Human Activity Recognition and Sleep Episode Classification Methods to Analyse Sleep Quality in COPD Patients

Isabel Martínez-Barona García

MInf Project Report
Master of Informatics
School of Informatics
University of Edinburgh

2025

Abstract

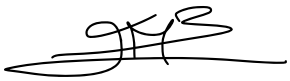
Sleep quality is critically affected in individuals with Chronic Obstructive Pulmonary Disease (COPD), contributing to poor health outcomes and diminished quality of life. This project explores the use of wearable sensor data—collected via the RESpeck device—for objectively assessing sleep quality in COPD patients. A hierarchical Human Activity Recognition (HAR) model was developed to classify static and dynamic activities, including fine-grained detection of lying down positions with over 95% accuracy. These classifications were then integrated into a sleep episode classification pipeline using supervised, unsupervised, and semi-supervised learning methods. This allowed the sleep metrics—including duration, efficiency, latency, WASO, and positional changes—to be automatically derived and analysed across multiple nights. The study reveals significant inter-individual variability and finds that sleeping in the supine (back) position correlates with an increased frequency of nocturnal coughing. Furthermore, statistically significant correlations were identified between objective sleep metrics (e.g., sleep efficiency, latency, positional changes) and CAT scores, supporting the clinical relevance of sensor-derived sleep monitoring. These results demonstrate the feasibility and potential of wearable sensing for objective sleep assessment in COPD, offering new possibilities for non-invasive, real-world patient monitoring and personalised intervention.

Research Ethics Approval

This project was planned in accordance with the Informatics Research Ethics policy. It did not involve any aspects that required approval from the Informatics Research Ethics committee.

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.



(*Isabel Martínez-Barona García*)

Acknowledgements

I would like to express my gratitude to my supervisor, Professor D.K. Arvind, for his guidance and support throughout my project. I am grateful for his encouragement, constant support and thoughtful feedback, which have helped me stay on track and always maintain motivation. I would also like to thank Passara Chanchotisatien and Andrew Bates for their support and invaluable insights, it has been a privilege to work with everyone.

I would also like to thank my friends and family, for their continuous support and advice through this very long year! It would not have been possible without your help.

Table of Contents

Chapter 1

Introduction

1.1 Motivation

Sleep quality is a fundamental determinant of human health, directly influencing cognitive performance, physiological regulation, and overall well-being [21]. In subpopulations with chronic respiratory conditions such as Chronic Obstructive Pulmonary Disease (COPD), sleep quality is frequently impacted [29]. COPD is a progressive and debilitating disease characterised by persistent airflow limitation and chronic inflammation of the airways, most commonly resulting from prolonged exposure to harmful airborne substances, including tobacco smoke [48]. It is recognised as a significant global health burden, affecting millions of people worldwide and becoming a leading cause of morbidity and mortality [48].

Patients with COPD often experience significant sleep disturbances, including frequent nocturnal awakenings, prolonged sleep latency, and reduced total sleep duration. These issues are frequently produced by symptoms such as chronic coughing, breathlessness, and nocturnal oxygen desaturation. As a result, sleep disruption contributes to daytime fatigue, impaired cognitive function, and diminished quality of life [29].

Traditionally, sleep quality is assessed using either subjective self-reported questionnaires or objective methods such as polysomnography (PSG). PSG remains the clinical gold standard, offering detailed measurements of brain activity, eye movements, muscle tone, heart rate, respiratory effort, and blood oxygen levels [45]. However, its utility is limited due to its cost, complexity, and requirement for specialised settings, making it unsuitable for long-term or routine monitoring. On the other hand, subjective instruments such as the Pittsburgh Sleep Quality Index (PSQI) [10] are convenient and widely used but prone to bias, variability in patient interpretation, and limited alignment with objective sleep disturbances.

Consequently, there is a need for continuous, objective and non-intrusive methods for assessing sleep quality among COPD patients. Wearable sensors have emerged as a promising solution by offering objective data collection in real-world environments without interfering with patient comfort or daily activities. This research specifically evaluates the feasibility and effectiveness of the RESpeck sensor—a lightweight, unob-

trusive wearable device—to provide accurate, continuous, and objective monitoring of sleep parameters in COPD patients.

1.2 Research Objectives

The main objectives that this dissertation aims to complete are the following:

- **Develop** and validate robust Human Activity Recognition (HAR) models capable of classifying static and dynamic physical activities using wearable sensor data.
- **Implement** methods to identify accurately and analyse sleep episodes in COPD patients, integrating supervised, unsupervised, and semi-supervised machine learning approaches.
- **Quantify** objective sleep quality metrics, such as sleep duration, efficiency, latency, awakenings (WASO), and positional changes.
- **Investigate** correlations between objective sleep measurements derived from wearable sensor data and subjective sleep quality evaluations provided through patient-reported COPD Assessment Test (CAT) scores.
- **Explore** predictive models to estimate CAT scores based on objective sleep metrics, offering insights for personalised clinical intervention.

1.3 Contributions and Limitations

The main contributions of this research include:

- A hierarchical HAR architecture, enhancing the accuracy of distinguishing lying down positions to 95%, which is essential for sleep analysis.
- The integration of HAR-derived categorical data with numerical sensor metrics, which improved sleep episode classification by reducing the amount of sleep fragmentation during the night.
- Objective quantification of clinically significant sleep metrics–sleep duration, sleep efficiency, sleep latency, WASO and number of positional changes– for COPD patients, using the RESpeck sensor.
- Identification of relationships between objective sleep data, patient-reported sleep quality (CAT scores), and cough frequency, highlighting the clinical utility of wearable sensors in respiratory disease management.

However, several limitations must be acknowledged. The absence of sufficient labelled sleep data enforced the use of unsupervised learning techniques, which, although practical, may limit the validation of the classification outcomes. In addition, potential discrepancies between the objective sensor data and subjective patient-reported experiences complicate the interpretation of correlations, as patient perceptions may not always align with their physiological signals. These limitations highlight the need for larger, labelled datasets for more refined analysis and to enhance model robustness.

1.4 Structure Overview

Before addressing the research objectives, it is essential to explain the basic concepts of sleep quality, Chronic Obstructive Pulmonary Disease (COPD), and Human Activity Recognition (HAR), as outlined in Chapter 2. Chapter 3 discusses the development and validation of three distinct HAR models, which are then applied to the COPD dataset to enable activity-based analysis. Chapter 4 explores various methodologies for sleep episode classification, including supervised, unsupervised, and semi-supervised techniques. The outcomes of these chapters form the basis for the analysis in Chapter 5, where objective sleep metrics are derived, and their correlations with patient-reported CAT scores and cough frequency are investigated. Finally, Chapter 6 summarises the key findings, addresses the study's limitations, and proposes directions for future research. The complete project is encapsulated in Figure 1.1.

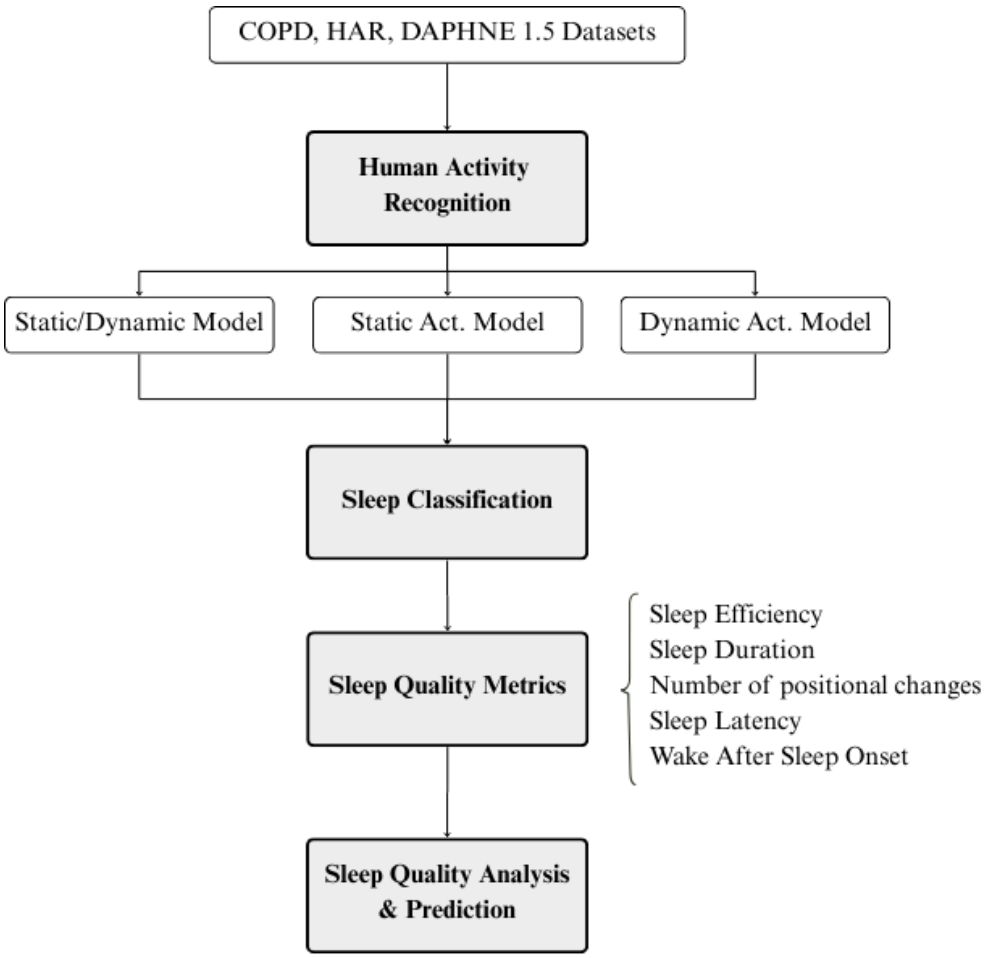


Figure 1.1: Summary of the project.

Chapter 2

Background

2.1 Sleep Quality

Sleep is a fundamental physiological process which is essential for human survival, and sleep quality is one of its critical components [21], as it influences overall health and well-being, impacting physical health, cognitive performance, and physiological functions [10]. Sleep quality is often subjectively defined, reflecting on each individual's feelings upon awakening and their daytime mental state following sleep. However, to achieve objectivity, several measurable sleep parameters have been employed to assess sleep quality, including total sleep time, sleep latency, sleep efficiency, and the number of awakenings [22][10][36]. Two widely recognised frameworks have used these metrics, the Pittsburgh Sleep Quality Index (PSQI) and the recommendations of the Sleep Consensus Panel [10][36].

2.1.1 Measurement of Sleep Quality

The gold standard for evaluating sleep quality is polysomnography, which monitors brain activity, eye movements, muscle activity, heart rate, breathing patterns and oxygen levels [45]. As a result, it provides enough information to extract the measurable sleep parameters and the percentage of time spent in each sleep stage (e.g., Rapid Eye Movement (REM) and non-REM stages). It is the gold standard for diagnosing sleep disorders by performing an overnight study in which the patient is monitored continuously by a professional [8]. However, this is an expensive and time-consuming procedure requiring a laboratory setting and technical training, which is impractical for assessing sleep quality in the general population [55].

Consequently, self-reported questionnaires have become the most common tool to assess sleep quality, as they are easy to administer, inexpensive, and applicable in epidemiological studies [16]. These questionnaires provide an assessment of the sleep quality experienced by the individual, capturing both quantitative and qualitative aspects of sleep.

2.1.2 Pittsburgh Sleep Quality Index (PSQI)

The Pittsburgh Sleep Quality Index (PSQI) consists of a self-report questionnaire designed to evaluate sleep quality over the previous month [10]. It comprises 24 questions which collectively address seven distinct dimensions, summarised in Table 2.1. Each component receives a score from 0 (no difficulty) to 3 (severe difficulty), with higher scores reflecting poorer sleep quality.

Component	Description
Subjective Sleep Quality	Self-reported sleep quality
Sleep Latency	Time taken to fall asleep
Sleep Duration	Total time spent asleep
Habitual Sleep Efficiency	Ratio of time asleep to time in bed
Sleep Disturbances	Frequency of waking during the night
Use of Sleep Medication	Frequency of sleep medication use
Daytime Dysfunction	Impact of sleep quality on daytime functioning

Table 2.1: Components of the Pittsburgh Sleep Quality Index (PSQI).

It currently serves as an accepted standard due to its extensive validation and possible everyday use in sleep research [16]. Although it consists of 24 items, only 19 items are self-reported, while the remaining five involve feedback from a bed or room partner. The 19 self-reported items include both numerical ratings (0–3) and open-ended responses, which are scored categorically based on patient responses to form the seven-component global score [16].

Revised versions of the PSQI have also been developed for specific populations or methodological improvements. For instance, a modified 16-item PSQI with a simplified 3-point Likert scale suitable for non-clinical populations has also been proposed [38]. Similarly, short-form versions have been developed (SC-PSQI), comprising nine items with a reduced scoring range (0–2), totalling up to 18 points [14].

Despite its strengths, the PSQI and other questionnaire-based sleep assessments have inherent limitations, such as potential inaccuracies due to unanswered or misinterpreted questions, inattentive or biased responses, and the need for partner responses for specific components, such as the number of positional changes during the night. Furthermore, the process of questionnaire completion itself can be burdensome for participants [4].

2.1.3 Wearable Sensors

Wearable sensors have become a promising alternative to traditional sleep assessment methods due to their ability to continuously and unobtrusively monitor sleep patterns in real-world settings. These devices can capture a wide range of physiological signals, including heart rate, respiratory rate, body movement, and skin temperature, providing valuable insights into sleep quality and sleep disorders.

Research has investigated the potential of wearable sensors for sleep monitoring, often employing advanced computational methods like deep learning. Arora et al. used commercial smartwatches and clinical actigraphy combined with deep learning to

estimate sleep quality objectively [4]. The study showed that parameters such as sleep efficiency, in-bed awake percentage and sleep variability were impactful features that could be used to predict sleep quality accurately. Similarly, Miwa et al. developed a method for detecting roll-over movements during sleep using a wearable armband device [31]. Sleep was classified into light or deep stages depending on the frequency of positional changes, and a novel Sleep Quality Score (SQS) was calculated based on the proportion of time spent in deep sleep. The study found that roll-over frequency can be a reliable, non-invasive method for evaluating sleep quality, with reduced SQS indicating poorer quality of sleep.

Wearable sensors integrating heart rate variability and accelerometry data have also shown promise for simultaneously monitoring stress and sleep quality. Majoe et al. developed a smart wearable sensor that combines heart rate variability analysis and physical activity detection to estimate stress levels and sleep quality [26]. Accelerometry was explicitly used to classify posture, detect awakenings, and evaluate sleep depth based on movement, demonstrating that it could accurately detect heartbeats and classify activities. The accuracy of postural changes and body acceleration data was also explored by Razjouyan et al., who evaluated the effectiveness of a chest-worn wearable sensor compared to a traditional wrist-worn device and gold-standard polysomnography (PSG) [41]. The study found that the chest sensor demonstrated superior accuracy and stronger correlation with PSG results, particularly in detecting wakefulness, an area where wrist-based methods underperform. The findings suggested that using positional data improves the precision of sleep quality assessments.

Despite the demonstrated potential of wearable devices for improving the accuracy of sleep assessments over self-reported questionnaires, existing studies have focused mainly on isolated sleep metrics such as total sleep time, sleep efficiency, or wake detection. While postural data and movement have been effectively used to enhance specificity and correlation with gold-standard polysomnography, an approach that employs a wearable sensor to estimate all major sleep parameters has not yet been studied.

2.1.3.1 RESpeck Sensor

The RESpeck is a compact, wireless wearable sensor developed by The University of Edinburgh Centre for Speckled Computing. Measuring 45 x 38 x 13 mm, it was designed to monitor respiratory activity and physical movement continuously [5][6]. It houses a tri-axial accelerometer and gyroscope, allowing it to detect subtle chest movements associated with breathing and general activity. It is meant to be worn on the lower rib cage and secured with tape to maintain consistent positioning. The device samples data at 12.5 or 25 Hz and transmits it in real-time via Bluetooth to a paired Android application, where information such as the timestamp, breathing rate and activity level are stored. Designed for long-term use, the RESpeck provides up to six months of battery life between wireless charges, making it an unobtrusive solution for monitoring, which was used for the data collection in this project.

2.2 Chronic Obstructive Pulmonary Disease

Chronic Obstructive Pulmonary Disease (COPD) is a primary global health concern and a leading cause of morbidity and mortality worldwide. It is a preventable and treatable disease characterised by persistent respiratory symptoms and airflow limitation that is usually progressive and associated with an abnormal chronic inflammatory response in the airways and lungs to harmful particles or gases, most notably tobacco smoke and air pollutants [48]. The airflow limitation in COPD arises from a combination of small airway disease (obstructive bronchiolitis) and parenchymal destruction (emphysema), which leads to symptoms such as chronic cough, sputum production, and breathlessness [48]. Risk factors for developing COPD include long-term exposure to tobacco smoke, indoor air pollution from biomass fuels, outdoor air pollution, respiratory infections, and occupational hazards, with specific populations—such as women, the elderly, and individuals with lower socioeconomic status—being more vulnerable [48].

As a progressive condition, COPD can result in frequent exacerbations and comorbidities that worsen the overall prognosis. Diagnosis is typically confirmed via spirometry, which assesses the degree of airflow obstruction and helps stage disease severity [48]. Management requires a combination of pharmacological and non-pharmacological approaches, including bronchodilators, corticosteroids, smoking cessation, pulmonary rehabilitation, and supplemental oxygen therapy. Wearable sensors to manage physical activity for COPD patients have been researched, showing that they can provide valuable insights into disease progression and showing which activities help [40][46]. This project investigates the potential of wearable sensor systems to measure sleep quality in COPD patients, which has been shown to be poor and has not been studied yet.

2.2.1 Sleep-Quality in COPD

Patients with COPD frequently experience disrupted sleep, reflected by diminished sleep efficiency, prolonged sleep latency, increased awakenings (WASO), and shorter total sleep durations. These disruptions are primarily driven by nocturnal symptoms such as persistent coughing, wheezing, shortness of breath, and oxygen desaturation episodes [29]. Poor sleep quality adversely affects disease outcomes, exacerbating respiratory symptoms, increasing the frequency and severity of exacerbations, and negatively influencing the quality of life and overall health status [29]. Managing sleep effectively in COPD is critical for improving symptom control and functional capacity.

Studies using wearable sensors have been conducted to monitor heart rate variability, detect exacerbation, or measure physical activity. However, the use of these sensors to monitor sleep quality in COPD patients is still to be studied.

2.3 Human Activity Recognition

Human Activity Recognition (HAR) involves identifying and classifying physical activities performed by individuals through sensor-based or vision-based systems [19]. HAR has significant implications for healthcare, fitness tracking, eldercare, and chronic disease management, providing essential insights into individuals' daily behaviours

and functional status [3]. Typical sensors employed in HAR include accelerometers, gyroscopes, magnetometers, GPS sensors, and physiological monitors such as heart rate and respiratory sensors [23].

2.3.1 State-of-the-Art

State-of-the-art HAR approaches utilise sophisticated machine learning and deep learning methods, such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs) including Long Short-Term Memory (LSTM) units, or hybrid CNN-LSTM architectures that exploit both spatial and temporal features of human movement data. LSTM-based architectures are compelling in modelling time-series data and capturing long-term temporal dependencies, which are essential for recognising complex daily activities. Studies have demonstrated that stacked LSTM networks and CNN-LSTM hybrids can achieve high accuracies in activity classification tasks. However, they tend to be computationally intensive and less suitable for deployment on energy-limited embedded systems [30][56][37].

To address the resource constraints typical of wearable HAR devices, several works have explored lightweight, efficient 1D CNN architectures. For instance, Reusch et al. evaluated and optimised a simple 1D-CNN model for HAR, showing that by extending the input to include multiple temporal steps, the model's accuracy improved from 74% to 88%, while keeping its computational footprint minimal [42]. Similarly, Chiang's work proposed a customisable 1D-CNN architecture tailored for time-series sensor input, achieving over 92% accuracy on the UCI-HAR dataset, and demonstrating superior generalisation and less overfitting compared to traditional machine learning methods like Random Forest and SVM [13]. Another study by Nguyen et al. built a high-performance 1D-CNN model capable of recognising complex, nuanced activities—such as mobile phone scrolling or gaming—from tri-axial accelerometer data, achieving an accuracy of 98.28% and proposing both software and hardware (FPGA) implementations for real-time HAR in wearable settings [34].

These recent developments highlight the promise of compact, low-power 1D-CNNs for real-time, embedded HAR applications. They demonstrate that with careful architectural tuning—such as the use of optimised kernel sizes, dropout, and extended temporal windows—deep learning models can balance performance with deployability, making them suitable for continuous monitoring in health scenarios.

2.4 Summary

This chapter provided an overview of sleep quality, COPD, and human activity recognition, highlighting the importance of sleep quality monitoring in COPD patients and the potential of wearable sensors for this purpose. The subsequent chapters will delve into the methodology and results of the project, focusing on the development of HAR and sleep classification models that can be used to monitor sleep quality in COPD patients.

Chapter 3

Human Activity Recognition

This chapter introduces the datasets and methodology for developing human activity recognition models. Accurately identifying and classifying human activities is essential in finding activity patterns for subsequent sleep analysis for COPD patients.

3.1 Datasets

Three datasets were used to develop and test the HAR models. Each dataset contains specific groups of subjects with recorded sensor data from the RESpeck devices. Table 3.1 summarises the activities included in each dataset.

Category	Activities Included
Static	Sitting/Standing, Lying (left side, right side, stomach, back)
Dynamic	Running, Walking, Shuffle Walking, Miscellaneous Movements, Descending Stairs, Ascending Stairs

Table 3.1: Summary of Activities in HAR Datasets

3.1.1 DAPHNE 1.5 Dataset

This dataset is comprised of 31 asthmatic adolescent subjects, who were recruited for the Delhi Air Pollution and Health Effects (DAPHNE) study [32]. The data collection protocol consisted of wearing the RESpeck sensors (configured at a sampling rate of 25 Hz) while reproducing the following process for each of the three hospital visits:

1. Collect continuous unlabelled data at home for a minimum of three weeks before each hospital visit.
2. At hospital visits, record labelled activity data for each of the 11 activities (30 seconds per activity).
3. Post-visit, collect two additional days of unlabelled data.

For this study, only the labelled data collected during hospital visits was used for model training and testing.

3.1.2 HAR Dataset

The HAR (Human Activity Recognition) labelled dataset includes activity data from 127 healthy adult subjects. It follows the same labelled data collection protocol as DAPHNE 1.5 (Step 2), where every subject recorded each activity for 30 consecutive seconds. This facilitates the combined use of both datasets for model training.

3.1.3 NHS Borders Dataset

This dataset is formed from 21 Chronic Obstructive Pulmonary Disease (COPD) patients who participated in pulmonary rehabilitation exercises between 2021 and 2022. The subjects wore the RESpeck sensors for an average of 42 days and produced three main folders:

- RESpeck: consists of the timestamped acceleration data, as well as the computed activity level and breathing rate values.
- Rehab: sensor data recorded during and after exercises.
- Diary: daily CAT scores, which reflect the patients' self-reported well-being.

The main focus of this chapter is the RESpeck folder, where developed HAR models will be applied in order to analyse the activity patterns of COPD patients. The following chapters will focus on the Diary folder for the sleep analysis.

3.2 Machine Learning Pipeline

Initially, an existing Convolutional Neural Network (CNN) model, developed by Yuxuan Wang [53], was tested, where classification inaccuracies were identified, particularly amongst static activities (e.g., the different lying positions). Accurate detection of static activities is essential for sleep analysis; as a result, a hierarchical classification approach was implemented, illustrated in Figure 3.1.

1. Data Acquisition and Preprocessing

The sensor data collected from the RESpeck sensors were stored as CSV files, each corresponding to a specific activity type. The preprocessing process involved:

- Loading all activity files from the dataset directory, only including the acceleration data.
- Segmenting the raw time-series data into sliding windows of 50 samples (equivalent to 2 seconds), using a stride of 50 (no overlap).
- Assigning each window a label corresponding to its activity class.

The CSV files were then randomly split into training and testing sets for each activity class using an 80/20 ratio. Then, each window of raw sensor data was used directly as input to the deep learning models. No manual feature engineering was required, as CNNs can automatically learn spatial-temporal patterns.

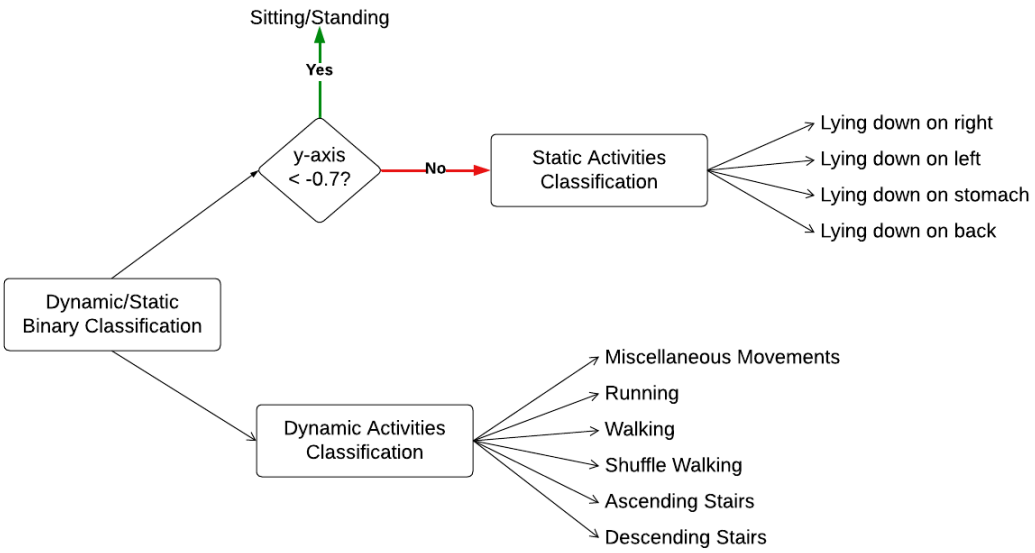


Figure 3.1: Summary of the Human Activity Recognition models.

2. Model Training

After preprocessing the data, three different models were trained, as shown in Figure 3.1:

1. Static/Dynamic Model: distinguish between static and dynamic activities.
2. Static Activities Model: classifies specific static activities.
3. Dynamic Activities Model: classifies specific dynamic activities.

3. Model Validation and Evaluation

To ensure generalisability, each model was evaluated using 5-fold cross-validation. This consists of splitting the entire dataset into five subsets or folds, where the model is then trained and evaluated five times. Four folds are used for training, and the remaining is used for validation for each iteration. Each time, a different fold is chosen for validation, and after the five runs, the results are averaged to provide a more reliable estimate.

To evaluate the test set performance, classification reports were produced showing precision, recall and F1-scores, and confusion matrices to visualise the prediction performance.

3.3 Static/Dynamic Classification

CNNs have been proven to work best in sequential data. Nonetheless, other machine-learning techniques were evaluated since classifying between static and dynamic is a binary task. The three approaches assessed are the following:

Method 1 – Random Forest

Random Forests [39] consists of an ensemble learning technique, which is popular for classification and regression tasks. During training, multiple decision trees are created and output the mode of the classes for the classification problem. It is a method which controls overfitting due to the averaging of the various decision trees. Nonetheless, it ignores temporal dependencies, which can affect the results of this human activity recognition task.

Method 2 – Support Vector Machines

Support Vector Machines (SVMs) [27] are supervised learning models that are known for performing effectively in binary classification tasks. This method finds a hyperplane in an N-dimensional space which best separates the points into their categories, where N represents the number of features. As a result, its objective is to maximise the distance between the hyperplane and the nearest point.

Method 3 - Convolutional Neural Networks

1D Convolutional Neural Networks were the main focus of the task due to their capability to process sequential data to find underlying patterns. Several CNN architectures were experimented with, with the optimum configuration consisting of two convolutional layers with 64 and 128 filters, respectively. Each convolutional layer has ReLU activation and L2 regularisation to avoid overfitting, followed by batch normalisation and a max pooling layer. The extracted features are then passed through a fully connected dense layer with 128 neurons and a softmax output layer, which is used to generate the class probabilities.

Results

Despite the efficacy of RFs and SVMs, CNNs demonstrated superior performance as they are better known for capturing spatial hierarchies in data, achieving a 90% validation accuracy across all folds. The results from the test set are summarised in Table 3.3.

Model	Accuracy	Precision	Recall	F1-Score
RF	93.99%	0.94	0.94	0.94
SVM	95.30%	0.95	0.96	0.95
CNN	97%	0.9723	0.9729	0.9716

Table 3.2: Test set accuracy of the machine learning models.

The classifications obtained from applying this model to the unseen DAPHNE 1.5 labelled dataset are illustrated in Figure 3.2. Further analysis of these misclassifications revealed that the majority stemmed from the *Miscellaneous Movements* category, which presents a challenge due to its variability and lack of well-defined patterns. While simpler models like RFs and SVMs offer certain advantages, especially concerning computational intensity, the deep learning approach provided more accurate results.

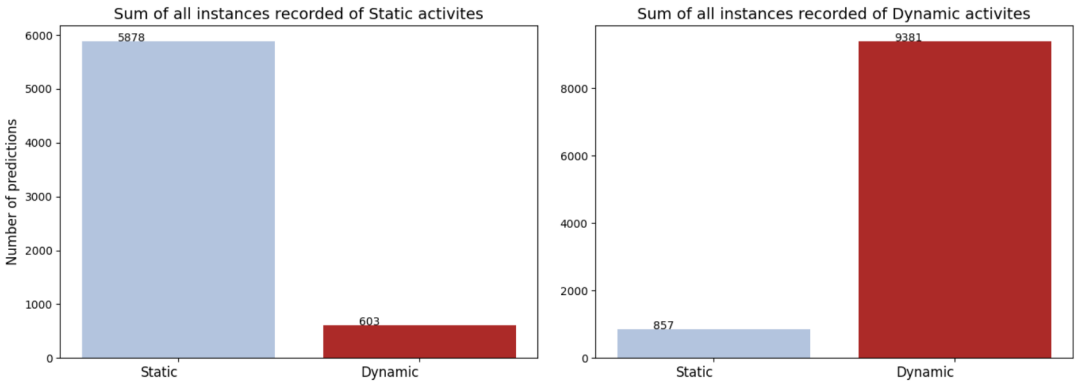


Figure 3.2: Static/Dynamic classification on unseen DAPHNE 1.5 data.

3.4 Static Activity Classification

Studies have demonstrated the effectiveness of CNNs in classifying sequential data [44], and the previous results motivated its use in classifying individual static activities. As the number of activities increases, the importance of modelling temporal relationships increases. A threshold-based criterion was applied to the accelerometer readings along the y-axis to enhance classification further. Typically, lying down positions exhibit y-axis accelerations around -1.0g; therefore, accelerations greater than -0.7g were classified as sitting/standing positions. This aligns with findings indicating that certain accelerometer values are associated with specific postures [49], which allows the differentiation between lying down and sitting/standing positions. However, a threshold was not applied to all lying-down activities due to the sensor's sensitivity to placement and orientation. Different placements could alter the readings for the same stance, making threshold use for all positions impractical [17].

The 1D CNN model developed also features two convolutional layers with 64 and 128 filters (kernel size 3) using ReLU activation and L2 regularisation. Each convolution is then followed by batch normalisation and max pooling. The model then flattens the output and passes it through a dense layer with 128 neurons, ReLU activation, L2 regularisation, and a 50% dropout before finishing with a softmax layer for multi-class prediction.

Results

Applying the model to the unseen labelled data from DAPHNE 1.5 showed the results illustrated in Figure 3.3, with the confusion matrix available in Figure A.1. Classification accuracy is shown to be high for lying-down postures (back, left, right, stomach), while minor misclassifications occur between sitting and lying-down activities, likely due to similar sensor readings for these positions. Overall, an accuracy of 95% was achieved, demonstrating the model's effectiveness.

By integrating this threshold-based criterion with the CNN architecture, the system effectively minimised misclassifications, particularly between lying and upright postures, therefore enhancing the overall reliability of activity recognition.

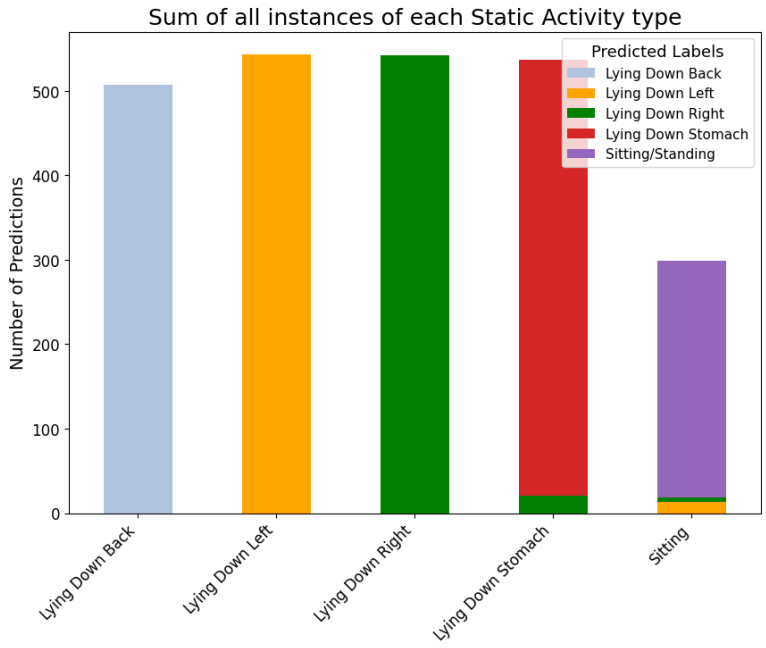


Figure 3.3: Classification counts for each static activity in the unseen DAPHNE 1.5 dataset.

3.5 Dynamic Activity Classification

For the same reason as above, CNNs were applied in the challenge of classifying dynamic activities. Unlike static activities, these are made of a broader range of motions, making the classification more challenging, as a result, two different techniques were evaluated:

Method 1 – Custom 1D-CNN

A CNN was trained from scratch using both, the DAPHNE 1.5 and HAR labelled datasets. This architecture is made of three 1D convolutional layers, with filter sizes of 64, 128 and 256, and kernel sizes of 5 and 3, each followed by batch normalisation to stabilise training, max pooling to reduce dimensionality, and dropout to mitigate overfitting. To capture temporal dynamics, a Long Short-Term Memory (LSTM) layer with 64 units is introduced, and its output is compressed using a GlobalMaxPooling1D layer. This is followed by two fully connected dense layers with 128 and 64 neurons, respectively, and a final softmax output layer to produce the class probabilities.

Method 2 – CNN with Transfer Learning

The same CNN from Method 1 was pre-trained only the HAR dataset, which contains data from healthy subjects. Subsequently, the model's weights were fine-tuned using the DAPHNE 1.5 labelled dataset to adjust to the specific characteristics of the target data, as healthy and asthmatic subjects may present physiological differences. The efficacy of this approach aligns with the existing literature, which highlights transfer learning's capability to improve learning from one domain by transferring information

from a related one [54].

For this architecture, the original output layer of the pre-trained model was removed, and a deeper, task-specific classification head was appended. This consisted of three fully connected dense layers, each followed by batch normalisation and dropout, with a final softmax layer for classification. Initially, all layers of the base-model are frozen while the classification head is trained. In the next phase, a fine-tuning process is applied, in which the layers are incrementally unfrozen. This allows the model to adapt to the new data while preserving the previously learnt features.

Results

The model without transfer learning performed accurately with the HAR dataset, achieving 92.47% mean validation accuracy across all folds and 81.75% accuracy on the test set. However, when tested on the DAPHNE 1.5 unseen data, the accuracy dropped to 50.70% (confusion matrix available in Figure A.2). This performance can be attributed to the use of data from both healthy and asthmatic subjects, which may exhibit different activity patterns due to physiological differences, meaning that the COPD dataset results could also be affected.

When using transfer learning, the accuracy increased, as shown in Table 3.3. The baseline model tended to over-predict shuffle-walking activities, leading to confusion with ascending and descending stairs, walking and miscellaneous movements. On the other hand, the refined model showed improvements in correctly classifying most classes, especially when it comes to shuffle-walking instances, as seen in the confusion matrix available in Figure A.3. These results show the potential of transfer learning to address the challenges related to activity variability in HAR tasks, giving a robust model to apply to the COPD dataset.

Model	Accuracy	Precision	Recall	F1-Score
Base-line CNN	50.70%	0.68	0.49	0.49
CNN with Transfer Learning	71%	0.74	0.61	0.62

Table 3.3: Test set accuracy of a CNN model with and without transfer learning.

3.6 COPD Dataset Results

The three models were then applied to the COPD dataset: one model to distinguish between static and dynamic, and the remaining two to classify the specific activities. This enables the analysis of the physical activity patterns of COPD patients, which is subsequently used for sleep classification.

In Figure 3.4, the activity distribution across all subjects is displayed, showing a predominance of sedentary behaviour, with sitting/standing representing 28% of all recorded activity. On the other hand, dynamic activities are scarce, where walking and shuffle-walking are low, and more demanding activities such as running, ascending, and descending stairs are almost absent ($< 0.3\%$). This aligns with the expected behaviour

of COPD patients, who can experience limited mobility and prolonged rest periods due to fatigue and breathlessness.

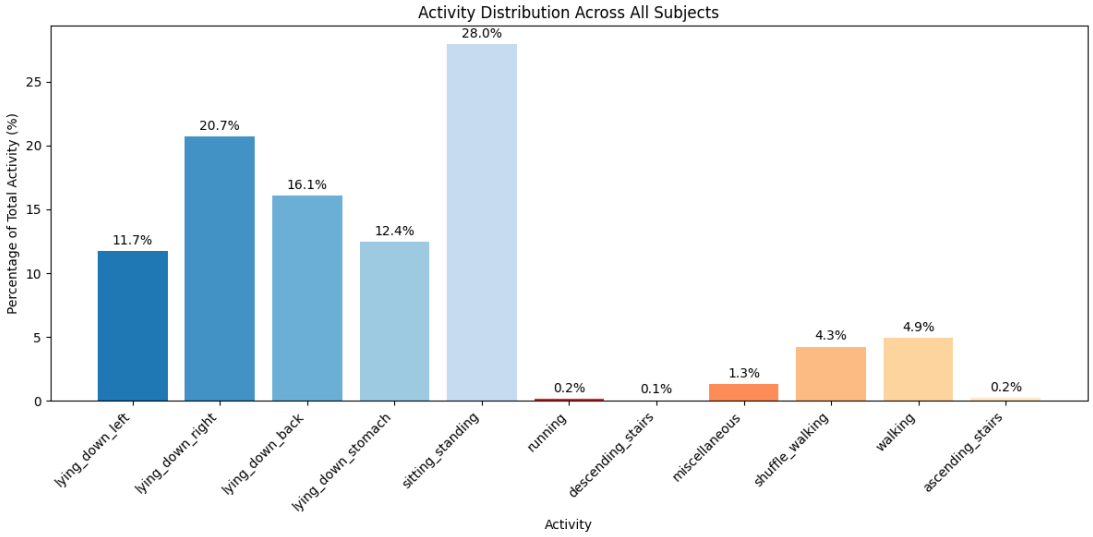


Figure 3.4: Activity distribution across all COPD patients.

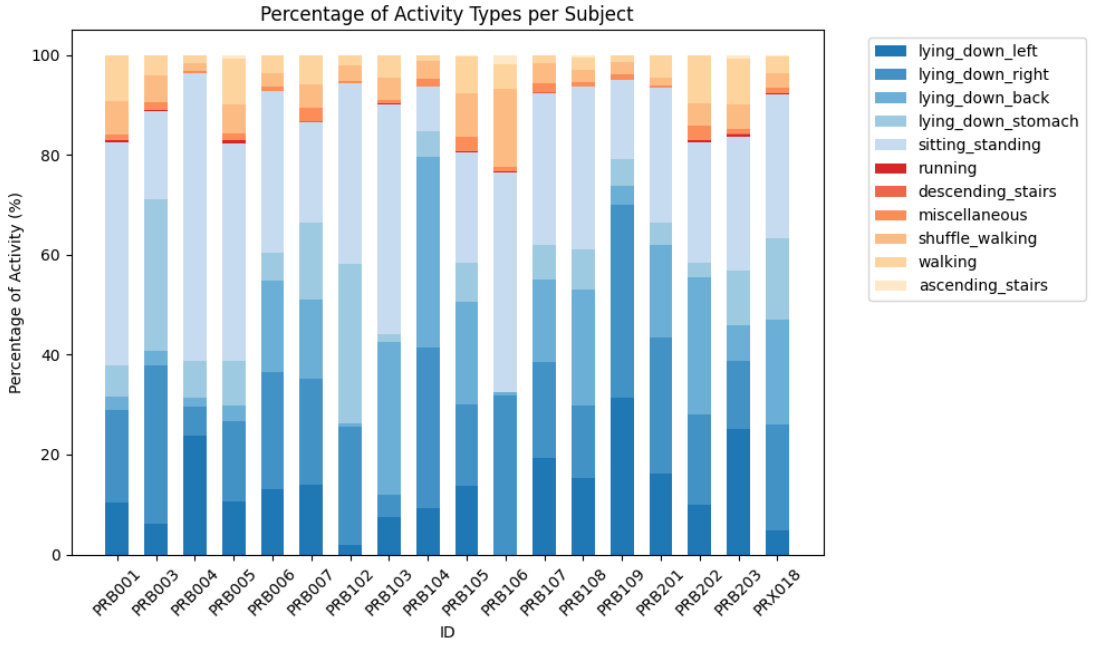


Figure 3.5: Activity distribution for each subject in the COPD dataset.

Figure 3.5 expands the perspective by showing the normalised activity distribution for each subject. As expected, most subjects dominate lying-down activities (blue shades), along with sitting/standing activities, with dynamic activities (red shades) being comparatively smaller. Subject PRB006 was chosen for further examination due to the large amount of night data analysed in subsequent chapters. This is shown in Figure 3.6, where a detailed visualisation of the patient's activity patterns over 24-hour periods is seen. Blue areas, representing static or lying positions, are most dominant,

particularly overnight, as expected, but also significantly throughout daytime hours, indicating potential daytime sleeping. Red areas, which indicate dynamic activities like walking, appear sporadically and in much smaller clusters. This pattern shows that while the patient does engage in some dynamic activities, these occur relatively infrequently and in short bursts.

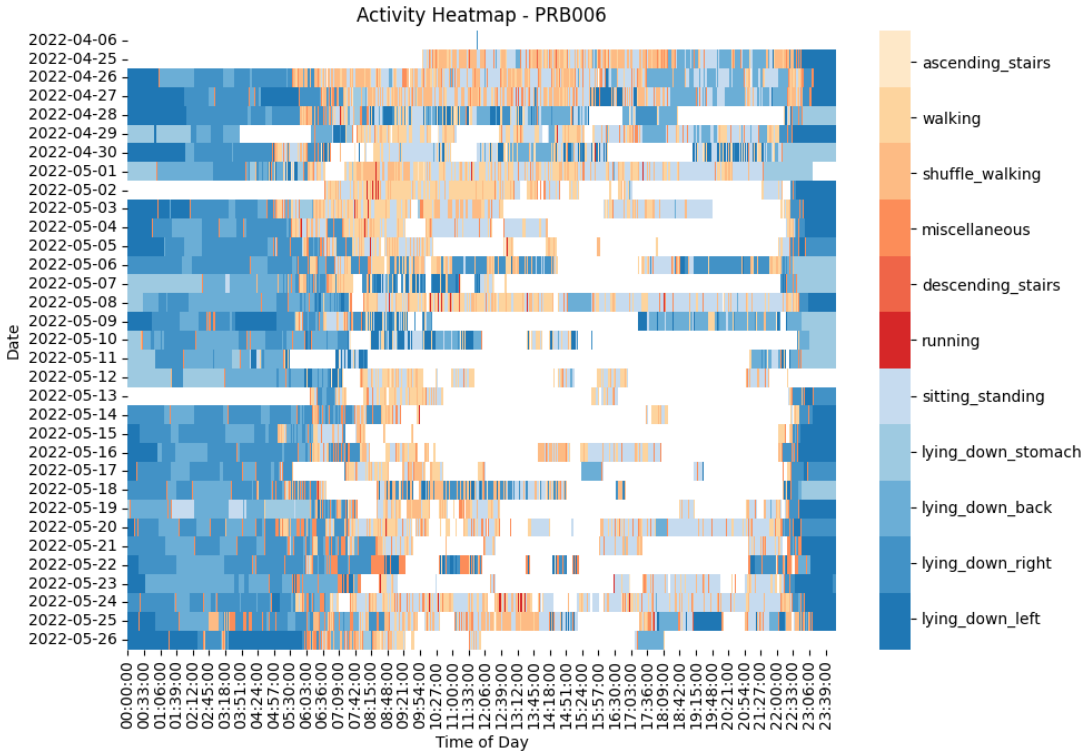


Figure 3.6: Activity patterns for subject PRB006.

These results illustrate the behaviour of a COPD patient with limited mobility, shown by the dominance of lying down or sitting. Walking, when it does occur, is the main form of dynamic activity, though still in low amounts compared to sedentary activities. The heatmap emphasises this by showing brief movement intervals between more extended periods of rest.

Chapter 4

Sleeping Episodes Classification

To perform a sleep quality analysis, it is essential to classify sleeping periods. A variety of machine learning approaches were employed to examine these sleep patterns, exploring supervised, unsupervised, and semi-supervised techniques. Each method was assessed for its effectiveness, suitability for the available data, and integration with human activity recognition (HAR) results.

4.1 Data Preprocessing

In order to perform sleep classification and analysis, the COPD dataset was resampled into 30-second resolutions which aligns with the value used in sleep studies using polysomnography– the clinical gold standard for sleep assessment. This enabled the extraction of the features listed in Table 4.1, which match the features found in the labelled dataset for sleep classification, used in the following section.

Feature	Explanation
RRV3ANN	An algorithm obtained from a previous study, which computes the respiratory rate variability after applying a 3-point moving average to smooth short-term fluctuations [7].
epochSDBR	The standard deviation of the breathing rate. Represents how much the breathing rate varies within a given time epoch (30 seconds).
BRSDCoeff	The breathing rate standard deviation coefficient. A normalisation of the standard deviation to quantify variability in a dimensionless form.
epochSDAL	The standard deviation of the activity level. Reflects the variation of the activity level within a 30-second window.

Table 4.1: Features used in the labelled dataset.

4.2 Supervised Learning

Initially, supervised learning techniques were implemented using a dataset from a prior study where 20 subjects wore a RESpeck sensor across 10 nights [7]. The valid labelled data included 318 wake episodes of 30 seconds each and 9,640 sleep episodes. The dataset was trimmed to 318 wake and 321 sleep episodes to address the class imbalance in training. The features included in the dataset are summarised in Table 4.1.

Among the supervised methods explored, Support Vector Machines (SVMs) were particularly effective due to their robust performance in binary classifications, achieving 92.9% accuracy on the test set. As a result, the features from the labelled dataset were extracted in the DAPHNE 1.5 and COPD datasets in order to apply these models.

Nonetheless, the results were not promising when visualised on the unlabeled data. Subject *PRB006* had consistent data during night times and has therefore been used to visualise the results of all machine learning methods. As illustrated in Figure B.1, no classification was made as all-time windows were classified as being awake. This shows that supervised methods present limitations in generalising to unlabelled datasets. This can be due to the scarce amount of training data available and the dependency on those four features of the dataset, which prevents it from using others that were not used in training.

4.3 Unsupervised Learning

To address the limitations encountered in supervised learning and to explore feature-independent techniques, unsupervised learning methods were tested. These can provide greater flexibility as they do not rely explicitly on labelled datasets or predefined features.

Method 1 – K-Means Clustering

The K-Means [25] clustering technique partitions the data into K clusters. For this experiment, $k = 2$ was used to capture the states *asleep* or *awake*. The algorithm starts by randomly assigning two centroids, and at each iteration, two steps are repeated until convergence:

1. Assigns each data point to the nearest centroid based on the Euclidean distance.
2. Updates the centroid of each cluster to the mean of its assigned points.

Its computational simplicity and scalability make it a good starting choice for clustering real-world activity data. Moreover, when features (e.g., respiratory rate variability or activity intensity) are continuously distributed, K-Means often converge quickly to a meaningful partition of the observations.

Method 2 – Expectation-Maximisation Clustering

Expectation-Maximisation (EM) [15] is a probabilistic clustering algorithm that extends Gaussian Mixture Models (GMMs). To show the *asleep* and *awake* states, a

two-component GMM was used, and the following two steps were repeated until convergence:

1. **Expectation (E-step)**: Based on the current model parameter, it computes the posterior probability that each point belongs to each Gaussian component.
2. **Maximisation (M-step)**: To maximise the expected log-likelihood of the data, the parameters of the Gaussian components (mean, covariance and mixture weights) are updated.

Unlike K-Means, which uses hard assignments, EM provides a soft, probabilistic interpretation, offering a richer representation of uncertainty. However, EM is more sensitive to initialisation and may converge more slowly, mainly when cluster overlap is high or the data dimensionality is large.

Method 3 – One-Class Support Vector Machines

One-Class SVM [47] is traditionally used for novelty or outlier detection, aiming to find a decision boundary around the majority of the data points. In the context of sleep analysis, it can be adapted to recognise the *asleep* pattern as the primary class and detect *awake* states (or other anomalies) as deviations. This approach can be advantageous if we consider sleep as a relatively homogeneous state, while *awake* segments vary more widely. However, as shown in the results, One-Class SVM can underperform if the data do not naturally map well onto the single-class boundary or if the feature distribution does not separate neatly into inliers vs. outliers.

4.3.1 Results

When these algorithms were tested on the labelled dataset, we obtained the results shown in Table 4.2.

Algorithm	Features	Accuracy
K-Means	RRV3ANN	0.929
	RRV3ANN, epochSDBR, epochSDAL, BRSDCoeff	0.840
EM	RRV3ANN	0.927
	RRV3ANN, epochSDBR, epochSDAL, BRSDCoeff	0.870
One-Class SVM	RRV3ANN	0.500
	RRV3ANN, epochSDBR, epochSDAL, BRSDCoeff	0.665

Table 4.2: Unsupervised learning accuracies on the labelled dataset.

From these results, it is evident that K-Means and EM perform better when classifying sleep episodes. Therefore, both methods were subsequently tested on the unlabeled dataset to cluster nocturnal episodes. Figure 4.1 illustrates the unsupervised clustering output for subject PRB006, revealing a clear improvement compared to the purely supervised approach. In particular, significant clusters corresponding to *sleep* emerge during nighttime, but the figure also highlights considerable fragmentation, implying multiple isolated *awake* classifications.

The isolated fragmentations in these algorithms, or the singleton *awake* instances, were further analysed as a method to validate the techniques; this consists of a single instance of a 30-second wake state, with sleep states on either side. In Table 4.5, the averaged number of singleton *awake* instances across all subjects is shown, with K-Means showing fewer singleton *awake* intervals than EM. This can also be seen in PRB006, where K-Means produced 675 singleton *awake* instances, while EM produced 716. This suggests that K-Means might be more robust for this dataset's distribution of features. However, as is often the case with unsupervised methods, neither clustering approach perfectly delineates true sleep segments.

Algorithm	Singleton Awake Instances
K-Means	752 (618)
EM	754 (622)

Table 4.3: Averaged number of singleton awake instances across all subjects, with the standard deviation in brackets.

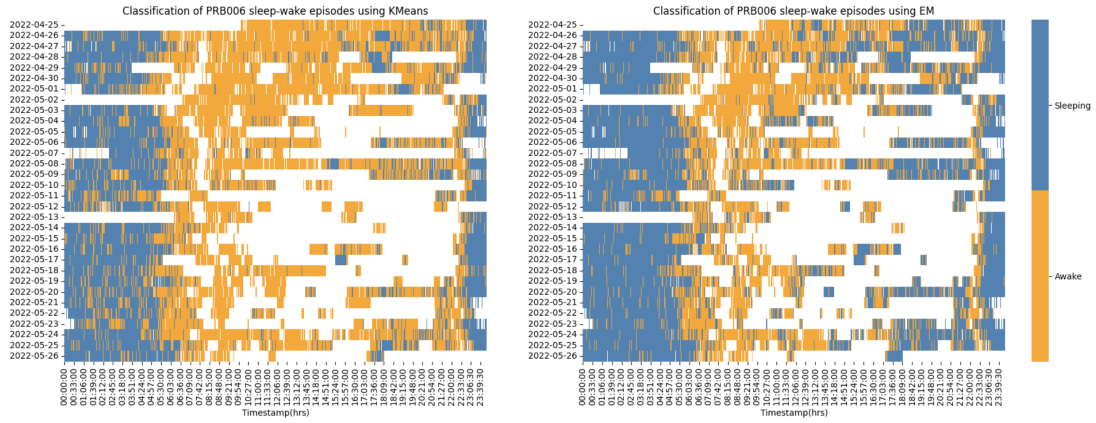


Figure 4.1: Unsupervised Clustering of Subject PRB006 Using K-Means and EM, respectively.

4.4 Semi-Supervised Learning

Given the success of supervised methods on the labelled dataset, semi-supervised methods were explored to address the limitations found when applied to the unlabelled data. These methods combine a small labelled dataset with a more extensive set of unlabelled data for better generalisation and performance.

Method 1 – Self-Training with SVMs

With this method, a classifier is trained on the labelled dataset, iteratively augmenting it with predictions from the classifier itself. Initially, an SVM is trained on the available labelled data. Then, the classifier predicts labels for the unlabelled data. The most confident predictions, over a threshold of 0.7, are added to the labelled set, and the classifier is retrained. This process is then repeated until the model converges or no new high-confidence samples are identified.

Method 2 – Gaussian Mixture Model (GMM)

GMMs assume that a mixture of multiple Gaussian distributions can model the data. Unlike hard clustering techniques, GMM assigns a probabilistic label to each data point, reflecting the uncertainty in the clustering process. For this task, two components were fitted on the scaled dataset (*asleep* and *awake*), resulting in soft probabilistic labels indicating the likelihood of each sample belonging to each cluster. These probabilistic labels were then converted into hard labels by selecting the cluster with the highest probability.

4.4.1 Results

The semi-supervised methods were tested on the labelled dataset, with the results shown in Table 4.4.

Algorithm	Features	Accuracy
SVMs	RRV3ANN, epochSDBR, epochSDAL, BRSDCoeff	0.950
GMM	RRV3ANN, epochSDBR, epochSDAL, BRSDCoeff	0.130

Table 4.4: Semi-supervised learning accuracies on the labelled dataset.

From the results, it is clear that the self-training SVM method outperforms the GMM approach. However, when applied to the unlabelled dataset, the results were not as promising. Figure 4.2 illustrates the classification output for subject PRB006, showing that the semi-supervised methods did not significantly improve the classification performance compared to the purely unsupervised approaches, obtaining over 1000 more singleton instances.

4.5 Integration with Human Activity Recognition (HAR)

It is evident from the previous results that unsupervised methods perform better than others, with the lowest fragmentation of singleton *wake* states at night. However, none of the unsupervised methods can integrate the categorical data obtained in Chapter 3 along with numerical data.

Method 1 – K-Prototypes

K-Prototypes [52] is an extension of the K-Means and K-Modes algorithms designed to handle data containing numerical and categorical features. Similar to K-Means, K-Prototypes iteratively refine clusters by minimising an overall cost function. However, it modifies both the distance measure and the centroid representation to accommodate mixed data types:

1. Numerical Features: For continuous variables, the squared Euclidean distance is calculated.

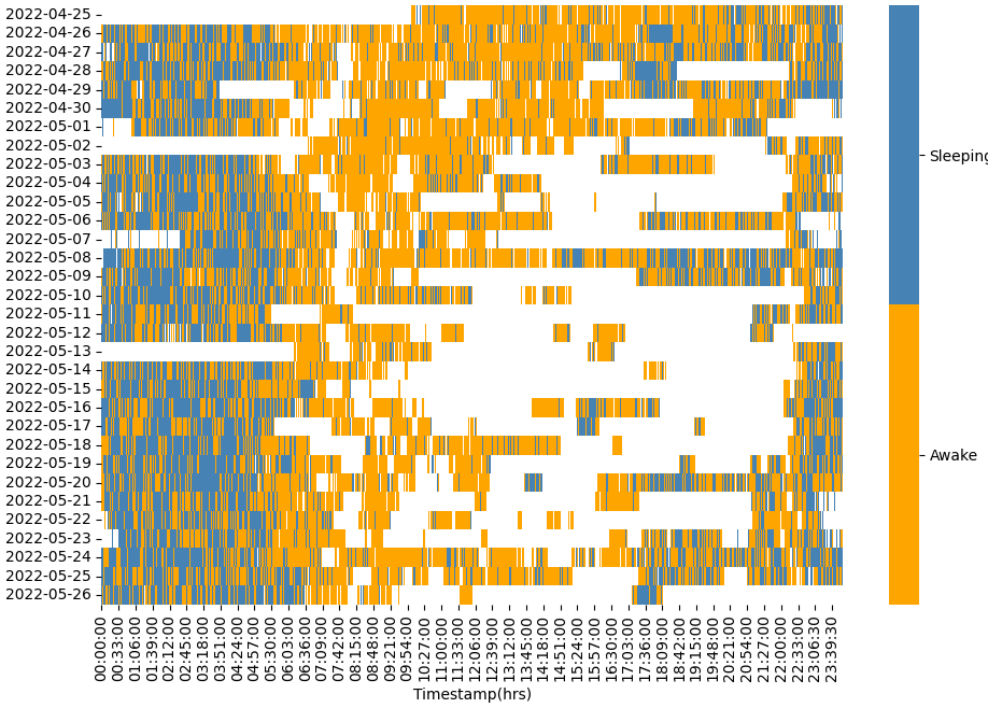


Figure 4.2: Semi-Supervised Clustering of Subject PRB006 Using Self-Training SVM.

2. Distance for Categorical Features: For categorical variables, the dissimilarity by counting mismatches is calculated: two categorical values are either the same or different, contributing a penalty for each mismatch.

Method 2 – FAMD with K-Means Clustering

This involves a combination of Factor Analysis of Mixed Data (FAMD) with K-Means clustering. FAMD is a dimensionality reduction technique that generalises Principal Component Analysis (PCA) and Multiple Correspondence Analysis (MCA) to support mixed data types. It projects both types of variables into a common low-dimensional space, preserving as much variance as possible. K-Means is then performed on the FAMD-transformed data. Since FAMD outputs continuous components, standard Euclidean distance is appropriate for K-Means to assign data points to both clusters.

4.5.1 Results

These methods could not be tested in the labelled dataset since it does not include an activity label. Therefore, the singleton awake instances were used to compare with the other methods. While K-prototype clustering did not outperform unsupervised techniques (813 average singleton awake instances across subjects), using FMAD with K-Means clustering led to substantial improvements when referring to a reduced number of fragmented classifications. This can be seen in Figure 4.3. When averaging the number of isolated waking instances across all subjects, the value was reduced to 746, indicating a more coherent segmentation of sleep cycles. This approach demonstrated superior practical relevance and interpretability compared to the purely supervised or un-

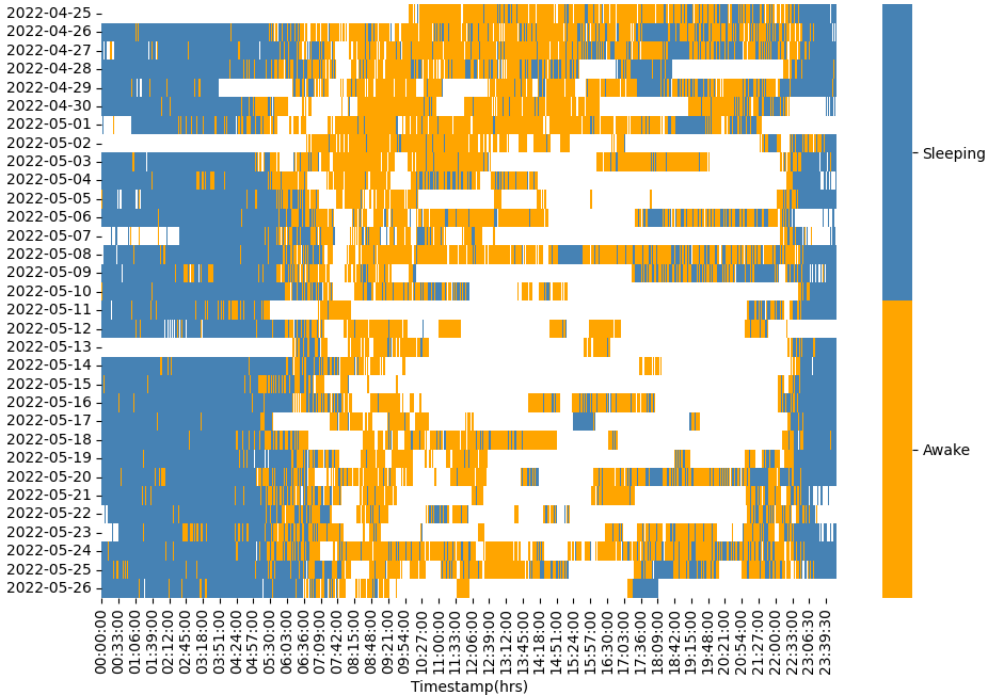


Figure 4.3: FAMD with K-Means Clustering of Subject PRB006.

supervised methods tested, by jointly considering continuous features (e.g., respiratory variability) and discrete HAR labels.

4.6 Summary

The results from the supervised, unsupervised, and semi-supervised methods were compared, with the FMAD with K-Means clustering method demonstrating the most promising results when comparing the singleton appearances, as seen in Table 4.5. Integrating HAR labels with continuous features, therefore, showed how it can help classify sleep episodes. As a result, this method was applied to the COPD dataset to do a sleep quality analysis. The results of this analysis are presented in the following chapter.

Technique	Singleton Awake Instances
Supervised	–
Unsupervised	752 (618)
Semi-Supervised	1798 (1233)
FMAD with K-means	746 (556)

Table 4.5: Averaged number of singleton awake instances across all subjects, with the standard deviation in brackets.

Chapter 5

Sleep Quality Analysis

Evaluating sleep quality involves examining multiple dimensions, including sleep duration, efficiency, awakenings, latency and position changes. Two recognised frameworks designed to measure sleep quality are the Pittsburgh Sleep Quality Index (PSQI) and the recommendations from the Sleep Consensus Panel [10][36]. Studies have explored these sleep metrics, relying on subjective reports or polysomnography, which can be resource-intensive [10][36]. However, the impact of these metrics on COPD patients has not been explored when measured by a wearable sensor.

The RESpeck datasets address these challenges by enabling the development of automated algorithms that can objectively quantify sleep metrics using sensor data. The algorithms developed target measurable metrics that have been highlighted in both the PSQI and the Sleep Consensus Panel, which are explained in detail in the following section. Automating these metrics can offer substantial benefits, including increased accuracy, scalability, and applicability in epidemiological studies or clinical practice. The COPD dataset was used to test the methodology since the subjects were requested to complete the COPD Assessment Test (CAT) daily, which includes a question about sleep quality. This allows for comparing the self-reported sleep quality and the objective sleep metrics obtained from the RESpeck sensor data.

5.1 Sleep Metrics

5.1.1 Sleep Duration

Sleep duration refers to the total amount of time slept during the night. Recommendations from the National Sleep Foundation suggest that adults should aim for 7–9 hours of sleep per night [33]. However, individuals with chronic respiratory conditions, such as COPD, may experience disrupted sleep patterns, leading to shorter or longer sleep durations [9]. This can worsen respiratory symptoms for COPD patients, impacting their overall physical health [18]. Algorithm 1 was developed to measure sleep duration. It groups nights between 21:00 and 10:00 hours, isolating the periods when the subject is asleep and lying down. It then detects continuous sleep periods and calculates the total sleep duration for each night. Nights with less than 4 hours of sleep are marked

as missing, as this is considered insufficient for adults [18]. This approach ensures an objective quantification of sleep duration, which is essential for monitoring and managing sleep disturbances in respiratory patients.

Algorithm 1 Calculate Sleep Duration Per Night

Require: DataFrame df with timestamps, sleep status, and lying down status.
Ensure: DataFrame *results* containing sleep duration per night for the given subject.

Initialise an empty DataFrame *results* with columns: `Date`, `ID`, `sleep_duration`.
for each unique night group *date* in df **do**
  if no night time data exists **then**
    Set sleep duration for this night group to NaN.
  else
    Filter records for sleeping and lying down activities.
    Calculate the time differences between consecutive records.
    Identify a new sleep period if the time difference exceeds 10 minutes.
    Assign a unique period ID to each sleep period.
    Determine the start and end times of each sleep period.
    Compute the duration of each sleep period (in hours).
    Sum the durations of all sleep periods for the current night group.
    if the total sleep duration is less than 4 hours **then**
      Set sleep duration for this night group to NaN.
    end if
  end if
  Append the current *date*, *ID*, and sleep duration to *results*.
end for
return *results*.

Results

Figure 5.1 illustrates the variation in sleep times between 21:00–10:00 across all individuals. Median values range from 6 to 10 hours, with some outliers suggesting unusually short sleep durations. This variability in each patient's sleep duration aligns with the literature, as it does not represent a regular sleep pattern manifested as exceptionally long or short sleep durations. While recommendations suggest 7–9 hours of sleep for adults, these results show that a significant portion of the cohort deviates from that range, either due to complications during sleep or compensatory behaviours that lead to prolonged rest periods. This highlights the need for personalised interventions and close monitoring of sleep disruptions in COPD management.

5.1.2 Sleep Efficiency

Sleep efficiency refers to the ratio of the time a person is asleep and the total time spent in bed, expressed as a percentage. Low sleep efficiency is typical in adults with COPD, often correlated with poor symptom management, increased respiratory discomfort and reduced overall well-being [28]. Algorithm 2 was developed to measure each subject's sleep efficiency. This algorithm calculates the total time spent in bed by identifying

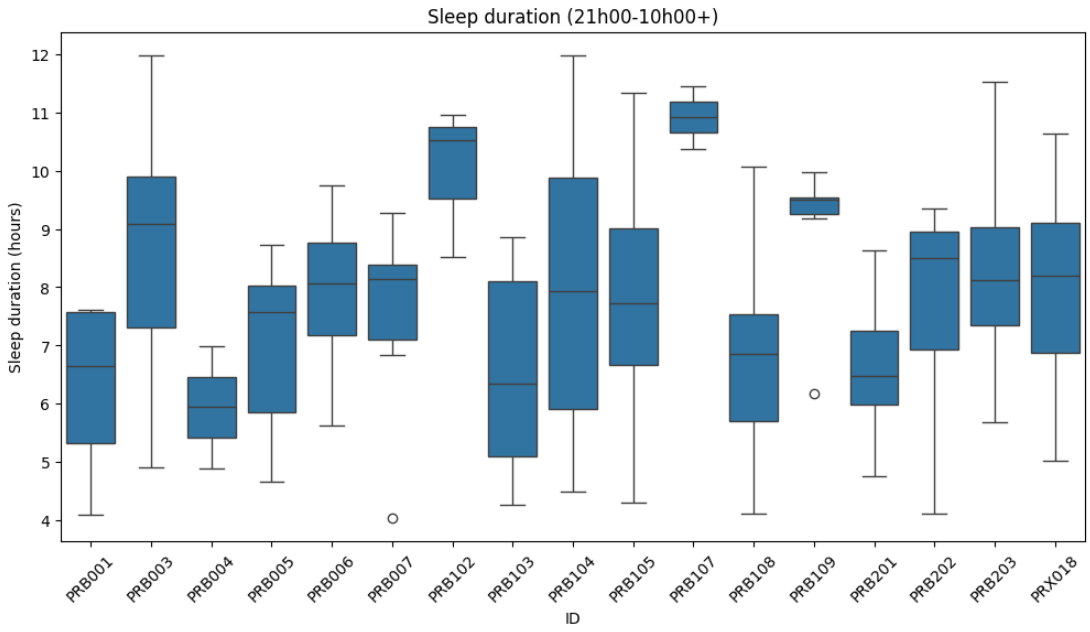


Figure 5.1: Sleep Duration per Night for All Subjects.

continuous lying-down periods at night. It then computes the total sleep duration and calculates the sleep efficiency ratio between sleep duration and total time in bed. Nights with no lying-down periods are marked as missing, as they do not provide valid sleep data. Consequently, an objective quantification of sleep efficiency can be obtained, which can be used for monitoring and managing sleep disturbances.

Results

Figure 5.2 illustrates the distribution of sleep efficiency across all subjects. It shows considerable inter-individual variation, with some subjects (e.g., PRB001) exhibiting high efficiencies, while others fall below 50% for certain nights. Healthy adults achieve sleep efficiencies of over 80% [1], and from the cohort, all patients' averages fall below this range considerably, except for subject PRB001. Several patients display narrower boxes (e.g., PRB001 or PRB109), indicating consistently low sleep efficiencies, whereas others with wider boxes (e.g., PRB006 or PRB203) suggest significant night-to-night fluctuations. Overall, these data highlight the impact of COPD on sleep, with some individuals approaching normative benchmarks of 85–95% sleep efficiency and others experiencing substantially lower values. This aligns with the literature, indicating that this respiratory condition disrupts standard sleep patterns.

5.1.3 Short Awakenings and WASO

Short awakenings, which refer to brief interruptions in sleep, can affect respiratory health. In sleep research, these events are commonly captured as Wake-After-Sleep-Onset (WASO), quantifying the total duration of these awakenings in minutes. WASO is a key metric in sleep quality assessment, as it reflects the ease of falling back asleep after waking up during the night [50]. In COPD populations, frequent WASO episodes

Algorithm 2 Calculate Sleep Efficiency

Require: DataFrame df with timestamps, night groups, sleep and activity status.
Ensure: DataFrame *results* containing sleep efficiency for each night of a subject.

for each unique night group $date$ in df **do**
  Filter the data to include only lying down activities.
  if no lying-down data exists for the current $date$ **then**
    Set sleep efficiency to NaN.
  else
    Calculate the time differences between consecutive records (in minutes).
    Identify a new lying-down period if the time difference exceeds 30 minutes.
    Assign a unique period ID to each lying-down period.
    Determine the start and end times of each lying-down period.
    Compute the duration of each lying-down period (in minutes).
    Sum these durations to obtain the total time in bed.
    Calculate sleep duration using Algorithm 1.
    if sleep duration is NaN or zero **then**
      Set sleep efficiency to NaN.
    else
      Compute sleep efficiency as:

$$\text{sleep efficiency} = \left(\frac{\text{sleep duration}}{\text{total time in bed}} \right) \times 100$$

    end if
  end if
  Append the current $date$, ID, and sleep efficiency to *results*.
end for
return *results*.

are associated with nocturnal symptoms, such as coughing, wheezing, and dyspnea, which can increase sleep fragmentation and reduce sleep quality [9]. Algorithm 3 was developed to quantify short awakenings: it identifies transitions from sleep to awake status, counting awakenings over 5 minutes since literature suggests that awakenings lasting less than 5 minutes are considered normal [36]. The algorithm then adds the total minutes of the night the subject was awake, providing the WASO metric.

Results

Figures 5.3 and 5.4 depict two complementary views of WASO episodes during each participant's nightly window. Figure 5.3 reveals the total minutes spent awake after initially falling asleep, highlighting variability within and across patients. Some individuals remain awake for only a few minutes each night, while others can exceed 100–200 minutes of wakefulness, indicating significantly fragmented sleep. Such extended periods of wakefulness are frequently associated with exacerbated respiratory symptoms, including nocturnal cough or dyspnea [9].

Meanwhile, Figure 5.4 shows the average number of WASO episodes for the same

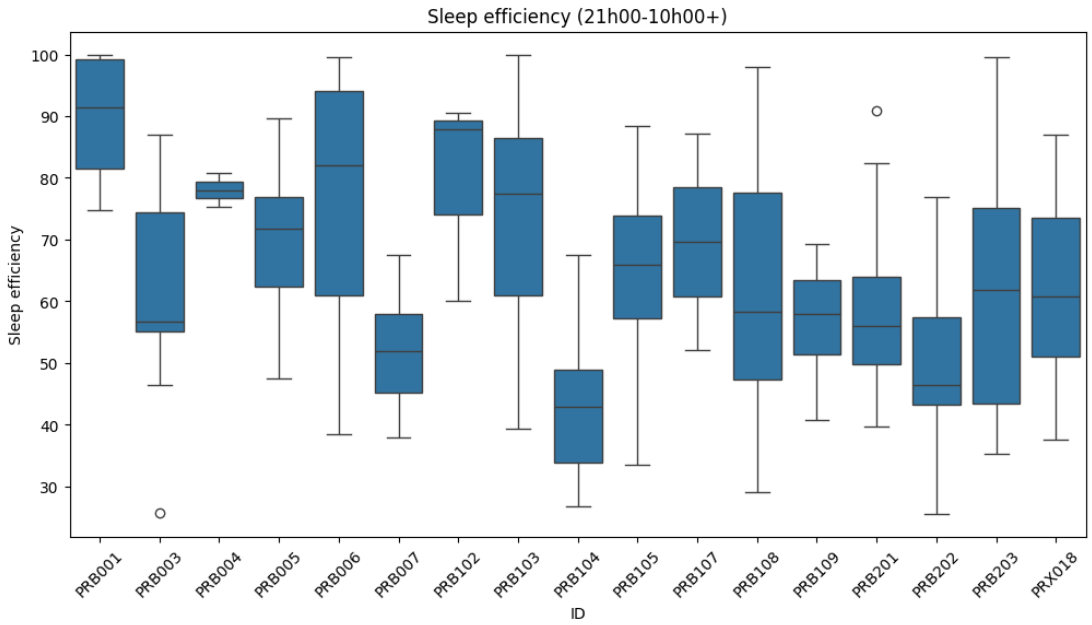


Figure 5.2: Sleep Efficiency per Night for All Subjects.

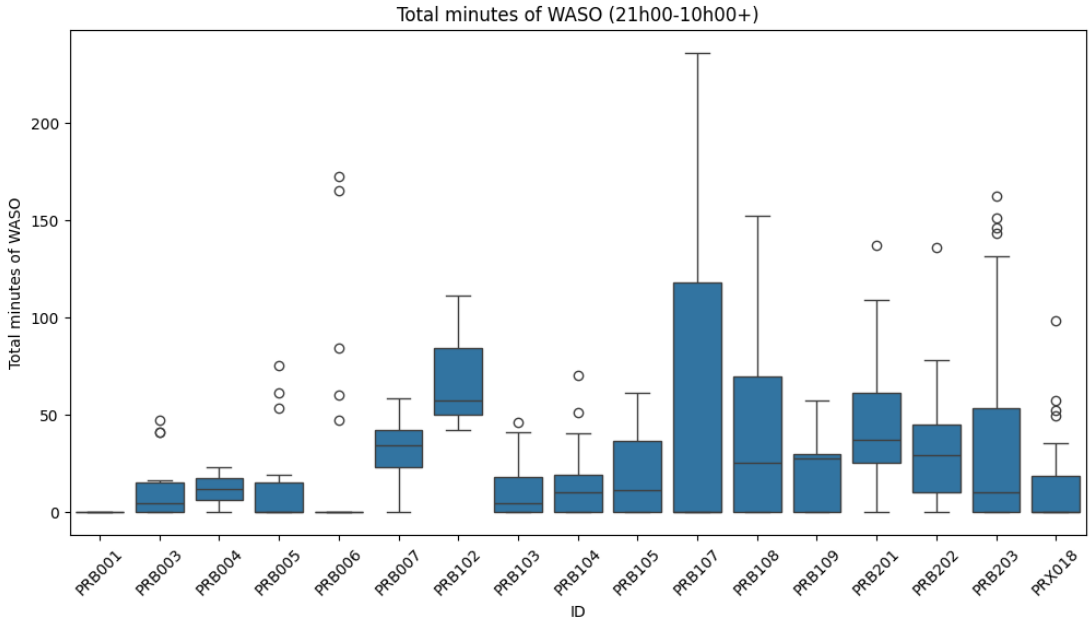


Figure 5.3: Calculated WASO in minutes for all subjects.

cohort. Although some participants exhibit relatively few brief awakenings, others experience multiple interruptions. This discrepancy underscores how shorter but more frequent awakenings can accumulate to produce varied levels of overall sleep disturbance. This aligns with prior literature suggesting that repeated nocturnal arousals contribute to daytime sleepiness and reduced quality of life in COPD populations [28].

These two metrics—minutes of wakefulness and number of awakenings—offer complementary insights into sleep fragmentation. Patients with both high total WASO duration

Algorithm 3 Short Awakenings and WASO

Require: DataFrame df with timestamps, night groups, sleep status, and activity status.

Ensure: DataFrame *results* containing the number of short awakenings and the WASO for a subject.

for each unique night group $date$ in df **do**
  Calculate sleep duration using Algorithm 1.
  Initialise short_awakenings $\leftarrow 0$ and total_minutes $\leftarrow 0$.
  if sleep duration ≥ 4 hours **then**
    Filter the data to include only lying down activities.
    if the filtered dataset is non-empty **then**
      Detect transitions between sleep and wake statuses.
      for each pair of indices $(start, end)$ corresponding to a short awakening **do**
        Calculate the duration (in minutes) between $start$ and end.
        if $5 \leq$ duration ≤ 20 **then**
          total_minutes $\leftarrow$ total_minutes + duration
          short_awakenings $\leftarrow$ short_awakenings $+ 1$
        end if
      end for
    else
      Set short_awakenings $\leftarrow$ NaN and total_minutes $\leftarrow$ NaN.
    end if
  else
    Set short_awakenings $\leftarrow$ NaN and total_minutes $\leftarrow$ NaN.
  end if
  Append $date$, ID, short_awakenings, and total_minutes to *results*.
end for
return *results*.

and frequent interruptions may require targeted interventions, such as improved symptom management or bedtime posture optimisation, to reduce the impact of disturbed sleep on respiratory health.

5.1.4 Sleep Latency

Sleep latency corresponds to how long it takes for an individual to transition from full wakefulness to the first sustained sleep episode [35]. Extended latency is a common phenomenon in COPD populations, linked to nocturnal symptoms such as dyspnea, coughing, and hypoxemia-related discomfort [9]. Longer latency reduces total sleep time and can fragment the early stages of rest, causing repeated awakenings or arousals before stable sleep is achieved. Therefore, monitoring and reducing latency is critical in managing patients' overall night quality. Algorithm 4 identifies the earliest lying-down timestamp as the bedtime and determines the first continuous sleep period. It then computes the difference between bedtime and sleep onset, providing the minute latency.

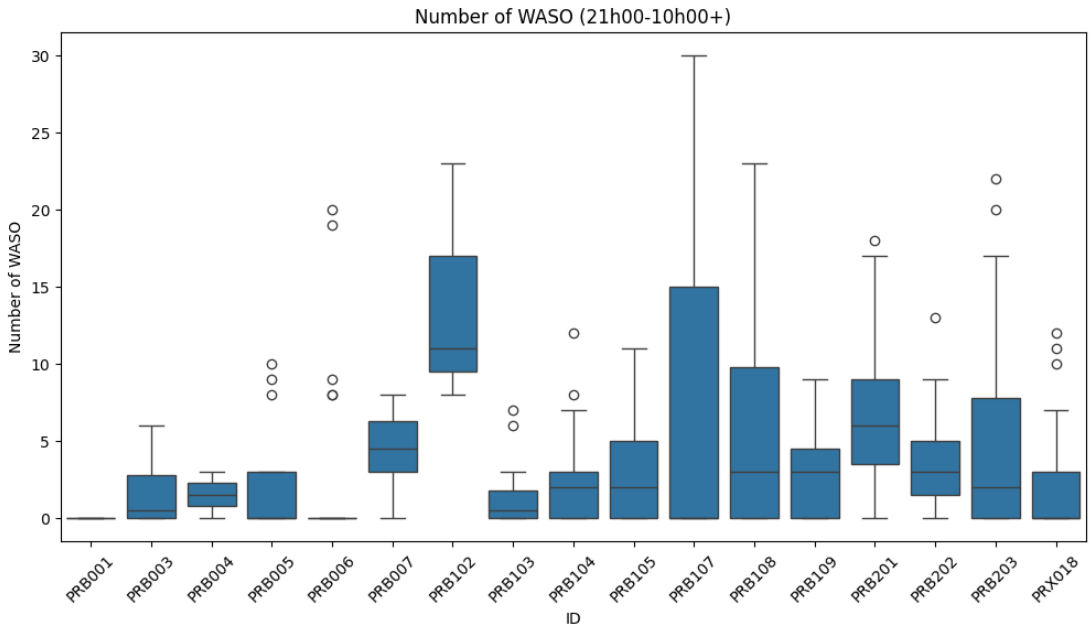


Figure 5.4: Number of short awakenings during the night for all subjects.

Results

Figure 5.5 summarises the distribution of sleep latency (in minutes) across the COPD cohort, illustrating the substantial range in the time it takes for individuals to transition from wakefulness to sustained sleep. While some participants consistently fall asleep within ten minutes, others experience extended latencies exceeding an hour. Such prolonged latency can be influenced by nocturnal respiratory symptoms, including dyspnea and frequent coughing, which are common in COPD and can lead to significant difficulties in initiating sleep [9]. Several outliers surpassing 80–100 minutes highlight potential exacerbations, further emphasising the need for individualised clinical management.

5.1.5 Positional Changes During Sleep

Frequent positional changes can indicate discomfort or respiratory difficulties, both commonly experienced by individuals with COPD [9]. Nocturnal dyspnea, coughing, and oxygen desaturation episodes can prompt patients to shift their posture in search of a more favourable breathing position [28]. While occasional changes are natural, excessive movements during the night may fragment sleep and reduce overall sleep efficiency, compounding daytime fatigue and diminishing quality of life [28]. Identifying patients with recurrent sleep-position fluctuations can guide targeted interventions—such as repositioning strategies to improve comfort. Algorithm 5 quantifies the number of position changes during sleep by identifying lying-down periods and counting transitions between different activity labels.

Algorithm 4 Calculate Sleep Latency

Require: DataFrame df with timestamps, activity labels, and sleep labels.
Ensure: DataFrame *results* containing the sleep latency for a subject.

for each night group *date* in df **do**
  Verify valid sleep duration using Algorithm 1.
  if a valid sleep duration exists **then**
    Determine bedtime as the earliest timestamp where the subject is recorded as lying down.
    Identify sleep onset as the start of the first continuous sleep period (e.g., five consecutive records indicating sleep).
    Compute sleep latency as the difference between sleep onset and bedtime.
  else
    Set sleep latency to NaN.
  end if
  Append the current *date*, *ID*, and sleep latency to *results*.
end for
return *results*.

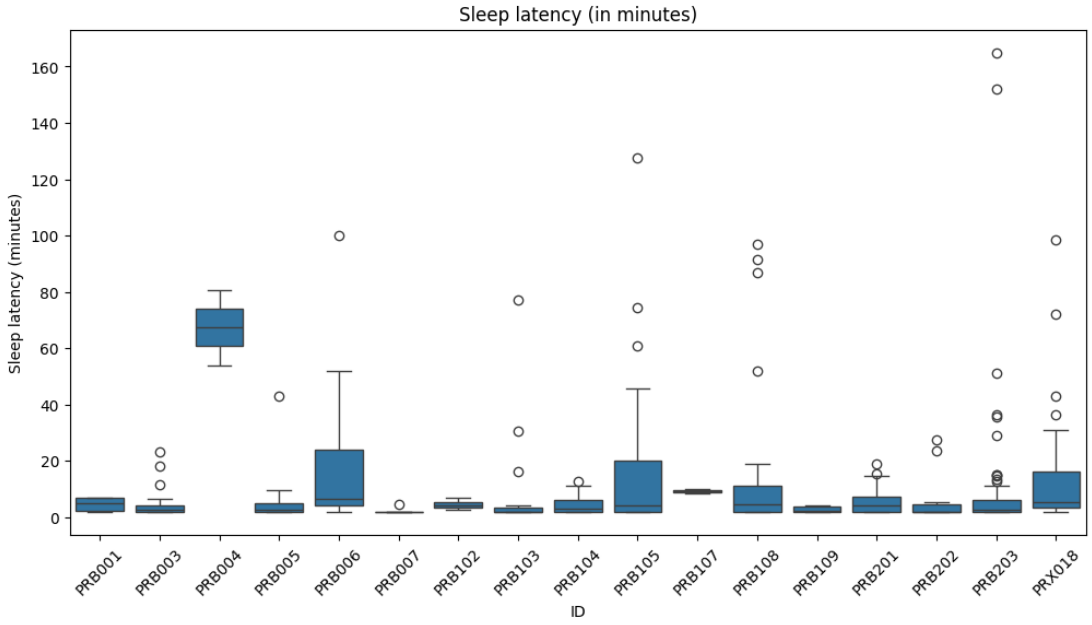


Figure 5.5: Sleep latency in minutes for all subjects.

Results

Figure 5.6 presents the distribution of nightly positional changes. Many patients show a moderate frequency of nightly position shifts, typically ranging from 10 to 30 movements per night, though specific individuals exhibit significantly higher rates—exceeding 50 shifts per night. Frequent position changes may reflect underlying sleep disruptions such as discomfort, respiratory disturbances, or difficulty maintaining stable breathing positions. Conversely, fewer position changes might indicate deeper or less disturbed sleep periods. Excessive movement and fragmented sleep positions have

been linked to impaired sleep quality and increased daytime fatigue.

Algorithm 5 Calculate Number of Positional Changes During Sleep

Require: DataFrame df with timestamps, activity labels, and sleep labels.
Ensure: DataFrame *results* containing the number of position changes for a subject.

for each night group $date$ in df **do**
  Filter records to include data between 10:00 PM and 10:00 AM.
  Verify valid sleep duration using Algorithm 1.
  if a valid sleep duration exists **then**
    Further filter the data to include only lying down activities during sleep periods.
    Count position changes as transitions between different activity labels.
  else
    Set the number of position changes to NaN.
  end if
  Append the current $date$, ID, and the number of position changes to *results*.
end for
return *results*.

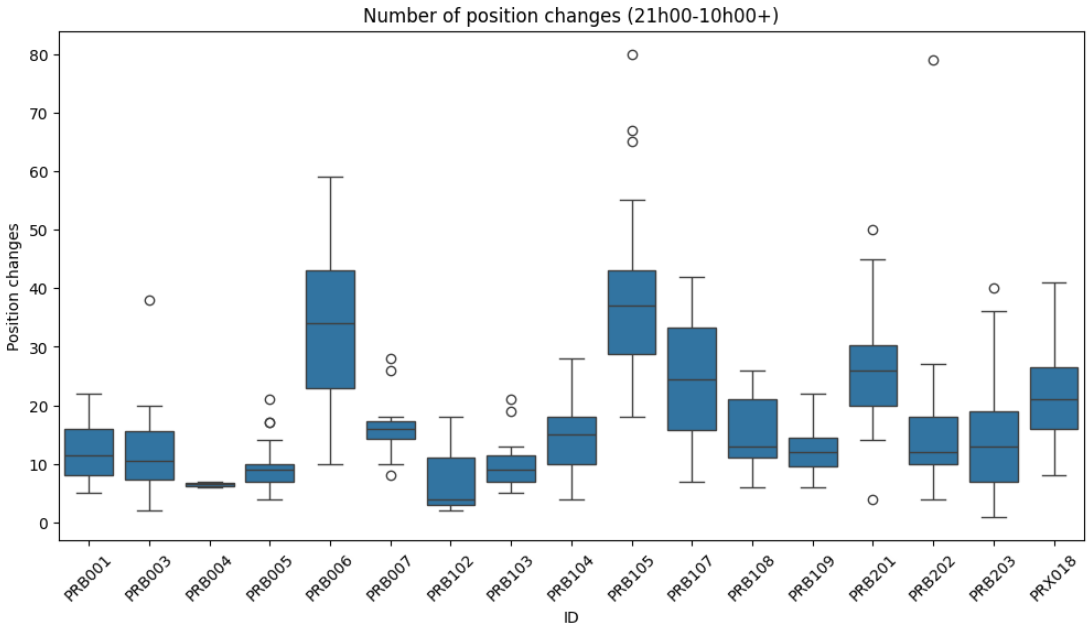


Figure 5.6: Number of position changes during the night for all subjects

5.2 Cough Frequency and Sleeping Position Analysis

A cough detection algorithm developed by Passara Chanchotsatien [11] was applied to the COPD dataset. This resulted in the hourly number of coughs for each subject, which was averaged to obtain the average number per night.

The existing literature recommends individuals with compromised respiratory function to sleep on their side [2]. This position can help with breathing and reduce the risk of

airway obstruction, which can be confirmed when analysing the sleeping positions of the cohort. In Figure 5.7, it is observed that out of the 17 subjects, 15 generally slept on their side, while only two slept in the supine and prone positions. This percentage is taken from the number of hours in each position while sleeping, which is also illustrated in Figure C.1. The dominance of sleeping on the side in COPD patients in the dataset corroborates the advice that they tolerate this sleeping position better.

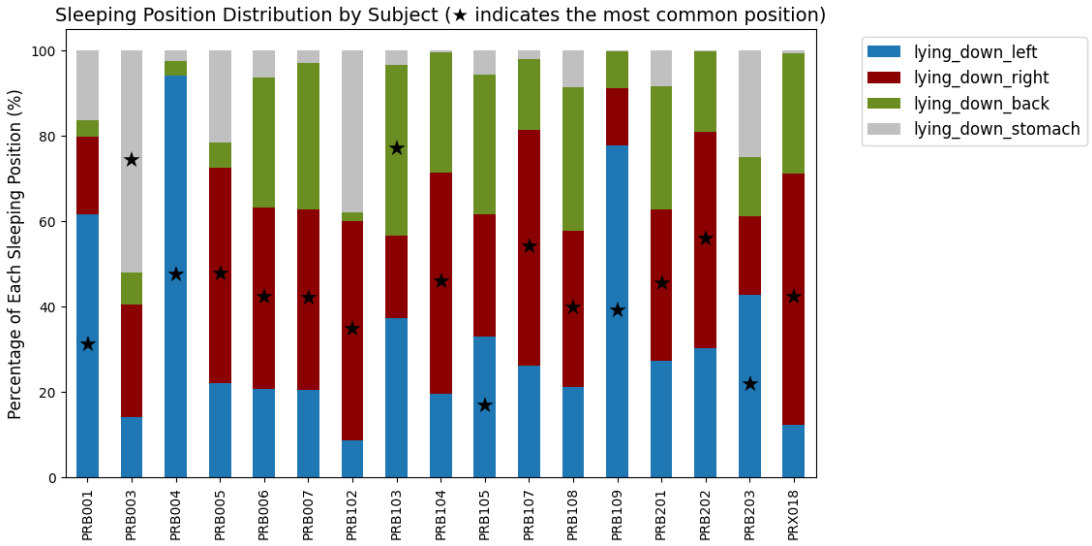


Figure 5.7: Sleeping position distribution across all subjects, where the ⋆ indicates the most common position.

The first graph in Figure 5.8 shows the average number of coughs per night for each sleeping position. The results suggest that the sleeping position can indeed influence the frequency of nocturnal coughing episodes. Among the four primary postures—back, left side, right side, and stomach—the highest average cough count per night is observed when sleeping on the stomach, followed by sleeping on the back. Side sleeping, particularly on the left side, is associated with slightly fewer coughs on average.

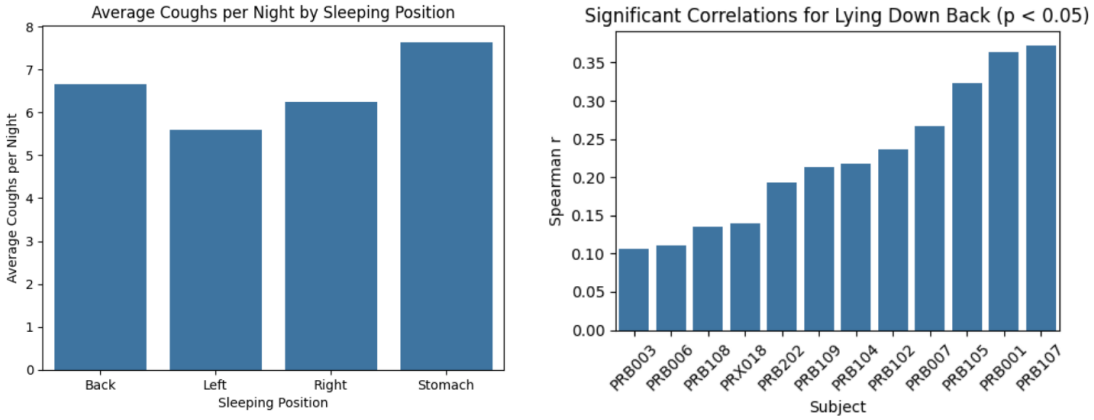


Figure 5.8: Left – The most common sleeping positions for each subject against the average number of coughs per night. Right – Correlations for the position lying down on the back, for subjects showing statistical significance.

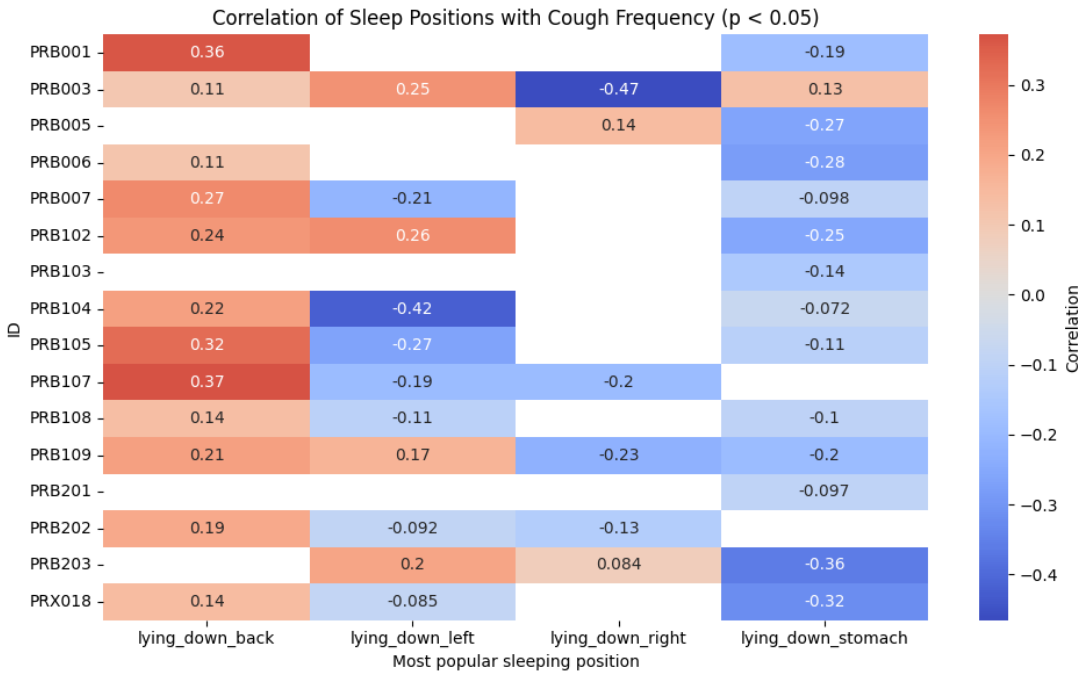


Figure 5.9: Statistically significant correlations between cough frequency and the sleeping position for each subject.

This can be analysed further by looking at the correlations between the number of coughs per hour and each sleeping position. Spearman's correlation, which consists of a non-parametric measure of rank correlation, was chosen to assess how well the relationship between two variables can be described with a monotonic function. The coefficients range from -1 to 1, where 1 indicates a perfect positive monotonic relationship, -1 is a perfect negative one, and 0 indicates no relationship. To do this, for each hour during nighttime, the mode of each lying position was chosen and joined with the corresponding cough counts. Then, the Spearman correlation coefficients were computed between each position and the hourly cough counts, both across the full cohort and on a per-subject basis, offering insights into how posture may influence symptom burden.

As seen on the right of Figure 5.8 and in Figure 5.9, results were especially significant for lying down on the back. Across all subjects, positive and negative correlations were found for lying on the sides and their stomach. However, lying on the back produced positive correlations for all subjects with a statistical significance, showing this position negatively affects their condition the most. Lying down on the sides produced more negative correlations than positive, suggesting that increased time spent sleeping on their side (left or right) is associated with fewer nocturnal cough episodes. Subjects PRB203, PRB109, PRB102 and PRB003 showed positive correlations, which might reflect the high prevalence of lateral sleeping positions in the cohort, making coughing episodes more likely in these positions simply due to longer recorded time. Lying on the back had mostly negative correlations, although this could be due to the small amount of data recorded in that position, where coughs were uncommon.

These correlations highlight the importance of personalised recommendations for sleep

posture among COPD patients. Identifying optimal positions that minimise cough frequency could provide valuable guidance in symptom management, potentially enhancing sleep quality and overall respiratory health.

5.3 COPD Assessment Test (CAT)

Every subject was asked to fill in the CAT questionnaire, which is a validated tool for assessing the impact of COPD on a patient's health status. The CAT is a simple, self-administered questionnaire comprising eight items, each scored from 0 to 5, with a total score ranging from 0 to 40. Higher scores indicate a more severe impact of COPD on the patient's health status. In particular, question 7 correlates to sleep, where they have to answer whether they sleep soundly or not due to their condition. However, as seen in Figure C.3, CAT scores were not recorded daily as intended, generating an incomplete dataset.

5.3.1 Poor Nighttime Sleep against Daytime Sleeping

Previous literature has established that COPD patients experiencing poor nighttime sleep quality often report increased daytime sleepiness, typically attributed to nocturnal respiratory difficulties such as coughing, dyspnea, or oxygen desaturation [9][57]. This relationship can be evaluated by comparing the CAT responses against the frequency of daytime sleep periods (naps) observed the subsequent day.

The number of naps was calculated by focusing on the daytime period, which consists of the times between 11:00 and 20:00. The algorithm identifies lying-down periods during this time frame when the subject is asleep. Periods, where the subject is asleep for more than 10 minutes, are considered naps since literature advises that naps should be at least this long to be beneficial [24]. The results are shown in Figure 5.10, accompanied by Figure C.4, which shows the duration of the naps against the CAT scores.

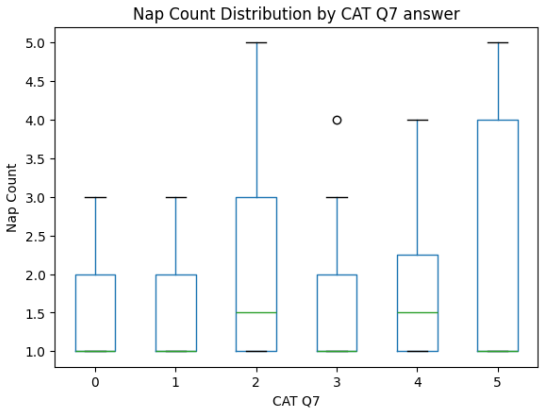


Figure 5.10: Number of naps after filling CAT question 7 for all subjects.

The figure shows a general trend where higher scores on CAT Question 7 (which specifically assesses sleep quality) correspond to an increased frequency of daytime napping. Specifically, subjects reporting the highest levels of nighttime sleep disturbance

(scores of 4 and 5) show notably higher median and maximum nap counts, suggesting a direct link between impaired nighttime sleep and compensatory daytime sleep episodes. These findings show the potential clinical significance of monitoring daytime sleepiness as an additional indicator of nighttime sleep disturbances. They also suggest that targeted interventions to improve nocturnal sleep quality could reduce reliance on daytime naps. This could improve overall daytime functioning, work productivity and quality of life for COPD patients.

5.3.2 CAT Score Correlation to Sleep Metrics

Each sleep metric was correlated to question 7 in the CAT score using Spearman's correlation coefficient as in Section 5.2, in order to measure the relationship between each variable. As a result, the sleep metrics from each night were joined to the CAT score diary from the day after, if available, to then compute the correlation. However, most subjects answered with a three in question 7, which is the neutral answer. This led to a low correlation between the sleep metrics and the CAT score in some cases. Additionally, it is important to note that sleep quality is a subjective matter, and the objective metrics obtained from the RESpeck data may not fully capture the subjective experience of sleep quality.

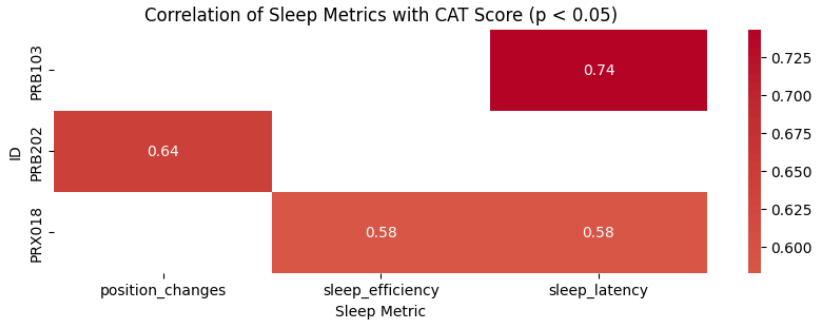


Figure 5.11: Statistically significant correlations between sleep metrics and the CAT score for each subject.

The sleep metrics that achieved a p-value lower than 0.05 are shown in Figure 5.11. Each block represents a statistically significant ($p < 0.05$) correlation between a participant's CAT score and a particular sleep metric, with all results indicating a positive (red) relationship. PRB103 and PRB102 exhibit the highest correlations for sleep latency and positional changes, respectively. This implies that when a COPD patient takes longer to fall asleep or has more frequent nighttime movements, the CAT score increases, suggesting higher symptom severity. Subject PRX018 shows correlations of 0.58 for both sleep efficiency and sleep latency, reinforcing the notion that multiple quality metrics can affect COPD patients' sleep.

5.3.3 CAT Score Prediction

Based on the collected sleep metrics, predictive models were developed to estimate the CAT scores. Initially, all available sleep metrics were included as predictors, but

preliminary results demonstrated limited predictive performance. Therefore, feature selection was applied by identifying metrics with statistically significant correlations to the CAT scores. Consequently, three primary metrics—sleep efficiency, sleep latency, and positional changes—were selected for subsequent modelling. Table 5.1 summarises the four main methods the prediction was tested with.

Algorithm	**Description**
Linear Regression	Models the relationship between a dependent variable and one or more independent variables by fitting a linear equation [51].
Random Forest	Constructs multiple decision trees during training and outputs the average prediction of the individual trees, can handle non-linear relationships and interactions between features [43].
Logistic Regression	Uses a logistic function to model the probability of a binary outcome. Used for binary classification tasks but is adapted for regression tasks by predicting continuous values [20].
XGBoost	A gradient-boosted decision tree framework that iteratively corrects residual errors to capture non-linear patterns and variable interactions [12].

Table 5.1: Summary of the four main methods used to predict the CAT scores.

Results

Algorithm	**Test Set Accuracy**
Linear Regression	0.28
Random Forest	0.60
Logistic Regression	0.70
XGBoost	0.50

Table 5.2: Prediction of CAT score using sleep metrics.

Table 5.2 summarises the predictive performance (expressed as accuracy on the test set) of each algorithm when predicting the CAT scores from the selected sleep metrics. The results indicate substantial variation in predictive accuracy among the methods tested. Linear regression, despite its simplicity and interpretability, showed the lowest predictive accuracy (28%), suggesting that linear relationships alone may not sufficiently capture the interactions between sleep parameters and subjective respiratory symptoms measured by the CAT score. Conversely, logistic regression provided the highest predictive accuracy (70%), demonstrating its effectiveness in capturing categorical relationships even when adapted for regression scenarios. Random Forest (60%) and XGBoost (50%) provided intermediate predictive accuracy, which, although lower than logistic regression, still highlights the importance of non-linear relationships captured by tree-based ensemble methods.

Nonetheless, it is important to note significant limitations affecting these predictions. Firstly, the dataset's size and variability are limited, constraining the predictive capabilities of all tested models. Furthermore, the subjective nature of the CAT score may not be fully represented by purely objective sleep metrics derived from the RESpeck sensor.

Chapter 6

Conclusions

6.1 Discussion

This research explored an objective approach to assessing sleep quality in patients with Chronic Obstructive Pulmonary Disease (COPD) through the use of wearable sensor technology. The initial step involved performing Human Activity Recognition (HAR) using labelled data from the DAPHNE 1.5 and HAR datasets. A hierarchical classification approach was adopted, in which an initial model distinguished between static and dynamic activities, followed by specialised models that classified the specific activity types. Among the tested architectures, Convolutional Neural Networks (CNNs) demonstrated superior performance, delivering high-accuracy activity classification, especially within lying-down activities, achieving an accuracy of 95%.

The following achievement was the development of a sleep classification model. A range of machine learning models–supervised, unsupervised and hybrid models–were investigated for sleep classification. Unsupervised methods achieved more robust results in this context, notably reducing the number of singleton (isolated) awake episodes during the night, which served as a key metric for evaluating model accuracy. However, it was the integration of categorical HAR outputs with numerical sensor data using FAMD with K-Means clustering that significantly enhanced the coherence of sleep episode classification. This finding emphasises that incorporating both categorical and numerical features outperforms traditional approaches relying solely on a single data type.

The HAR and sleep classification results enabled the objective quantification of key sleep metrics—such as total sleep duration, sleep efficiency, sleep latency, number of awakenings, and nocturnal position changes—offering insights into the sleep disturbances experienced by COPD patients. Correlations between these objective metrics and subjective patient-reported outcomes (CAT scores) emerged, confirming the clinical value of wearable-derived data. Notable findings included a positive correlation (0.64) between increased nocturnal positional changes and poorer self-reported sleep quality. A strong association between prolonged sleep latency and efficiency with higher CAT scores was found (0.74 and 0.58, respectively), indicating more significant discomfort or symptom burden. High correlations were also found between the number of coughs

and the chosen sleeping position, where sleeping in a supine position always resulted in a higher number of coughs, whereas side sleeping postures minimised symptoms.

Finally, predictive modelling using logistic regression achieved promising results, with up to 70% accuracy in estimating CAT scores from sleep metrics—despite the limitations of the small dataset size. These findings provide evidence supporting the role of wearable sensors in clinical settings, particularly for the continuous, objective monitoring of sleep in COPD patients.

6.2 Limitations and Future Work

Despite the promising results, several limitations were encountered. One of the primary challenges was the limited size and inconsistency of the dataset, particularly regarding the CAT (COPD Assessment Test) scores. Not all patients completed the questionnaire daily, resulting in sparse data that hindered the model's ability to effectively learn and predict CAT scores. Furthermore, as CAT answers are based on subjective self-assessment, they may not fully align with the objective sleep disturbances captured by the RESpeck wearable sensor. This discrepancy between subjective reports and sensor-derived data likely impacted the accuracy and generalisability of the predictive models, contributing to varying correlation patterns between sleep metrics and CAT scores across different patients. Future studies with a larger and more consistently collected CAT dataset could offer deeper insights into these relationships and enhance model performance.

In terms of sleep classification, supervised learning techniques achieved high accuracy on the labelled test set, demonstrating the potential of these models. For future research, it would be beneficial to apply these models to a larger labelled dataset in order to ensure more robustness.

Additionally, the proposed HAR and sleep classification models, along with the derived sleep metrics algorithms, present an opportunity for future integration into real-time applications. Deploying these models in live settings could enable immediate clinical feedback and facilitate personalised intervention strategies for managing COPD, especially during the night.

In conclusion, this research highlights the value of wearable sensor technologies in objectively identifying activity and sleep patterns, offering a foundation for monitoring sleep quality in COPD patients. With further validation and expansion, these tools could enhance patient care in real time.

Appendix A

Human Activity Recognition

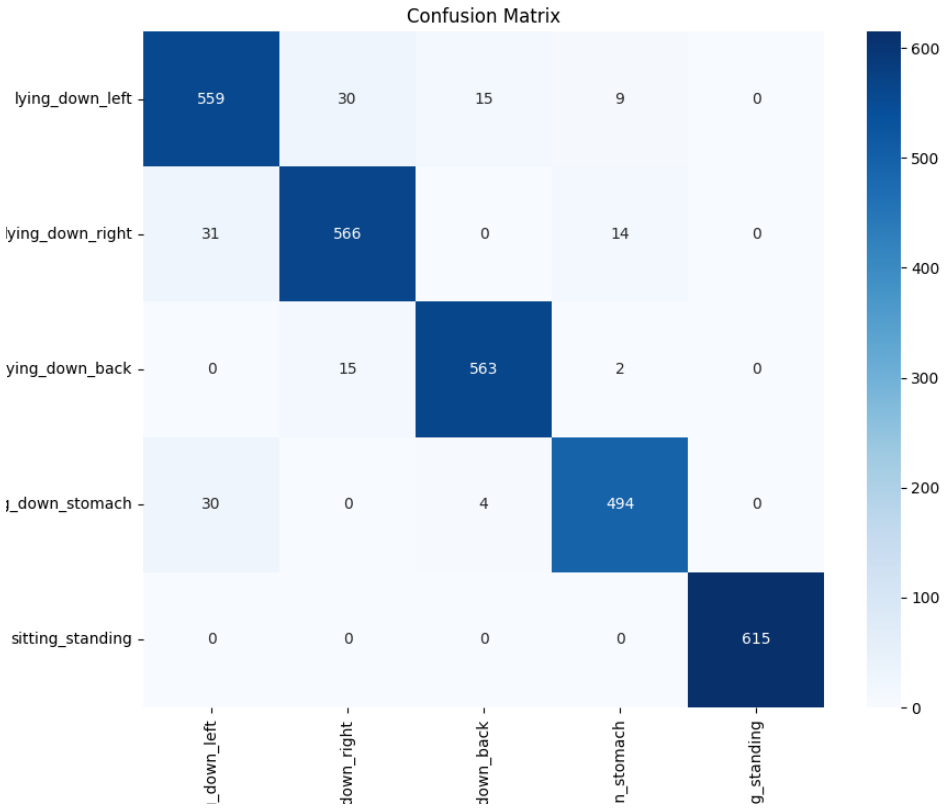


Figure A.1: Confusion Matrix for static activities classification on unseen DAPHNE 1.5 data.

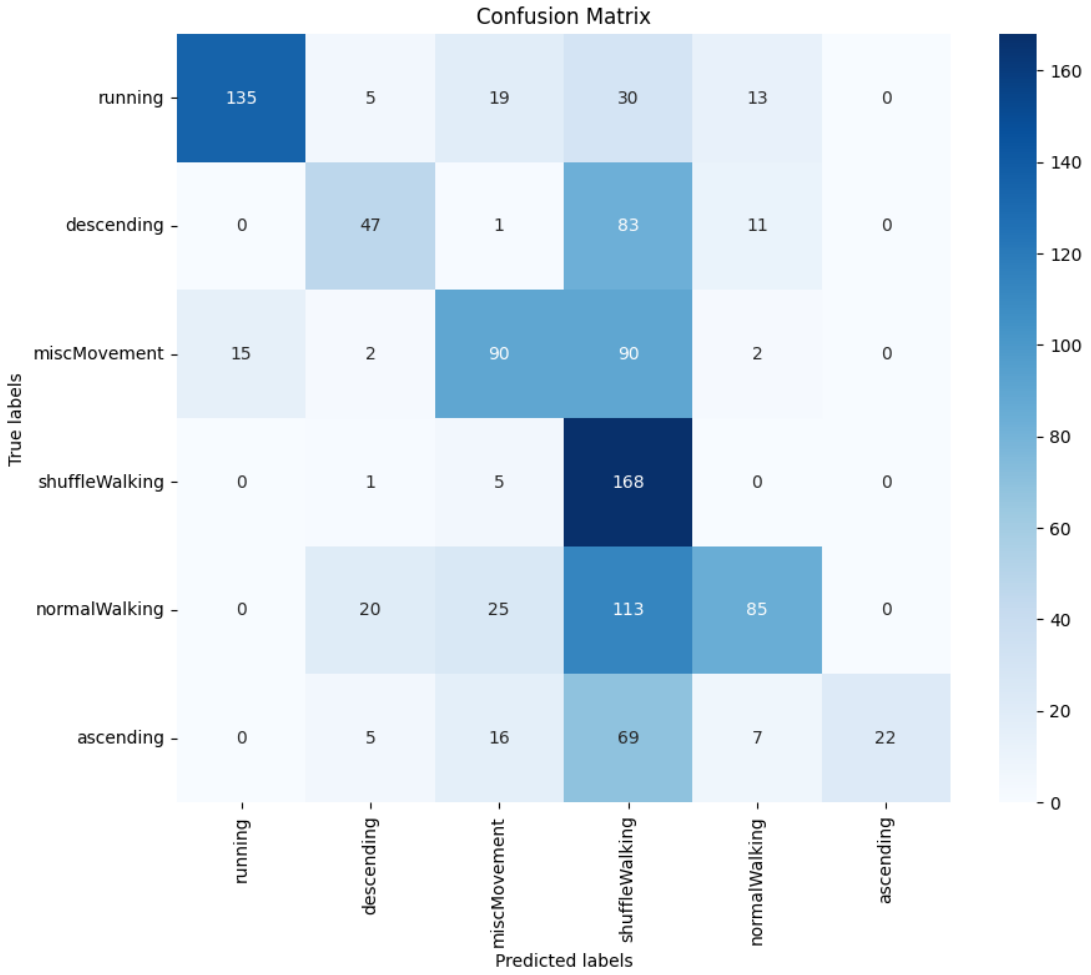


Figure A.2: Confusion Matrix for dynamic activities classification on unseen DAPHNE 1.5 data, without transfer learning.

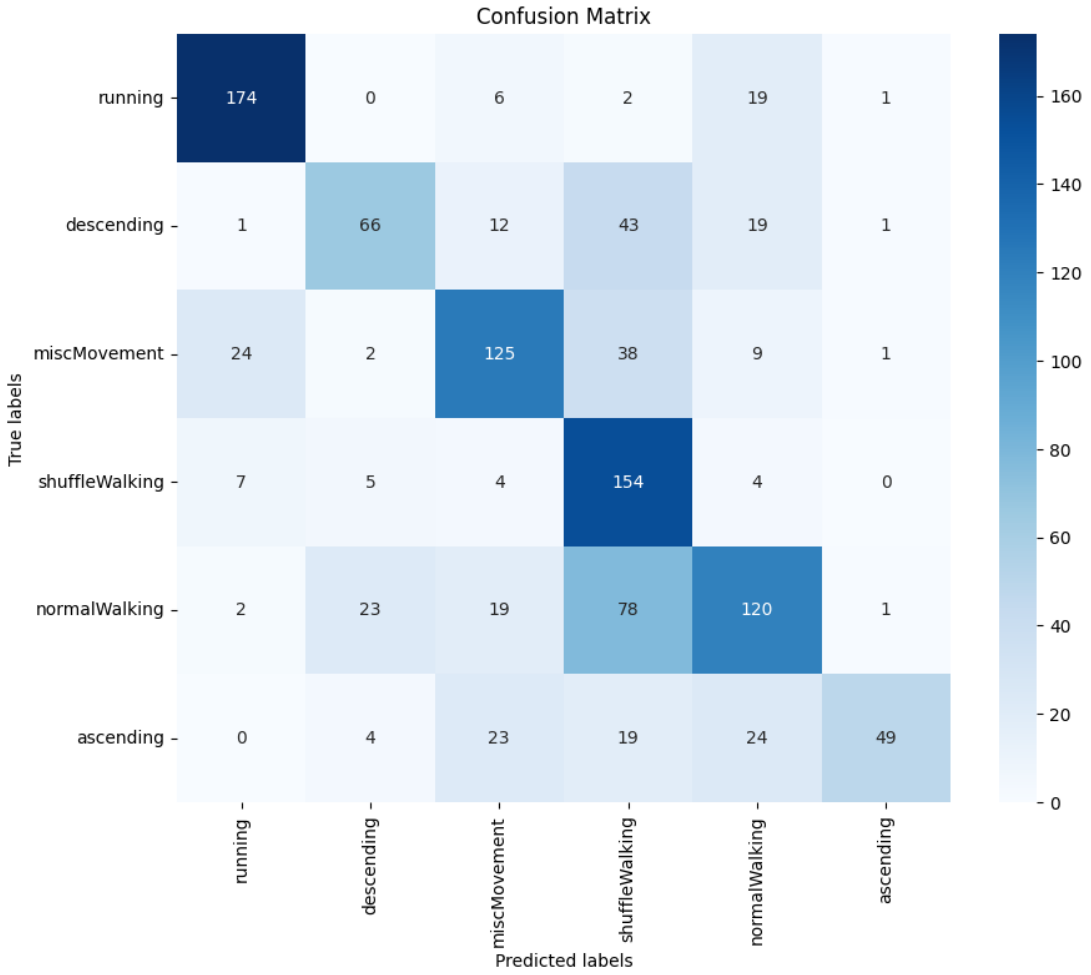


Figure A.3: Confusion Matrix for dynamic activities classification on unseen DAPHNE 1.5 data, with transfer learning.

Appendix B

Sleep Episodes Classification

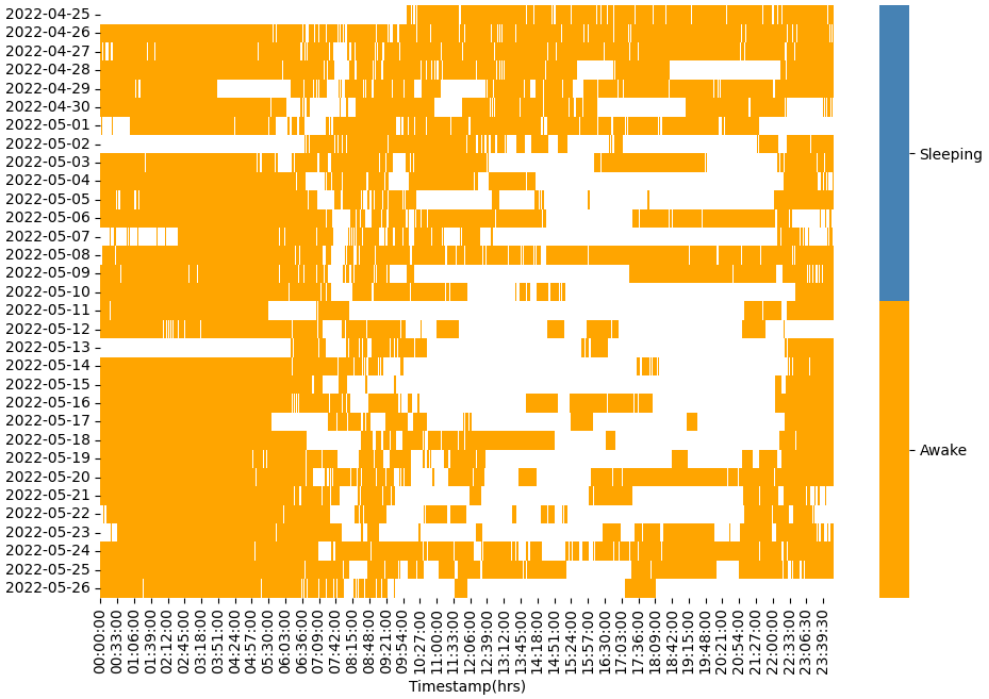


Figure B.1: Supervised Classification with SVM of Subject PRB006.

Appendix C

Sleep Quality Analysis

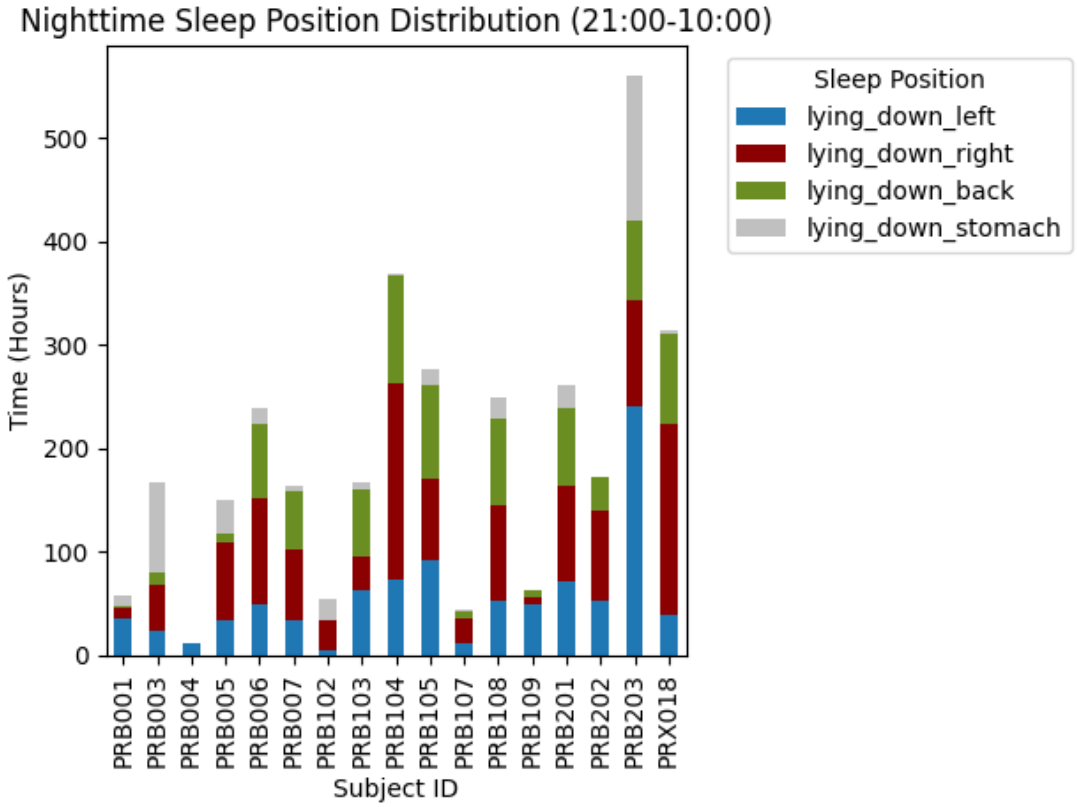


Figure C.1: Number of hours spent in each position across the COPD cohort.

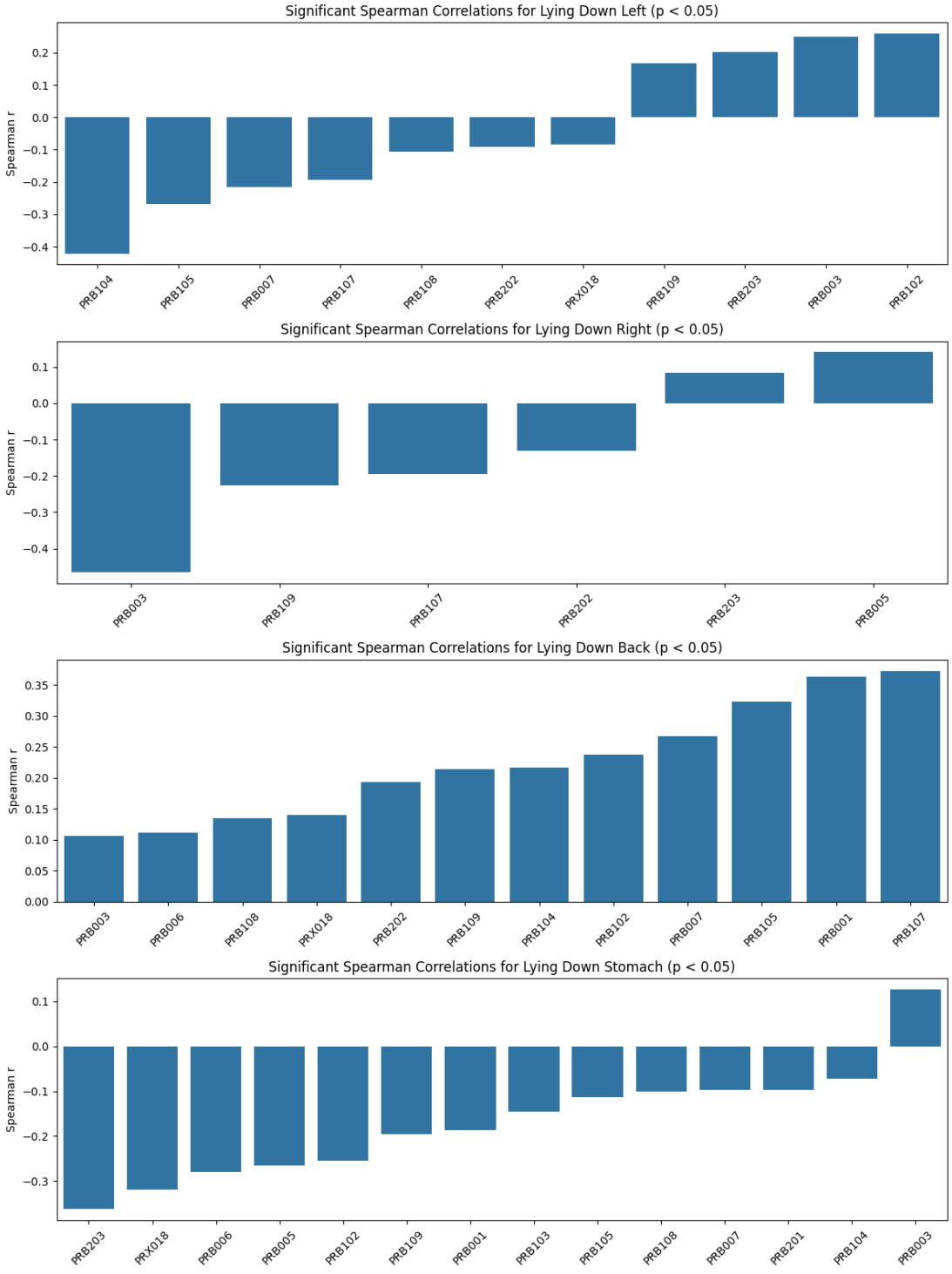


Figure C.2: Correlations between all sleeping positions and cough frequency.

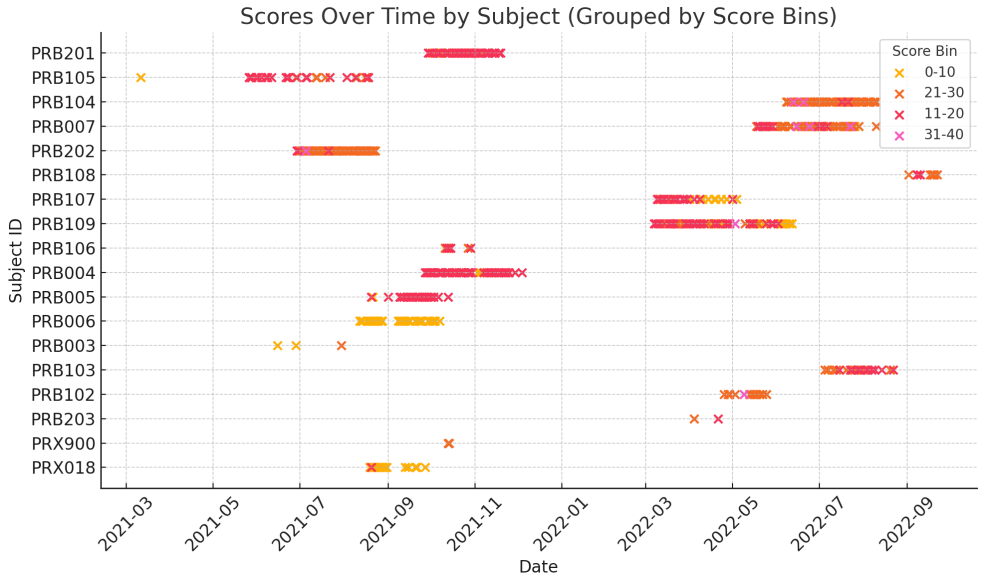


Figure C.3: CAT scores completion over time for all subjects.

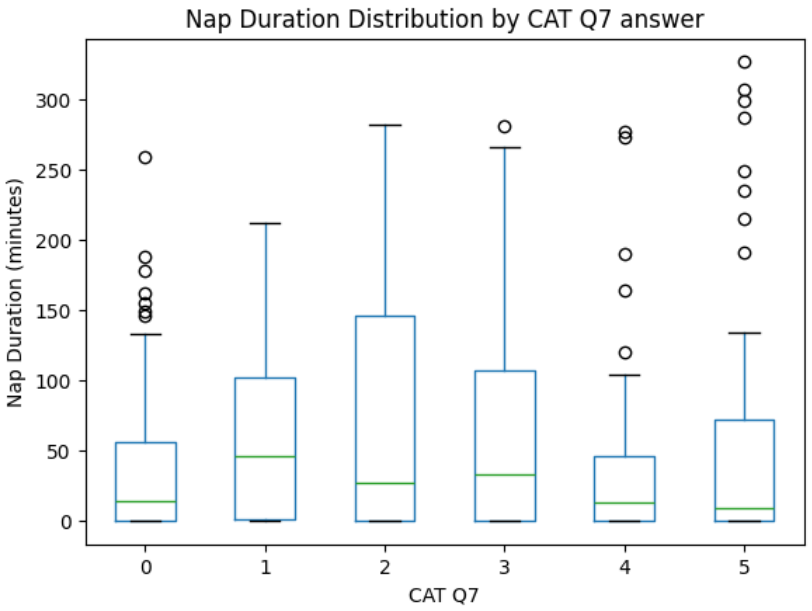


Figure C.4: Nap duration distribution by CAT Q7 score.

Bibliography

[1] Sleep efficiency. https://www.hypersomniafoundation.org/glossary/sleep-efficiency/. Accessed: 2025-03-21.

[2] Y. Akita, T. Nishimura, N. Shibata, K. Kawakatsu, C. Hattori, and M. Yagisawa. Sleeping position in sleep-related respiratory disorders. *Nihon Kikan Shokudoka Gakkai Kaiho*, 51:28–35, 2000.

[3] Ong Chin Ann and Lau Bee Theng. Human activity recognition: A review. In *2014 IEEE international conference on control system, computing and engineering (ICCSCE 2014)*, pages 389–393. IEEE, 2014.

[4] Anshika Arora, Pinaki Chakraborty, and MPS Bhatia. Analysis of data from wearable sensors for sleep quality estimation and prediction using deep learning. *Arabian Journal for Science and Engineering*, 45:10793–10812, 2020.

[5] DK Arvind, DJ Fischer, CA Bates, and S Kinra. Characterisation of breathing and physical activity patterns in the general population using the wearable respeck monitor. In *Body Area Networks: Smart IoT and Big Data for Intelligent Health Management: 14th EAI International Conference, BODYNETS 2019, Florence, Italy, October 2-3, 2019, Proceedings 14*, pages 68–78. Springer, 2019.

[6] DK Arvind, T Georgescu, CA Bates, D Fischer, and Q Zhou. Home-based pulmonary rehabilitation of copd individuals using the wearable respeck monitor. In *EAI International Conference on Body Area Networks*, pages 176–191. Springer, 2021.

[7] DK Arvind and Zac Teo. Machine-learning based classification of sleep-wake states using a respeck patch. In *2024 20th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, pages 266–270. IEEE, 2024.

[8] Konrad E Bloch. Polysomnography: a systematic review. *Technology and health care*, 5(4):285–305, 1997.

[9] Rohit Budhiraja, Tauseef A Siddiqi, and Stuart F Quan. Sleep disorders in chronic obstructive pulmonary disease: etiology, impact, and management. *Journal of Clinical Sleep Medicine*, 11(3):259–270, 2015.

[10] Daniel J Buysse, Charles F Reynolds, Timothy H Monk, Susan R Berman, and

David J Kupfer. The pittsburgh sleep quality index: a new instrument for psychiatric practice and research. *Psychiatry Research*, 28(2):193–213, 1989.

[11] P. Chanchotisatien and D. K. Arvind. Improved cough classification with symmetric projection attractor reconstruction. Submitted to EMBC 2025, 47th IEEE International Conference of the Engineering in Medicine and Biology, Copenhagen, July 2025.

[12] Tianqi Chen, Tong He, Michael Benesty, Vadim Khotilovich, Yuan Tang, Hyunsu Cho, Kailong Chen, Rory Mitchell, Ignacio Cano, Tianyi Zhou, et al. Xgboost: extreme gradient boosting. *R package version 0.4-2*, 1(4):1–4, 2015.

[13] Justin Chiang. Simple 1d cnn approach to human-activity-recognition (har) in pytorch. https://github.com/jchiang2/Human-Activity-Recognition/, 2022. Accessed: 2025-03-26.

[14] Tsair-Wei Chien, Sen-Yen Hsu, Chein Tai, How-Ran Guo, and Shih-Bin Su. Using rasch analysis to validate the revised psqi to assess sleep disorders in taiwan's hi-tech workers. *Community Mental Health Journal*, 44:417–425, 2008.

[15] Arthur P Dempster, Nan M Laird, and Donald B Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, 39(1):1–22, 1977.

[16] Marco Fabbri, Alessia Beracci, Monica Martoni, Debora Meneo, Lorenzo Tonetti, and Vincenzo Natale. Measuring subjective sleep quality: a review. *International journal of environmental research and public health*, 18(3):1082, 2021.

[17] Hristijan Gjoreski, Mitja Lustrek, and Matjaz Gams. Accelerometer placement for posture recognition and fall detection. In *Proceedings of the 7th International Conference on Intelligent Environments*, pages 47–54. IEEE, 2011.

[18] Max Hirshkowitz, Kaitlyn Whiton, Stephanie M. Albert, Cathy Alessi, Oliviero Bruni, Lydia DonCarlos, Nancy Hazen, Joseph Herman, Emily S. Katz, Leila Kheirandish-Gozal, et al. National sleep foundation's sleep time duration recommendations: methodology and results summary. *Sleep Health*, 1(1):40–43, 2015.

[19] Charmi Jobanputra, Jatna Bavishi, and Nishant Doshi. Human activity recognition: A survey. *Procedia Computer Science*, 155:698–703, 2019.

[20] David G Kleinbaum, K Dietz, M Gail, Mitchel Klein, and Mitchell Klein. *Logistic regression*. Springer, 2002.

[21] Jun Kohyama. Which is more important for health: sleep quantity or sleep quality? *Children*, 8(7):542, 2021.

[22] Andrew D Krystal and Jack D Edinger. Measuring sleep quality. *Sleep medicine*, 9:S10–S17, 2008.

[23] Oscar D Lara and Miguel A Labrador. A survey on human activity recognition using wearable sensors. *IEEE communications surveys & tutorials*, 15(3):1192–1209, 2012.

[24] Nicole Lovato and Leon Lack. The effects of napping on cognitive functioning. *Progress in brain research*, 185:155–166, 2010.

[25] J.B. MacQueen. Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, volume 1, pages 281–297. University of California Press, 1967.

[26] Dennis Majoe, Philipp Bonhof, Thomas Kaegi-Trachsel, Jurg Gutknecht, and Lars Widmer. Stress and sleep quality estimation from a smart wearable sensor. In *2010 4th International Conference on Pervasive Computing Technologies for Healthcare*, pages 1–8. IEEE, 2010.

[27] Alessia Mammone, Marco Turchi, and Nello Cristianini. Support vector machines. *Wiley Interdisciplinary Reviews: Computational Statistics*, 1(3):283–289, 2009.

[28] Walter T. McNicholas, Johan Verbraecken, and Jose M. Marin. Sleep disorders in copd: the forgotten dimension. *European Respiratory Review*, 22(129):365–375, 2013.

[29] David G Mcsharry, Silke Ryan, Peter Calverley, J Colin Edwards, and Walter T Mcnicholas. Sleep quality in chronic obstructive pulmonary disease. *Respirology*, 17(7):1119–1124, 2012.

[30] Suphakant Mekruksavanich and Aitthipat Jitpattanakul. Lstm networks using smartphone data for sensor-based human activity recognition in smart homes. *Sensors*, 21(5):1636, 2021.

[31] Hiroyasu Miwa, Shin-ichiro Sasahara, and Toshihiro Matsui. Roll-over detection and sleep quality measurement using a wearable sensor. In *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 1507–1510. IEEE, 2007.

[32] W. Mueller, P. Wilkinson, J. Milner, M. Loh, S. Vardoulakis, Z. Petard, and D. K. Arvind. The relationship between greenspace and personal exposure to pm2.5 during walking trips in delhi, india. *Environmental Pollution*, 305:119294, 2022.

[33] National Sleep Foundation. Recommended sleep duration. urlhttps://www.sleepfoundation.org/how-sleep-works/how-much-sleep-do-we-really-need, 2023. Accessed: March 17, 2025.

[34] Thien Nguyen and Tuan Dinh. Complex human activities recognition based on high performance 1d cnn model. *Procedia Computer Science*, 198:601–608, 2022.

[35] Gillian M Nixon, John MD Thompson, Dug Yeo Han, David MO Becroft, Phillipa M Clark, Elizabeth Robinson, Karen E Waldie, Chris J Wild, Peter N Black, and Edwin A Mitchell. Falling asleep: the determinants of sleep latency. *Archives of disease in childhood*, 94(9):686–689, 2009.

[36] Maurice Ohayon, Emerson M. Wickwire, Max Hirshkowitz, Stephanie M. Albert, Stephen Auerbach, Francis J. Daly, Kathy A. Dudley, James M. Parish, and ...

National sleep foundation's sleep quality recommendations: first report. *Sleep Health*, 3(1):6–19, 2017.

[37] Francisco Javier Ordóñez and Daniel Roggen. Deep convolutional and lstm recurrent neural networks for multimodal wearable activity recognition. *Sensors*, 16(1):115, 2016.

[38] Panayiotis Panayides, Marios Gavrielides, Christodoulos Galatopoulos, and Mikaella Gavriilidou. Using rasch measurement to create a quality of sleep scale for a non-clinical sample based on the pittsburgh sleep quality index (psqi). *Europe's Journal of Psychology*, 9(1):113–135, 2013.

[39] Aakash Parmar, Rakesh Katariya, and Vatsal Patel. A review on random forest: An ensemble classifier. In *International conference on intelligent data communication technologies and internet of things (ICICI) 2018*, pages 758–763. Springer, 2019.

[40] Shyamal Patel, Chiara Mancinelli, Jennifer Healey, Marilyn Moy, and Paolo Bonato. Using wearable sensors to monitor physical activities of patients with copd: A comparison of classifier performance. In *2009 Sixth International Workshop on Wearable and Implantable Body Sensor Networks*, pages 234–239. IEEE, 2009.

[41] Javad Razjouyan, Hyoki Lee, Sairam Parthasarathy, Jane Mohler, Amir Sharafkhaneh, and Bijan Najafi. Improving sleep quality assessment using wearable sensors by including information from postural/sleep position changes and body acceleration: A comparison of chest-worn sensors, wrist actigraphy, and polysomnography. *Journal of Clinical Sleep Medicine*, 13(11):1301–1310, 2017.

[42] Rafael Schild Reusch, Leonardo Rezende Juracy, and Fernando Gehm Moraes. Assessment and optimization of 1d cnn model for human activity recognition. In *Brazilian Symposium on Computing Systems Engineering (SBESC)*. IEEE, 2022.

[43] Steven J Rigatti. Random forest. *Journal of Insurance Medicine*, 47(1):31–39, 2017.

[44] Charissa Ann Ronao and Sung-Bae Cho. Human activity recognition with smartphone sensors using deep learning neural networks. *Expert Systems with Applications*, 59:235–244, 2016.

[45] Jessica Vensel Rundo and Ralph Downey III. Polysomnography. *Handbook of clinical neurology*, 160:381–392, 2019.

[46] Sebastian Rutkowski, Joren Buekers, Anna Rutkowska, Błażej Cieślik, and Jan Szczegielniak. Monitoring physical activity with a wearable sensor in patients with copd during in-hospital pulmonary rehabilitation program: A pilot study. *Sensors*, 21(8):2742, 2021.

[47] Bernhard Schölkopf, John C Platt, John Shawe-Taylor, Alexander J Smola, and Robert C Williamson. Estimating the support of a high-dimensional distribution. *Neural Computation*, 13(7):1443–1471, 2001.

[48] Dave Singh, Alvar Agusti, Antonio Anzueto, Peter J Barnes, Jean Bourbeau, Bartolome R Celli, Gerard J Criner, Peter Frith, David MG Halpin, Meilan Han,

et al. Global strategy for the diagnosis, management, and prevention of chronic obstructive lung disease: the gold science committee report 2019. *European Respiratory Journal*, 53(5), 2019.

[49] Jørgen Skotte, Mette Korshøj, Jesper Kristiansen, Christoph Hanisch, and Andreas Holtermann. Detection of physical activity types using accelerometers and machine learning. *Scandinavian Journal of Medicine & Science in Sports*, 24(5):e110–e117, 2014.

[50] Sleep Foundation. Wakefulness after sleep onset (waso). *Sleep Foundation*, 2023.

[51] Xiaogang Su, Xin Yan, and Chih-Ling Tsai. Linear regression. *Wiley Interdisciplinary Reviews: Computational Statistics*, 4(3):275–294, 2012.

[52] H. Wang and J. Mi. Intuitive-k-prototypes: A mixed data clustering algorithm with intuitionistic distribution centroid. *Pattern Recognition*, 158:111062, 2025.

[53] Yuxuan Wang. Classification of human activity based on transfer learning. Master's thesis, The University of Edinburgh, 2024.

[54] Karl Weiss, Taghi M Khoshgoftaar, and DingDing Wang. A survey of transfer learning. *Journal of Big data*, 3:1–40, 2016.

[55] Hyeryeon Yi, Kyungrim Shin, and Chol Shin. Development of the sleep quality scale. *Journal of sleep research*, 15(3):309–316, 2006.

[56] Tony Zebin, Matthew Sperrin, Niels Peek, and Alexander J Casson. Human activity recognition from inertial sensor time-series using batch normalized deep lstm recurrent networks. In *Proceedings of the IEEE EMBC*, pages 1–4. IEEE, 2018.

[57] Mohammad Ali Zohal, Zohreh Yazdi, and Amir Mohammad Kazemifar. Daytime sleepiness and quality of sleep in patients with copd compared to control group. *Global journal of health science*, 5(3):150, 2013.