MULTISENSOR DATA FUSION FOR
SLEEP APNEA MONITORING AND CLASSIFICATION

by

Swapna Devdinie Premasiri

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES
(Mechanical Engineering)

THE UNIVERSITY OF BRITISH COLUMBIA
(Vancouver)

July 2018

© Swapna Devdinie Premasiri, 2018
The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, a thesis/dissertation entitled:

**Multisensor Data Fusion for Sleep Apnea Monitoring and Classification**

submitted by Swapna Devdinie Premasiri in partial fulfillment of the requirement for the degree of Master of Applied Science in Mechanical Engineering.

**Examinining Committee:**

Clarence W. de Silva, Mechanical Engineering  
*Supervisor*

Lalith B. Gamage, Sri Lanka Institute of Information Technology  
*Co-supervisor*

Boris Stoeber, Mechanical, Electrical and Computer Engineering  
*Examiner*

Mu Chiao, Mechanical Engineering  
*Examiner*
Abstract

Sleep apnea is the most common type of sleep disorder that is related to breathing, amongst the adult population. Although laboratory polysomnography is the gold standard for the detection of apnea, wearable sleep monitoring devices are preferred due to many reasons such as comfort, monitoring in a familiar sleep environment, accessibility without delay, and low cost. It is important then to extract suitable features for the classification of apneic events as suitable for a wearable sleep monitoring device while maintaining the same accuracy levels as in laboratory polysomnography.

This thesis first identifies suitable preprocessing and feature extraction techniques for standard biomedical signals monitored in laboratory polysomnography. Then it develops a feature-extraction technique and designs and implements a neural network for sleep apnea detection and classification of sleep stages. Composite Multiscale Sample Entropy (CMSE) is used as the feature extraction technique, in view of its desirable characteristics. The performance of the developed methodology is evaluated using true clinical data of sleep monitoring. The designed neural networks in the present work is found to have the ability to process and classify apneic events and sleep stages while maintaining the accuracy levels of sleep scoring in clinical polysomnography, which is the existing gold standard of sleep monitoring and scoring. The neural network used for classification of sleep stages may be subsequently incorporated as the input to a neural network for classifying apneic events. In addition, the thesis demonstrates the extent to which each individual signal that is monitored in polysomnography has the ability to
independently detect apneic events. This would be useful in the implementation of a portable wearable device with clinical capability for sleep monitoring, which is the end objective of the current project.
Lay Summary

Comprehensive sleep monitoring is performed in sleep clinics for monitoring common forms of sleep disordered breathing such as sleep apnea. This thesis identifies suitable techniques for pre-processing and feature extraction of biomedical signals that are usually monitored in sleep clinics. Then it implements a sensor fusion method and two neural networks, one having the ability to accurately diagnose apneic events and the other to classify sleep stages. Identification of apneic events through sensor fusion has proven to provide better accuracy than other existing methods. A system for sleep stage classification is developed, which shows higher accuracy levels compared to similar existing systems. The ability of individual biomedical signals to classify apneic events using independent neural networks is analyzed, since this could be subsequently incorporated in to the main neural network, in a wearable sleep monitoring device that would match the performance of clinical sleep monitoring and diagnosis.
Preface

The work presented in this thesis was performed by Swapna Premasiri, in the Industrial Automation Laboratory (IAL) of The University of British Columbia (UBC), under the supervision of Prof. Clarence W. de Silva, Professor of Mechanical Engineering, UBC, and Prof. Lalith B. Gamage, Vice Chancellor and CEO, Sri Lanka Institute of Information Technology (also, an Affiliate Professor in the Department of Mechanical Engineering, UBC). The presented content; i.e., “Multisensor Data Fusion for Sleep Apnea Monitoring and Classification” was conducted independently by the author at the Industrial Automation Laboratory, The University of British Columbia, as a sub-task of the comprehensive project “Sensory Information Technologies and Implementation for Sleep Disorder Monitoring,” funded by the Natural Sciences and Engineering Research Council (NSERC) of Canada through the Strategic Partnership Grants project STPGP 493908, whose Principal Investigator is Prof. Clarence W. de Silva. Clinical Sleep Solutions Inc., Burnaby, Canada is the industrial collaborator of the project.

All data used in this study was obtained from the polysomnography recordings provided by St. Vincent’s University Hospital/University College, Dublin, Ireland, to the open source database (Physionet), which is maintained by the Massachusetts Institute of Technology (MIT) [1].

Chapter 1: presents an introduction to sleep apnea monitoring. Issues of existing systems and methods are discussed based on a thorough review of the literature related to
apnea monitoring devices and the problems indicated by the industrial collaborator. An overview of the methodology is presented in the last section of this chapter where sub-tasks for the main tasks: diagnosis of apneic events and classification of sleep stages, are summarized.

Chapter 2: and 3 present the steps followed in data preprocessing and feature extraction followed by techniques used for apnea event detection and the classification of sleep stages. The main contributions of these two chapters have been published as the conference paper:


Chapter 4: presents the tasks carried out in the classification of sleep stages, based on the feature extraction and classification techniques developed in Chapter 2: and the analysis done in Chapter 3: i.e., apneic event detection. The associated main contributions will be published in a journal as:

**Swapna Premasiri, Lalith B. Gamage, and Clarence W. de Silva, “An improved Multi-sensor Data Fusion Approach for Sleep Apnea Monitoring” (in preparation).**
Chapter 5: summarizes the main contributions of the previous chapters and suggests possible future work to improve the presented work. The improved technology may be incorporated into a comprehensive and efficient wearable device for sleep disorder monitoring and sleep apnea event detection. Such a device is being developed by the team working in the NSERC Strategic Partnership Grants project “Sensory Information Technologies and Implementation for Sleep Disorder Monitoring,” of which the current thesis forms a sub-project.
Table of Contents

Abstract .................................................................................................................. iii

Lay Summary .......................................................................................................... v

Preface .................................................................................................................... vi

Table of Contents .................................................................................................. ix

List of Tables .......................................................................................................... xii

List of Figures ......................................................................................................... xiii

List of Symbols ....................................................................................................... xiv

List of Abbreviations ............................................................................................ xv

Acknowledgements ................................................................................................ xvii

Dedication ................................................................................................................ xviii

Chapter 1: Introduction .......................................................................................... 1

1.1 Motivation and Background ........................................................................... 1

1.2 Related Work .................................................................................................. 3

1.3 Problem Definition ......................................................................................... 7

1.4 Thesis Objectives ........................................................................................... 8

1.5 Overview of Methodology ............................................................................. 9

1.6 Thesis Organization ....................................................................................... 12

Chapter 2: Data Preprocessing and Feature Extraction ....................................... 14
2.1 The Input Signals ........................................................................................................... 14
2.2 Normalization ............................................................................................................... 16
2.3 Feature Extraction Using Composite Multiscale Sample Entropy (CMSE) ............. 16
  2.3.1 Acquisition ............................................................................................................. 19
  2.3.2 Coarse Graining .................................................................................................. 19
  2.3.3 Calculating the Composite Multiscale Sample Entropy Value ......................... 21
2.4 Scaling the Extracted Features ................................................................................... 23

Chapter 3: Detection of Sleep Apnea Events ............................................................... 25
3.1 Multi-sensor Data Fusion Approach using Neural Networks ......................... 25
  3.1.1 Neural Network for Event Detection ................................................................. 25
    3.1.1.1 Scaled Conjugate Gradient Backpropagation (BP) ............................ 28
    3.1.1.2 Bayesian Regularization (BR) ................................................................. 29
  3.1.2 Simulation and Results ....................................................................................... 30
    3.1.2.1 Feature Extraction .................................................................................... 30
    3.1.2.2 Outputs from Neural Network ................................................................. 31
  3.1.3 Conclusion ........................................................................................................... 34
3.2 Multi-sensor Data Fusion using Support Vector Machines ............................ 34
  3.2.1 The Support Vector Machine ........................................................................... 35
  3.2.2 Simulation and Results ....................................................................................... 37
  3.2.3 Conclusion ........................................................................................................... 39
3.3 The Ability of Individual Signals to Classify Sleep Apnea ............................... 39
  3.3.1 Simulation and Results ....................................................................................... 40
List of Tables

Table 1.1 – AASM standard compliance: PSG vs. existing wearable SMDs. ............ 4
Table 2.1 – Types of signals used, sensor placement and units. ................................ 15
Table 3.1 – Annotations for apneic events. ............................................................. 25
Table 3.2 – NN Results for classification of apneic events using fused data. ............ 33
Table 3.3 – SVM results for apneic event classification using fused data ............... 37
Table 3.4 – The ability of each independent signal to classify apneic events .......... 40
Table 3.5 – Comparison of accuracy levels of apneic event classification systems... 43
Table 4.1 – Sleep stages and their respective annotations....................................... 46
Table 4.2 – Comparison of classification of sleep stages for different feature types. 48
Table 4.3 – Comparison of accuracy levels of SS classification systems ............... 49
Table A.1 – Basic patient information.................................................................. 70
Table A.2 – Criteria for classification of sleep stages (adopted from [12]). .......... 73
List of Figures

Figure 1.1 – Images of Laboratory PSG (a) [15] (b) [16] ............................................. 2
Figure 1.2 – Examples of existing portable sleep monitoring devices. ......................... 3
Figure 1.3 – An example of signal recordings for a laboratory PSG [1] ................. 8
Figure 1.4 – Classification of fused data from all signals. ...................................... 10
Figure 1.5 – Classification of independent signals followed by fusion............... 11
Figure 2.2 – An example of template matching in CMSE computation. ............... 22
Figure 3.1 – A representation of the NN architecture(s) used in this study. .......... 26
Figure 3.2 – An example of features extracted from a 30s window......................... 30
Figure 3.3 – Confusion Matrices for trained data using algorithms: (a) BP (b) BR... 31
Figure 3.4 – An example of visualization of feature distribution for the 4 classes. ... 35
Figure 3.5 – Hyperplanes for feature separation (a) OVA (b) OVO approach. ....... 36
Figure 3.6 – Confusion matrices for SVM trained data......................................... 38
Figure 4.1 – Confusion matrices for SS classification ......................................... 47
Figure A.1 – Schematic positioning of electrodes (10–20 system) [79]. ................. 72
List of Symbols

\( \alpha \)  
An arbitrary scaling factor which may vary between 5 to 10

\( A^m \)  
The number of repetitive sequences for templates with \( m \) elements

\( A^{m+1} \)  
The number of repetitive sequences for templates with \( m+1 \) elements

\( C \)  
Coarse grained time series (with scale factor \( S \))

\( E \)  
Original time series element

\( Feature_i \)  
Element of the feature vector

\( i \)  
Index (position) of the original time series element

\( Input_i \)  
Element of the input vector to be fed into the NN

\( j \)  
Index (position) of the coarse-grained time series element

\( jS \)  
Total number of elements in the coarse-grained time series

\( n \)  
Total number of elements in the original time series

\( N_h \)  
Number of neurons in the hidden layer

\( N_I \)  
Number of neurons in the input layer

\( N_O \)  
Number of neurons in the output layer

\( N_S \)  
Number of samples in the training dataset

\( Q_i \)  
25\(^{th}\) quantile of (the feature) dataset

\( Q_3 \)  
75\(^{th}\) quantile of (the feature) dataset

\( r \)  
Threshold (20 \% of the \( \sigma \) of the original time series)

\( S \)  
Scale factor

\( \text{SpO}_2 \)  
Peripheral Oxygen Saturation

\( \mu \)  
Mean

\( \sigma \)  
Standard deviation
### List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>AASM</td>
<td>American Academy of Sleep Medicine</td>
</tr>
<tr>
<td>AHI</td>
<td>Apnea Hypopnea Index</td>
</tr>
<tr>
<td>ANN</td>
<td>Artificial Neural Network</td>
</tr>
<tr>
<td>ApEn</td>
<td>Approximate Entropy</td>
</tr>
<tr>
<td>BP</td>
<td>Scaled Conjugate Gradient Backpropagation</td>
</tr>
<tr>
<td>BR</td>
<td>Bayesian Regularization</td>
</tr>
<tr>
<td>CMSE</td>
<td>Composite Multiscale Sample Entropy</td>
</tr>
<tr>
<td>CSA</td>
<td>Central Sleep Apnea</td>
</tr>
<tr>
<td>ECG</td>
<td>Electrocardiogram</td>
</tr>
<tr>
<td>EDR</td>
<td>ECG Derived Respiratory Signals</td>
</tr>
<tr>
<td>EEG</td>
<td>Electroencephalography</td>
</tr>
<tr>
<td>EMG</td>
<td>Electromyography</td>
</tr>
<tr>
<td>EOG</td>
<td>Electrooculography</td>
</tr>
<tr>
<td>MSampEn</td>
<td>Multiscale Sample Entropy</td>
</tr>
<tr>
<td>MSE</td>
<td>Multiscale Sample Entropy</td>
</tr>
<tr>
<td>NN</td>
<td>Neural Network</td>
</tr>
<tr>
<td>OSA</td>
<td>Obstructive Sleep Apnea</td>
</tr>
<tr>
<td>PSG</td>
<td>Polysomnography</td>
</tr>
<tr>
<td>REM</td>
<td>Rapid Eye Movement</td>
</tr>
<tr>
<td>RFC</td>
<td>Random Forest Classifier</td>
</tr>
<tr>
<td>SA</td>
<td>Sleep Apnea</td>
</tr>
<tr>
<td>SampEn</td>
<td>Sample Entropy</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>SDB</td>
<td>Sleep Disordered Breathing</td>
</tr>
<tr>
<td>SMDs</td>
<td>Sleep Monitoring Devices</td>
</tr>
<tr>
<td>SS</td>
<td>Sleep Stages</td>
</tr>
<tr>
<td>SVM</td>
<td>Support Vector Machines</td>
</tr>
<tr>
<td>TrD</td>
<td>Training Data</td>
</tr>
<tr>
<td>TstD</td>
<td>Test Data</td>
</tr>
<tr>
<td>ValD</td>
<td>Validation Data</td>
</tr>
</tbody>
</table>
Acknowledgements

My gratitude is extended to my supervisor, Prof. Clarence W. de Silva, for his rigorous guidance, advice and cooperation in my research and academic activities, and my co-supervisor, Prof. Lalith Gamage, Vice Chancellor and CEO of Sri Lanka Institute of Information Technology (SLIIT), for giving me the opportunity to work with Prof. de Silva and for guiding me throughout this research.

I acknowledge the Natural Sciences and Engineering Research Council (NSERC), Canada for funding this research through the Strategic Partnership Grants Project, “Research in Sensory Information Technologies and Implementation in Sleep Disorder Monitoring,” whose Principal Investigator is Prof. Clarence W. de Silva.

The valuable feedback given by my colleagues in the Industrial Automation Lab (IAL) and Sleep Disorder Monitoring group throughout my research is highly appreciated. I am also grateful to Clinical Sleep Solutions Inc, the industrial collaborator for this research, for their contribution.

A big thank you to my friends: particularly but not limited to; Hiroshan Gunawardane, Anna Mittelholz, Lucas Falch and Thilini Amarasooriya for the encouragement and fruitful conversations.

I am deeply indebted to my family at UBC: Mayumi and Sahan Ranamukhaarachchi for their love and encouragement, Hasitha Jayatilleka for inspiring me to work harder, Thamali Kariyawasam for all the advices and for being the sister I could share anything with and, Primal Wijesekera for always finding time to listen to me and being the perfect mentor. I would not have had a great time at UBC or gotten through the most difficult times if it weren’t for them.

Last but certainly not the least, I would like to extend my utmost gratitude to Amma and Thaththa (my parents) – Sirikumari and Saman Premasiri, for believing in me, encouraging me, their endless love and for everything they have done, for me to become who I am today.
To my parents
Chapter 1: Introduction

1.1 Motivation and Background

Sleep is a state where sensors of the body remain partially active and voluntary muscles or motors of the body remain inactive, in a restricted state of awareness [3], [4]. According to studies, a majority of adults show symptoms such as snoring, hypertension, restless legs syndrome, bruxism, choking, and cessation of breathing [5], [6], [7], [8], [9] during their sleep, as a result of underlying sleep disorders, particularly Sleep Disordered Breathing (SDB). As a result of these symptomatic disorders, they tend to suffer from excessive daytime sleepiness or fatigue.

Sleep Apnea (SA) is the most common disorder in the context of SDB and this can be categorized into Obstructive Sleep Apnea (OSA), which occurs due to collapsing of the upper airway; Central Sleep Apnea (CSA), which is a result of temporary cessation of the pacemaker that generates breathing rhythm; and Hypopnea if symptoms of both OSA and CSA are present [7], [8], [9].

Issues of sleep disordered breathing (SDB) such as sleep apnea, increases vulnerability to cardiovascular and metabolic diseases and if untreated for a prolonged period of time, risks of stroke may result. Therefore, it is important to monitor for symptoms of SA in order to improve the quality of life and to diminish such undesirable outcomes as morbidity and mortality rates caused by SA [10].
According to the American Academy of Sleep Medicine (AASM), laboratory polysomnography (PSG); i.e., sleep monitoring carried out in a professional sleep clinic through an overnight stay remains the most accurate technique and the gold standard to date [11], [12]. Generally, a sleep clinic would cost a patient an amount between $500 to $5000 per night [13], [14] which may vary depending on the technologies used and facilities provided. For proper diagnosis of sleep apnea, a patient is expected to visit the sleep clinic more than once. Hence, this procedure is expensive, and economically rather infeasible. Furthermore, according to some sleep centers, a full PSG procedure at home (in the presence of a physician) would cost $1300 [13] per night. Also, even though the diagnosis system is reliable and accurate, the process and the equipment used maybe highly uncomfortable (Figure 1.1 [15], [16]). The unfamiliar environment at a sleep clinic may also cause the monitored signals to vary from the usual signals that may have been recorded on a usual night and even if the procedure is conducted in a familiar environment to the patient, it may result in similar negative effects due to the discomfort. Besides, it is known that long waiting lists exist of patients who require clinical testing.

![Laboratory PSG (a)](image1.png) ![Laboratory PSG (b)](image2.png)

Figure 1.1 – Images of Laboratory PSG (a) [15] (b) [16]
Therefore, a wearable sleep disorder monitoring device that can be used in a familiar home environment and that is economically feasible is preferred. Nevertheless, it is essential that the standards of laboratory PSG procedures are preserved to maintain the accuracy levels of scoring and diagnosis when developing a wearable sleep monitoring device with the capability to monitor sleep disorders.

1.2 Related Work

Sleep may be monitored for the evaluation of SA using various existing devices such as ApneaLink Plus, Apnea Risk Evaluation System [17], Jawbone, Fitbit [18], Zeo [19], and iSleep [20], which are leading contenders. These devices cannot acquire all signals that are monitored in laboratory PSG (Signal details in Table 2.1). Typically, they monitor only two to three signal types while at least eight types of signals are monitored in a full PSG. Moreover, in some cases the standards of the American Academy of Sleep Medicine (AASM) are not complied with in terms of ADC resolution and minimum number of electrodes to be used (Figure 1.2 (b) [21], (c) [19], [22]) and in other cases, the devices lack comfort (Figure 1.2 (a) [17], [23] (b) [21], (c) [19], [22]) and accuracy. A summary of these issues of existing sleep monitoring devices is given in Table 1.1.

Figure 1.2 – Examples of existing portable sleep monitoring devices.
<table>
<thead>
<tr>
<th>Monitoring Technique</th>
<th>Measured Parameters</th>
<th>Issues and Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AF</td>
<td>EOG</td>
</tr>
<tr>
<td>Conventional PSG</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Wearable Sleep Monitoring Devices (SMDs)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ApneaLink Plus Device</td>
<td>✓</td>
<td>✗</td>
</tr>
<tr>
<td>ARES [17]</td>
<td>✓</td>
<td>✗</td>
</tr>
<tr>
<td>Jawbone &amp; Fitbit [17]</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>Zeo [18]</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>Beddit [19]</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>iBrain [24]</td>
<td>✗</td>
<td>✗</td>
</tr>
<tr>
<td>iSleep [20]</td>
<td>✗</td>
<td>✗</td>
</tr>
</tbody>
</table>

AF – Air flow  
EOG – Electrooculography  
S – Sound / Snoring  
EMG – Electromyography  
EEG – Electroencephalography  
ECG – Electrocardiogram  
SpO₂ – Oxygen Saturation  
M – Movement/Body position
It follows that, according to the AASM, PSG remains the most accurate technique and the gold standard to date, even though sleep scoring and apneic diagnosis are done manually by medical professionals.

As mentioned before, the current sleep apnea monitoring devices and automated techniques for apnea detection fail to use all signals used in laboratory PSG. However, they have the ability to reach a considerably high level of accuracy, even though the design or the measurement technique do not completely comply with the AASM standards as the preceding work has discussed. Varon et al. has proposed an apnea detection technique based on measurements from a single lead ECG, which guarantees an accuracy level for the prediction of apneic events of more than 85% by using least squares support vector machines (SVM) where the features used are QRS complexes [25], heartrate, standard deviation and serial correlation [26]. A similar method has been proposed by Song et al. where eight features from ECG-derived respiratory (EDR) signals and R wave to R wave (RR) intervals of the ECG recording are fed in to a Hidden Markov Model to obtain accuracy levels greater than 85% [27]. A system that uses the amplitude ratio and the frequency ratio of the abdominal and thoracic movements obtained via wearable piezoelectric bands to classify apneic events using SVMs has been suggested by Lin et al. where the accuracy levels of over 70% are possible [28].

Other biomedical applications that have high accuracy levels in performing their tasks but not commonly used in sleep apnea monitoring are studied to identify their techniques of feature extraction and classification. In other biomedical applications such
as depth of anesthesia monitoring during surgery by Liu et al. [29] and unsupervised eye
blink artifact denoising of EEG data by Mahajan et al. [30], multiscale sample entropy
(MSE) has been employed as the feature extraction technique, giving successful accuracy
rates once classified or analyzed using artificial neural networks (ANNs). Azami et al. has
proposed a method to successfully use MSE to characterize EEG signals from Alzheimer
patients [31].

Algorithms that are used for the calculation of Sample Entropy allow quantification
of irregularities of a time series [32], [33], it allows the identification of repeating patterns.
Most conventional algorithms used for Sample Entropy calculations (E.g., Shannon
Entropy) are efficient for a univariate time series and allows the quantification of
regularities of the signal, which is also an estimate of predictability. In the field of
biomedical signal processing, Approximate Entropy (ApEn), Sample Entropy (SampEn)
and Multiscale Sample Entropy (MSE) are frequently used, of whom MSE is proven to
have the highest accuracy, based on literature. Moreover, there is no serious impact even
if the data from different sensors (time series) have different lengths when using an MSE
approach unlike other techniques based on entropy. This means the technique is tolerant
to unavoidable technical errors that may affect the data acquisition of one or more sensors
or unintentional disconnection of one or more sensors.

According to the AASM scoring manual [12], sleep stage classification does not
play a significant role in in apneic event classification. However, sleep stage classification
is essential to determine the quality of sleep, as indicated, for example, in [34], [35], [36],
which is an indispensable component in any sleep monitoring device regardless of whether it is wearable or not. Sleep stage identification could also be used as an additional input in the detection of sleep disordered breathing such as sleep apnea, as mentioned in [39] and [40], and it could also play an important role in the case that the device is extended to detect other sleeping disorders such as narcolepsy, rapid eye movement and parasomnia [39], [41].

It is clear that there is a need for a sleep monitoring device that has the ability to detect apneic events and sleep stages with the same quality and accuracy as laboratory polysomnography, which will overcome the issues and inadequacies of existing sleep monitoring devices, as described before.

### 1.3 Problem Definition

In the expected sleep monitoring device, all biomedical signals that are measured in a laboratory PSG, according to the AASM standards, need to be acquired and processed. An example of signals that are recorded in this process, which are at present manually scored by professionals for necessary diagnosis, is shown in Figure 1.3.

These signals are disparate and the time series may be irregular. For example: one or more sensors may be disconnected as a result of unexpected movements by the patient or due to a faulty sensor. In such instances, the process should remain uninterrupted.

Therefore, an automated system is required to process the signals in order to bring disparate signals to a common basis and fuse the information for classification of sleep
stages and identification of apneic events, such that the system could continue functioning despite any interruptions within certain limits of tolerance. The system will initially act as an assisting resource to sleep scoring professionals and may function independently as a stand-alone device at a later stage.

![Signal recordings](image)

**Figure 1.3** – An example of signal recordings for a laboratory PSG [1].

### 1.4 Thesis Objectives

In this thesis a multi-sensor data fusion approach for accurate classification of apneic events and sleep stages using innovative feature extraction and neural network implementation are presented. Specifically, the neural networks are trained based on the
features extracted from fourteen signals using composite multiscale sample entropy (CMSE).

The accuracy levels for the neural network implemented for the classification of apneic events is compared with the accuracy levels obtained from a support vector machine (a commonly used technique used for sleep apnea even classification) that is implemented and optimized for the same task, in order to determine which most commonly used technique is particularly suitable for the application.

Furthermore, the ability of each independent signal to detect apneic events is analyzed so that at the stage where the final device is being implemented, it could be decided which signals should have low impact on the final decision in terms of weights or whether certain signals could be eliminated from the system while maintaining the same accuracy levels as the implemented system.

1.5 Overview of Methodology

This section gives an overview of steps followed for the diagnosis of apneic events and classification of sleep stages which will be discussed in chapters 2–4. An overview of the process is shown in Figure 1.4, where the classification is done after the features extracted from the signals are fused.

As shown in Figure 1.4, input data which are recordings obtained from the Physionet data base provided to MIT by the University College Dublin [1] that have been externally filtered, are normalized to reduce or eliminate any redundancy of data. Features
are extracted from the filtered original signals for subsequent use because too much information can reduce the effectiveness of classification. Therefore, meaningful information is extracted so that key information is retained. Fusion will allow the incorporation of information from multiple sensors to produce more consistent and accurate information. In this case, we perform feature level fusion prior to classification. In the presented work, there are two classification systems:

- Apneic event classification
- Sleep stage classification

The intention of the classification of sleep stages is to incorporate the system into the apneic event classification so as to improve the accuracy levels of classification.

![Diagram]

Figure 1.4 – Classification of fused data from all signals.
It is also possible to fuse the decisions made using features extracted from each of the individual signals that have been monitored or recorded, in order to classify the apneic events, as shown in Figure 1.5. In this case, the classification or decision is fused at the end of the process. This is suitable if certain signals prove to have a higher impact on the decision. This method will result in a complex system because each signal will require a separate neural network. Therefore, this method in not chosen as a suitable approach.

In this thesis, even though the approach presented in Figure 1.5 is not chosen as a suitable approach for the apnea monitoring device that is being developed by us, due to its complex nature, the ability of each signal type to classify the apneic events independently is analyzed as well, so that at a later stage the neural network within the approach presented in Figure 1.4 or any other apnea evaluation system can be improved. This may be done based on the identification of which signals should have a greater impact on the final

Figure 1.5 – Classification of independent signals followed by fusion.
decision in the process of classifying apneic events. Also, in a situation where optimization is required and the number of signals monitored in the wearable device needs to be reduced, this analysis is appropriate.

1.6 Thesis Organization

The remaining thesis is organized as follows:

In Chapter 2: preprocessing and feature extraction techniques that are typically used for biomedical signals are discussed. The process of normalizing the (externally) filtered biomedical signals involves preprocessing, and the feature extraction is performed using composite multiscale sample entropy (CMSE). The extracted features are standardized (feature-wise) using a robust scaling algorithm in an attempt to eliminate any outlying features prior to being used as the inputs for training or testing the implemented neural networks, as discussed in the preceding chapters.

Chapter 3: presents the process of implementation of a suitable neural network for sleep apnea event detection. The classification accuracy results are compared with the results of an optimized support vector machine (SVM) whose implementation is discussed in the same chapter in order to determine a more suitable technique for apneic event classification. Furthermore, the ability of each independent signal to classify apneic events is discussed in this chapter.
In Chapter 4: classification of sleep stages based on the preprocessing and feature extraction techniques presented in Chapter 2: while taking in to account the analysis given in Chapter 3: In particular, the classification of apneic events is discussed.

Chapter 5: summarizes the previous chapters while highlighting the main contributions of the thesis and proposes possible future work that would improve the presented work. Such improvements could be efficiently incorporated in to the wearable sleep disorder monitoring device which is being developed by the team working on “Sensory Information Technologies and Implementation for Sleep Disorder Monitoring” in our laboratory, for sleep apnea event detection.
Chapter 2: Data Preprocessing and Feature Extraction

2.1 The Input Signals

In order to maintain the standards of laboratory PSG, it is essential that eight or more signals are measured. In the present work, recordings that fulfill the requirements of the AASM are obtained from the polysomnography recordings provided by St. Vincent’s University Hospital/University College Dublin to the open source database “Physionet,” which is maintained by the Massachusetts Institute of Technology (MIT) [1]. This database is chosen as suitable for the application since there are overnight PSG recordings of 25 patients, while there are over 1 million data points for each signal. Furthermore, in comparison to other available databases, the database of concern comprised signals and their standards to suit the standards of the AASM.

Fourteen different biomedical signal recordings measured in the process of laboratory PSG’s are available in the mentioned database, from twenty-five patients. All the signals are used in the present study. The medical terminology of the monitored signal or process, the location of placement of the necessary sensors to measure and monitor the respective signals, and their respective units are given in Table 2.1.

For each patient, the start and end times of sleep monitoring and the basic patient details such as height, weight, gender, age, number of recordings, BMI and Apnea Hypopnea Index (AHI) as determined by professionals based on the manually scored data, are available in a separate sheet (See appendix A.1). Based on all fourteen signals for each
patient, the respective timestamps for manually-called sleep events are also available as separate files.

The sleep stages for each patient have been determined manually for intervals of 30 seconds. A portion of this information or recordings from the database has been used to train the system that has been developed and the remaining (unused) information is used to verify or cross check the developed system’s performance in the testing stage.

<table>
<thead>
<tr>
<th>Signal and Sensor Placement</th>
<th>Abbreviation</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Electroocculography – Left Eye</td>
<td>EOG-Left</td>
<td>nV</td>
</tr>
<tr>
<td>Electroocculography – Right Eye</td>
<td>EOG-Right</td>
<td>nV</td>
</tr>
<tr>
<td>Electromyography (Leg/Submental muscle)</td>
<td>EMG</td>
<td>nV</td>
</tr>
<tr>
<td>Electroencepholography (See appendix A.2)</td>
<td>EEG - C3A2</td>
<td>nV</td>
</tr>
<tr>
<td>Electroencepholography (See appendix A.2)</td>
<td>EEG – C4A1</td>
<td>nV</td>
</tr>
<tr>
<td>Electrocardiogram</td>
<td>ECG</td>
<td>nV</td>
</tr>
<tr>
<td>Oxygen Saturation (Peripheral)</td>
<td>SpO2</td>
<td>%</td>
</tr>
<tr>
<td>Tracheal Sound (Snoring)</td>
<td>Sound</td>
<td>µPa</td>
</tr>
<tr>
<td>Oronasal airflow</td>
<td>flow</td>
<td>l/min</td>
</tr>
<tr>
<td>Plethysmography (Chest)</td>
<td>Chest</td>
<td>µV</td>
</tr>
<tr>
<td>Plethysmography (Chest and Abdomen)</td>
<td>Sum</td>
<td>µVcm</td>
</tr>
<tr>
<td>Plethysmography (Abdomen)</td>
<td>Abdomen</td>
<td>µVcm</td>
</tr>
<tr>
<td>Pulse (Finger)</td>
<td>Pulse</td>
<td>/min</td>
</tr>
<tr>
<td>Body Position</td>
<td>Body Position</td>
<td>By observation</td>
</tr>
</tbody>
</table>
2.2 Normalization

The raw data or the original time series are completely disparate and their ranges vary. Therefore, prior to feature extraction, they are brought to a common ground normalizing the (externally) filtered data.

This process is necessary as it reduces, and may be in some cases eliminate, redundancy amongst the used data. The data is first normalized with respect to the mean and the standard deviation of a signal for data points from 1 to $t$ (given $t$ is the number of data points in the time series), given by,

$$\text{Signal}_{Normalized}(t) = \frac{\text{Signal}(t) - \mu}{\sigma}$$  \hspace{1cm} (2.1)

where;
- $\mu$ – Mean of the original time series, signal $(t)$
- $\sigma$ – Standard deviation of the original time series, signal $(t)$

When a dataset is normalized with respect to its mean and the standard deviation, if a particular dataset has a smaller standard deviation, it automatically receives a higher weight, which is an added advantage of using this method of normalization. Subtraction of the mean results in a zero-mean data set, which is a convenience of reference.

2.3 Feature Extraction Using Composite Multiscale Sample Entropy (CMSE)

Feature extraction allows mining of useful information without impairing key information of a signal, as suitable for a classification model. This is necessary because
presenting a system with original signals (time series) for processing or classification may not be suitable since too much information will reduce the effectiveness of data.

Sleep disorder monitoring involves information derived from several sensors over time. Hence, multivariate time series have to be taken into account in the process of information fusion. The signals may comprise repetitive patterns and they may also behave unexpectedly depending on the subject who is being monitored and the medical conditions of the subject which may vary with time as well as from person to person.

Algorithms used for calculation of the Sample Entropy allows quantification of irregularities of the time series [32], [42]. Most conventional algorithms used for Sample Entropy calculations (e.g., Shannon Entropy) are efficient for univariate time series and allow the quantification of regularities of the signal, which is also an estimate of predictability. In the field of biomedical signal processing, Approximate Entropy (ApEn), Sample Entropy (SampEn), and Multiscale Sample Entropy (MSE) are frequently used, and MSE is proven to have the highest accuracy based on previous work. Moreover, there is no serious impact even if the data from different sensors (time series) have different lengths, unlike other techniques. This means, the technique is tolerant to unavoidable technical errors that may affect the data acquisition of one or more sensors or in a case where a sensor is disconnected. Therefore, in the present work, Multivariate Sample Entropy (MSE), which is based on the fundamentals of Sample Entropy, is used as the feature extraction algorithm for the polysomnography signals.
There are improved methods that are derived based on the fundamental algorithm of MSE such as the work for the derivation of the algorithms for: Refined Composite Multiscale Entropy (RMSE) [43], Modified Multiscale Entropy for short-term time series [44], Intrinsic Mode Entropy [45], Hierarchical Entropy analysis [46], Adaptive Multiscale Sample Entropy [47], and Composite Multiscale Sample Entropy (CMSE) [32].

Based on cited literature and previous work on biomedical applications, CMSE is concluded to be most suitable method of feature extraction for this application of sleep apnea event detection and is used in the present work.

The signals used for the evaluation of sleep apnea may comprise repetitive patterns and unique readings. Identifying similar patterns corresponding to various apneic categories among different patterns will allow the predictability of apneic events. Therefore, CMSE is used in the present work to extract features from the time series of the signals that will be fed into the neural network(s) in the next stage.

Obtaining CMSE from time series data involves mainly three stages listed below.

(i) Acquisition
(ii) Coarse-graining
(iii) Calculating the sample entropy

The elaborated processes performed in each stage are described now.
2.3.1 Acquisition

This involves the acquisition of raw data from the sensors, where the data used from the database are (externally) filtered and preprocessed, as described in section 2.2, according to AASM standards.

2.3.2 Coarse Graining

Coarse-graining of each signal, which comprises a large set of time series data, is done using equation 2.2, which is a generic equation used to determine elements in a coarse-grained time series (CGTS) in the process of determining the MSE for a signal [33], [48], [49].

\[
C_j^S = \frac{1}{S} \sum_{i=(j-1)S+1}^{jS} E_i
\]  

(2.2)

Where,

\( C \) – Coarse grained time series (with scale factor \( S \))

\( S \) – Scale factor

\( E \) – Original time series element

\( i \) – Index (position) of the original time series element

\( j \) – Index (position) of the coarse-grained time series element (where \( 1 \leq j \leq n/S \))

\( n \) – Total number of elements in the original time series
The coarse graining process would decrease and progressively eliminate random uncorrelated elements in the time series, which would result in a smoothing effect on the time series [50], [51]. However, for CMSE, which is adapted from MSE, if the number of original time series elements are not divisible by the scale factor, the window that is coarse grained is shifted [32] as shown in Error! Reference source not found.. In such cases, when more than one CMSE value is present for a single scale factor, the CMSE values are averaged after the entropy value calculation is complete, i.e. after the process explained in section 2.3.3 is complete.

Although the scale factor should ideally vary from 1 to the number of elements in the original time series, scale factors ranging from 1 to 10 are used in the present work. This is consistent with the observations made from the work by Chang et al. [52], Azami et al. [31], and Wei et al. [53] who have successfully applied MSE and observed that the mentioned scale factor range is sufficient to obtain accurate results using the algorithm.

For every 30 second window for each signal that is monitored or recorded, the described coarse graining process will result in 10 coarse-grained time series (CGTS) corresponding to each scale factor. According to the mentioned previous studies, repetitive patterns for each CGTS will allow retaining of the important features while excluding uncorrelated items over the coarse graining process. The process to achieve the requirements is the process of calculating the CMSE value, which is explained in the following section.
2.3.3 Calculating the Composite Multiscale Sample Entropy Value

Calculating the sample entropy for each coarse-grained time series using the technique described in section 2.3.2, is done using:

\[ MSampEn(C^S) = -\ln \left( \frac{A^{m+1}(r)}{A^m(r)} \right) \]  

(2.3)

Where,

- \( C^S \) – Coarse grained time series with respect to the scale factor \( S \)
- \( r \) – Threshold
- \( A^m \) – The number of repetitive sequences for templates with \( m \) elements
- \( A^{m+1} \) – The number of repetitive sequences for templates with \( m+1 \) elements

The repetitive sequences are found with respect to a reference template, which is a window of size \( m \) or \( m+1 \) elements, shifted through the coarse-grained time series [32], [48], [49]. Generally, \( m \) is considered to be a template of 2 elements [50], [54]. A pattern is identified to be repetitive with respect to a time series if the element of the reference template and the corresponding element of the template that is being compared with is within the threshold \( r \). The threshold \( r \) is usually defined to have a value within 10-20 percent of the standard deviation of the original signal. This percentage has been established empirically, based on applications where the algorithm has been successfully applied. The method for determining the threshold value and the need for empirical derivation of the value are established based on the work of Wei et al. [53], Li et al. [54],
and Humeau-Heurtier [55]. In the present work, the threshold used is 20% in order to increase the rate of inclusion of repetitive patterns.

An illustration of a time series is given in Figure 2.2. Here it can be seen that for the template C₁ and C₂ of size \( m = 2 \), the threshold values are \( r₁ \) and \( r₂ \) (coincidentally, in this case the values are the same for both elements). Hence, \( r₁ \) and \( r₂ \) would be \( C₁ + r \) and \( C₁ - r \), respectively. (In a case where the values for \( C₁ \) and \( C₂ \) are different, another set of threshold values have to be considered; i.e., an upper threshold of \( C₂ + r \) and a lower threshold of \( C₂ - r \), in addition to \( C₁ + r \) and \( C₁ - r \)).

With reference to Figure 2.2, there are two sets of consecutive elements that fall within the threshold range \( r₁-r₂ \), corresponding to \( C₁ \) and \( C₂ \): \( P₁-P₂ \) and \( P₃-P₄ \). This value is \( A^m \), which is the number of repetitive sequences for a template with \( m \) elements, where
\( m = 2. \) The value for \( A^{m+1} \) is calculated in a similar manner by determining the number of repetitive sequences for a template with \( m+1 \) elements, where \( m = 3. \)

The composite multiscale sample entropy (CMSE) value also denoted as the \( MSamp(En) \) for each CGTS is expressed by a relationship between the occurrence of the number of two consecutive data points that are repetitive and the number of three consecutive data points that are repetitive, using equation 2.3. In laboratory PSG, scoring is performed by medical experts in 30 second intervals. Therefore, in order to maintain the same standards, features are extracted in a window of 30 seconds for each time series, in the present study.

2.4 Scaling the Extracted Features

The input data to the neural network (NN) and the support vector machine (SVM) (i.e., the features extracted from the signal) are scaled feature-wise using the robust scaling algorithm given by [56, 57]:

\[
Input_i = \frac{Feature_i - Q_2(Feature)}{Q_3(Feature) - Q_1(Feature)} \tag{2.4}
\]

Where,

- \( Q_3 - Q_1 \) – Interquartile range (\( Q_1 \) and \( Q_3 \) are the 25\(^{th} \) and 75\(^{th} \) quantiles, respectively)
- \( Q_2 \) – Median (50\(^{th} \) Quantile)
- \( Feature_i \) – Element of the feature vector
- \( Input_i \) – Element of the input vector to be fed into the NN
Standardizing the inputs for the NN or SVM using the robust scaling algorithm is desirable particularly in the presence of outliers since distributions are brought to the same scale and are also made to overlap that outliers will not be present within the range of the bulk of the new distributions in the data inputs [56, 58]. By this process, the features of the same standard are fused into one information set, to be presented in to a neural network.
Chapter 3: Detection of Sleep Apnea Events

3.1 Multi-sensor Data Fusion Approach using Neural Networks

The features extracted from each of the signals are fused and input to a single neural network to diagnose the presence of apneic events in intervals of 30 seconds. The implementation of the neural network and the simulation and results are elaborated in the sub-sections below.

3.1.1 Neural Network for Event Detection

The architecture of the neural network (NN) used in the present work is illustrated in Figure 3.1. The features obtained from the feature extraction process (10 scale factors or features per signal) are fed into the neural network. The fused features are to be classified into the four categories given in Table 3.1. Since the target outputs are known in the available scored records of the professionals, supervised learning is appropriate for the task.

<table>
<thead>
<tr>
<th>Signal and Sensor Placement</th>
<th>Annotation</th>
<th>Target Outputs to NN</th>
</tr>
</thead>
<tbody>
<tr>
<td>No apneic event (NE)</td>
<td>1</td>
<td>[1 0 0 0]</td>
</tr>
<tr>
<td>Central Sleep Apnea (CSA)</td>
<td>2</td>
<td>[0 1 0 0]</td>
</tr>
<tr>
<td>Obstructive Sleep Apnea (OSA)</td>
<td>3</td>
<td>[0 0 1 0]</td>
</tr>
<tr>
<td>Hypopnea (HYP)</td>
<td>4</td>
<td>[0 0 0 1]</td>
</tr>
</tbody>
</table>
The input features and the corresponding outputs, which comprise recordings of 25 subjects, each with 140 features (10 features per signal) are then split into the training, validation, and testing datasets in the ratios of 70%, 15% and 15%, respectively. The training data set is used to train the NN, where the NN “learns” while updating the weights of the system. A validation dataset is used to test if the trained NN is being generalized accurately. The training of the NN is stopped if the accuracy of predicting the outputs of the validation dataset is decreasing, in order to avoid overfitting. The test data set is a set

Figure 3.1 – A representation of the NN architecture(s) used in this study.
of data that the NN has not used yet, which when tested on, will determine the overall prediction capability of the trained NN. The implemented NN therefore comprises an input layer with neurons equal to the number of input features, a hidden layer and an output layer with four neurons, as classification is to be done into four classes. The NN is trained using the two commonly used algorithms: scaled conjugate backpropagation algorithm and the Bayesian regularization algorithm. Results obtained using the respective algorithms are discussed and compared in 3.1.3, to identify the most suitable algorithm for using in the final implementation of the system.

Only one hidden layer is used since the neural network need not be unnecessarily complex if the required performance can be achieved with a single hidden layer. The number of neurons in the hidden layer is chosen such that overfitting does not take place during the training process of the NN. The maximum number of hidden neurons that will not cause overfitting ($N_h$) is determined using a rule of thumb given by [57]:

$$N_h = \frac{N_s}{\alpha(N_I - N_O)} \quad (3.1)$$

Where,

- $N_h$ – Number of neurons in the hidden layer
- $N_I$ – Number of neurons in the input layer
- $N_O$ – Number of neurons in the output layer
- $N_s$ – Number of samples in the training dataset
- $\alpha$ – An arbitrary scaling factor which may vary between 2 and 10
Also, the optimal value for $N_h$ can be determined by equation 3.2 [59], where the definitions of the parameters are the same as for equation 3.1.

$$N_h = \sqrt{N_I N_O} \quad (3.2)$$

The number of neurons in the hidden layer is determined such that it is less than or equal to $N_h$, satisfying equations 3.1 and 3.2, and $\alpha = 3$ is used, which has been determined empirically, to achieve the best possible performance.

### 3.1.1.1 Scaled Conjugate Gradient Backpropagation (BP)

In this supervised learning algorithm (based on the original work in [60]), 70% of the data which is designated as the training dataset is used for training the neural network. After every iteration in the training process, the validation dataset, which comprises 15% of the original dataset, is introduced to the system to test its overall ability of prediction. The process is continued until at least one of the following conditions is met [61]:

(i). When the training epoch is reached; i.e., when the maximum number of epochs defined (1000 in the present case).

(ii). Training time exceeds the maximum time allowed for training by software.

(iii). When performance is minimized to the defined goal.

(iv). When performance gradient falls below the defined minimum gradient.

(v). When validation performance increases more than the number of maximum fail times defined (6 in the present case) consecutively.
If training does not stop, the errors (the difference between the corresponding targets for the training data set and the outputs resulting from training the data), are backpropagated to adjust the weights accordingly, such that the system is trained in order to reduce the error in the next iteration of training.

At the end of the training process, the test dataset, which also comprise 15% of the original dataset, to which the NN has not used thus far, is employed to test the prediction accuracy of the system to previously unexposed data.

If the results are satisfactory, the trained NN is saved for subsequent use; i.e., classification or diagnosis of real-time sleep recordings where the diagnosis will be done for periodic intervals of 30 seconds.

### 3.1.1.2 Bayesian Regularization (BR)

A linear combination of squared errors and weights is minimized in this algorithm while modifying the relationship in such a way that the system reaches a considerably high generalization ability once the training process is completed. The algorithm is based on the original work of MacKay [62]. The conditions for terminating the training is somewhat similar to that in the BP algorithm, and the conditions are listed below [61].

(i). When the training epoch is reached; i.e., when the maximum number of epochs defined (1000 in this case).

(ii). Training time exceeds the maximum time allowed for training by software.

(iii). When performance reaches the defined goal.
(iv). When performance gradient falls below the defined minimum gradient.

(v). When \( \mu \) (damping factor) exceeds the defined maximum \( \mu \).

A validation data set is not required for this process. Nevertheless, to make the two algorithms comparable, the percentage distributions of the sample data set is maintained same for both cases.

3.1.2 Simulation and Results

3.1.2.1 Feature Extraction

Each 30 second window of the fourteen signals will contain 10 features corresponding to each scaling factor. The scaling factors may increment from 1 to 10, based on the factors stated in section 2.3.2. Altogether, there will be a total of 140 features

![Figure 3.2 – An example of features extracted from a 30s window.](image-url)
fused for every window of 30 seconds. An example of the features extracted from a 30 second window for each of the signals is shown in Figure 3.2.

### 3.1.2.2 Outputs from Neural Network

![Confusion Matrices for trained data using algorithms: (a) BP (b) BR.](image)

The confusion matrices for the trained neural networks for each of the algorithms: Scaled Conjugate Backpropagation (BP) and Bayesian regularization (BR) are shown in Figure 3.3.

The lowermost row (gray) presents the prediction accuracy levels; i.e., the true positive and the false negative values. This represents the percentages of accurate classifications for each target class. The diagonal of the confusion matrices (dark blue) shows the percentage that a class is accurately classified. The rightmost column (gray)
shows the accurate classification per class. The remaining cells (light blue and light purple), represent the inaccurate classification of output classes for each target class. The bottom right column (white) in each case represents the overall prediction accuracy.

For example, in Figure 3.3 (a), 74.6% of non-apneic events (NE) are classified as non-apneic events. 0.2%, 0.4% and 0.1% of non-apneic events are classified in accurately as CSA, OSA, and HYP, respectively. Considering all non-apneic event data, 99.1% of the features have been classified accurately (which means 0.9% of non-apneic event data have been classified inaccurately) and out of all the features that have been classified as non-apneic events, 93.9% of features have been classified accurately. The overall prediction accuracy for the scaled conjugate backpropagation algorithm is denoted in the bottom right cell as 94.2%.

In Figure 3.3 (b), the confusion matrix is read in a similar manner as described earlier. The comparison of the resulting confusion matrices in the two algorithms is further described in section 3.1.3.

The prediction accuracies for the two algorithms of concern, using the neural network that is designed to classify fused data with reference to Figure 1.4 and Figure 3.1, are given in Table 3.2. The prediction accuracies for Training Data (TrD), Validation Data (ValD) and Test Data (TstD) are compared for each case.
Table 3.2 – NN Results for classification of apneic events using fused data.

<table>
<thead>
<tr>
<th>Event</th>
<th>Apneic Event Category</th>
<th>Overall Prediction Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NE</td>
<td>CSA</td>
</tr>
<tr>
<td></td>
<td>BP</td>
<td>BR</td>
</tr>
<tr>
<td>TrD</td>
<td>99.1%</td>
<td>100%</td>
</tr>
<tr>
<td>ValD</td>
<td>98.5%</td>
<td>–</td>
</tr>
<tr>
<td>TstD</td>
<td>98.5%</td>
<td>95.4%</td>
</tr>
</tbody>
</table>

For each category of sleep apnea, it can be seen that accuracy levels of classification of training data is higher when using BR algorithm. For test data, the prediction accuracy levels when using BR is higher for cases of CSA and HYP, and lower for NE and OSA compared to when the BP algorithm is used. The overall average accuracy for test data are only slightly different when the two algorithms are compared. Therefore, with respect to the overall average accuracy and the accuracy of training data, performance when the BR algorithm is used seems to be better with respect to prediction accuracy levels.

The Bayesian regularization (BR) algorithm does not require the use of a validation data set. Therefore, the table will not indicate the percentage values for validation data (ValD) for the BR algorithm.
3.1.3 Conclusion

Section 3.1 presents a technique of sleep apnea event classification using neural networks, comparing the accuracy of diagnosis using the two learning algorithms during training: scaled conjugate backpropagation and Bayesian regularization. CMSE, a variant of MSE, which has been successfully used in other biomedical applications [29], [30], [31], [32] is used to extract features from the signal recordings of a laboratory PSG.

The neural network where the extracted features were trained using the Bayesian regularization algorithm, yielded a higher overall average accuracy compared to the same set of features trained using the scaled conjugate backpropagation algorithm. Comparing the results from the two algorithms, the classification accuracies for the test data are close. Therefore, considering the classification accuracies for training data, it is seen that the neural network when the BR algorithm was used is higher. The overall average accuracy of the final NN is 94.5%. Therefore, the final system will be implemented with the NN trained using the Bayesian regularization algorithm.

3.2 Multi-sensor Data Fusion using Support Vector Machines

Support Vector Machines (SVMs) have been popular in classification applications [63] such as the text classification [64], [65] and predicting risks in pits in infrastructure projects [66], and in particular, classification applications of biomedicine such as the work done by: Hsu et al. for EEG Classification of Imaginary Lower Limb Stepping Movement Based on Fuzzy SVM [67], Spilka et al. for Fetal heart-rate classification [68], and Varon
et al. [26] for detection of sleep apnea. In this backdrop, the classification of apneic events is done here using an SVM to compare the results obtained from the NN implemented for the same task in section 3.1. The inputs to the SVM are the same as those used for the NNs.

3.2.1 The Support Vector Machine

The features provided to the support vector machine (SVM) are the same set of standardized features that were used as the inputs to the neural network. As shown in Figure 3.4, the features to be supplied to the current system are difficult to separate.

In cases where the features are linearly separable (as shown in Figure 3.5), the separation of classes is done by using linear hyperplanes as shown (adopted from [69], [70]).

![Figure 3.4 – An example of visualization of feature distribution for the 4 classes.](image)
The task of the SVM is to determine an optimal hyperplane for the patterns of features, maximizing the margin surrounding the hyperplane separating the classes. Thus, it is possible to use the following two approaches when using a support vector machine:

(i) One-against-all (OVA) \((\text{See Figure 3.5} (a))\)

(ii) One-against-one (OVO) \((\text{See Figure 3.5} (b))\)

For features that are difficult to separate, a nonlinear SVM is required. Therefore, Medium Gaussian SVM with the Gaussian kernel is chosen as the classification learner because of the nature of separation of the classes as shown in the example in Figure 3.4, and also based on the information from [71]. The more suitable approach from OVO and OVA, for data classification is determined empirically by analyzing the derived outputs for systems based on the results in section 3.2.2.
3.2.2 Simulation and Results

Results of the implemented SVM is given in Table 3.3 and the respective classification accuracies for the trained data are presented using the confusion matrix in Figure 3.6.

![Table 3.3 – SVM results for apneic event classification using fused data](image)

Based on the prediction accuracies of Table 3.3, it can be seen that the accuracy levels for training data when using both approaches differ only by 0.5%. Therefore, the suitable approach is decided based on the prediction accuracy levels for test data and the overall average accuracy. Accordingly, the OVA approach is chosen.
The confusion matrices are read as explained in section 3.1.2.2. For example, in Figure 3.6 (a), 74.3% of non-apneic events (NE) are classified as non-apneic events. 1.9%, 3.1% and 0.1% of non-apneic events are classified as accurately as CSA, OSA and HYP, respectively. Considering all the non-apneic event data, 93.6% of the features have been classified accurately (which means 6.4% of non-apneic event data have been classified inaccurately) and out of all the features that have been classified as non-apneic events, 100% of features have been classified accurately. The overall prediction accuracy for the scaled conjugate backpropagation algorithm is denoted in the bottom right cell as 94.4%.

In Figure 3.6 (b), the confusion matrix is read in a similar manner as described for Figure 3.6 (a).
3.2.3 Conclusion

Based on the results of section 3.2.2, it is seen that the SVM using the one vs. one (OVO) approach yielded slightly lower prediction accuracy levels compared to the SVM that used the one vs. all (OVA) approach for training data. However, when the system was tested on the test dataset, the OVO approach did not seem to have the ability to classify all the classes efficiently; i.e., the SVM that used the OVO approach did not have the ability to classify three of the four classes at all.

Thus, the approach for the SVM was chosen as OVA, since it demonstrated a prediction accuracy level of 90% once tested on the training dataset. Accordingly, the results from the OVA are compared with the finalized NN design. The overall average accuracy of the finalized SVM is 91.2%. This is also consistent with the conclusion derived from the work done by Milgram et al. [72].

3.3 The Ability of Individual Signals to Classify Sleep Apnea

The ability of each individual signal to classify apneic events is analyzed since it is useful to identify which signals would have a greater impact on the final decision. In this particular study, the information from all signals are fused following the feature extraction process for apneic event diagnosis. However, even though the technique provides promising results, the system could be further improved by changing the impact that more influencing signals have on the final diagnosis; i.e., by changing the weights corresponding to the input layer of the NN. The analysis is carried out such that the suggested
improvement could be carried out at a later stage. It was decided that the use of neural networks was suitable for this task as well, based on the results and analysis in the previous sections, 3.1 and 3.2.

The neural network used for each signal and the number of neurons in the hidden layer are determined using the same technique as used for the diagnosis process of the fused features. Specifically, the architectures are based on Figure 3.1 and the number of neurons in the hidden layers are determined using equations 3.1 and 3.2.

### 3.3.1 Simulation and Results

The classification accuracy levels for each independent signal is tabulated in section 3.3.1 in Table 3.4 and the results are discussed in section 3.3.2.

<table>
<thead>
<tr>
<th>Signal Type</th>
<th>Classification Accuracy for Apneic Event Categories</th>
<th>Overall Average Prediction Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NE</td>
<td>CSA</td>
</tr>
<tr>
<td>EOG-Left</td>
<td>95.2%</td>
<td>87.5%</td>
</tr>
<tr>
<td>EOG-Right</td>
<td>95.2%</td>
<td>87.5%</td>
</tr>
<tr>
<td>EMG</td>
<td>84.2%</td>
<td>47.6%</td>
</tr>
<tr>
<td>EEG–C3A2</td>
<td>81.5%</td>
<td>45.8%</td>
</tr>
<tr>
<td>EEG–C4A5</td>
<td>76.2%</td>
<td>66.7%</td>
</tr>
<tr>
<td>ECG</td>
<td>66.7%</td>
<td>62.5%</td>
</tr>
<tr>
<td>SpO₂</td>
<td>9.5%</td>
<td>12.5%</td>
</tr>
</tbody>
</table>
### Signal Type

<table>
<thead>
<tr>
<th>Signal Type</th>
<th>Classification Accuracy for Apneic Event Categories</th>
<th>Overall Average Prediction Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>NE</td>
<td>CSA</td>
<td>OSA</td>
</tr>
<tr>
<td>Sound</td>
<td>95.2%</td>
<td>66.7%</td>
</tr>
<tr>
<td>Flow</td>
<td>4.8%</td>
<td>33.3%</td>
</tr>
<tr>
<td>Chest</td>
<td>0.0%</td>
<td>54.2%</td>
</tr>
<tr>
<td>Sum</td>
<td>28.6%</td>
<td>70.8%</td>
</tr>
<tr>
<td>Abdomen</td>
<td>38.1%</td>
<td>25.0%</td>
</tr>
<tr>
<td>Pulse</td>
<td>0.0%</td>
<td>16.7%</td>
</tr>
</tbody>
</table>

### 3.3.2 Analysis of Results

The ability of each signal to classify apneic events based on the process in Figure 1.5 is given in Table 3.4. It is observed that the features extracted from EEG, EMG, EOG, and ECG signals can be used to classify an apneic event with an accuracy of over 60%. The signals detect OSA and hypopnea with a greater accuracy than CSA. EOG signals have the greatest ability to detect CSA as well as the highest overall accuracy. Even though SpO$_2$ signals do not have a high accuracy in classifying all apneic events, it has an accuracy of 94.7% to detect OSA, probably because obstructions in the airway significantly reduce the level of oxygen saturation. Similarly, features extracted from the microphone input and the airflow sensors have an accuracy of nearly 90% in detecting OSA, even though other apneic events have not been successfully classified using these signals alone. All plethysmography-based recordings allow accurate classification of OSA. Also, the combined plethysmography readings obtained from the abdomen and the chest allow
detection of CSA as well. The tabulated information can be used to subsequently improve the overall neural network by adjusting its weights. This analysis may be used to compare the optimized weights of the trained system as well. Furthermore, this analysis may also be used if there is a situation where the number of signals used for monitoring are to be reduced for optimization purposes while maintaining the accuracy levels of lab PSG, when implementing the wearable sleep monitoring device.

3.4 Conclusion – Sleep Apnea Event Detection

Based on the results from this chapter it can be observed that the use of all signals in laboratory PSG for classification of apneic events, is more successful than using one or just several sensor signals for the task; i.e., using fused features with a total of 140 features will allow the system to have higher accuracy levels for prediction compared to when features extracted from one or few sets of features are used.

The NN designed using the Bayesian regularization algorithm yielded better results with respect to accuracy, compared to when the scaled conjugate backpropagation was used as the learning algorithm, based on the results obtained in section 3.1.2. Results shown and discussed in section 3.2.2 demonstrated that the one vs. all approach yielded better performance in terms of accuracy rather than when the one vs. one approach was used, when a SVM was used to classify apnea events.

Thus, when comparing results from the two techniques; i.e., when the classification system was designed using SVM and NN, it could be concluded that the NN showed a
higher ability in terms of accuracy to classify the sleep data for detecting apneic events from the extracted features, as it showed an overall classification accuracy of 99.1% whereas the finalized SVM approach only yielded an accuracy of 94.4% for the training dataset. The classification accuracies for the test dataset were the same in both cases (90.0%). However, the overall average accuracies for the finalized NN and SVM were 94.8% and 91.2%, respectively.

From the foregoing, it is concluded that the neural network designed as described in section 3.1.1 is the suitable technique to classify apneic events based on the processed and extracted features. Based on this conclusion, a comparison of the accuracy levels of the present work with other apneic event classification systems discussed in section 1.2, are summarized in Table 3.5.

From this comparison it is observed that the average classification accuracy level of the present work guarantees a higher accuracy level compared to other apneic event classification systems and that it complies with the standards of the AASM, unlike other compared systems, initially.

<table>
<thead>
<tr>
<th>Classification System</th>
<th>Overall Avg. Accuracy</th>
<th>Complies with AASM standards?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present work</td>
<td>&gt; 94 %</td>
<td>Yes</td>
</tr>
<tr>
<td>Varon et al. [73]</td>
<td>&gt; 85 %</td>
<td>No</td>
</tr>
<tr>
<td>Song et al. [74]</td>
<td>&gt; 85 %</td>
<td>No</td>
</tr>
<tr>
<td>Lin et al. [75]</td>
<td>&gt; 70%</td>
<td>No</td>
</tr>
</tbody>
</table>
The present work confirmed the aforementioned fact by comparing the demonstrated results with devices or techniques developed or proposed for sleep monitoring without multi-sensor data fusion [17], [20], [18], [19], and it could be observed that these techniques were not as accurate as using fused information obtained from all sensor signals. Furthermore, it was shown that this system guarantees accuracy levels higher than with other systems that have been proposed for apneic event classification. This system may be further improved using the information obtained on prediction accuracies of individual sensor signal features by changing the weights of a particular sensor based on its importance when implementing the neural network to train fused feature information.
Chapter 4: Classification of Sleep Stages

4.1 Multi-sensor Data Fusion for Classification of Sleep Stages

Sleep stage classification influences the classification of apneic events, according to the AASM scoring manual [12]. Besides, the process is a crucial task when determining the quality of sleep. A few examples of related work are [34], [35], [36], [37], [38]. This information will also be useful in situations where other sleep disorders such as parasomnia need to be identified [39], [41].

According to [12], scoring of sleep stages is predominantly based on EEG recordings and occasionally on EOG recordings (See Appendix A.3). Therefore, it is possible that the classification of sleep stages can be done based only on the features extracted from EEG signals. Thus, classification of sleep stages based on features extracted from the following signals will be analyzed in this chapter:

(i). All fourteen biomedical signals used in laboratory PSG
(ii). Both EEG and EOG recordings
(iii). Only EEG recordings
(iv). Only EOG Recordings

From Chapter 3: it was established that for the features extracted from the signals, the best technique for the classification of sleep stages would be neural networks, where
Bayesian regularization is suited best as the training algorithm. Hence, the same techniques are used for classification of sleep stages as well.

The sleep stages are classified as given in Table 4.1 and the most suitable set(s) of input features are decided based on the results shown and described in section 4.1.1. For each annotation, a corresponding binary output is defined for training purposes of the neural network, as shown in Table 4.1

<table>
<thead>
<tr>
<th>Sleep Stage</th>
<th>Annotation</th>
<th>Target Outputs to NN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wake (W)</td>
<td>0</td>
<td>[ 1 0 0 0 0 0 ]</td>
</tr>
<tr>
<td>REM</td>
<td>1</td>
<td>[ 0 1 0 0 0 0 ]</td>
</tr>
<tr>
<td>Stage I (SI)</td>
<td>2</td>
<td>[ 0 0 1 0 0 0 ]</td>
</tr>
<tr>
<td>Stage II (SII)</td>
<td>3</td>
<td>[ 0 0 0 1 0 0 ]</td>
</tr>
<tr>
<td>Stage III (SIII)</td>
<td>4</td>
<td>[ 0 0 0 0 1 0 ]</td>
</tr>
<tr>
<td>Stage IV (SIV)</td>
<td>5</td>
<td>[ 0 0 0 0 0 1 ]</td>
</tr>
</tbody>
</table>

(Technically, there are “Indeterminate” and “Artifact” classes according to the annotations given in the database. However, these stages are not present in the available scored data. Thus, these two classes are excluded in the process of classification of sleep stages as performed in the present study).
4.1.1 Simulation and Results

The confusion matrices for training data resulting from the classification of sleep stages using the features extracted from all data, EEG signals, EOG signals and both EEG and EOG Signals, are given in Figure 4.1. The results and the average prediction accuracies for the aforementioned dataset and the test dataset are summarized in Table 4.2.

![Confusion Matrices](image)

**Figure 4.1 – Confusion matrices for SS classification using features extracted from: (a) All, (b) EEG, (c) EOG, (d) EEG and EOG signals.**
Table 4.2 – Comparison of classification of sleep stages for different feature types.

<table>
<thead>
<tr>
<th>Features</th>
<th>Accuracy Classification of Sleep Stages</th>
<th>Overall Prediction Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Wake</td>
<td>REM</td>
</tr>
<tr>
<td>All</td>
<td>98.7%</td>
<td>99.3%</td>
</tr>
<tr>
<td>EEG</td>
<td>98.6%</td>
<td>94.6%</td>
</tr>
<tr>
<td>EOG</td>
<td>96.1%</td>
<td>95.6%</td>
</tr>
<tr>
<td>EEG &amp; EOG</td>
<td>98.0%</td>
<td>99.0%</td>
</tr>
</tbody>
</table>

4.1.2 Conclusion

Based on the results shown in the previous section, it is observed that the neural networks which used features from EEG and EOG or all the features have better accuracy levels compared to the NN that used EEG signals or EOG signals only. It is also observed that even though the accuracy levels when using all features to classify sleep stages yielded a slightly higher prediction accuracy compared to when both EEG and EOG signals were used, the accuracy levels on predicting sleep stages for the test dataset is considerably higher. Therefore, it is concluded that all features should be used for the task of sleep stage classification with the given settings and training as described, for the wearable device that is being developed in the overall project. Nevertheless, in the final stage, if the device requires optimization with respect to the number of channels, based on the results, it is observed that satisfactory performance can be obtained using only EEG and EOG channels.
Furthermore, the results of the finalized system for the classification of sleep stages are compared with existing recent devices or proposed systems that perform the same task, in order to explore the classification accuracy levels of the present system. The comparisons are presented in Table 4.3.

<table>
<thead>
<tr>
<th>Proposed by</th>
<th>Classification Model</th>
<th>Signals/Channels Used</th>
<th>Overall Average Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present work</td>
<td>NN</td>
<td>14 signals (<em>See Table 2.1</em>)</td>
<td>&gt; 85 %</td>
</tr>
<tr>
<td></td>
<td>NN</td>
<td>2 EEG channels</td>
<td>&gt; 80 %</td>
</tr>
<tr>
<td></td>
<td>NN</td>
<td>2 EEG + 1 EOG channels</td>
<td>&gt; 60 %</td>
</tr>
<tr>
<td>Gouveris et al. [76]</td>
<td>SVM</td>
<td>2 EEG channels</td>
<td>≈ 85 %</td>
</tr>
<tr>
<td>Nakamura et al. [77]</td>
<td>SVM</td>
<td>1 EEG channel</td>
<td>&gt; 93 %</td>
</tr>
<tr>
<td>Alickovic et al. [78]</td>
<td>SVMs</td>
<td>1 EEG channel</td>
<td>&gt; 84 %</td>
</tr>
<tr>
<td>Memar et al. [79]</td>
<td>RFC</td>
<td>2 EEG channels</td>
<td>&gt; 95 %</td>
</tr>
<tr>
<td>da Silveira et al. [80]</td>
<td>RFC</td>
<td>1 EEG channel</td>
<td>&gt; 92 %</td>
</tr>
</tbody>
</table>

The overall average accuracy levels are the highest in the present work when fused features of all 14 signals are used. Although the training accuracy levels are higher when compared to other existing systems for classification of sleep stages, the average accuracy of prediction is comparatively low. This could be due to the limitations of scored data that are available for training. While there are recordings of 25 patients where sleep stages are classified for every 30 second window, when the percentages are split for training, validation and testing, certain sleep stages may occur less frequently and hence they may
be not included in the categories. Therefore, accuracy levels for test data may be lower than expected.
Chapter 5: Conclusion and Further Work

5.1 Conclusion and Contributions

In this thesis, four main contributions have been presented; namely, implementation of a neural network for classification of apneic events, comparing the results of the finalized neural network design with an optimized support vector machine designed for the same task, analysis of the capability of independent biomedical signals recorded in a laboratory PSG to classify apneic events using neural networks and classification of sleep stages using neural networks. To carry out these tasks, the filtered raw monitored signals were normalized, features were extracted using the composite multiscale sample entropy (CMSE) algorithm, and the extracted features were standardized prior to using as inputs to the neural network or the support vector machine. The results of the final neural network, which has been designed for apneic event detection and classification of sleep stages, show that the pre-processing and feature extraction techniques used in the present thesis is appropriate for applications involving biomedical signals and for a multi-sensor data fusion approach.

For the classification of apneic events using neural networks, appropriate analysis and comparison have been made for different learning algorithms (specifically, scaled conjugate gradient backpropagation and Bayesian regularization). The results and comparison show that the neural network has a higher prediction accuracy when trained using the Bayesian regularization algorithm rather than when the scaled conjugate gradient
backpropagation algorithm is used as the learning algorithm. Hence, the final neural network comprises a single hidden layer, while Bayesian regularization is used as the learning algorithm.

A Gaussian kernel is used for the support vector machine (medium Gaussian SVM) designed for apneic event classification. It was shown that, the overall average accuracy in predicting apneic events was higher when the one versus all approach was used, compared to when the one versus one approach was used. This conclusion that is based on results seems to be consistent with other studies that involve similar comparisons.

By the analysis performed in this thesis to explore the abilities of independent signals to classify apneic events, it is observed that certain signals such as EEG and EOG signals have a higher ability to classify apneic events. Moreover, the results show that a system using a multi-sensor data fusion approach is more consistent and tends to yield higher accuracy levels in terms of prediction, compared to systems that use one of few sensors. This observation is made since the prediction accuracy of the corresponding neural network to any independent signal classifying apneic events is not higher than the overall prediction accuracy of the final neural network.

Based on the results obtained for the neural network designed for classification of sleep stages, it is observed that accuracy levels are the highest when all features are used. However, it is observed that satisfactory results can be obtained even if features extracted
only from EEG and EOG signals are used, given that a situation arises where all channels are not used in the final device that will be implemented.

This thesis has identified suitable preprocessing and feature extraction techniques for laboratory polysomnography recordings which may be used as inputs to the two neural networks presented in this work. For each case, thorough study, comparisons and analysis have been made in deciding the algorithms and techniques that are suitable for each process.

The limitations of this work are mostly based on the limitations of the available data. This is because the database contains scored sleep data of only patients who suffer from sleep apnea as the patient would have visited the sleep clinic for overnight recordings only in the presence of symptoms of sleep apnea. Therefore, recordings from healthy subjects who do not suffer from sleep apnea are not readily available. Also, the occurrence of certain apneic events is uncommon within the database. In such cases, when the data is split into training, validation and test data, the features corresponding to the uncommon classes may not be adequately available within the datasets. This can be overcome by including recordings of more patients; i.e., by using a larger database.

5.2 Possible Future Work

The work presented in this thesis shows that future research based on this work could be focused in two directions: First, improving the accuracy levels based on the adjustment of corresponding weights to the inputs based on the analysis done on the
classification accuracies using neural networks implemented for individual signals. Second, a health quality index could be developed which would give an interpretation of the severity of sleep apnea and the quality of sleep.

The work done in this thesis used the open access database, available in [1]. Therefore, the developed systems must be optimized as suited for the signals obtained from the sensors that are being developed by other members of the wearable sleep monitoring device project. Once this is done, the accuracy and completeness of the data needs to be estimated. Furthermore, the confidence level of the final decision needs to be interpreted. These tasks may be achieved by comparing and determining the percentage errors between the outputs by the implemented system with scoring done by several experts of sleep scoring.
Bibliography


[72] J. M. M. Cheriet and R. Sabourin, "‘One against one’ or ‘one against all’: Which one is better for handwriting recognition with SVMs?,” in *tenth international workshop on Frontiers in handwriting recognition*, 2006.


Appendices

Appendix A : Information Relevant to Input Signals Used in This Study

A.1 Available Basic Patient Information

Basic information of patients whose recordings have been entered in the polysomnography database used in this study is given below in Table A.1.

<table>
<thead>
<tr>
<th>Patient No.</th>
<th>Height (cm)</th>
<th>Weight (kg)</th>
<th>Gender</th>
<th>PSG Start Time</th>
<th>AHI</th>
<th>BMI</th>
<th>Age</th>
<th>Duration (hr)</th>
<th>Data Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>172</td>
<td>100.3</td>
<td>M</td>
<td>00:11:04</td>
<td>23</td>
<td>33.9</td>
<td>54</td>
<td>6.2</td>
<td>22470</td>
</tr>
<tr>
<td>02</td>
<td>179</td>
<td>102.0</td>
<td>M</td>
<td>23:07:50</td>
<td>51</td>
<td>31.8</td>
<td>48</td>
<td>7.3</td>
<td>26478</td>
</tr>
<tr>
<td>03</td>
<td>176</td>
<td>100.4</td>
<td>M</td>
<td>23:28:42</td>
<td>13</td>
<td>32.4</td>
<td>65</td>
<td>6.9</td>
<td>24798</td>
</tr>
<tr>
<td>04</td>
<td>185</td>
<td>103.5</td>
<td>M</td>
<td>23:57:14</td>
<td>31</td>
<td>30.2</td>
<td>52</td>
<td>6.7</td>
<td>24267</td>
</tr>
<tr>
<td>05</td>
<td>183</td>
<td>084.0</td>
<td>M</td>
<td>23:30:22</td>
<td>12</td>
<td>25.1</td>
<td>47</td>
<td>6.8</td>
<td>24405</td>
</tr>
<tr>
<td>06</td>
<td>145</td>
<td>059.8</td>
<td>F</td>
<td>23:29:11</td>
<td>05</td>
<td>28.4</td>
<td>63</td>
<td>6.4</td>
<td>23041</td>
</tr>
<tr>
<td>07</td>
<td>180</td>
<td>101.5</td>
<td>M</td>
<td>22:35:22</td>
<td>12</td>
<td>31.3</td>
<td>52</td>
<td>7.7</td>
<td>27759</td>
</tr>
<tr>
<td>08</td>
<td>174</td>
<td>119.0</td>
<td>M</td>
<td>22:51:18</td>
<td>34</td>
<td>39.3</td>
<td>38</td>
<td>7.6</td>
<td>27211</td>
</tr>
<tr>
<td>09</td>
<td>188</td>
<td>101.0</td>
<td>M</td>
<td>22:47:38</td>
<td>08</td>
<td>28.6</td>
<td>51</td>
<td>7.5</td>
<td>27030</td>
</tr>
<tr>
<td>10</td>
<td>179</td>
<td>097.5</td>
<td>M</td>
<td>23:23:21</td>
<td>25</td>
<td>30.4</td>
<td>51</td>
<td>7.2</td>
<td>25941</td>
</tr>
<tr>
<td>11</td>
<td>153</td>
<td>080.0</td>
<td>F</td>
<td>23:44:00</td>
<td>16</td>
<td>34.2</td>
<td>62</td>
<td>6.8</td>
<td>24333</td>
</tr>
<tr>
<td>12</td>
<td>177</td>
<td>091.0</td>
<td>M</td>
<td>23:37:59</td>
<td>36</td>
<td>29</td>
<td>56</td>
<td>6.4</td>
<td>23239</td>
</tr>
<tr>
<td>13</td>
<td>170</td>
<td>083.9</td>
<td>M</td>
<td>23:02:45</td>
<td>06</td>
<td>29</td>
<td>28</td>
<td>7.6</td>
<td>27488</td>
</tr>
<tr>
<td>14</td>
<td>176</td>
<td>117.0</td>
<td>M</td>
<td>23:16:05</td>
<td>12</td>
<td>37.8</td>
<td>53</td>
<td>6.6</td>
<td>23684</td>
</tr>
<tr>
<td>Patient No.</td>
<td>Height (cm)</td>
<td>Weight (kg)</td>
<td>Gender</td>
<td>PSG Start Time</td>
<td>AHI</td>
<td>BMI</td>
<td>Age</td>
<td>Duration (hr)</td>
<td>Data Points</td>
</tr>
<tr>
<td>------------</td>
<td>-------------</td>
<td>-------------</td>
<td>--------</td>
<td>----------------</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>15</td>
<td>171</td>
<td>077.0</td>
<td>M</td>
<td>23:49:02</td>
<td>02</td>
<td>26.3</td>
<td>35</td>
<td>6.8</td>
<td>24685</td>
</tr>
<tr>
<td>16</td>
<td>178</td>
<td>097.8</td>
<td>M</td>
<td>23:30:33</td>
<td>16</td>
<td>30.9</td>
<td>49</td>
<td>7.1</td>
<td>25573</td>
</tr>
<tr>
<td>17</td>
<td>179</td>
<td>108.8</td>
<td>M</td>
<td>23:48:21</td>
<td>15</td>
<td>34</td>
<td>52</td>
<td>6.3</td>
<td>22586</td>
</tr>
<tr>
<td>18</td>
<td>161</td>
<td>087.0</td>
<td>F</td>
<td>22:52:05</td>
<td>13</td>
<td>33.6</td>
<td>41</td>
<td>7.6</td>
<td>27409</td>
</tr>
<tr>
<td>19</td>
<td>166</td>
<td>080.7</td>
<td>M</td>
<td>23:35:05</td>
<td>07</td>
<td>29.3</td>
<td>34</td>
<td>6.6</td>
<td>23640</td>
</tr>
<tr>
<td>20</td>
<td>165</td>
<td>089.0</td>
<td>F</td>
<td>22:55:51</td>
<td>39</td>
<td>32.7</td>
<td>68</td>
<td>7.2</td>
<td>25850</td>
</tr>
<tr>
<td>21</td>
<td>172</td>
<td>099.9</td>
<td>M</td>
<td>22:58:02</td>
<td>24</td>
<td>33.8</td>
<td>54</td>
<td>7.6</td>
<td>27250</td>
</tr>
<tr>
<td>22</td>
<td>174</td>
<td>128.6</td>
<td>M</td>
<td>0:25:37</td>
<td>91</td>
<td>42.5</td>
<td>52</td>
<td>5.9</td>
<td>21350</td>
</tr>
<tr>
<td>23</td>
<td>175</td>
<td>084.0</td>
<td>M</td>
<td>22:58:13</td>
<td>14</td>
<td>27.4</td>
<td>49</td>
<td>7</td>
<td>25160</td>
</tr>
<tr>
<td>24</td>
<td>182</td>
<td>093.0</td>
<td>M</td>
<td>22:56:30</td>
<td>55</td>
<td>28.1</td>
<td>45</td>
<td>7.4</td>
<td>26791</td>
</tr>
<tr>
<td>25</td>
<td>172</td>
<td>088.9</td>
<td>M</td>
<td>0:29:08</td>
<td>46</td>
<td>30.1</td>
<td>50</td>
<td>6</td>
<td>21660</td>
</tr>
</tbody>
</table>

*Apnea Hypopnea Index (AHI) is the average number of occurrences of apneas or hypopneas (abnormal events) that takes place during a person’s sleep.*
A.2 The 10–20 System Electrode Placement for EEG Monitoring

Various systems are used for EEG electrode placement, out of which the 10–20 system shown in Figure A.1, is the most commonly used one. The combinations where electrodes had been placed for the EEG signals to be recorded in this particular study were C3–A2 and C4–A1, with reference to the 10–20 system, where A1 and A2 were considered as reference points.

Figure A.1 – Schematic positioning of electrodes (10–20 system) [81].
A.3 Criteria for Classification of Sleep Stages According to AASM

It was discussed in Chapter 4: that the classification of sleep stages is done predominantly based on EEG recordings and occasionally, EOG recordings. Therefore, the criteria followed in manual scoring based on the AASM scoring manual [12] is summarized below.

<table>
<thead>
<tr>
<th>Sleep Stage</th>
<th>Characteristics</th>
</tr>
</thead>
</table>
| **W** Waking State | If the following takes place for more than 50% of the epoch:  
Dominant alpha rhythm trains in EEG recordings  
Vertical eye movements (Blinking) between 0.5–2 Hz, SEM’s followed by REM’s (Reading EM’s) in EOG recordings |

| N1 Drowsiness | If alpha rhythm is replaced by low amplitude mixed frequency (LAMF) EEG signals (4–7 Hz)  
OR if there is no dominant alpha rhythm in the EEG signal,  
EEG activity in 4–7 Hz range and slowing of background frequencies to >1 Hz compared to stage W  
V waves – i.e., waves that are clearly distinguishable from background activity due to its sharp nature, usually lasts less than 0.5 seconds.  
SEMS  
The majority of the epoch should comprise characteristics of N1 and until there is no indication of initiation of another sleep stage.  
Other occurrences where N1 scoring is necessary are given below and will be considered as N1 until initiation of another sleep stage is initiated.  
The successive segments if stage N2 is interrupted by an arousal  
The segments of the recordings with EM’s in a case where stage R is interrupted by an arousal, followed by LAMF EEG signals and SEM
<table>
<thead>
<tr>
<th>Sleep Stage</th>
<th>Characteristics</th>
</tr>
</thead>
</table>
| **N2** Light Sleep | An epoch is scored as N2 when:  
One or more sleep K complexes take place without arousal and/or when one or more sleep spindles take place when there is no sign of N3 characteristics.  
Even if an epoch with LAMF signals does not comprise the characteristics, that epoch may be scored as N2 if the preceding epochs fulfill the criteria.  
The majority of an epoch comprises aforementioned characteristics.  
When a K complex and an arousal take place in the same epoch, the preceding epoch is scored as N2.  
When the characteristics of an epoch fail to meet the criteria for N3 and if it does not meet or initiate stage W or R as well.  
Scoring of N2 is halted when:  
There is a transition to stages W, N3 or R  
An arousal takes place followed by LAMF (scored as N1 till K complex occurs)  
A major body movement followed by SEMs and LAMF EEG signals without K complexes or sleep spindles  
(In the absence of SEMs, the epoch is scored as N2 and the following epoch is scored as N1). |
| **N3** SWS | An epoch is scored as N3 during Slow Wave Sleep (SWS) when:  
Slow Wave Activity (SWA) is observed; i.e., waves with frequencies ranging from 0.5 to 2 Hz and amplitudes greater than 75 μV in EEG signals of frontal regions  
20 percent or more of an epoch comprises SWA |