methodologies to assess driver drowsiness," *J. Psychosom. Res.*, vol. 61, no. 3, pp. 335–342, Sep. 2006.

[28] E. A. Schmidt, M. Schrauf, M. Simon, M. Fritzsche, A. Buchner, and W. E. Kincses, "Drivers' misjudgement of vigilance state during prolonged monotonous daytime driving," *Accid. Anal. Prev.*, vol. 41, no. 5, pp. 1087–1093, Sep. 2009.

[29] C. R. H. Innes, G. R. Poudel, and R. D. Jones, "Efficient and Regular Patterns of Nighttime Sleep are Related to Increased Vulnerability to Microsleeps Following a Single Night of Sleep Restriction," *Chronobiol. Int.*, vol. 30, no. 9, pp. 1187–1196, 2013.

[30] A. Anund, C. Fors, J. Ihlström, and G. Kecklund, "An on-road study of sleepiness in split shifts among city bus drivers," *Accid. Anal. Prev.*, vol. 114, pp. 71–76, May 2018.

[31] V. Riethmeister, U. Bültmann, M. Gordijn, S. Brouwer, and M. de Boer, "Investigating daily fatigue scores during two-week offshore day shifts," *Appl. Ergon.*, vol. 71, pp. 87–94, Sep. 2018.

[32] M. Di Muzio *et al.*, "Not only a Problem of Fatigue and Sleepiness: Changes in Psychomotor Performance in Italian Nurses across 8-h Rapidly Rotating Shifts," *J. Clin. Med.*, vol. 8, no. 1, p. 47, Jan. 2019.

[33] A. B. Buriro, R. Shoorangiz, S. J. Weddell, and R. D. Jones, "Predicting Microsleep States

Using EEG Inter-Channel Relationships," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 12, pp. 2260–2269, Dec. 2018.

[34] T. Nguyen, S. Ahn, H. Jang, S. C. Jun, and J. G. Kim, "Utilization of a combined EEG/NIRS system to predict driver drowsiness," *Sci. Rep.*, vol. 7, no. 1, p. 43933, Dec. 2017.

[35] S. K. L. Lal and A. Craig, "Electroencephalography Activity Associated with Driver Fatigue: Implications for a Fatigue Countermeasure Device. - PsycNET," *J. Psychophysiol.*, vol. 15, pp. 183–189, 2001.

[36] C.-T. Lin, R.-C. Wu, T.-P. Jung, S.-F. Liang, and T.-Y. Huang, "Estimating Driving Performance Based on EEG Spectrum Analysis," *EURASIP J. Appl. Signal Processing*, vol. 19, pp. 3165–3174, 2005.

[37] M. Abo-Zahhad, S. M. Ahmed, and S. N. Abbas, "A New EEG Acquisition Protocol for Biometric Identification Using Eye Blinking Signals," *Int. J. Intell. Syst. Appl.*, vol. 7, no. 6, pp. 48–54, May 2015.

[38] N. Gurudath and H. B. Riley, "Drowsy Driving Detection by EEG Analysis Using Wavelet Transform and K-means Clustering," *Procedia Comput. Sci.*, vol. 34, pp. 400–409, Jan. 2014.

[39] L. Yang, R. Ma, H. M. Zhang, W. Guan, and S. Jiang, "Driving behavior recognition using EEG data from a simulated car-following experiment," *Accid. Anal. Prev.*, vol. 116, pp. 30–40, Jul. 2018.

[40] B. T. Jap, S. Lal, and P. Fischer, "Comparing combinations of EEG activity in train drivers during monotonous driving," *Expert Syst. Appl.*, vol. 38, no. 1, pp. 996–1003, Jan. 2011.

[41] E. Abe, K. Fujiwara, T. Hiraoka, T. Yamakawa, and M. Kano, “Development of drowsy driving accident prediction by heart rate variability analysis,” in *Signal and Information Processing Association Annual Summit and Conference (APSIPA), 2014 Asia-Pacific*, 2014, pp. 1–4.

[42] R. L. Burr, “Interpretation of normalized spectral heart rate variability in- dices in sleep research: a critical review.,” *Sleep*, vol. 30, no. 7, pp. 913–919, 2007.

[43] G. Li and W.-Y. Chung, “Detection of Driver Drowsiness Using Wavelet Analysis of Heart Rate Variability and a Support Vector Machine Classifier,” *Sensors*, vol. 13, no. 12, pp. 16494–16511, 2013.

[44] J. Vicente, P. Laguna, A. Bartra, and R. Bailon, “Detection of driver’s drowsiness by means of HRV analysis,” *2011 Comput. Cardiol.*, pp. 89–92, 2011.

[45] A. Zhang and F. Liu, “Drowsiness detection based on wavelet analysis of ECG and pulse signals,” *2012 5th Int. Conf. Biomed. Eng. Informatics, BMEI 2012*, no. 30670529, pp. 491–495, 2012.

[46] G. Ernst, “The Autonomic Nervous System,” in *Heart Rate Variability*, London: Springer London, 2014, pp. 27–49.

[47] G. D. Furman, A. Baharav, C. Cahan, and S. Akselrod, “Early detection of falling asleep at the wheel: A Heart Rate Variability approach,” in *2008 Computers in Cardiology*, 2008, pp. 1109–1112.

[48] E. Michail, A. Kokonozi, I. Chouvarda, and N. Maglaveras, “EEG and HRV markers of sleepiness and loss of control during car driving,” in *2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2008, pp. 2566–2569.

[49] J. Pan and W. J. Tompkins, “A Real-Time QRS Detection Algorithm,” *IEEE Trans. Biomed. Eng.*, no. 3, 1985.

[50] N. Selvaraj, A. Jaryal, J. Santhosh, K. K. Deepak, and S. Anand, “Assessment of heart rate variability derived from finger-tip photoplethysmography as compared to electrocardiography,” *J. Med. Eng. Technol.*, vol. 32, no. 6, pp. 479–484, 2008.

[51] R. Mukkamala *et al.*, “Toward Ubiquitous Blood Pressure Monitoring via Pulse Transit Time: Theory and Practice,” *IEEE Trans. Biomed. Eng.*, vol. 62, no. 8, pp. 1879–1901, Aug. 2015.

[52] J. Allen and A. Murray, “Variability of photoplethysmography peripheral pulse measurements at the ears, thumbs and toes,” *IEE Proc. - Sci. Meas. Technol.*, vol. 147, no. 6, pp. 403–407, Nov. 2000.

[53] M. Bolanos, H. Nazeran, and E. Haltiwanger, “Comparison of heart rate variability signal features derived from electrocardiography and photoplethysmography in healthy individuals,” *Annu. Int. Conf. IEEE Eng. Med. Biol. - Proc.*, pp. 4289–4294, 2006.

[54] G. Lu, F. Yang, J. A. Taylor, and J. F. Stein, “A comparison of photoplethysmography and ECG recording to analyse heart rate variability in healthy subjects,” *J. Med. Eng. Technol.*, vol. 33, no. 8, pp. 634–641, Nov. 2009.

[55] S. Lu *et al.*, “Can photoplethysmography variability serve as an alternative approach to obtain heart rate variability information?,” *J. Clin. Monit. Comput.*, vol. 22, no. 1, pp. 23–29, 2008.

[56] Y. Punsawad, S. Aempedchr, Y. Wongsawat, and M. Parnichkun, “EEG-based mental fatigue alarm system using weighted-frequency index,” *Asia-Pacific Signal Inf. Process. Assoc. Annu. Summit Conf. APSIPA ASC*, pp. 193–196, Dec. 2010.

[57] J. Arnin *et al.*, "Wireless-based portable EEG-EOG monitoring for real time drowsiness detection," in *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2013, pp. 4977–4980.

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

[59] W. W. Wierwille, S. S. Wreggit, C. L. Kirn, L. A. Ellsworth, and R. J. Fairbanks, "Research on Vehicle-Based Driver Status/Performance Monitoring; Development, Validation, and Refinement of Algorithms for Detection of Driver Drowsiness, Final Report," *U.S. Dep. Transp. Natl. Highw. Traffic Saf. Adm.*, pp. 1–247, Dec. 1994.

[60] E. Başar, "Brain oscillations in neuropsychiatric disease.," *Dialogues Clin. Neurosci.*, vol. 15, no. 3, pp. 291–300, Sep. 2013.

[61] D. Millett, "Hans Berger: From Psychic Energy to the EEG," *Perspect. Biol. Med.*, vol. 44, no. 4, pp. 522–542, 2001.

[62] K. Blinowska and P. Durka, "Electroencephalography (EEG)," in *Wiley Encyclopedia of Biomedical Engineering*, Hoboken, NJ, USA: John Wiley & Sons, Inc., 2006.

[63] G. M. Rojas, C. Alvarez, C. E. Montoya, M. de la Iglesia-Vayá, J. E. Cisternas, and M. Gálvez, "Study of resting-state functional connectivity networks using EEG electrodes position as seed," *Front. Neurosci.*, vol. 12, no. APR, pp. 1–12, 2018.

[64] A. Delorme and S. Makeig, "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *J. Neurosci. Methods*, vol. 134, no. 1, pp. 9–21, Mar. 2004.

[65] Chin-Teng Lin, Ruei-Cheng Wu, Sheng-Fu Liang, Wen-Hung Chao, Yu-Jie Chen, and

Tzyy-Ping Jung, “EEG-based drowsiness estimation for safety driving using independent component analysis,” *IEEE Trans. Circuits Syst. I Regul. Pap.*, vol. 52, no. 12, pp. 2726–2738, Dec. 2005.

[66] C. Wang *et al.*, “Spectral Analysis of EEG during Microsleep Events Annotated Via Driver Monitoring System to Characterize Drowsiness,” *IEEE Trans. Aerosp. Electron. Syst.*, pp. 1–13, 2019.

[67] Y. Kang, J. Escudero, D. Shin, E. Ifeachor, and V. Marmarelis, “Principal Dynamic Mode Analysis of EEG Data for Assisting the Diagnosis of Alzheimer’s Disease,” *IEEE J. Transl. Eng. Heal. Med.*, vol. 3, pp. 1–10, 2015.

[68] A. Phinyomark, S. Thongpanja, H. Hu, P. Phukpattaranont, and C. Limsakul, “The Usefulness of Mean and Median Frequencies in Electromyography Analysis,” in *Computational Intelligence in Electromyography Analysis - A Perspective on Current Applications and Future Challenges*, InTech, 2012.

[69] Y. Wada, Y. Nanbu, Y. Koshino, Y. Shimada, and T. Hashimoto, “Inter-and lntrahemispheric EEG Coherence During Light Drowsiness,” *Clin. Electroencephalogr.*, vol. 27, no. 2, pp. 84–88, 1996.

[70] B. T. Jap, S. Lal, and P. Fischer, “Inter-hemispheric electroencephalography coherence analysis: Assessing brain activity during monotonous driving,” *Int. J. Psychophysiol.*, vol. 76, no. 3, pp. 169–173, Jun. 2010.

[71] W. Kong, W. Lin, F. Babiloni, S. Hu, and G. Borghini, “Investigating Driver Fatigue versus Alertness Using the Granger Causality Network.,” *Sensors (Basel).*, vol. 15, no. 8, pp. 19181–98, Aug. 2015.

[72] J.-P. Lachaux *et al.*, “Estimating the time-course of coherence between single-trial brain

signals: an introduction to wavelet coherence," *Neurophysiol. Clin. Neurophysiol.*, vol. 32, no. 3, pp. 157–174, Jun. 2002.

[73] R. Saab, M. J. McKeown, L. J. Myers, and R. Abu-Gharbieh, "A Wavelet Based Approach for the Detection of Coupling in EEG Signals," in *Conference Proceedings. 2nd International IEEE EMBS Conference on Neural Engineering, 2005.*, pp. 616–620.

[74] A. Catarino, A. Andrade, O. Churches, A. P. Wagner, S. Baron-Cohen, and H. Ring, "Task-related functional connectivity in autism spectrum conditions: an EEG study using wavelet transform coherence," *Mol. Autism*, vol. 4, no. 1, p. 1, Jan. 2013.

[75] M. Malik, "Heart Rate Variability.," *Ann. Noninvasive Electrocardiol.*, vol. 1, no. 2, pp. 151–181, Apr. 1996.

[76] A. Hernando *et al.*, "Autonomic Nervous System Measurement in Hyperbaric Environments Using ECG and PPG Signals," *IEEE J. Biomed. Heal. Informatics*, vol. 23, no. 1, pp. 132–142, Jan. 2019.

[77] W.-H. Lin, D. Wu, C. Li, H. Zhang, and Y.-T. Zhang, "Comparison of Heart Rate Variability from PPG with That from ECG," Springer, Cham, 2014, pp. 213–215.

[78] N. Pinheiro *et al.*, "Can PPG be used for HRV analysis?," in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2016, pp. 2945–2949.

[79] T. Tigges *et al.*, "Assessment of In-ear Photoplethysmography as a Surrogate for Electrocardiography in Heart Rate Variability Analysis," Springer, Singapore, 2019, pp. 293–297.

[80] N. Wilson, B. Guragain, A. Verma, L. Archer, and K. Tavakolian, "Blending Human and Machine: Feasibility of Measuring Fatigue Through the Aviation Headset," *Hum. Factors*

J. Hum. Factors Ergon. Soc., p. 001872081984978, Jun. 2019.

[81] J. Franklin, "The elements of statistical learning: data mining, inference and prediction," *Math. Intell.*, vol. 27, no. 2, pp. 83–85, Mar. 2005.

[82] S. Arlot and A. Celisse, "A survey of cross-validation procedures for model selection," *Stat. Surv.*, vol. 4, pp. 40–79, 2010.

[83] A. B. Rad *et al.*, "ECG-Based Classification of Resuscitation Cardiac Rhythms for Retrospective Data Analysis," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 10, pp. 2411–2418, Oct. 2017.

[84] S. R. Jagannathan, A. Ezquerro-Nassar, B. Jachs, O. V. Pustovaya, C. A. Bareham, and T. A. Bekinschtein, "Tracking wakefulness as it fades: Micro-measures of alertness," *Neuroimage*, vol. 176, pp. 138–151, Aug. 2018.

[85] M. Lenne, G. Grubb, R. Myers, and J. Da Cruz, "Seeing Machines Data Description Document," vol. 1, pp. 1–3, 2018.

[86] C. . Dussault, C., Jouanin, J.C., Philippe, M., Guezennec, "EEG and ECG changes during simulator operation reflect mental workload and vigilance," *Aviat. Sp. Environ. Med.*, vol. 76, no. 4, pp. 344–351, 2005.

[87] B. T. Jap, S. Lal, P. Fischer, and E. Bekiaris, "Using EEG spectral components to assess algorithms for detecting fatigue," *Expert Syst. Appl.*, vol. 36, no. 2, pp. 2352–2359, Mar. 2009.

[88] "ANR 2888 Aviation Headset." [Online]. Available: http://www.flightstore.com.au/anr-2888-aviation-headset.

[89] C. A. L. Jr. *et al.*, "Practice parameters for the indications for polysomnography and related procedures," *Sleep*, vol. 20, no. 6, pp. 406–422, 1997.

[90] R. Acharya U., O. Faust, N. Kannathal, T. Chua, and S. Laxminarayan, "Non-linear analysis of EEG signals at various sleep stages," *Comput. Methods Programs Biomed.*, vol. 80, no. 1, pp. 37–45, Oct. 2005.

[91] K. Hartman, J. Strasser, and C. Systematics, "Saving lives through advanced vehicle safety technology : intelligent vehicle initiative," *U.S. DOT's Intelligent Transportation Systems (ITS)*. United States. Federal Highway Administration, 01-Sep-2005.

[92] M. M. Ohayon *et al.*, "The National Sleep Foundation's Sleep Satisfaction Tool," *Sleep Heal.*, vol. 5, no. 1, pp. 5–11, Feb. 2019.

[93] S. K. L. Lal and A. Craig, "Electroencephalography Activity Associated with Driver Fatigue: Implications for a Fatigue Countermeasure Device," *J. Psychophysiol.*, vol. 15, pp. 183–189, 2001.

[94] C. Vidaurre, N. Krämer, B. Blankertz, and A. Schlögl, "Time Domain Parameters as a feature for EEG-based Brain–Computer Interfaces," *Neural Networks*, vol. 22, no. 9, pp. 1313–1319, Nov. 2009.

[95] D. P. Subha, P. K. Joseph, R. Acharya U, and C. M. Lim, "EEG Signal Analysis: A Survey," *J. Med. Syst.*, vol. 34, no. 2, pp. 195–212, Apr. 2010.

[96] C. Papadelis *et al.*, "Monitoring sleepiness with on-board electrophysiological recordings for preventing sleep-deprived traffic accidents," *Clin. Neurophysiol.*, vol. 118, no. 9, pp. 1906–1922, 2007.

[97] M. Simon *et al.*, "EEG alpha spindle measures as indicators of driver fatigue under real traffic conditions," *Clin. Neurophysiol.*, vol. 122, no. 6, pp. 1168–1178, 2011.

[98] C. Papadelis *et al.*, "Indicators of Sleepiness in an ambulatory EEG study of night driving," *Annu. Int. Conf. IEEE Eng. Med. Biol. - Proc.*, pp. 6201–6204, 2006.

[99] F.-C. Lin, L.-W. Ko, C.-H. Chuang, T.-P. Su, and C.-T. Lin, "Generalized EEG-Based Drowsiness Prediction System by Using a Self-Organizing Neural Fuzzy System," *IEEE Trans. Circuits Syst. I Regul. Pap.*, vol. 59, no. 9, pp. 2044–2055, Sep. 2012.

[100] M. Jagannath and V. Balasubramanian, "Assessment of early onset of driver fatigue using multimodal fatigue measures in a static simulator," *Appl. Ergon.*, vol. 45, no. 4, pp. 1140–1147, Jul. 2014.

[101] Z. Guo, Y. Pan, G. Zhao, S. Cao, and J. Zhang, "Detection of Driver Vigilance Level Using EEG Signals and Driving Contexts," *IEEE Trans. Reliab.*, vol. 67, no. 1, pp. 370–380, Mar. 2018.

[102] Y.-K. Wang, T.-P. Jung, and C.-T. Lin, "Theta and Alpha Oscillations in Attentional Interaction during Distracted Driving.," *Front. Behav. Neurosci.*, vol. 12, p. 3, 2018.

[103] H. Wang, A. Dragomir, N. I. Abbasi, J. Li, N. V. Thakor, and A. Bezerianos, "A novel real-time driving fatigue detection system based on wireless dry EEG," *Cogn. Neurodyn.*, vol. 12, no. 4, pp. 365–376, Aug. 2018.

[104] G. Li, B. Li, G. Wang, J. Zhang, and J. Wang, "A New Method for Human Mental Fatigue Detection with Several EEG Channels," *J. Med. Biol. Eng.*, vol. 37, no. 2, pp. 240–247, Apr. 2017.

[105] 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," *Inf. Fusion*, vol. 43, pp. 66–76, Sep. 2018.

[106] C.-S. Wei, Y.-T. Wang, C.-T. Lin, and T.-P. Jung, "Toward Drowsiness Detection Using Non-hair-Bearing EEG-Based Brain-Computer Interfaces," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 2, pp. 400–406, Feb. 2018.

[107] M. Golmohammadi, A. H. Harati Nejad Torbati, S. Lopez de Diego, I. Obeid, and J. Picone, “Automatic Analysis of EEGs Using Big Data and Hybrid Deep Learning Architectures,” *Front. Hum. Neurosci.*, vol. 13, p. 76, Mar. 2019.

[108] M. Simon *et al.*, “EEG alpha spindle measures as indicators of driver fatigue under real traffic conditions,” *Clin. Neurophysiol.*, vol. 122, no. 6, pp. 1168–1178, Jun. 2011.

[109] T. Åkerstedt and M. Gillberg, “Subjective and Objective Sleepiness in the Active Individual,” *Int. J. Neurosci.*, vol. 52, no. 1–2, pp. 29–37, Jan. 1990.

[110] C. Dussault, J. Jouanin, and C. Guezennec, “EEG and ECG Changes During Selected Flight Sequences,” *Aviat. Sp. Environ. Med.*, vol. 75, no. 10, pp. 889–897, 2004.

[111] T. Brown, R. Johnson, and G. Milavetz, “Identifying periods of drowsy driving using EEG.,” *Ann. Adv. Automot. Med. Assoc. Adv. Automot. Med. Annu. Sci. Conf.*, vol. 57, pp. 99–108, 2013.

[112] C.-S. Hung *et al.*, “Local Experience-Dependent Changes in the Wake EEG after Prolonged Wakefulness,” *Sleep*, vol. 36, no. 1, pp. 59–72, Jan. 2013.

[113] 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 Syst.*, vol. 80, pp. 143–152, May 2015.

[114] F. Gharagozlou, G. N. Saraji, A. Mazloumi, A. M. Nasrabadi, A. R. Foroushani, and A. Arab, “Detecting Driver Mental Fatigue Based on EEG Alpha Power Changes during Simulated Driving,” *Iran J Public Heal.*, vol. 44, no. 12, pp. 1693–1700, 2015.

[115] T. Torsvall, L, Åkerstedt, “Sleepiness on the job: continuously measured EEG changes in train drivers,” *Electroencephalogr. Clin. Neurophysiol.*, vol. 66, pp. 502–511, 1987.

[116] H. J. Eoh, M. K. Chung, and S. H. Kim, “Electroencephalographic study of drowsiness in

simulated driving with sleep deprivation," *Int. J. Ind. Ergon.*, vol. 35, no. 4, pp. 307–320, 2005.

[117] G. N. Boldyreva and L. A. Zhavoronkova, "Interhemispheric asymmetry of EEG coherence as a reflection of different functional states of the human brain.," *Biomed. Sci.*, vol. 2, no. 3, pp. 266–70, 1991.

[118] M. Arthuis *et al.*, "Impaired consciousness during temporal lobe seizures is related to increased long-distance cortical–subcortical synchronization," *Brain*, vol. 132, no. 8, pp. 2091–2101, Aug. 2009.

[119] I. C. A. O. (ICAO), "Fatigue Risk Management Systems -Manual for Regulators," Montreal, Quebec, Canada, 2012.

[120] A. Jerman, "How To Measure Fatigue Among Pilots ?," *Zb. Veleučilišta u Rijeci*, vol. 6, no. 1, pp. 13–22, 2018.

[121] J. Velazquez and Jonathan, "The Presence of Behavioral Traps in U.S. Airline Accidents: A Qualitative Analysis," *Safety*, vol. 4, no. 1, p. 2, Jan. 2018.

[122] G. Yang, Y. Lin, and P. Bhattacharya, "A driver fatigue recognition model based on information fusion and dynamic Bayesian network," *Inf. Sci. (Ny).*, vol. 180, no. 10, pp. 1942–1954, May 2010.

[123] C. Varghese and U. Shankar, "Passenger Vehicle Occupant Fatalities by Day and Night – A Contrast," *Traffic Saf. Facts - Res. Note*, no. HS-810 637, May 2007.

[124] A. E. Rogers, "The Effects of Fatigue and Sleepiness on Nurse Performance and Patient Safety," in *Patient Safety and Quality: An Evidence-Based Handbook for Nurses. Agency for Healthcare Research Quality (US)*, Agency for Healthcare Research and Quality (US), 2008.

[125] C. Jacobé de Naurois, C. Bourdin, A. Stratulat, E. Diaz, and J.-L. Vercher, "Detection and prediction of driver drowsiness using artificial neural network models," *Accid. Anal. Prev.*, Dec. 2017.

[126] M. Awais, N. Badruddin, and M. Drieberg, "A Hybrid Approach to Detect Driver Drowsiness Utilizing Physiological Signals to Improve System Performance and Wearability," *Sensors*, vol. 17, no. 9, p. 1991, Aug. 2017.

[127] E. Abe, K. Fujiwara, T. Hiraoka, T. Yamakawa, and M. Kano, "Development of drowsy driving accident prediction by heart rate variability analysis," in *Signal and Information Processing Association Annual Summit and Conference (APSIPA), 2014 Asia-Pacific*, 2014, pp. 1–4.

[128] T. Gohara *et al.*, "Heart rate variability change induced by the mental stress: the effect of accumulated fatigue," *Proc. 1996 Fifteenth South. Biomed. Eng. Conf.*, pp. 367–369, 1996.

[129] J. S. Richman and J. R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," *Am. J. Physiol. Circ. Physiol.*, vol. 278, no. 6, pp. H2039–H2049, Jun. 2000.

[130] Z. Mardi, S. N. M. Ashtiani, and M. Mikaili, "EEG-based Drowsiness Detection for Safe Driving Using Chaotic Features and Statistical Tests.," *J. Med. Signals Sens.*, vol. 1, no. 2, pp. 130–7, May 2011.

[131] A. D. McDonald, J. D. Lee, C. Schwarz, and T. L. Brown, "Steering in a Random Forest," *Hum. Factors J. Hum. Factors Ergon. Soc.*, vol. 56, no. 5, pp. 986–998, Aug. 2014.

[132] F. Rundo *et al.*, "An Innovative Deep Learning Algorithm for Drowsiness Detection from EEG Signal," *Computation*, vol. 7, no. 1, p. 13, Feb. 2019.

[133] P. R. Davidson, R. D. Jones, and M. T. R. Peiris, "EEG-Based Lapse Detection With High Temporal Resolution," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 5, pp. 832–839, May 2007.

[134] W. Liu *et al.*, "Convolutional Two-Stream Network Using Multi-Facial Feature Fusion for Driver Fatigue Detection," *Futur. Internet*, vol. 11, no. 5, p. 115, May 2019.

[135] M. K. Kiymik, M. Akin, and A. Subasi, "Automatic recognition of alertness level by using wavelet transform and artificial neural network," *J. Neurosci. Methods*, vol. 139, no. 2, pp. 231–240, Oct. 2004.

[136] M. Yang *et al.*, "Toward Wearable EEG-based Alertness Detection System Using SVM with Optimal Minimum Channels," *MATEC Web Conf.*, vol. 214, p. 03009, Oct. 2018.

[137] J. Hu and J. Min, "Automated detection of driver fatigue based on EEG signals using gradient boosting decision tree model," *Cogn. Neurodyn.*, vol. 12, no. 4, pp. 431–440, Aug. 2018.

[138] J. Hu, "Comparison of Different Features and Classifiers for Driver Fatigue Detection Based on a Single EEG Channel," *Comput. Math. Methods Med.*, vol. 2017, pp. 1–9, 2017.

[139] I. Belakhdar, W. Kaaniche, R. Djemal, and B. Ouni, "Single-channel-based automatic drowsiness detection architecture with a reduced number of EEG features," *Microprocess. Microsyst.*, vol. 58, pp. 13–23, Apr. 2018.

[140] F. Mohamed, S. F. Ahmed, Z. Ibrahim, and S. Yaacob, "Comparison of Features Based on Spectral Estimation for the Analysis of EEG Signals in Driver Behavior," in *2018 International Conference on Computational Approach in Smart Systems Design and Applications (ICASSDA)*, 2018, pp. 1–7.

[141] T. Nakamura, Y. D. Alqurashi, M. J. Morrell, and D. P. Mandic, "Automatic detection of

drowsiness using in-ear EEG," in *2018 International Joint Conference on Neural Networks (IJCNN)*, 2018, pp. 1–6.

[142] R. Chai *et al.*, "Improving EEG-Based Driver Fatigue Classification Using Sparse-Deep Belief Networks," *Front. Neurosci.*, vol. 11, p. 103, Mar. 2017.

[143] D. Qian *et al.*, "Drowsiness Detection by Bayesian-Copula Discriminant Classifier Based on EEG Signals during Daytime Short Nap," *IEEE Trans. Biomed. Eng.*, vol. 64, no. 4, pp. 743–754, 2017.

[144] G. Li, B.-L. Lee, and W.-Y. Chung, "Smartwatch-Based Wearable EEG System for Driver Drowsiness Detection," *IEEE Sens. J.*, vol. 15, no. 12, pp. 7169–7180, Dec. 2015.

[145] C. Jacobé de Naurois, C. Bourdin, C. Bougard, and J.-L. Vercher, "Adapting artificial neural networks to a specific driver enhances detection and prediction of drowsiness," *Accid. Anal. Prev.*, vol. 121, pp. 118–128, Dec. 2018.

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

[147] A. Zhang and Y. Chen, "EEG feature extraction and analysis under drowsy state based on energy and sample entropy," in *2012 5th International Conference on BioMedical Engineering and Informatics*, 2012, pp. 501–505.

[148] J.-R. Huang *et al.*, "Application of Multivariate Empirical Mode Decomposition and Sample Entropy in EEG Signals via Artificial Neural Networks for Interpreting Depth of Anesthesia," *Entropy*, vol. 15, no. 12, pp. 3325–3339, Aug. 2013.

[149] A. B. Rad *et al.*, "Automatic cardiac rhythm interpretation during resuscitation," *Resuscitation*, vol. 102, pp. 44–50, May 2016.

[150] C. Jacobé de Naurois, C. Bourdin, A. Stratulat, E. Diaz, and J.-L. Vercher, "Detection and prediction of driver drowsiness using artificial neural network models," *Accid. Anal. Prev.*, Dec. 2017.

[151] M. Golz, D. Sommer, M. Holzbrecher, and T. Schnupp, "Detection and prediction of driver's microsleep events," in *14th International Coneference Road Safety on Four Continents, vol. 11*, 2007, pp. 1–11.

[152] Y. Liang *et al.*, "Prediction of drowsiness events in night shift workers during morning driving," *Accid. Anal. Prev.*, vol. 126, no. November, pp. 105–114, 2019.

[153] T. Damrongwatanapokin and K. Mikami, "Fatigue prediction and intervention for continuous play in video games," *Proc. - 2018 Int. Conf. Cyberworlds, CW 2018*, pp. 451–453, 2018.

[154] D. Naito, R. Hatano, and H. Nishiyama, "Prediction of Drowsy Driving Using EEG and Facial Expression by Machine Learning," in *Proceedings of 34th International Conference on Computers and Their Applications*, 2019, vol. 58, pp. 225–235.

[155] Y. Shen, E. Olbrich, P. Achermann, and P. F. Meier, "Dimensional complexity and spectral properties of the human sleep EEG. Electroencephalograms.," *Clin. Neurophysiol.*, vol. 114, no. 2, pp. 199–209, Feb. 2003.

[156] B. Guragain *et al.*, "EEG-based Classification of Microsleep by Means of Feature Selection: An Application in Aviation," 2019, pp. 4060–4063.

[157] H. S. Kim, D. Yoon, H. S. Shin, and C. H. Park, "Predicting the EEG Level of a Driver Based on Driving Information," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 4, pp. 1215–1225, 2019.