SOMNUS: An Ultra-Wideband Radar-Based Approach for Neonatal Sleep State Classification

M. Silos Viu

Technische Universiteit Delft

TU Delft Delft University of Technology

Challenge the future

SOMNUS: AN ULTRA-WIDEBAND RADAR-BASED APPROACH FOR NEONATAL SLEEP STATE CLASSIFICATION

by

M. Silos Viu

in partial fulfillment of the requirements for the degree of

Master of Science
in Biomedical Engineering

in

Bioelectronics Group
Microelectronics Department
Delft University of Technology,

to be defended publicly on Thursday November 9, 2017 at 2:00 PM.

Supervisor:	Prof. dr. ir. W. Serdijn	
Thesis committee:	Prof. dr. ir. W. Serdijn,	TU Delft
	Ir. C. J. Bees,	TU Delft
	Dr. ir. A. C. Schouten,	TU Delft
	Dr. J. Dudink,	UMC Utrecht

This thesis is confidential and cannot be made public until November 9, 2018.

TU Delft Delft University of Technology

ABSTRACT

Recently, there has been an increasing awareness of the critical role of sleep for the brain development of young infants. During the early neonatal stages of human development, the basic activity of the brain is to develop itself during sleep. Prolonged sleep is required in infants for further development of the nervous system. Sleep also plays an important role in body temperature regulation and energy saving. Neonatal sleep is divided into two main different sleep stages: Rapid Eye Movement (REM) and Non-Rapid Eye Movement (NREM). As the infant develops, sleep stages vary in maturity, length and distribution, thus underlining the importance of the quantification of these stages, which could eventually lead to new biomarkers of neonatal brain development.

Nowadays, the golden standard in sleep monitoring is Polysomnography (PSG), in which vital signs as well as EEG and muscle activity are recorded during a whole-night study and subsequently sleep stages are classified by an expert. However, the high obtrusiveness of the multiple electrodes involved in PSG and its high associated costs make it impossible to use PSG as a routine monitor system. The SOMNUS project had two main targets: (i) to accurately measure respiration signals from patients using an ultra-wideband radar module, and (ii) to detect differences in respiration between REM and NREM phases in order to unobtrusively and automatically score sleep states of infants without the need of any electrode attached to the patient. The system has been developed using a training dataset of 23 patients from 3 months to 14 years old. It is for the first time UWB radar technology is used to monitor sleep in young patients. Moreover, this work provides a new data analysis algorithm to suppress motion artifacts from radar signals and increase the robustness of respiration monitoring.

The results from the breathing detection algorithm developed in the present work provided an average mean absolute error of 3.25 Respirations per Minute in respiration rate. Amount of movement was employed to estimate Sleep/Awake events, with an average error in percentage of Total Sleep Time of 7.48%. Regarding sleep-state classification, several classification analyses were performed in order to study the best classification variable to detect REM/NREM, along with analyses regarding the best classification technique. Respiration variability was the main feature determining REM/NREM state, with an overall sleep classification accuracy of around 80% when using linear classifiers such as Support Vector Machines and Fisher Linear Discrimination.

We need women at all levels, including the top,
to change the dynamic, reshape the conversation,
to make sure women's voices are heard and heeded,
not overlooked and ignored.

SHERYL SANDBERG

PREFACE

I can't believe it's been two years since I left Spain and started this new adventure in such a great (and rainy) country as the Netherlands. There are so many people who I want to thank for all I have learned from them and for constantly making me feel at home in Delft.

I would like to especially thank my supervisor, Prof. Wouter Serdijn, for all the guidance throughout the project and for giving me the opportunity to join the Bioelectronics department. It was really awesome to work in such a great environment.

To Dr. Jeroen Dudink for coming up with such an amazing project, for all your enthusiasm, which motivated me along the whole duration of the project and for making SOMNUS possible.

To Robbin de Goederen for your endless contributions to this project, for sharing all these fruitful discussions about SOMNUS and for all your help with the data acquisition process.

To *Novelda AS*, especially Olav Liseth, for all your help making me understand *XeThru* technology.

To my parents and brother for supporting all my decisions from the very beginning. Os quiero.

To Atul, for understanding me at levels no one possibly can, I cannot find enough words to thank you for everything. To Jon, Meg and Tim, for always being there and for making me smile every time, I never imagined I would manage to find such a great group of people. To Mònica, for being there in all my best memories from Delft.

M. Silos Viu
Delft, October 2017

CONTENTS

List of Figures

List of Tables

ABBREVIATIONS

AASM	American Academy of Sleep Medicine
AS	Active Sleep
BCG	Ballistocardiography
BPM	Beats Per Minute
CA	Corrected Age
CW	Continuous Wave
ECG	Electrocardiography
EIRP	Equivalent Isotropically Radiated Power
EEG	Electroencephalography
EM	Electromagnetic
EMG	Electromyography
EOG	Electrooculography
ETSI	European Telecommunications Standards Institute
FCC	Federal Communications Commission
FMCW	Frequency-Modulated Continuous Wave (Radar)
GA	Gestational Age
HR	Heart Rate
IR	Infrared
IS	Intermediate Sleep = Transitional Sleep (TS)
HRV	Heart Rate Variability
LDA	Linear Discriminant Analysis
LO	Local Oscillator
LOOCV	Leave-One-Out Cross Validation
MAE	Mean Absolute Error
ME	Mean Error
NICU	Neonatal Intensive Care Unit
PCA	Principal Component Analyses

PRF Pulse Repetition Frequency

PSG Polysomnography

QS Quiet Sleep

RE Relative Error

REM Rapid Eye Movement

RF Radio-frequency

RIP Respiratory Inductance Plethysmography

RMSE Root Mean Squared Error

RPM Respirations Per Minute

RR Respiration Rate

SCSB Static Charge Sensitive Bed

STD Standard Deviation

SVM Support Vector Machines

SoC System on a chip

TS Transitional Sleep = Intermediate Sleep (IS)

TST Total Sleep Time

UWB Ultra-Wideband

1

INTRODUCTION

In recent years, there has been an increasing awareness on the critical role of sleep, and more importantly REM sleep for brain development in young infants [16]. During the early neonatal stages of human development, the basic activity of the brain is to sleep. Prolonged sleep is required in infants for further development of brain systems such as hippocampus, pons, brainstem, and midbrain, as well as neurosensory systems [16]. Sleep also plays an important role in body temperature regulation and as energy saving system [17].

As Dudink and van de Hoogen state in their systematic review [18], even though there are many studies reporting procedures to promote and optimize neonatal sleep on the neonatal intensive care unit (NICU), there are great differences among them regarding sleep assessment methods, targeted sleep behaviors and study populations. These prevent medical community from recommending a concrete intervention to promote neonatal sleep on the NICU. Hence, evidently requiring a great need to improve and standardize sleep research procedures.

Neonatal sleep is divided into two main different sleep stages: Rapid Eye Movement (REM) and Non Rapid Eye Movement (NREM). The cycling of these two states has shown to be critical for brain maturity [16]. As the infant develops, sleep stages vary in maturity, length and distribution, thus the importance of the quantification of these stages that could eventually lead to new biomarkers of neonatal brain development.

Polysomnography (PSG) studies currently represent the gold standard in sleep monitoring. During this procedure, several biosignals such as Electrocardiogram (ECG), Electroencephalogram (EEG), Electrooculogram (EOG), Electromyogram (EMG), respiration patterns and heart rate (HR), among others, are monitored for a long sleeping period, after which qualified experts perform the subsequent sleep stages classification manually. However, the high obtrusiveness of the multiple electrodes involved in the PSG and its high associated cost make it impossible to be used in sleep longitudinal studies[1] in young infants. It is also an inconvenient method for long-term ambulatory sleep monitoring.

Despite of being highly resource demanding and obtrusive, sleep monitoring plays a critical part of the diagnostic process. As such, it is desirable to develop an automated method for continuously determining sleep states, which ideally requires very few non-invasively extracted physiological signals.

[1] Longitudinal studies are referred to observational studies performed during a long period of time.

In order to be able to unobtrusively classify sleep stages in neonates, there are two main challenges to address: first, to non-invasively acquire physiological signs which determine sleep phases; and second, to extract the right features that will be used for the subsequent classification of awake/ REM/ NREM states. Several contactless vital sensing approaches which may potentially be used as neonatal sleep monitoring, have recently been investigated for neonatal breathing and heart rate monitoring. Some of these approaches include thermal imaging [3, 19], ballistocardiography [2] and radar [20–22] methods, which were thoroughly discussed and compared during the development of this project. Due to the availability of a stand-alone, respiration detector radar module, the focus of this project was on validating its applicability as a sleep monitoring system. To do so, the radar module, first acquires patient's vital signs to subsequently identify different patterns of each sleep state. This radar module is market available at a relatively low cost. Its software provides information of movement, breathing pattern and breathing rate. However, during the validation study, a need for a more robust, less noisy readings system was noted. This work aims at employing new algorithms to increase robustness of recordings and decrease movement noise. As a result, we could then employ pattern recognition algorithms in order to address the sleep stages classification problem.

1.1. THESIS OBJECTIVES

In this thesis, the focus was on acquiring movement, respiration and heart rate signals by means of an Ultra-Wideband (UWB) radar module, and subsequently extract a set of features employed for sleep states classification. The final results off this study were compared to the gold standard (PSG) to ultimately classify sleep stages in young infants. Figure 1.1 represents the diagram of the major tasks to be addressed in this thesis. The objectives of the present work were divided in several tasks:

- To understand UWB radar and to study the feasibility of using a commercial UWB radar sensor on neonatal intensive care unit.
- To extract movement quantity, breathing pattern and breathing rate from radar module data.
- To improve robustness and noise level of the radar module used for respiration detection.
- To adapt conventional respiration detection algorithms for every patient automatically.
- To automatically detect motion artifacts that prevent correct breathing readings.
- To study the system accuracy for vital signs detection in comparison with gold standard (PSG).
- To perform feature extraction and selection from the previous vital signs.
- To detect Sleep-Awake events based on patient's movement.
- To make use of a suitable classification algorithm to successfully classify REM and NREM states of patients.
- To evaluate final classification performance compared to PSG sleep state annotations.

1.2. THESIS SCOPE BOUNDARIES

The project described in this thesis was previously discussed with neonatologist experts who provided their opinion on the current sleep monitoring procedures on hospital conditions. Hence, the scope of the project was adjusted accordingly to their needs. Time constraints also played a big role

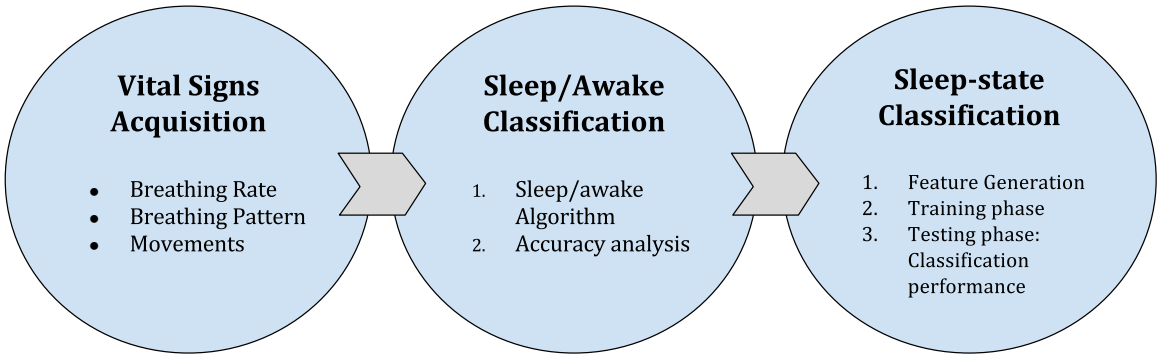


Figure 1.1: Diagram of Thesis Objectives.

in the definition of the project scope. The boundaries that will not be addressed in this thesis are established bellow:

- Classification of sleep states does not need to be highly time-accurate as the transition between sleep states is a slow process. According to the American Academy of Sleep Monitoring (AASM), sleep periods should be classified in 30 seconds-epochs [23], therefore the time-resolution will not be higher than that.
- No electronic components of the employed radar modules will be altered or modified. Its circuitry will remain the same as bought.
- The system will not be implemented to work as a real-time final product designed for sleep monitoring. The present work aims to study the feasibility of such system in the future.
- The system is not intended to be used as a clinical device that would substitute a PSG study, but rather a complementary device that would give an overall idea of sleep information of patients at NICU that do not require a complete PSG study.
- Due to high processing complexity for HR detection, the system will not aim at acquiring HR information.
- Age and characteristics of the patients (healthy or not) does not depend on the researchers but on the availability of patients used as benchmark of the system.

1.3. THESIS OUTLINE

Background problem statement, project objectives and scope were stated in this first chapter of this work.

Chapter 2 and 3 comprises the literature research assignment undergone for this Master Thesis project.

In Chapter 2, background about neonatal sleep monitoring is given. Moreover, discussion of the most outstanding studies in the sleep research literature regarding sleep states classification in both adults and young infants is presented.

Chapter 3 describes the three main methods for vital signs detection, which consequentially represent potential methods for sleep monitoring. A further and more detailed explanation is given for

radar technology which fulfills the aims of this project.

Chapter 4 explains the main characteristics of the radar modules used, such as radar parameters, raw data and internal processing algorithms.

Chapter 5 comprises the aspects of data acquisition, including patients characteristics, study protocol and set-up, data pre-processing and PSG and radar data acquisition.

Chapter 6 explains the details behind the breathing detection algorithm as well as accuracy analyses results compared to the reference data (PSG).

Chapter 7 describes the different aspects to take into account in the sleep classification problem. The performance of the classification system is studied when using different features and classifiers.

Finally, Chapter 8 discusses the findings of this project while establishing different conclusions to take into account for future improvements of the system.

2

COMPUTER-AIDED NEONATAL SLEEP MONITORING

Age is considered by sleep experts as the most decisive factor that determines the way humans sleep. Development and maturation define the progress of the human body, and our sleep cycle patterns change accordingly. The low maturity of the neonate's nervous system gives rise to very different sleep patterns compared to the ones of adults. In order to properly monitor neonatal sleep, we require a previous knowledge about how sleep works in young infants. In this section the different characteristics of the neonatal sleep states will be explained, in terms of prevalent vital signs attributes and their distribution within the total sleep time (TST). Furthermore, information about state-of-the-art technologies in automatic sleep classification in both adults and children will be given.

2.1. NEONATAL SLEEP STATES

Many different staging criteria have been used in the past to classify neonatal sleep states. The current accepted guidelines for infant sleep-staging criteria are defined by the American Association of Sleep Medicine (AASM) [24]. These guidelines divide neonatal sleep into two basic phases called *Rapid Eye Movement* (REM) and *Non-Rapid Eye Movement* (NREM), each of which with unique characteristics. In addition, when the brain development has not reached its full maturity, another phase called Transitional Sleep (TS), which is also called Intermediate Sleep (IS), may be characterized by the presence of several attributes of both REM and NREM sleep. The AASM states that REM and NREM phases are roughly equivalent to Active Sleep (AS) and Quiet Sleep (QS) phases, respectively, defined by Anders criteria [25] when used in infants with less than 2 months gestational age [2]. However, as the infant gets older, sleep patterns develop and the NREM phase is divided in four subphases: NREM1, NREM2, NREM3 and NREM4 [24], therefore differentiation between AS and QS cannot be performed with older infants (from 4 months on).

The golden standard of sleep monitoring is *Polysomnography* (PSG). In this procedure, clinicians record all vital signs that will subsequently be analyzed by technicians for sleep classification. Signals recorded during a common PSG study are: EEG, ECG, EOG, EMG, respiration (thoracic, abdominal and airflow), oxygen saturation (SaO2) and carbon dioxide saturation (SaCO2). Audio and video recordings are also common in this type of studies.

Until now, sleep researchers can be divided into two main groups. The first group focuses on the sleep state's differences of behavioral patterns, such as facial, eyelid and full body movements, heart rate patterns, and respiration. The second group of researchers rely on EEG patterns and changes to identify each sleep state. Since the aim of the present work was to find a suitable sleep scoring system that avoids electrode-patient contact, focus on monitoring behavioral patterns was the main priority. Moreover, as the AASM states [26], *"regularity (or irregularity) of respiration is the single most useful PSG characteristic for scoring sleep states at this age"*, referring to sleep scoring methods for infants of 0 to 2 months of age.

In terms of behavioral patterns, there are a few notorious differences between REM and Non-REM (NREM) sleep states. The REM phase owes its name to the characteristic Rapid Eye Movements (REMs), which only occur during this phase of sleep. In REM state, there is an increase in sympathetic activation relative to non-REM sleep [27] in both adults and children. Hence, the irregular patterns of heart rate, respiration and respiration rate are characteristic of the REM phase [26, 28]. Apneas[1] caused due to sleep disorders normally happen during this state of sleep [29]. The eyes are closed but might occasionally open [30]. Particularly for preterm or very young full-term infants, a wide range of motor activity can be found sporadically occurring in bursts of 5-60 s [31–33], with low muscle tone between these movements. Also in these young patients, specific facial expressions can be observed during REM, including frowns, smiles and grimaces, and random bursts of sucking movements as well as small sporadic twitches [34]. Given the wide range of movements occurring during this phase, REM can be equally named "Active Sleep" according to Anders criteria. These terms can only be interchanged for infants of 0 to 3 months of age [25] .

In contrast to REM, NREM is characterized by more regular sleep patterns. Both respiration and respiration rate are quite regular [26], [28], with very little or no apneas in between respirations. Also the heart rate normally stays very steady in this sleep phase, adopting rather lower values. Little or no motor activity is appreciable [31], only appearing as occasional brief agitations or sighs. No REMs can be detected in this phase. Muscles are governed by a tonic motor level [35] and mouth movements or sucking can be observed [25]. The steady state of most of the body vital signs during this state gave rise to the term "Quiet Sleep" used to define this phase by Anders criteria [25]. Table 2.1 summarizes these huge differences of behavioral patterns between REM and NREM sleep, excluding parameters that are not interesting for this work such as facial movements or EEG patterns.

Table 2.1: Most characteristic behavioral patterns of REM and NREM sleep for infants of 0 to 3 months of age.

	REM	**NREM**
Heart Rate	High, Irregular	Lower, Irregular
Breathing Pattern	Irregular	Regular
Motion	Sporadic bursts of motion	No movements
REMs	Yes	No
Apneas	Relatively frequent	Absent
Anders criterion	Active Sleep	Quiet Sleep

From the previously mentioned sleep patterns of REM and NREM sleep, it is important to mention that motion patterns do not normally apply to patients older than 3 months post-term. This is something important to take into account since the present work has been performed with patients of ages ranging between 2 months and 12 years.

[1] *Apneas* are defined as transient absence of breathing.

The distribution of sleep stages changes over the course of the infancy due to the developed maturity in sleep patterns. The appearance of differentiated behavioral patterns in REM and NREM sleep states can be seen as early as 25-30 weeks [26] gestational age (GA), when EEG is still undifferentiated [36]. Recognizable EEG patterns of REM vs. NREM can be seen from 32-34 weeks GA.

The change in total sleep time (TST) associated with neonatal development occurs at the same time as the change in sleep organization [37]. In the initial phases of sleep development, REM sleep governs most of the sleep cycle [38]. Due to the immaturity of sleep development in these early stages, it is common to find very diffusive patterns with characteristics of both REM and NREM sleep, which would be defined as Transitional Sleep (TS). From the period when infant approaches term until the early weeks of life, REM and NREM sleep distribution becomes nearly equal. By the last months of the infant's first year, NREM sleep occupies about 80% of sleep time [16], without barely any trace of TS. This change in sleep states' organization is a crucial manifestation of the maturation and evolution of the brain development process. Classifying each sleep state and determining its proportion within TST would allow assessment whether the developmental process is following the anticipated path or not.

2.2. PREVIOUS WORK ON SLEEP-STATE CLASSIFICATION

In the past decades, a vast amount of clinical research has been focused on achieving a better understanding of the biological processes happening in the human body during sleep, giving rise to a whole new research branch identified as Sleep Research. With the increased awareness of the critical part of REM in humans, and more importantly in the neonatal sleep cycle, a considerable amount of resources have been allocated to the specific task of sleep-staging using automatic methods. While some researches first focused on using the whole set of PSG vital signs to automatic the sleep-staging process, others aimed at first identifying the best vital parameters for this task with the end goal of decreasing invasiveness. In this section, some of the work regarding automatic sleep-staging will be presented by dividing it into developed methods for sleep-staging in adults and the ones only focused on young infants. Specifically, previous work regarding studies in neonates is summarized in Table 2.2. Rigorous comparisons between the reported classification accuracies of these studies cannot be made since they differ in recording conditions, validation procedures, patient age and sleep states classification nomenclature. Moreover, some studies were developed using healthy subjects, subjects with sleep disorders or both together.

In every sleep classification study, total sleep time is divided into specif time windows with fixed duration, each of them constituting an *epoch*. The duration of the epochs is an important parameter to take into account regarding the accuracy of the sleep-state classification. The smaller this epoch size, the less time the scorer has to recognize patterns between sleep states. According to the AASM, the optimal epoch size to classify sleep events is 30 seconds [23].

Automatic sleep-staging methods have been widely developed with adult subjects. The majority of studies have been focused on the analysis of EEG spectral features with a favorable average accuracy of up to 96% [39]. Koley et al [39], Shimada et al [40], Kaplan et al [41], Ebrahimi et al [42] and Gudmumndsson et al [43] are just some of the examples of the extensive literature of EEG-based sleep-state classification. Another great part of sleep-staging studies focused on the analysis of EEG along with other PSG electrical signals, i.e., EMG and EOG signals [44–47]. Khalighi et al [46] reached a classification accuracy as high as 87%. The analyses of EOG signal alone has also reported good classification results with an agreement level of around 72% [48]. The use of pneumatic sensors located under the bed of the patients has been quite common in sleep research for the extraction of

cardiorespiratory and body motion signals. Watanabe et al [49], Migliorini et al [50] and Kortelainen et al [51] are good examples of these methods, with an accuracy ranging from 38 to 82% by Watanabe et al [49]. A very common ECG feature called Heart Rate Variability (HRV), has also been extensively used for addressing sleep-staging. The interest behind using HRV for sleep classification relies on the concept that HRV provides a good evaluation of the autonomous nervous system state level, which increases during REM state. Ebrahimi et al [52] only evaluated HRV to classify sleep states with an accuracy ranging between 80-86%. To conclude, three other sleep-classification approaches were found in which only respiratory information was used [53–55], among which Chung et al [53] reached an average accuracy as high as 89%.

Classification algorithms vary widely among the previously mentioned studies. Some of the most popular ones in which EEG features are involved were neural networks [40, 42] and Support Vector Machines (SVM) [43, 46]. For cardiorespiratory features-based classification methods, approaches such as Linear and Quadratic Discrimination [50, 52], Hidden Markov model (HMM) [51], respiration variance- thresholding [53] and dissimilarity measures [55] were used. Robert et al [56] provided a broad review of the use of neural networks in sleep-staging classification methods.

The extensive research on sleep-state classification on adults has lead to a wide range of commercial solutions, which are described by Kelly et al [57]. However most of them lack validation studies and have not been tested in young infants or children. As explained in Kelly et al[57], quite a big limitation of most of these products is the overall tendency of misclassifying REM sleep as Awake state or viceversa, due to the assumption that REM sleep is characterized by irregular breathing and higher amount of movements.

A novel radar sensor called SleepMinder was developed to classify sleep-awake states in adults [58]. The same sensor was used to detect apnea events in adults [59] with a high accuracy level. This radar sensor was afterwards developed as a commercial product called *ResMed S+* (http://sleep.mysplus.com/). From the research provided using this sensor, one can observe that motion is the only parameter used to perform the sleep-classification, with an average accuracy of around 80%. The use of this commercial sensor for the present work was not possible since it is not possible to access the raw data of the sensor, hence only real time studies can be performed.

Regarding automatized sleep-classification research on neonates, no such extensive work has been done in comparison with the one on adults. The work found during this literature study can be summarized in Table 2.2. None of these infant subjects were older than 1 year old.

Several studies only employed EEG features for the sleep classification [63, 64, 69–71]. Among them, only Piryantiska et al [63] and Čić et al [64] are mentioned in Table 2.2 since EEG is not the monitoring parameter this present work is focused on. Čić et al [64] reached a classification accuracy of 80 and 90% by using an SVM classifier. EEG, EOG and EMG features have also been used for sleep state classification by Pfurtscheller et al [61] and Estevez et al [62]. The latter combined EEG, EOG and EMG with body movements information, but both studies provided average accuracies of around 60-80%. Scher et al [60] used all possible available PSG signals to perform linear discrimination, reaching an accuracy of 90-97.4%. Harper et al[65] also performed linear discriminant analyses to address sleep-classification by means of cardiorespiratory measures, with an overall mean accuracy of 84.8%. They also studied the classification performance using only cardiac and respiratory features with accuracies of 82% and 80%, respectively. To finalize, only respiration-based classification methods have been developed by Haddad et al [66], Isler et al [67] and Terrill et al [68]. Algorithms used for the classification relied on a decision-based system [66], thresholding-rules [67] based on respiration rate variance, and linear discriminant analysis [68], with a maximum accuracy of 93% reached by Haddad et al [66]. It is important to observe that classification approaches using only

Table 2.2: State-of-the-art Sleep-Classification Research performed on neonate subjects.

Scoring Parameters	Classification Classes	Classification Process	Quality Evaluation	Author(s), year
All PSG channels	Awake/AS/QS	Linear Discrimination	Accuracy Range: 90-97.4%	Scher et al, 1996 [60]
EMG EEG EOG	Movements/AS/ QS/IS/Awake	Neural Networks	Accuracy Range: 65-80%	Pfurtscheller et al, 1992 [61]
EMG EEG EOG Body Movements	Awake/REM/ NREM1/NREM2/ NREM3/NREM4	Own algorithm based on thresholding	Accuracy Range: 60-80%	Estevez et al, 2002 [62]
EEG	AS/QS	Cluster Analysis	Accuracy Range: 70-90%	Piryatinska et al, 2009 [63]
	REM/NREM	Support Vector Machines (two examinees)	Mean Accuracy: 80%- 90%	Čić et al 2013 [64]
ECG Respiration	Awake/REM/ QS	Linear Discrimination	Mean Accuracy: 84.8 %	Harper et al 1987 [65]
Respiration	REM/QS	Decision-rule based on coefficients of variation	Mean Accuracy : 93%	Haddad et al, 1987 [66]
	AS/QS	Thresholding based on instantaneous breathing rate variance	Concordance Range: 80-87%	Isler et al, 2016 [67]
	Awake/REM/ NREM	Linear Discrimination	Mean Agreement Range: 79.2 - 84.9%	Terril et al, 2012 [68]

respiratory features used relatively simple algorithms.

Static Charge Sensitive Bed (SCSB) methods have been used in the 90s in a couple of studies [72, 73] to extract ballitocardiogram, respiration and movement signals in neonates subjects from sensors placed under the bed. They showed an accuracy for sleep-classification as high as 85% [73]. However, these two studies were not performed using a computer-based classification. Instead, examinees did a manual classification based on visual hints from SCSB signals. SCSB technology will be further explained in the next chapter.

Most of the previously mentioned studies used as reference the sleep classification performed by a technician during the PSG studies, which is indeed the golden standard of sleep classification. However, it is important to take a closer look into differences in the performance of sleep-staging algorithms when different technicians are scoring the sleep states. Čić et al [64] showed that changes in average accuracy of sleep classification using two different technicians to score sleep were as high as 10%.

2.3. CONCLUSIONS ON SLEEP-CLASSIFICATION LITERATURE REVIEW

Sleep patterns for young infants have been discussed in this section. Both EEG and behavioral patterns can be used to differentiate sleep states in adults and infants. Adults have particularly been the most common population of sleep research when developing automatic, computer-based approaches for sleep classification.

Literature on previous work on computer based sleep-state classification has been thoroughly analyzed for this literature review, getting into a more detailed analysis of literature using young infant populations. However, a strict comparison between studies presents some difficulties due to the fact that they present a great heterogeneity in terms of sleep states under analysis, amount of sleep recordings, classification algorithm and study population, among other differences.

Sleep-classification studies discussed in this section used a wide range of vital sign parameters, such as EEG, EMG, EOG, body movements, HR and breathing patterns. As observed in Table 2.2, accuracy seems to not be affected by the chosen physiological parameter under analysis, since most of those studies had a relatively good accuracy in the classification.

As established by the AASM [26], behavioral patterns, specially regularity of respiration, can better determine sleep state in neonates within the first three months of life. This and the fact that that respiration is a physiological parameter that can be acquired through contactless methods relatively more easily compared to EEG signals, lead the researchers of the present work to focus on respiration as the main physiological parameter to monitor sleep states.

In conclusion, this literature analysis has presented the past usage of behavioural patterns such as respiratory, cardiac and motion features, to achieve an automatic classification method for sleep monitoring. Their low invasiveness and the possibility to be extracted without high disturbance of the sleep make them attractive parameters for classifying sleep phases. However, finding an unobtrusive method to extract these sleep behavioral patterns has been another important and big dilemma of sleep monitoring. In the next chapter, several contactless techniques for vital sign monitoring will be presented due to their potential in sleep monitoring application in preterm babies.

3

Contactless Methods for Vital Signs Detection

Just as many other organs of premature neonates, the skin needs a long time to fully develop. One of the main concerns hospital personnel have to deal with in the NICU, is the distinctive thin skin of preterm neonates which is extremely sensitive to the continuous use of electrodes and plays a huge role in their undeveloped thermoregulatory systems. Even the smallest electrodes used to monitor these young patients can cause deep damage to their fragile skin (Figure 3.1). Thus, finding a contactless system to monitor the vital signs of this young population of infants is the highest priority when trying to create a sleep monitoring system for them. Just as with sleep research, contactless monitoring approaches have been studied much more in detailed in adult applications. The main contactless vital signs monitoring methods can be divided into three major classes: ballistocardiography, optical and radar approaches. Their working principles along with some previous work using these techniques, are defined in this chapter. Specific details will be given for radar methods, as the present work is based on this specific approach.

Figure 3.1: Photo illustrating the irritation effect on the skin of a preterm neonate due to ECG electrodes. Picture taken from [1].

3.1. BALLISTOCARDIOGRAPHY

Ballistocardiography (BCG) techniques measure the slight movements of the body caused by the ejection of the blood at each cardiac cycle [74], i.e., the ballistic forces of the heart. This mechanical signal of heart movements is often retrieved from small sensors placed in the mattress or seat in which the patient is lying. To extract a BCG signal, chest movements during respiration must be often removed from the total mechanical signal retrieved. For this reason, BCG techniques are generally used to extract breathing information as well, as shown in several studies [75, 76]. Zao et al [76] successfully applied BCG in sleep apnea syndrome detection. Also Austin et al [77] made use of it as a sleep-awake classification method. Some commercial products for adult sleep monitoring using BCG technology have arisen in the past few years, such as the *Beddit Sleep Monitor* (www.beddit.com) from *Apple*.

Atallah et al [2] used this BCG principle to unobtrusively monitor cardiac activity signals from capacitive electrodes placed in the small mattress of the preterm baby incubator (Figure 3.2). In another study [78], capacitive electrodes are placed bellow the wheels of the baby's incubator to monitor both HR and breathing rate. This study shows very good results in the NICU environment, reaching correlations of 0.74 and 0.91 between breathing and HR signals with their respective reference signals.

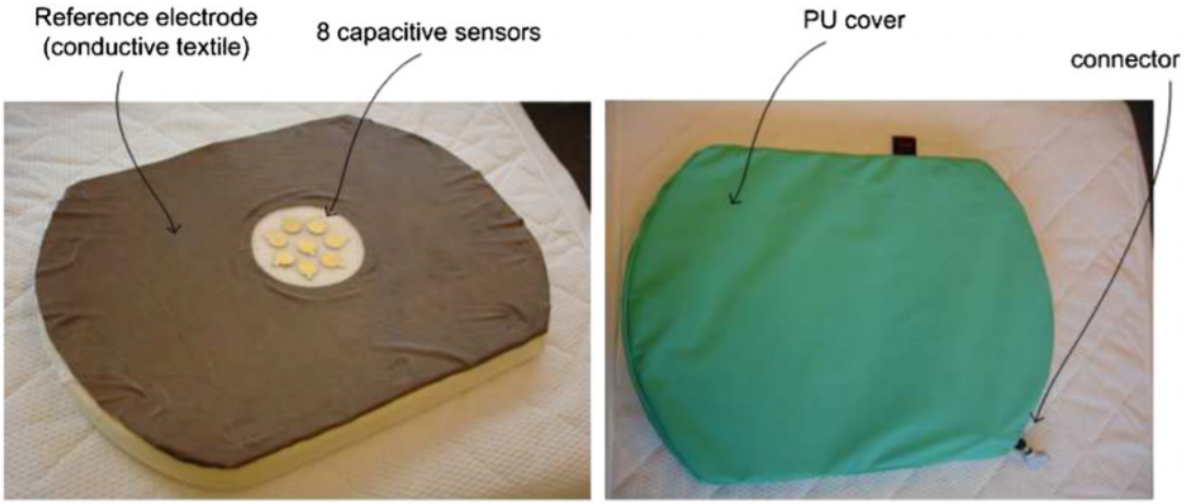


Figure 3.2: Neonatal mattress with integrated capacitive electrodes used to extract BCG signals at NICU, covered by Polyurethane (PU). Described in [2].

As mentioned in the previous section, approaches using Static Charge Sensitive Bed (SCSB) have shown successful results in studies developed in the 90s [72, 73] as a sleep monitoring method. SCSB can be used to acquire BCG signals. This technology was developed in the late 70s [79]. It consisted of two sheets of material with different dielectric constants, enclosed within the mattress [80]. The scrubbing movement between these sheets against each other caused by respiration and heart rate movements generates an electric charge that is picked up and analyzed to extract the corresponding vital signs of the patient lying on top of the mattress. The device requires neither an external power source nor the need to be recharged, as it only relies of these electrostatic charges generated by the sensor embedded in the mattress.

There are two main limitations of applying BCG technology while monitoring preterm babies: firstly, the particularly small signals acquired in these small patients due to the weak forces exerted by their

heart and chest at these early stages of life make it difficult for these signals to be detected; and secondly, the high sensitivity of this technology to other natural body movements of the patient.

3.2. IMAGING APPROACHES

Another big cluster of contactless vital signs monitoring is comprised of imaging approaches, mainly differentiated into conventional video-based and infrared (IR) thermal approaches.

In the first place, video-based approaches capture normal ambient light images using videocameras in order to detect the slight motions produced by the respiration and cardiac activity. Eulerian Video Magnification [81] is one of the most outstanding and relatively recent techniques. It uses input video frames to magnify both motion caused by respiration and changes in skin color due to blood perfusion. This information is hidden to the human eye at first instance, but can be extracted from video frames by means of spatial decomposition. The system first decomposes the input video into different spatial frequency bands and applies the same temporal bandpass filter to each of them (cutoff frequencies adapted according to application, 0.5-2 Hz for respiration and heart rate), which are then amplified by a given factor and added back to the original signal. This process gives rise to a final output video where the magnified motions can be now visually detected. Wu et al. have a very interesting website (http://people.csail.mit.edu/mrub/evm/) which provides some videos and examples of this technique. Furthermore, some open code is available for any user to employ a normal webcam to detect breathing and HR. Some of their example videos include very young babies.

Besides Eulerian Video Magnification, other video techniques have reported good results in detecting heart and breathing rate information using adult patients [1, 82–84]. Some other studies have further applied their video-based vital sign monitoring to populations of premature babies [85, 86] with very high correlations (0.87- 0.99) with the reference breathing rate reached by Koolen et al [86].

Despite of the great accuracy of video-based methods for both respiration and heart rate detection, it is important to notice that the efficiency of these methods is highly dependent on ambient illumination of the patient. However, when monitoring preterm babies inside the incubator, the light is dimmed in order not to damage the eyes of the baby and to try simulating the womb of the mother. Furthermore, they are often treated with Ultraviolet-phototherapy in order to lower their bilirubin levels and prevent jaundice [87]. Both these two situations highly affect the image quality of video recordings, making cameras unable to extract the vital signs of the patient.

Infrared (IR) thermal imaging looks like a good alternative to the previous standard video-based technique in the sense of preterm babies monitoring. It relies on the change of temperature phenomena that occur in the areas of the face proximal to the nose or mouth orifices during periods of inhalation and exhalation. Several studies provided good results when measuring breathing rate on adults using standard IR imaging [88–90]. The work of Abbas et al [3], is specially outstanding, since it applied this technology in NICU conditions. They show the capability of the system to track these slight temperature differences around the mouth of the baby as shown in Figure 3.3, even though there is some background noise due to the temperature fluctuations inside the incubator of the patient. Some further studies must be performed in order to improve this technique.

Besides of the need to further investigate IR thermal imaging within premature populations in order to improve breathing detection and provide heart rate data, there are some other limitations of this

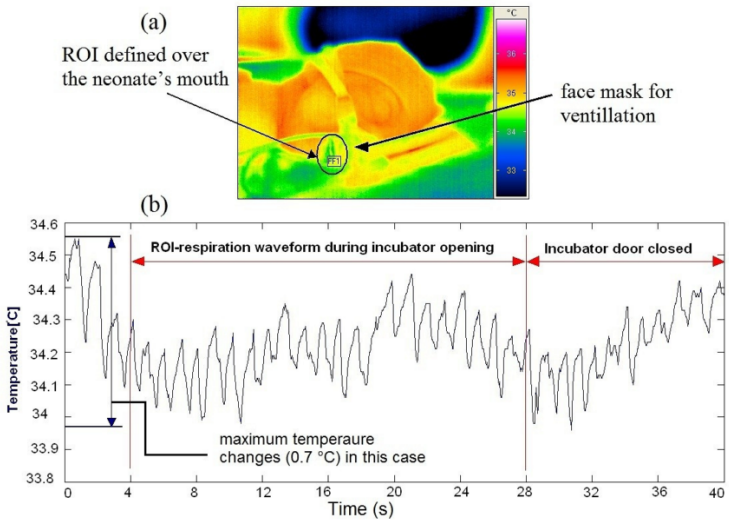


Figure 3.3: Respiration detection of a patient inside the incubator by means of IR thermal imaging. From [3]. ROI: Region Of Interest.

novel technique. The high cost of the thermal cameras used in these studies makes it more difficult for the technique to be adopted as a patient monitoring system. Furthermore, both IR and normal video-based approaches normally involve a quite high computational cost due to the advanced algorithms performed to detect vital signs, especially in IR imaging where the algorithm needs to detect the face as a region of interest first. Finally, a big limitation of both techniques is the privacy issue that may arise because of the patients' recordings.

3.3. RADAR APPROACHES

Radar, an acronym for *RA*dio *D*etection *A*nd *R*anging, is a system that emits electromagnetic (EM) waves, more specifically radiofrequency (RF) signals, from a transmitter (TX) and studies their corresponding reflections, or *echos*, in a receiver (RX). This is done in order to measure certain characteristics of specific targets, such as range[1], height, velocity or angular direction. Nowadays, radar systems exist for a variety of applications from weather observation to military and patient monitoring applications. The recent application of radar systems in patient monitoring translates the position of patient within a time interval into body motion, respiration rate and even heart rate.

The concept behind using RF signals for medical applications is to make use of the echo reflections at the tissue boundaries of the person under study. When RF signals are transmitted towards a patient, the majority of reflections (around 70% for 5 GHz UWB radar signals [91]) are produced due to the different dielectric constants between air and skin. Another contribution to the reflected echo energy comes from the part of the signal penetrating the skin that has been reflected at a next tissue boundary. These echo signals reflected by the human body are used to detect the overall movement of the chest. In this way, one can extract breathing patterns and even HR due to the small contribution of cardiac activity to the whole chest motion (about 10 times smaller than breathing respiration). Figure 3.4 shows the schematic of a radar module used for baby monitoring.

The time-domain characteristics of the transmitted EM wave defining its waveform are of great im-

[1] In radar terms, *range* is referred to the distance between radar antenna and reflecting object (or target).

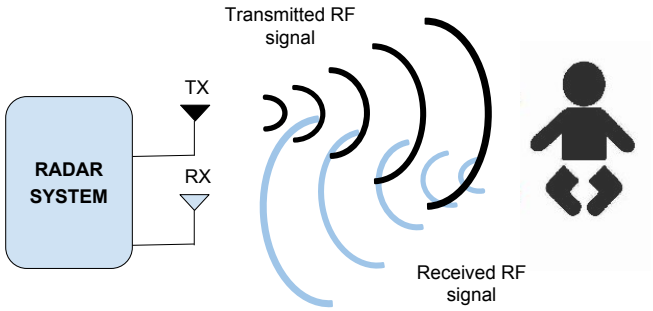


Figure 3.4: Schematics of a radar system applied to baby monitoring. Dimensions are not in scale.

portance. A common way of classifying radar systems is according to this waveform, as shown in Figure 3.5. Radar systems can make use of either a continuous wave or a pulse waveform depending on whether the RF signal is continuously transmitted or not. These most common modalities are further discussed in the following subsections in order to provide the reader with a greater background of radar technologies used in short-range applications.

Regarding the characteristics of radars within frequency domain, most traditional radar systems are characterized by working in a narrow frequency band. However there has been an increasing interest in radar technologies operating in wider frequency bands. The radar module used for this project is an *Ultra-Wideband* module implemented using a pulse Doppler architecture. The basics of this radar modality will be further explained in this section, however the specifications of the radar sensor along with some safety concerns will be analyzed in Chapter 4.

3.3.1. CLASSIFICATION BY WAVEFORM

As previously mentioned, there are two big radar modalities according to the time-domain characteristics of the waveform of a radar system as shown in Figure 3.5.

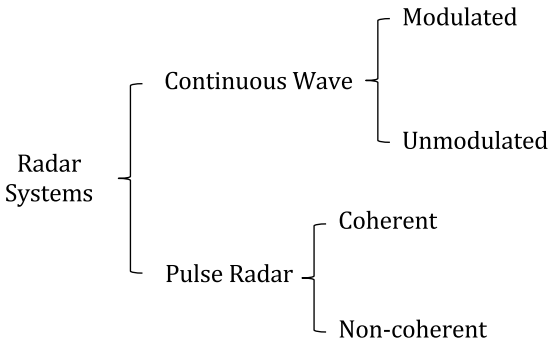


Figure 3.5: Radar classification according to time-domain waveform.

CONTINUOUS WAVE RADAR (DOPPLER RADAR)

As its name suggests, *Continuous Wave* (CW) radar broadcasts a continuous RF sinusoid that is both transmitted and received at the same time, thus transmitter and receiver are always switched on simultaneously. This radar modality makes use of the *Doppler effect*, by which a non-stationary

target moving towards the radar system causes a frequency shift between transmitted and received signal, which is employed to infer the velocity of the target. For this reason, this kind of radar is often referred to as *Doppler radar*. The relation between this frequency shift, referred to as the Doppler frequency f_D, and the radial velocity v of a target is given by Equation 3.1:

$$f_D = f_r - f_t = \frac{2vf_t}{c}, \tag{3.1}$$

where c is the speed of light in free space. It is important to notice that the radial velocity v is pointing towards the reference point, or in this case the radar receiver. v will be positive when the target is moving towards the reference point, giving rise to a positive f_D in the received signal, and viceversa when the target moves away from the reference [92].

In its basic form, unmodulated CW radar is only capable of detecting moving objects, as stationary targets will not cause any frequency shift. This modality may involve a very cheap hardware design, however it does not provide information about the range and it cannot differentiate between more than one target. Unmodulated CW radar has been widely applied for motion detection systems, presence alarms and speed gauges for traffic control.

In order to be able to determine range, a CW radar system should undergo either a frequency or phase modulation. In particular, the so-called *Frequency Modulated Continuous Wave* (FMCW) radar has gained popularity in the recent years, as its latest hardware advances have allowed this radar modality to determine both velocity and range of targets with an extremely high accuracy.

In FMCW radar, a continuous signal is transmitted with a specific reference frequency f_1 that is linearly increased up to a frequency f_2 periodically [93]. Meanwhile, the echo signal is received with a time delay Δt that directly translates to range R by Equation 3.2:

$$R = \frac{c|\Delta t|}{2} = \frac{2c|\Delta f|}{2(\partial f/\partial t)} \tag{3.2}$$

Just as with unmodulated CW, if the target object has a radial speed towards the radar antenna, a Doppler frequency f_D can be observed between the transmitted and received signal. Figure 3.6 describes the operation principle of a FMCW radar.

Even if CW radar was not originally designed to determine range but only velocities, FMCW is able to resolve both range and velocity of targets with a very high range accuracy [93].

PULSE RADAR

In pulse radar, the transmitter emits high powerful and short pulses received with a time delay Δt that can be translated to range as previously shown in Equation 3.2. A pulse radar waveform is characterized by parameters such as pulse repetition frequency PRF, inter-pulse interval IPI, carrier frequency f_c and pulse-width τ. Figure 3.7a represents the basic parameters of a pulse radar waveform.

Contrary to CW radar, the system transmitter and receiver are never on simultaneously. When a pulse is transmitted during a time τ, the receiver is off. As soon as this pulse ends, the transmitter goes off and the receiver automatically switches on, and so on and so forth. Figure 3.7brepresents this working principle. Having a receiver switched on and off periodically involves some limitations regarding the minimum and maximum ranges the radar can detect.

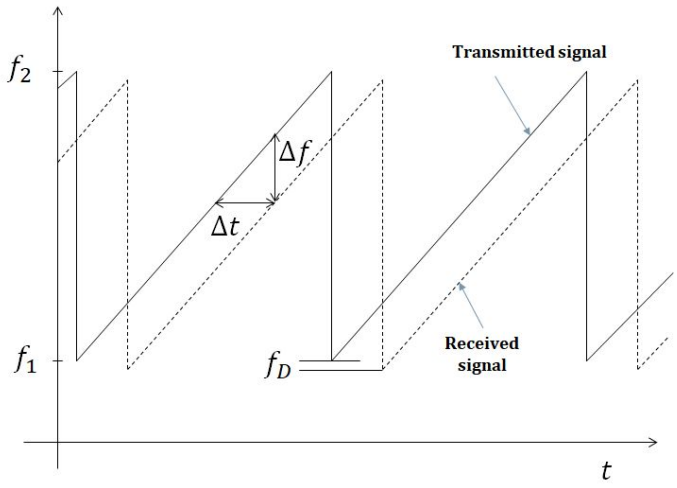
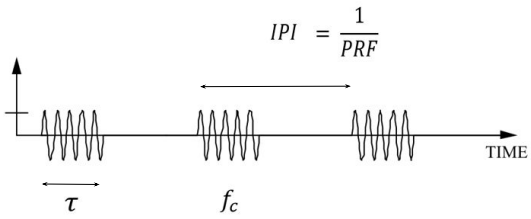
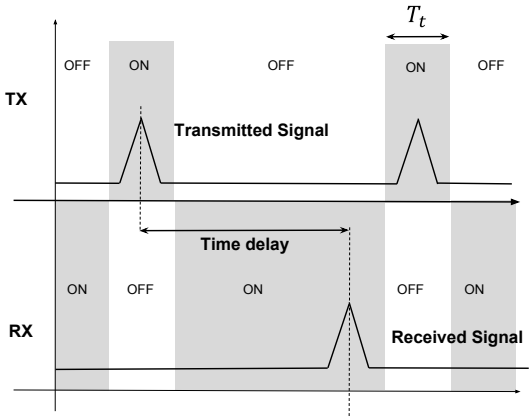


Figure 3.6: FMCW radar working principle of frequency modulation.



(a) Pulse radar waveform parameters.



(b) Working principle pulse radar transmitter and receiver. For simplicity, pulses are represented with triangles. T_t is the tranmission time.

Figure 3.7: Pulse radar schematics.

On the one hand, during the time T_t, called *transmission time* (Figure 3.7b), in which the transmitter is emitting the pulse the radar cannot receive; a "blind range" in the radar is given rise to. The target must be located further than this minimum detectable range in order to be detected. This minimum range R_{min} is given by Equation 3.3:

$$R_{min} = \frac{c \cdot T_t}{2} \tag{3.3}$$

Moreover, there is another concern to take into account in pulse radar systems: the so-called range ambiguity. Since this radar assumes an echo pulse will be received within a specific time window before the next pulse is emitted, the target must be located within a specific range in order to be detected by the receiver. This is the so-called maximum unambiguous range, and is given by Equation 3.4:

$$R_{max} = \frac{c \cdot IPI}{2} = \frac{c}{2 \cdot PRF} \tag{3.4}$$

Range resolution in short-range[2] radar systems is a very important parameter. It represents how far apart two targets should be in range with the radar, so they can be resolved as two different targets. In pulse radars, range resolution highly depends on the pulse width. If two targets are too close to each other, their two corresponding echos will superimpose and both targets will be detected as one object (Figure 3.8). Pulse radars range resolution is expressed by Equation 3.5:

$$\Delta R = \frac{c\tau}{2} \tag{3.5}$$

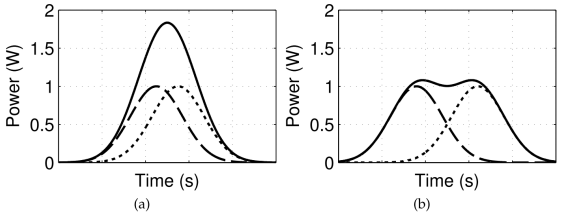


Figure 3.8: Schematics of range resolution resolvability in two scenarios. In (a), two pulses are superimposed when two targets are too close to each other, being non-resolvable. In (b), two peaks can be detected from two separated targets [4].

Pulse radars can be subdivided into coherent and non-coherent systems. In non-coherent systems, each pulse is transmitted with a random phase, whereas pulses in coherent pulse radars share a common phase. This allows for coherent receivers with information about the phase variation in the received signal, that is phase modulated caused by motion of the target. This phenomena provides extremely high precision to coherent systems. Phase variation can be used to detect some really small target movements, according to Equation 3.6:

$$\Delta\theta = \frac{4\pi\Delta x}{\lambda} = \frac{4\pi\Delta x f_c}{c}, \tag{3.6}$$

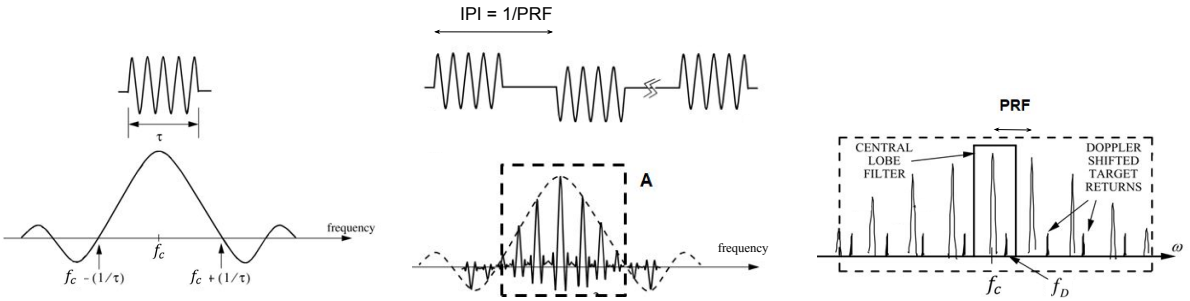
$\Delta\theta$ being the phase shift between emitted and transmitted pulses, Δx the displacement of the target and λ the radar wavelength that can be calculated from the pulse carrier frequency f_c and speed of light ($\lambda = \frac{c}{f_c}$). This phase modulation concept is also exploited by some short-range CW systems for vital signs detection.

Pulse radar is not only useful to acquire an accurate range of the target but also to determine its velocity. The simplest versions of pulse radars calculate this velocity by simply detecting a change in range within a specific time window. However, when dealing with high velocities of several targets this method offers some limitations that can be resolved with a modality of pulse radars, the so-called *pulse Doppler radar*.

Pulse Doppler radar systems make use of both (i) the pulse radar principle to resolve the target range by means of the time between emitted and transmitted pulse and (ii) the Doppler radar principle to resolve the velocity of the target by relying on the Doppler effect.

In pulse Doppler radars, the Doppler frequency cannot be as easily obtained from the frequency spectrum of the echo signal as with Doppler radar. If we start form the Fourier transform of a rect-

[2] In this work we will refer to short-range radar as radar systems with a maximum range of less than 10 m.



(a) Fourier transform of a single pulse.

(b) Fourier transform of a train of pulses.

(c) Zoom area A from **(b)**.

Figure 3.9: Frequency spectrum schematics of pulse Doppler radar adapted from [5].

angular pulse, we would obtain a sinc function ($sinc(t) = \frac{sin(t)}{t}$) centered at the carrier frequency of the pulse f_c with a frequency bandwidth inversely proportional to the pulse width τ (Figure 3.9a). However, when we have a train of pulses as with pulse Doppler radar, the resulting Fourier transform consists of several sinc functions under the envelope of the previous sinc function (Figure 3.9b). If we zoom in onto the central lobe (Figure 3.9c), it can be observed that the resulting spectrum of the train of pulses are sinc functions whose peaks are a distance equal to 1/PRF between each other. Furthermore, small peaks coming from the echo signals appear at the Doppler frequency f_D. By filtering this central lobe these Doppler frequencies can be extracted and subsequently used to calculate the velocity of the target. It is important to take into account that the maximum Doppler frequency $f_{D,max}$ that can be detected is equal to half of the PRF. This frequency is known as the maximum unambiguous Doppler frequency [5].

In pulse Doppler radars, both maximum unambiguous range R_{max} and maximum unambiguous Doppler frequency $f_{D,max}$ depend on PRF. By increasing PRF, R_{max} would decrease and $f_{D,max}$ increase, and the other way round for a decreased PRF. Depending on the application, PRF is designed to correct for both range and frequency ambiguities. Furthermore, there has been an increasing number of new techniques to resolve these limitations.

FMCW and pulse Doppler radars are the two most common architectures when designing short-range radars.

In short-range radar applications, the minimum detectable range is highly important. As previously mentioned, FMCW radar does not have this limitation since the receiver is always on. Hence, only pulse (Doppler) radar systems need to deal with this limitation that depends on the transmission time. Most recent pulse Doppler radar systems make use of impulse waveforms with extremely short transmission times that make it possible to drastically decrease the minimum detectable range by increasing range resolution simultaneously.

In terms of maximum range, pulse radars are also limited to a specific maximum range that may vary by changing the PRF value depending on the application. Similarly, FMCW radar also has a maximum range that depends on the periodic increment of frequency. Nevertheless, having a low maximum range when working in short-range applications, such as vital signs monitoring, is not an important factor to take into account.

Generally, hardware of CW radars is easier to implement than with pulse radars since CW are always uninterruptedly emitting. However, CW radars face a significant challenge when isolating transmit-

ter (TX) and receiver (RX), as the power ratio can reach values on the order of $P_{TX} : P_{RX} = 10^9$, decreasing the probability of detection [4]. Since pulse radar always have either TX or RX on at a time, it is much simpler to detect an echo signal at the expense of increased hardware and signal complexity. Circuit complexity of FMCW is much higher than simple CW radar due to the implementation of frequency modulation.

In terms of power consumption, on-off mode of transmitter and receiver make pulse radars in general much less power demanding than CW radar systems, since they are uninterruptedly emitting RF signals [94].

Finally, even though FMCW radar systems are capable of resolving both the velocity and location of a target, pulse Doppler radars are considered to resolve the location of two different targets much easier than FMCW radars [94].

3.3.2. FREQUENCY-DOMAIN RADAR CHARACTERISTICS

Frequency-domain characteristics of a radar system are also important to consider. Fractional bandwidth BW_f is a measure of the relative width of a radar and is given by Equation 3.7:

$$BW_f(\%) = \frac{f_h - f_l}{f_c} \cdot 100, \tag{3.7}$$

f_h and f_l being the highest and lowest points of the frequency spectrum and f_c the center frequency, which is the carrier frequency for pulse radars. In radar systems, f_h and f_l are normally expressed at -10 dB gain, whereas other applications commonly use -3 dB.

Conventional radar systems, such as CW radars are considered to be narrow band, as their fractional bandwidth BW_f normally does not go further than 1%. In contrast, another radar type has gained more popularity over the recent decades, the so-called *Ultra-Wideband* radar.

ULTRA-WIDEBAND RADAR

Ultra-Wideband (UWB) radars are characterized by having a fractional bandwidth greater than 20%, or a minimum absolute BW of 500 MHz [4]. Figure 3.10 illustrates the frequency spectra of UWB radar in comparison with narrowband radar and thermal noise unintentionally emitted from home appliances. These radars are commonly referred to as Impulse Radio-UWB (IR-UWB) radar since they transmit extremely short pulses on the order of hundreds of picoseconds, giving rise to a frequency spectrum in the gigahertz scale [6], since pulse width is the inverse of bandwidth as follows:

$$\tau = \frac{1}{BW} \tag{3.8}$$

The impulse waveform of UWB radars provides them with really high range resolutions, as it increases with decreased pulse width as shown by Equation 3.5. Furthermore, UWB systems have distinctive penetrating capabilities due to their wide frequency spectrum. The emitted signals can go through clothing, plastic casings and even walls, which has allowed them to gain increasing popularity in military and biomedical applications [95–97].

In 2002, the U.S. Federal Communications Commission (FCC) regulation allocated a frequency band from 3.1 to 10.6 GHz as an unlicensed band for the general public, therefore most UWB sys-

tems operate within this frequency band. The FCC allows the coexistence of UWB radiation with present radio services with the condition that UWB systems operate at a maximum mean effective isotropic radiated power (EIRP) density level of 41.3 dBm/MHz, EIRP being the total radiation power emitted from the antenna [98].

Figure 3.10 also shows how low power density UWB frequencies do not interfere with other narrowband devices. In the same way, one may think environmental thermal noise prevents the echo pulses from UWB radar to be detected. However, in the output of the receiver a *coherent pulse integration* is performed over a certain time period, by which the UWB system increases the Signal-to-Noise Ratio (SNR) of the system, reaching high accuracy in the measurements. During this pulse integration, i.e. summation, random noise tends to be canceled out, thus reducing the noise power variance. In contrast, the interest signal will integrate constructively, improving the probability of detection and increasing SNR [99].

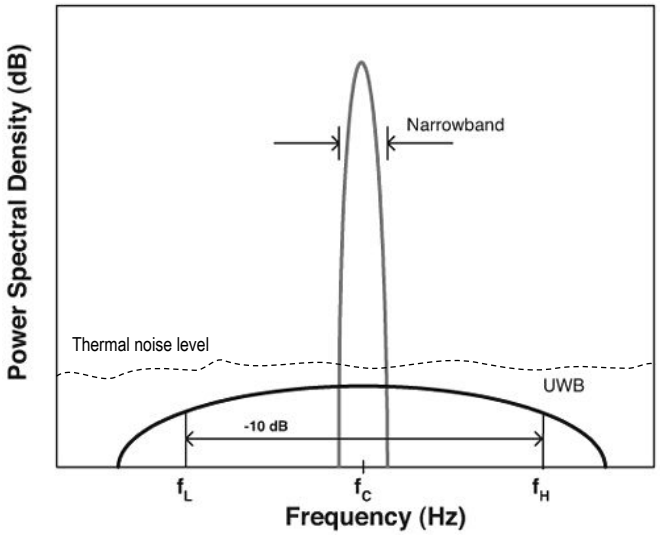


Figure 3.10: Photo illustrating the frequency spectrum of UWB compared to conventional narrow band radar and environmental thermal noise. Adapted from[6].

3.3.3. PREVIOUS WORK ON RADAR-BASED VITAL SIGNS MONITORING

Radar approaches have gained great popularity as contactless methods for patient monitoring and have been broadly used over the recent few decades. The research work is divided into two main radar architectures: CW and UWB radar [20–22, 91, 100–116]. As UWB are normally implemented by a pulse-doppler radar arquitecture, both radar systems CW and UWB rely on Doppler frequencies to resolve target velocities.

CW radar has been broadly used in the literature due to its simple hardware implementation and low development cost. However, UWB is well known to have better spatial resolution. Despite their differences, both radar systems can be used to detect the movement of the chest due to respiration and heart beats by measuring the Doppler shift from the echo signal (either pulse or continuous wave) [117].

CW radar was first used in 1972 by Johnson & Guy in order to sense physiologic movement [118]. This first original work was done with bulky and expensive components, with a limited research environment application. Recent advances in microfabrication and in wireless technology have enabled the fabrication of compact and low cost CW radar systems [119].

A high variety of studies have shown that CW radar has reached surprisingly high accuracy in respiration detection systems with adult subjects [100, 104, 106, 107]. Dei et al [104] compared the respiration waveform from their Doppler radar system and reference data from a spirometer, obtaining values of the coefficient of determination R^2 ranging between 0.95 and 0.98 .

Besides respiration, other researchers have employed CW radar to both respiration and HR detection [101–103, 105, 108–110], most of them using adult subjects as well. A noticeable piece of work was performed by Hafner et al [100], where they extracted both respiration and HR information with neonates reaching very low errors in both HR (maximum error of 1.6 beats per minute) and respiration (maximum error of 0.29 breaths per minute) measurements.

A similar technology based on the Doppler effect has been occasionally used to measure respiration [111] as well as HR [120], the so-called Laser Doppler Vibrometer (LDV). In this technique, a laser beam is directed to the chest of the patient, the echo signal is received by a detector that measures changes in phase shift to determine the instantaneous displacement. Marchionni et al [120] applied their approach to a preterm population showing very low error rates (maximum of 3% and 6% error for respiration and HR, respectively). The limitation of this method is that it is highly dependent on the position of the patient being measured so it would need a continuous location update system in order to be applicable at hospital conditions.

Using UWB radars as a measurement method for vital signs detection has quite some advantages compared to CW radars [96, 117] such as:

- Decreased EM radiation level due to low power spectral density of transmitted impulses.
- Increased spatial resolution due to higher time-domain resolution.
- Higher penetration capabilities.
- Ability to perform through-wall measurements.
- Better capability to differentiate between multiple targets.

UWB has been broadly used for respiration detection with adult subjects in multiple studies [112–114], and as respiration and heart rate detector [91, 115, 116] with subjects of similar characteristics. It is important to notice the existence of three studies in which UWB radar is applied to a neonate population [20–22]. Huan et al [22] uses the radar system as apnea detector and two other studies [20, 21] as respiration and HR detector. Ziganshin et al [20] and Huang et al [22] do not show accuracy or error analyses in their studies. On the other hand, Immoreev et al [21] show an agreement level of around 95-98% compared to the reference data. The UWB system used in this study was relatively bulky compared to nowadays compact, lightweight UWB radar systems on the market.

3.4. CONCLUSIONS

Premature babies are particularly sensitive to electrodes or any other measurement technique in contact with their thin, weak skin. For this reason, finding a contactless approach to monitor their vital signs is a priority in sleep research with preterm populations. In the previous chapter, the importance of breathing characteristics for sleep scoring has been shown, therefore monitoring this vital sign is the aim of the present work.

Imaging, BCG and radar approaches have been discussed in this chapter as potential contactless methods for neonatal sleep monitoring. Even though studies have reported good results with each

of these approaches, there are some limitations in imaging and BCG methods when applied in the NICU conditions, which may be overcome with radar approaches.

Imaging approaches based on normal video cameras may present some difficulties regarding image quality due to the dark conditions of the incubator and the UV treatments commonly used in neonates. Camera approaches are also highly sensitive to position of the patient and may involve a significant level of computer cost. Lastly, IR thermal imaging may also involve an overall higher cost due to its expensive equipment.

BCG approaches have been extensively used in adults subjects providing a relatively good sleep classification accuracy (up to 85% in Kirjavainen's et al work [73] using SCSB). However, this technique may give rise to large difficulties when being applied to such small patients in the NICU conditions. This technique is also extremely sensitive to motion noises of both patient and environment.

Radar approaches represented an excellent alternative as contactless vital signs monitoring method and potential technique for sleep monitoring. The principles and classification of different radar systems have been explained in this chapter. UWB and CW radar systems are the most common systems used in the literature regarding vital signs extraction. In contrast to CW, UWB radars have numerous advantages, such as high penetration, increased spatial resolution, low power consumption and radiation, among others.

UWB systems are commonly built following a pulse Doppler radar architecture. The radar module employed for the development of this project is an IR-UWB pulse Doppler radar. However, the Doppler effect was not part of the processing part of this system. In this work we use the high spatial resolution of UWB radar to measure changes in the position of the chest in time due to respiration, and not its actual Doppler velocity.

Previous research work has been done in the past using UWB radar on neonates [112–114]. However, lack of information regarding accuracy in respiration detection is encountered in various studies [20, 22]. In contrast Immoreev and Tao [21], show very good results (95-98% accuracy) with respect the golden standard, with no details about processing procedures for the respiration detection.

4

RADAR SENSORS

The radar modules used in the present work are part of the so-called *XeThru* ("See-through") technology developed by *Novelda AS*. These sensors are CMOS Impulse Radio Ultra Wideband (IR-UWB) sensors that make use of Pulse-Doppler processing. The use of these radar modules was a matter of convenience since the choice of the radar was not optimized. The main deciding factor to use this technology was the easiness of data acquition without a need of deep knowledge about radar circuitry and data processing. The sensors are commercially advertised as respiration detectors since they are ready-to-use sensors that are capable of tracking the presence of a person and provide breathing information. Furthermore, *Novelda AS* provides a great amount of documentation about the XeThru technology and its application in several case studies.

In this section, a general overview of data processing occurring in these radar modules will be given. Moreover, the key features of the two XeThru sensors employed in the present work will be provided.

Output signals of each sensor as well as specific software used to record these signals are further explained in more detail in Section 5.4 *Radar Data Acquisition*.

4.1. RADAR SIGNAL PROCESSING

In pulse-doppler systems, a local oscillator (LO) generates coherent pulses at the pulse repetition frequency (PRF). These are propagated through the environment until they face an object that will make them reflect in multiple directions, some of which being reflected back into the radar's receiver. Received pulses are phase modulated caused by motion according to Equation 3.6.

Each reflected pulse is then sampled. This sampling step occurs in between the transition of two pulses. In case of X2M200, the number of samples between two transmitted pulses is 256 for 1 meter of detection space [8], with an equivalent sampling rate of around 40 GHz. The set of samples for a specific time window is called the radar's *frame*. In this way, a matrix samples-frames can be built, having as columns the samples of one frame and as rows the set of frames acquired along a certain reading period (Figure 4.1).

The received RF signals, which have certain amplitude and phase, are normally represented in radar terms by means of two sinusoids that are offset by 90 degrees with respect to each other. These are the so-called *In-Phase* (I) and *Quadrature* (Q) signals, which are mixed with the received RF signal

	Sampler 1	**Sampler 2**	**...**	**Sampler N**
Frame 1 (t_1)	R (1, 1)	R (1, 2)	...	R (1, N)
Frame 2 (t_2)	R (2, 1)	R (2, 2)	...	R (2, N)
...	...	...	...	...
Frame M (t_M)	R (M, 1)	R (M, 2)	...	R (M, N)

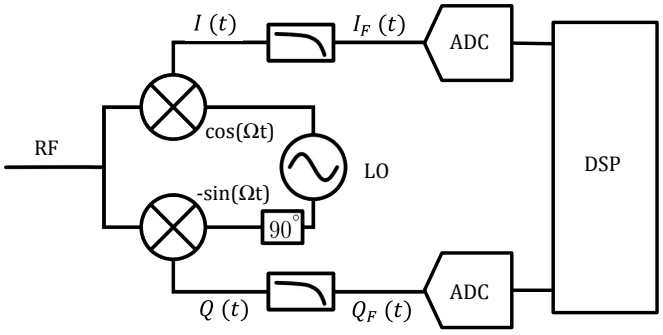
Figure 4.1: Schematics of radar sampling of N samples during a time window of M frames.

in a process commonly denominated I/Q demodulation [121]. In this process, the LO that generates I/Q brings down the frequencies of the received signal to baseband signals, after that, I and Q are filtered in order to suppress the remaining high frequencies (Figure 4.2a) [122]. This process in known as *baseband downconversion*. Baseband data is much easier to be processed in further blocks of the radar system. Another great advantage of I/Q demodulation is avoiding null points of the original A/P signal, since I and Q are never zero at the same time due to their 90 degrees offset. I/Q signals can be easily converted to amplitude (A) and phase (P) terms by taking into account the IQ plane graph represented in Figure 4.2b, following the equations below:

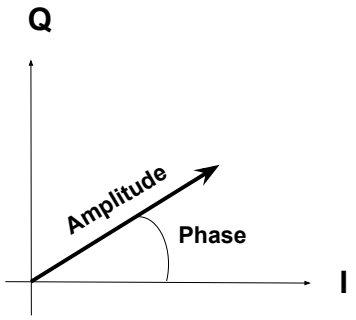
$$I = A cos\theta, \tag{4.1}$$

$$Q = A sin\theta, \tag{4.2}$$

θ being the phase of the signal and A its amplitude.



(a) I/Q mixer analog implementation. I_F and Q_F represent the final baseband I/Q signals. ADC: Analog-to-Digital converter. DSP: Digital Signal Processor [122].



(b) IQ plane diagram for conversion between I/Q and A/P signals.

Figure 4.2: Schematics IQ demodulation.

The I/Q demodulation performed in the XeThru modules used in the present work converts the RF signals into final baseband frequencies of 20 and 17 Hz in the X2M200 and X4M200 modules respectively.

As explained in previous chapters, UWB makes use of coherent pulse integration. This techniques consists of a summation of coherent pulses over a specific time-window in order to increase the probability of detection of low power pulses. The Signal-to-Noise Ratio of final signal increases by $10 \log N$, where N is the number of integrated pulses [4]. Figure 4.3 (adapted from [7]) shows the

signal detection capabilities differences between an integration done with 16 pulses (up) and 512 pulses (down). The higher the integration interval, thus the number of pulses, the larger the SNR.

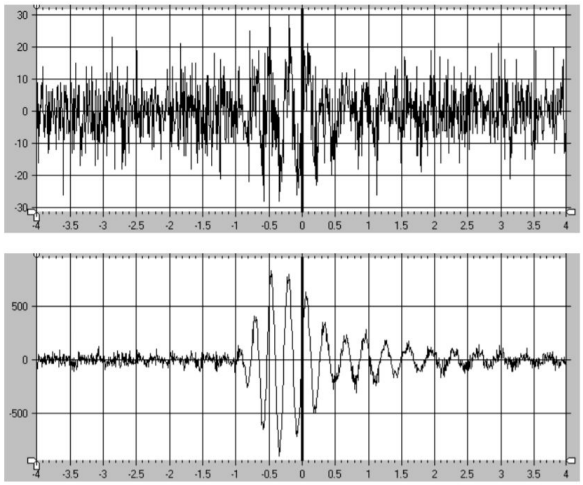
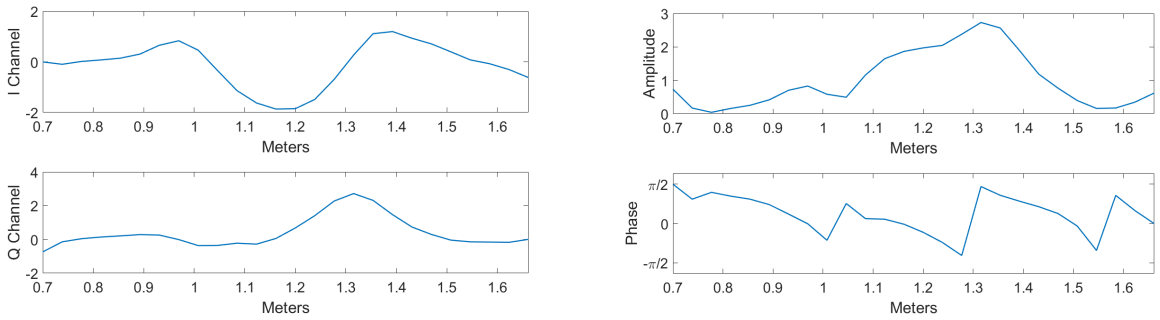


Figure 4.3: SNR increase by pulse integration of 16 pulses (up) and 512 pulses (down) [7].

After data processing and baseband down-conversion, the output signal for a specific radar frame looks similar to Figure 4.4, where both representations I/Q and Amplitude/Phase of radar baseband data can be seen along a specific distance range divided in *distance bins*. Signal amplitude in Figure 4.4b shows some reflectors along a radar detection area of 0.7 - 1.7 m. The strongest reflection is located at about 1.3 m and it is caused by a sleeping person. Units in y axes for amplitude, I and Q signals have arbitrary units from the radar outputs and phase is represented in radians. Both radar modules can be set to provide either I/Q or A/P data. As both representations can be interchangeable, in this work A/P representation was chosen, as these magnitudes are the most comprehensible. This baseband data was what it is referred to as the radar *raw data* in the present work. Further processing of this raw data is what was employed to determine movement and breathing information of the patient, to eventually classify sleep states.



(a) Output I/Q baseband radar data.

(b) Output Amplitude/Phase baseband radar data.

Figure 4.4: Example of baseband radar output data of one frame.

4.2. XETHRU RESPIRATION DETECTOR ALGORITHM

XeThru modules extract breathing information through the use of a *sliding Fast Fourier Transform* (sFFT), what they call *Pulse-Doppler processing*. The overall block diagram of the processing tasks developed in order to achieve respiration detection in XeThru sensors is shown in Figure 4.5.

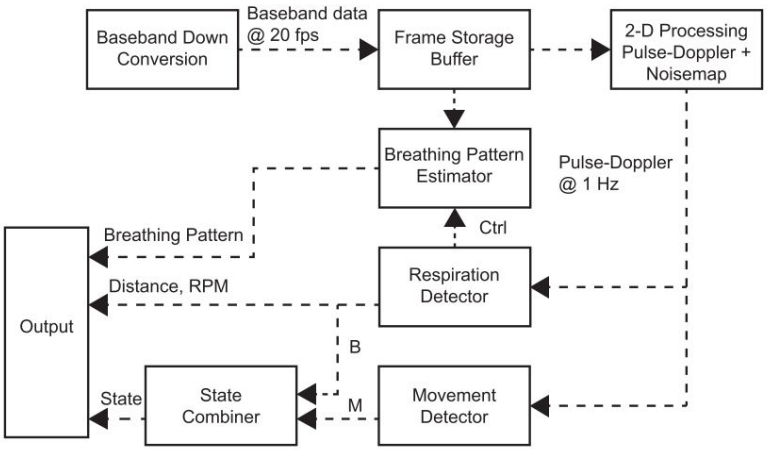


Figure 4.5: Block diagram of data processing tasks from raw baseband data to respiration detection [8].

A set of baseband frames during a specific time window (normally 15 sec) is analyzed with a sFFT. The resulting matrix, which is called the *range-Doppler matrix*, is comprised by each of the frequency spectrums at every distance bin. The representation of the range-Doppler matrix can be seen in Figure 4.6, illustrating the power peaks of the spectrum in the frequency corresponding to respiration [9]. Symmetrical peaks are found in the spectrum due to expansion and compression of the chest. Also the harmonics to this chest movement can be observed at multiple values of the respiration frequency. This pulse-Doppler processing step is performed every 1 second (1 Hz).

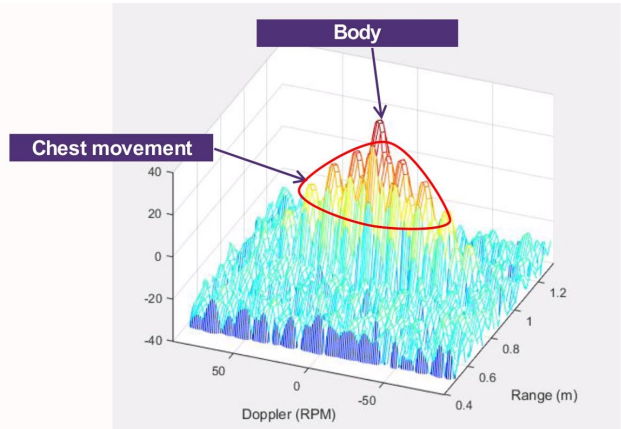


Figure 4.6: Range-Doppler matrix of a breathing person located 6 m away [9].

After this frequency analysis step, by removing low frequency signals around 0 Hz caused by static body reflections, a clear peak can be located at a specific frequency value and at a specific distance bin. In the previous example, the peak is located around 20 Respirations per minute (RPM) and at 1.2 m from the radar antennas.

Each XeThru module has different operating modes that allow the peak finding process to focus on specific frequency windows. For instance, for adult monitoring, respiration rates are found in the 8-15 RPM range, hence the peak finding algorithm focuses on this specific range. The exact frequency ranges for each module are further explained Section 4.3 *X2M200 Module* and 4.4 *X4M200 Module*.

Furthermore, while pulse-Doppler processing takes place, reflections coming from static objects are eliminated through the use of a *noisemap* computed in the beginning of the processing. This

noisemap allows to keep the radar frames during the fist instants of recording and create a map of static reflections within these frames that will be removed during pulse-Doppler processing.

Once frequency and range values are located within the range-frequency matrix, the algorithm focuses on the distance bin within the baseband phase data in order to study the relative movements within this bin with a sub-mm accuracy. Figure 4.7 shows the breathing pattern output for the respiration movements of the chest. This output is only available in module X2M200.

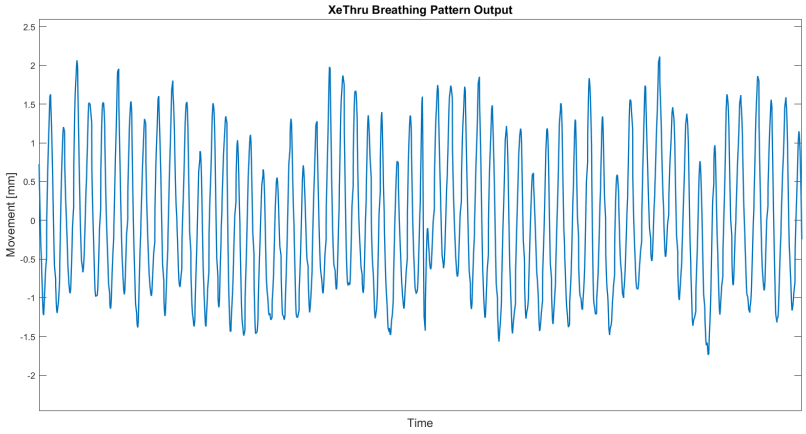


Figure 4.7: Breathing Pattern output of the XeThru X2M200 module.

It is important to notice that respiration detection via radar is not possible when subjects under study are constantly moving. Therefore, radar sensors do not output any information when too much movement is present.

XeThru sensor has three possible states depending on whether respiration detection has been successful or not. These states can be:

- *Breathing*: Respiration detection has been successful. Breathing outputs are enabled.
- *Breathing Tracking*: Respiration peaks have not been detected yet (because of 15 sec window in the FFT). No breathing output is given.
- *Movement*: Movement is present and respiration detection is not possible.

4.3. X2M200 MODULE

X2M200 module was the first sensor used for the present work. A photo of the front and rear sides of this module are shown in Figure 4.8. The module comprises an IR-UWB radar X2 system on a chip (SoC), with built-in transmitter (TX) and receiver (RX) antennas. For signal processing, the module also counts with an on-board ARM Cortex M4 MCU [10].

Two possible interfaces can be used to communicate with the module: UART serial interface or USB. In the present work, only USB communication was used. The module was connected to a small laptop with a specific software installed on it. Further details of the software will be given in Section *5.4 Radar Data Acquisition*.

In order to achieve the most robust and accurate recordings, the front side must be facing the chest of the patient. The key characteristics of this module are summarized in Table 4.1.

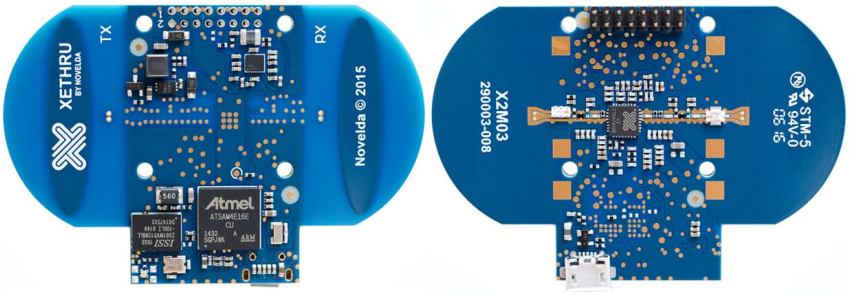


Figure 4.8: Front and rear sides of X2M200 module.

Before recordings take place, several settings need to be predefined by the user. Firstly, the mode can be set to *Respiration* or *Sleep*. Sleep is intended to measure respiration (rate and pattern) and movement of an adult sleeping through the night. Frequency ranges for this mode are limited to adult ones (8-30 RPM). Also detection area and sensitivity can be set. Total detection area has a total of 1 m for *Respiration* mode and 2 m for *Sleep* mode, both modes measuring up to a range of 2.5 m. Even though the user may set as detection area a small range, the actual radar frame will cover from 10 cm before the minimum range established by the user until 1 or 2 m more (depending on the mode). For instance, for an input of detection area equal to 0.6 - 0.8 m, the real radar detection area would be 0.5 - 1.5 in *Sleep* mode and 0.5 - 2.5 in *Respiration* mode.

In the X2M200 module, the pulse-Doppler processing described in *Section 4.2* analyzes the radar frames comprised within a window of 15 seconds.

Sensitivity values, which range from 0 to 9, can also be set by the user before the reading. High sensitivity values allow for detection of smaller objects, while low values give rise to more robust recordings. Sensitivity was set to 5 for most of the patient recordings.

The most important features and electrical parameters can be found in Table 4.1 [10].

Even though *Sleep* mode provides information about movement of the subject, its frequency range is very limited and does not include the normally high respiration rate values of children (up to 50 RPM in babies). Thus, *Respiration* was always set for the sleep studies in the present work.

Regarding antenna characteristics, Figure 4.9 shows the antenna lobe plots of RX at a center frequency of 7.5 GHz. Lobe plots represent the antenna gain as a function of its elevation (Figure 4.9a) and azimuth (Figure 4.9b) angles. The reference position of 0 ° azimuth and 0 ° elevation is represented in Figure 4.10. The manufacturer claims the X2M200 module has a field of view of 60 °, however RF signals can also be detected with less intensity outside this field of view, as shown in Figure 4.9.

4.4. X4M200 MODULE

X4M200 is the new generation XeThru sensor with similar characteristics as the X2M200 and small differences involving data processing and data outputs [14]. In contrast to X2M200, this module has an IR-UWB radar X4 SoC. The photo of the X4M200 module is shown in Figure 4.11, with its main characteristics summarized in Table 4.2. Also operational mode, sensitivity and detection area are required from the user in order for the sensor to start recording.

This module has different operational modes compared to previous module, these are *Adult* and

Table 4.1: Key characteristics X2M200 module [10].

Feature	Value	Units
Module Size	7 x 4.7 x 0.4	cm x cm x cm
Frames per Second (baseband)	20	fps
PRF	⩽ 30	MHz
Pulse Width	0.6	ns
Range resolution	7.5	cm
Operational Modes	*Sleep* / *Respiration*	
Detection Area	1 (*Respiration*), 2 (*Sleep*)	m
	min: 0.4, max: 2.5	m
Output distance bin length	3.8	cm
Frequency Range	8 – 65 (*Respiration*)	RPM
	8 – 30 (*Sleep*)	RPM
Supply Voltage VCC	3.3 - 5.5	V
Supply Voltage USB	4.5 - 5.5	V
Radar SoC	Novelda UWB X2	
TX bandwidth (-10 dB)	6.1 - 8.4	GHz
RX sampling rate	39	GHz
Power consuption (X2 / X2M200)	120 / 435	mW
Operating Temp. Range	-40 to +85	° C

fps: frames per second

Baby Monitoring. The only difference between the two modes is the breathing frequency ranges. Both modes provide respiration and movement data from the reading, however breathing pattern is not provided, in contrast with X2M200 module. X4M200 module has the possibility to output the moving information of other targets within its field of view. Also radar cross section[1] of the detected objects and pulse-Doppler processing outputs are provided.

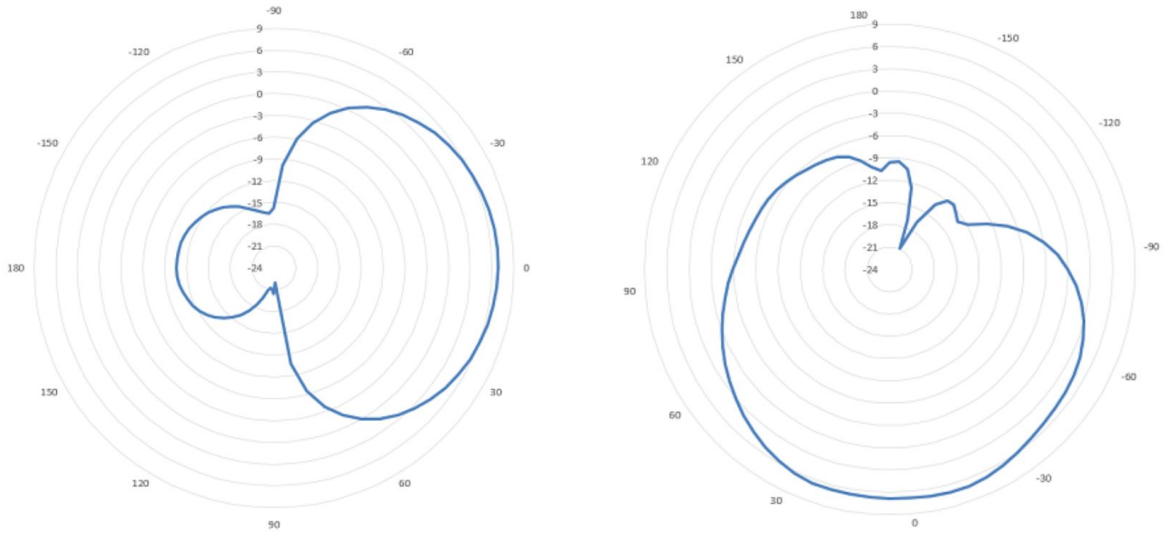
Another important difference between both modules relies on the pulse-Doppler processing step. In the X4M200 module, two pulse-Doppler processing steps are running in parallel, one using a time window of 20 sec and the other one of 6 sec. Both processes allow the module to provide information of immediate (last 6 sec) and non-immediate (last 20 sec) movement at every distance bin. Breathing detection is performed with the pulse-Doppler processing of 20 sec window.

Both the X2M200 and the X4M200 share the same communication protocols. USB connection with a laptop with a specific software was used for data acquisition of X4M200, just as with the previous module.

The change of X2M200 to X4M200 sensor was only caused by the availability of more output data in the new generation sensor. However, all the information provided by both sensors can be externally computed only using the baseband signals. Therefore no practical differences are found between the data from the subjects recorded with X2M200 and the ones recorded with X4M200 module.

No lobe plots are provided in the documentation of X4M200 datasheet. However, the manufacturer claims both modules have similar fields of view.

[1] Radar Cross Section (RCS) is a measure of how likely a target is to be detected by the receiver antenna of a radar system.



(a) RX elevation lobe plot (0 Azimut). **(b)** RX azimuth lobe plot (0 Elevation).

Figure 4.9: RX lobe plots [10].

Table 4.2: Key characteristics X4M200 module. [14]

Feature	Value	Units
Module Size	6.7 x 4.2 x 0.35	cm x cm x cm
Frames per Second	17	fps
PRF	16	MHz
Pulse Width	0.5-1	ns
Range resolution	<15	cm
Operational Modes	*Baby* / *Adult*	
Detection Area	1.4	m
	min: 0.4, max: 4.6	m
Output distance bin length	5	cm
Frequency Range	8 - 30 (*Adult*)	RPM
	15 - 65 (*Baby*)	RPM
Supply Voltage VCC	3.3 - 5.5	V
Supply Voltage USB	4.5 - 5.5	V
Radar SoC	Novelda UWB X4	
TX bandwidth (-10 dB)	6.5 - 7.9	GHz
RX sampling rate	23	GHz
Power consuption (X4 / X4M200)	120/600*	mW
Operating Temp. Range	0 to +85	° C

* power consumption information has not been made available yet for X4M200. It is expected 600 mW. fps: frames per second,

4.5. TRANSMISSION POWER REGULATIONS

Two main worldwide entities regulate the use of UWB devices in order to prevent this modality of radio frequency devices from interfering with already existing communication devices. These entities are the Federal Communications Commission (FCC) in USA and the European Telecommunications Standards Institute (ETSI) in Europe [15]. The maximum allowed mean equivalent isotropically ra-

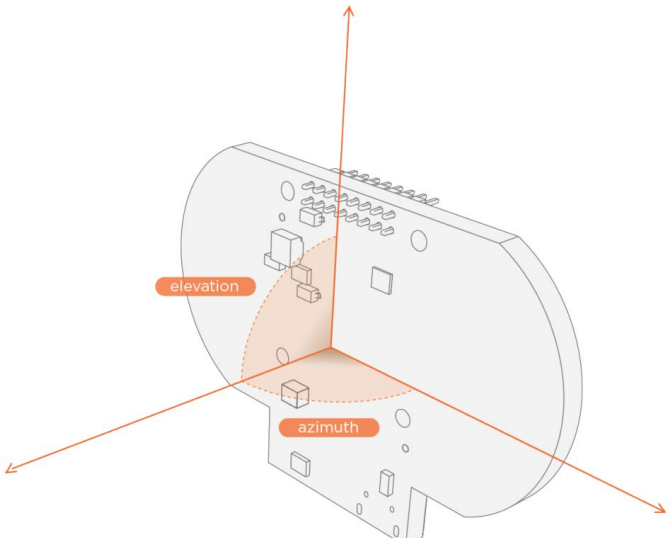


Figure 4.10: Reference Azimuth and Elevation for X2M200.

Figure 4.11: X4M200 module.

diated power (EIRP) spectral density is -41.3 dBm/MHz within a frequency mask of 6.0-8.5 GHz for ETSI [123] and 3.1 – 10.6 for FCC [98]. Therefore, the corresponding total allowed EIRP values for each entity would be:

$$EIRP_{FCC} = -41.3dBm/MHz \cdot 7.5GHz = 74.13nW/MHz = 556\mu W \tag{4.3}$$

$$EIRP_{ETSI} = -41.3dBm/MHz \cdot 2.5GHz = 74.13nW/MHz = 185.3\mu W \tag{4.4}$$

Experiments using X2M200 have shown its maximum PRF value in order to comply with the ETSI compliance is 30 MHz [15]. Therefore PRF of the sensor was limited to 30 MHz. Figure 4.12 shows the EIRP values of X2M200 within the frequency mask of EIRP, giving a total transmitted power of 84 μW withing the ETSI mask bandwidth of 6.5 - 8 GHz.

Even though these experiments are only reported for X2M200, both modules X2M200 and X4M200 declare they comply with the CE/ETSI compliance and its content of hazardous substances [10, 14].

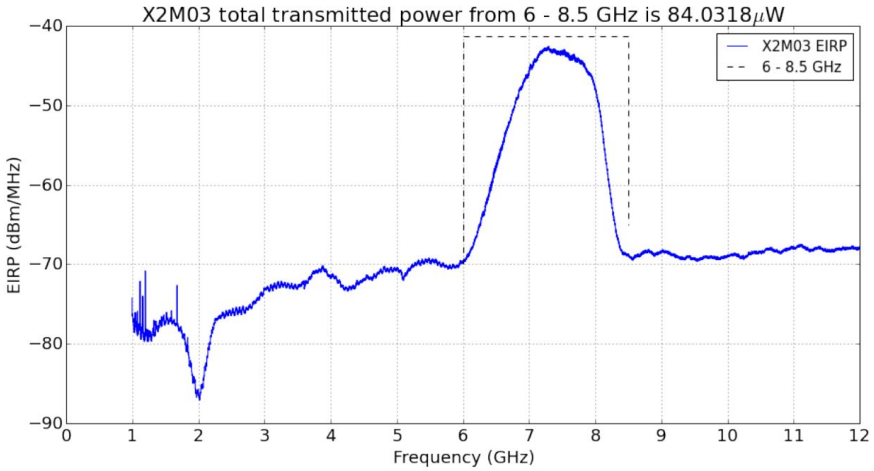


Figure 4.12: EIRP values for X2M200.

4.6. SAFETY CONCERNS

Since these radar modules were intended to be used in hospital conditions with patients, safety concerns are very important when using radio frequency devices. The energy that they emit is regulated by agencies around the world.

The International Commission on Non-Ionizing Radiation Protection (ICNIRP) defines Specific Absorption Rate (SAR) limits for electromagnetic fields, in order to prevent exposure to high levels of non-ionizing radiation [124]. SAR limits, defined as the power absorbed per mass of tissue (watts per kilogram), are set to 0.08 W/kg for whole body average, 2 W/kg for head and trunk and 4 W/kg for limbs.

ICNIRP provides reference levels of fields generated by the RF devices in the range of 2-300 GHz in order to comply with previously mentioned SAR limits. These reference levels are enumerated in Table 4.3. Since electric and magnetic field are strictly related by $H = \frac{E}{Z_0}$, being Z_0 the vacuum impedance (Z_0= 120 π), only one of these measures needs to be reported.

Table 4.3: ICNIRP reference levels fields for public exposure [15].

Field	**Limit**	*Units*
Electric Field strength	61	V/m
Magnetic Field strength	0.16	A/m
Equivalent Plane wave power density	10	W/m^2

In order to calculate the electric field generated by XeThru sensors, the power density S at a distance R is calculated by Equation 4.5.

$$S = \frac{EIRP}{4\pi \cdot R^2} \tag{4.5}$$

Whereas electric field is dependent on power density as below:

$$E = \sqrt{S \cdot Z_0} \tag{4.6}$$

The power density of a XeThru sensor with a transmitted power of 84 μW at a distance of 0.4 m would be equal to 42 $\mu W/m^2$ (reference = 10 W/m^2), with an electric field equal to 0.126 V/m, 486 lower than the ICNIRP levels. When comparing this value to other common RF devices , X2M200 module produces an electric field 34 and 108 times lower than a 2.4 GHz and a 5 GHz WiFi devices, respectively [15].

ICNIRP has as an additional requirement for pulsed systems, such as XeThru sensors; the field strength average over the pulse width power should not be higher than 32 times the reference field levels. Peak power can be calculated as:

$$P_{peak} = \frac{P_{average}}{\tau \cdot PRF}, \tag{4.7}$$

being τ the pulse width of the system. A power emission of 84 μW at a PRF of 30 MHz and a τ equal to 0.64 ns, give rise to a P_{peak} equal to 4.4 mW with an electric field strength of 0.91 V/m, much lower than $61 V/m \cdot 32 = 1952 V/m$ reference level.

In conclusion, XeThru radar emission levels are very much lower than ICNIRP reference levels for public exposure. Both sensors X4M200 and X2M200 modules are in compliance with CE/ETSI and its content of hazardous substances.

4.7. CONCLUSIONS

The radar modules employed for the present work have been explained in detail within this chapter in terms of their signal processing, their respiration detection algorithm, the key features of each module and the safety and power transmission issues. Once we have a clear idea of how the technology used in the present research work, now we can continue with how the data was acquired and preprocessed, which is explained in the next chapter.

5

Data Acquisition

In this section, a detailed description of the data acquired during the study will be provided, along with information about patient selection, study protocol, study set-up, and the needed preprocessing processes that were required before the subsequent vital signs extraction and sleep classification tasks.

5.1. Patients

23 patients were recorded during the months of May and September of 2017 with a total of around 200 study hours. Interruptions during which parents or nurses were in contact with the patient comprise about 15% of the each recording.

Even though the final goal of the system is to develop a sleep monitoring approach for babies born premature, due to the lack of PSG studies prescribed on these young patients, older babies and children were employed as the study population for the feasibility study of the system.

All the data acquired during the lifespan of the SOMNUS study came from pediatric patients who were already scheduled by their doctor to perform a PSG study. This type of studies is normally prescribed to patients with conditions that affect respiratory patterns or who have suffered from any kind of respiratory complications in the past. Table 5.1 comprises the different patients included in the study with their respective patient information, including their diagnosis.

Patient selection was not necessary when testing the accuracy of the breathing detector, so all the patients were included in this analyses. In contrast, the sleep classification task is highly dependent on the data employed for a classifier to be trained, thus it is very important to use a good training dataset, which is as representative as possible for the normal sleep patterns of children with similar age to these study population. Therefore, a patient selection was performed in order to remove those patients with sleep abnormalities. This task was undergone with the help of a neonatologist sleep researcher. The following patients were removed from the sleep classification study:

- **Subject 4**: Phsycomotor retardation and severe neurological disorder that cause a very fragmented hypnogram (constant change of sleep phase) with many central apneas present along the study.

- **Subject 11** : Severe brain damage because of which not a single REM event is present in the hypnogram [1].
- **Subject 18**: The radar sensor data recording was stopped 3 hours after the beginning of the study due to a failure in the software.

Table 5.1: Demographic characteristics of SOMNUS Study Patients.

Subject	Gender	Group	Age	Diagnose
Somnus1	M	B	9 mo	Bronchopulmonary Dysplasia
Somnus2	F	E2	5 y, 1 mo , 1 wk	Spinal muscular atrophy (type 2)
Somnus3	M	D	11 mo	Robin sequence
Somnus4	F	C	9 mo, 3 wk	Chromosomal disorder
Somnus5	M	E1	2 y, 1 mo, 1 wk	Neuroendocrine cell hyperplasia of infancy
Somnus6	M	D	10 mo	Bronchopulmonary Dysplasia
Somnus7	M	E1	1 y, 2 mo	Cleft lip, Robin sequence
Somnus8	F	D	9 mo	Tracheomalacia
Somnus9	F	E	13 y, 2 mo, 3 wk	Seizures, cerebral sinus thrombosis, acute lymphatic leukemia
Somnus10	F	A	2 mo, 1 wk	Crouzon syndrome
Somnus11	M	E2	6 y, 8 mo, 2 wk	Hypoxic brain damage
Somnus12	M	B	4 mo, 1 wk	Robin sequence
Somnus13	M	E2	4 y, 9 mo	Spinal muscular atrophy (type 2)
Somnus14	M	E2	7 y, 9 mo, 2 wk	Unknown
Somnus15	M	E1	1 y, 7mo, 3 wk	Down syndrome
Somnus16	F	E2	5 y, 3 mo	Robin sequence
Somnus17	F	E3	12 y	Spinal muscular atrophy (type 2)
Somnus18	M	E2	4 y, 5 mo	Isolated cleft palate
Somnus19	F	E2	11 y, 11 mo	Spinal muscular atrophy (type 2)
Somnus20	F	E3	14 y	Crouzon syndrome
Somnus21	M	D	11 mo, 3 wk	Prader-Willi syndrome
Somnus22	M	B	5 mo	Unknown syndrome
Somnus23	M	C	8 mo, 3 wk	Bronchopulmonary Dysplasia

mo: months, y: years, wk: weeks.

Table 5.2: Summary of patients groups' characteristics.

Group	Group Age Criteria	No. Patients
A	Age < 3 mo.	1
B	3 mo. < Age < 6 mo.	2
C	6 mo. < Age < 9 mo.	3
D	6 mo. < Age < 1 y	4
E1	1 y < Age < 4 y	4
E2	4 y < Age < 12 y	6
E3	Age > 12 y	3

mo: months, y: years, wk: weeks.

Patients were subsequently organized by 7 groups according to their age. There were four groups, A-D, for each trimester of the first year of life, and 3 other groups, E1-E3, for patients with an age

[1] An hypnogram is the resulting final graph comprising the sleep classification along the duration of the PSG study.

ranging from 1 to 15 years old. This patient classification was suggested by the doctors in Erasmus MC, who took into account similarities in developmental features within each group. Table 5.2 summarizes this group age criteria. Unfortunately, there was a slightly unbalanced number of patients in each group. Ideally, we would require a high number of patients in group A since their developmental features are the closest to preterm babies. However availability of patients scheduled for a PSG study was not dependent on the researchers involved in the present work.

5.2. STUDY PROTOCOL

The study protocol (Figure 5.1) was comprised by several tasks included within a time span of 3-5 days. This study had the approval of the hospital's Ethical Committe (EC) of Erasmus MC. According to the EC, since the radar sensor used in the study was not an obstacle for the PSG procedures of the patient, no signed consent was needed from the patient's parents in order to be included in our study. However, the researchers always asked them for permission before setting up the sensor in the PSG room.

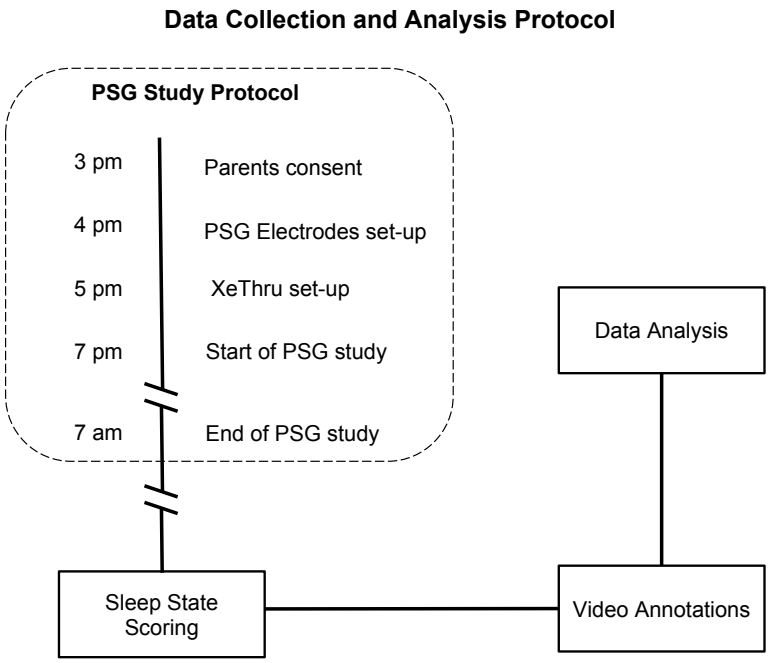


Figure 5.1: SOMNUS Study Protocol Schematics.

In the first place, patients along with their parents were scheduled during the afternoon in order to start preparing for the study. Patients were called around 3 pm in order for the doctor to place the EEG electrodes on the patient. After that, the researchers asked parents for consent, and if positive, XeThru sensor was placed inside the room at the head of the patient's crib (further details on the experimental set-up in Section 5.3). At around 7-8 pm, nurses prepared the patients and started the recording of every sensor. Meanwhile, the researcher, with remote access to the computer connected to the radar sensor, used to start the recordings around that time as well. Each patient's sleep study was performed for a whole-night recording, so the study normally ended at 7-8 am of the next day. Parents could stay with their children inside the same room during the study, and that was what mostly all parents did with the exception of two patients.

In the subsequent two to three days, a subsequent sleep state scoring was performed by a qualified technician in the hospital. There are three technicians in total, so not every hypnogram was scored by the same technician.

Afterwards, video recordings from the PSG study were inspected in order to identify irregularities in the study, i.e. nurse and/or parents interruptions.

Finally, after all results were ready, data was properly anonymized in order to be subsequently analyzed by the researcher.

5.3. EXPERIMENTAL SETTINGS

A summary of the experimental settings for each study can be seen in Table 5.3.

Table 5.3: Experimental Settings per study

Subject	Module	Room	Recording Time
Somnus1	X2M200	Er-A	11 h 7 min
Somnus2	X2M200	Er-A	10 h 57 min
Somnus3	X2M200	Er-B	10 h 59 min
Somnus4	X2M200	Er-A	10 h 0 min
Somnus5	X2M200	Er-A	10 h 14 min
Somnus6	X2M200	Er-B	11 h 18 min
Somnus7	X2M200	Er-A	12 h 1 min
Somnus8	X2M200	Pb-B	12 h 44 min
Somnus9	X2M200	Pb-B	11 h 18 min
Somnus10	X4M200	Pb-A	10 h 43 min
Somnus11	X4M200	Pb-B	9 h 42 min
Somnus12	X4M200	Pb-A	9 h 53 min
Somnus13	X4M200	Pb-A	10 h 55 min
Somnus14	X4M200	Pb-B	9 h 41 min
Somnus15	X4M200	Pb-B	10 h 19 min
Somnus16	X4M200	Pb-B	11 h 39 min
Somnus17	X4M200	Pb-B	8 h 26 min
Somnus18	X4M200	Pb-B	5 h 1 min
Somnus19	X4M200	Pb-B	7 h 20 min
Somnus20	X4M200	Pb-B	10 h 51 min
Somnus21	X4M200	Pb-B	11 h 11 min
Somnus22	X4M200	Pb-B	8 h 23 min
Somnus23	X4M200	Er-A	9 h 10 min

POSITION OF THE SENSOR WITH RESPECT TO THE PATIENT

As explained in the previous chapter, the XeThru sensor achieves its highest accuracy when being placed pointing towards the chest of the patient to be monitored. For our study, the sensor was located at the head of the patient's crib, in an elevated position. Distance between patient and sensor was kept as constant as possible along the measurements. During the set-up of each study, the radar sensor was located at a height of 90 cm with respect to the mattress of the patient, however, during the study normally nurses used to change the elevation of the crib so it was not possible to

know the real height at each study. Also, horizontal distance use to change since patients move during the night. A photo of the radar sensor next to the patient's crib is shown in Figure 5.2.

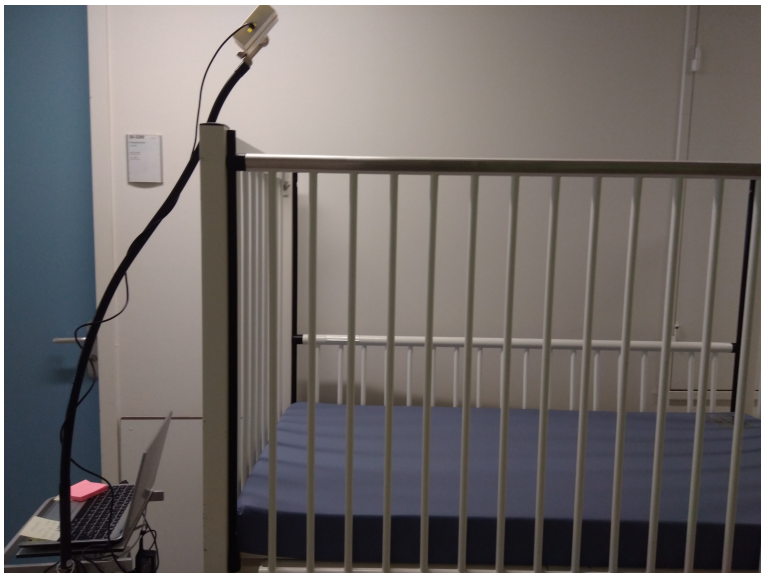


Figure 5.2: SOMNUS location at the head of the patient's crib.

The radar sensor was connected via USB to a computer used to record the raw radar data of the study. As shown in Figure 5.2, the sensor is attached to a stiff stick mounted on a portable, small table on which the computer was placed at the head of the crib. In this way, the sensor acquires a high elevation that allows it to have a clear detection area without no objects in its line of sight with the patient.

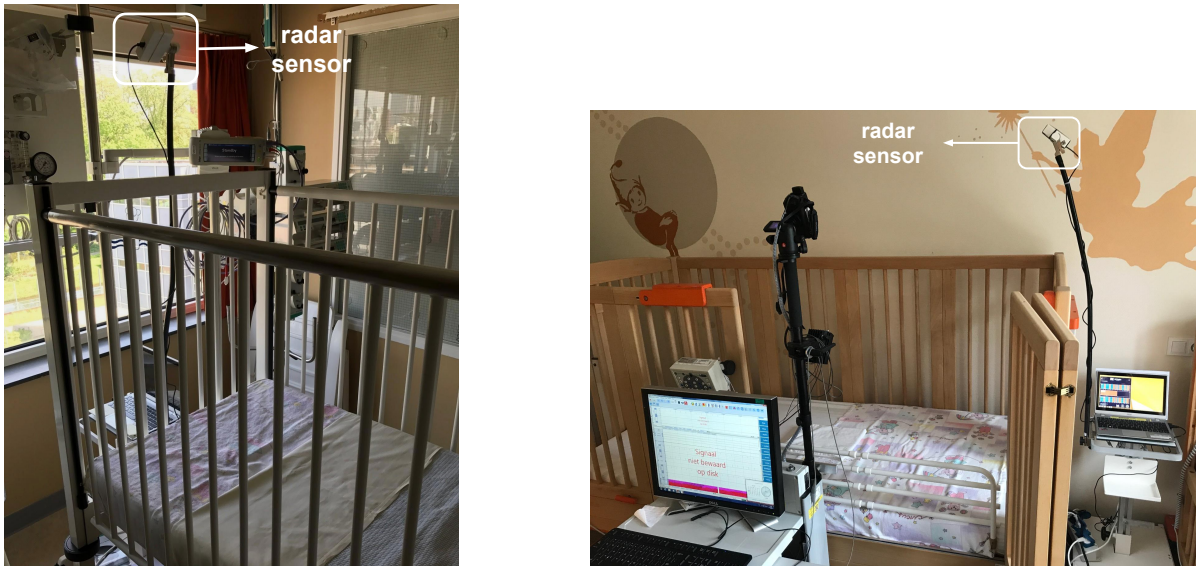
When the patient was older than around 8 years old, the crib was replaced by a real bed but the sensor settings were kept the same. This only happened with four subjects.

LOCATION INSIDE THE ROOM

The studies were done in two centers: Erasmus MC, specifically i the Sophia Kinderziekenhuis, and Pallieteburght, where the PSG procedures took place during the summer period of 2017. Subjects 1 to 9 were measured in Erasmus MC and the rest in Pallieteburght. Equipment and PSG procedures were exactly the same at each center.

Regarding Eramus MC, there were two different rooms in which PSG studies normally took place, referred to in the present work as *Er-A* and *Er-B*. Both rooms have the same distribution of medical equipment and very similar dimensions. In this center, the radar sensor was always positioned at the head of the patient's crib/bed, with only a window at the back side of the radar sensor. In the room there was only the crib for the patient and some chairs for the parents. The area reserved for the parents was far away from the radar detection area. The exact location where the parents stayed during the complete night was not known to the researchers. This information can not be acquired from the video data since the camera is always pointing at a small part of the patient's crib, having a very limited sight of the patient's environment. Sometimes, information about patient's environment (nurses or parents present) can only be inferred from audio data. A photo of the study setup at Erasmus MC can be seen in Figure 5.3a .

In the case of Pallieteburght, PSG could be also performed in two different rooms with similar dimensions. In contrast to the previous center, this one had an extra bed next to the patient available



(a) Sensor setup at Erasmus MC (room A).

(b) Sensor setup at Pallieteburght (room Pb-B).

Figure 5.3: Somnus set-up within the PSG room.

for one of the parents. In one of the two rooms of this center, which was referred to as room *Pb-A*, the parent bed was too far away from the radar to be detected and interfere with the recording. However, in the other room, which was be referred to as room *Pb-B*, the radar sensor could only be placed in between the bed and the crib of the patient. As seen in the previous chapter, even though the sensor is always placed pointing towards the patient, its antenna can still emit and detect signals from objects located at its back side, with lower magnitude signals. Since patient and parent were at a very similar range to the radar when located in room B, this caused a concern to be arisen by the researchers, thinking both signals could get mixed by the sensor. However, a posteriori analysis comparing radar and PSG data showed the radar could accurately detect breathing of patients with very few artifacts, even if most of the recordings in Pallieteburght were done in room Pb-A. A photo from the study setup at Pallieteburght can be seen in Figure 5.3b .

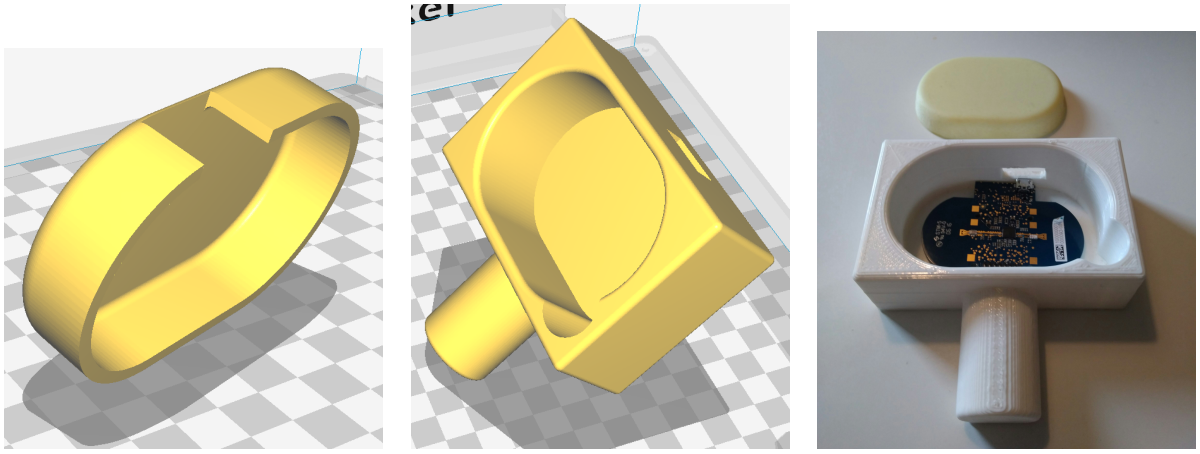
The variation of rooms used for the study (4 in total) was initially a concern raised by the researchers. However, the high accuracy shown by the radar sensor for breathing detection (Chapter 6) with almost every patient was the deciding factor to assume the differences in study rooms were negligible.

SENSOR CASE

Two cases were 3D printed for each radar module. Both were built with the same material using PLA polymer. Figure 5.4a and 5.4b show the schematics of the first case and its lid printed for the X2M200 module using *Solid Edge ST9* software. Figure 5.4c shows the real pictures of the sensor inside the case.

X4M200 case was printed with the help of *Novelda AS*. The design schematics and a real photo of the final case can be seen in Figure 5.5.

The possibility of using a thin radio-absorbent material to be placed at the back and lateral sides inside the case was studied in order to focus the radar's field of view. A possible adhesive material was found, however the supplier stopped its distribution in the past. To the knowledge of the author, no other suitable radio-absorbent material could be found that could served as a coating of the inside of the case.

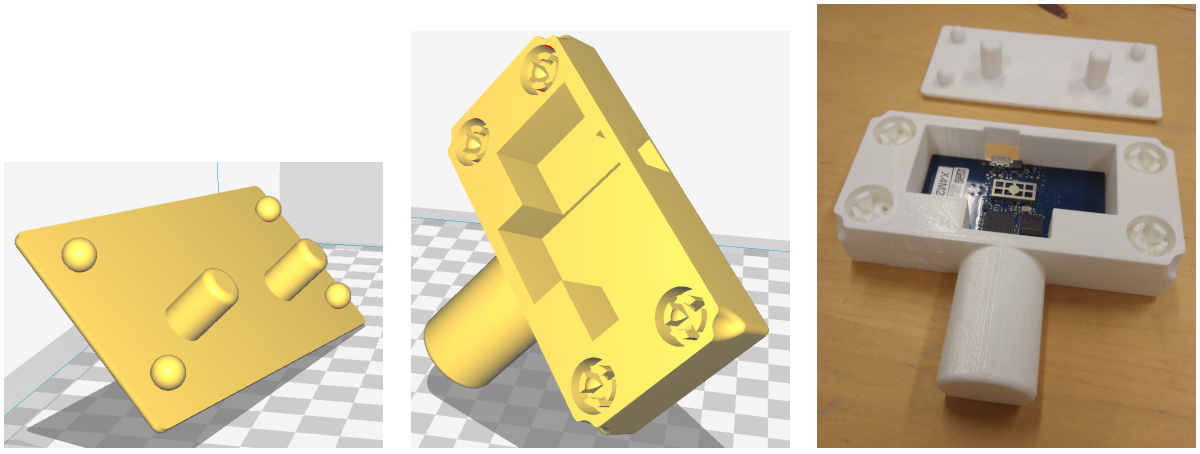


(a) 3D print design of the X2M200 lid

(b) 3D print design of the X2M200 case

(c) Photo of the X2M200 module inside its case (down) and the lid (up)

Figure 5.4: X2M200 module case.



(a) 3D print design of the X4M200 lid

(b) 3D print design of the X4M200 case

(c) Photo of the X4M200 module inside its case (down) and the lid (up)

Figure 5.5: X4M200 module case.

5.4. RADAR DATA ACQUISITION

Radar sensor data was acquired with the use of specific software used for the Xethru sensors called *XeThru Explorer*. This software was installed in the small laptop connected to the radar module and remotely controlled by the researcher via *TeamViewer* software. This made the data acquisition very easy, as the researcher could start the recordings at night and stop them in the morning without the need to be present in the PSG study room. Data transfer was also made possible via *TeamViewer*.

The *XeThru Explorer* software allows the user to input the radar detection area that the sensor will focus on. For the X2M200 module this area had a total coverage of 1 m, so it was always set to 0.8 - 1.8 m, allowing for full coverage of the patient's bed. For the X4M200 module, this detection area was a bit wider and was set to 0.7 - 1.9 m. This software was also used to visualize radar recordings in real-time. Figure 5.6 shows an example of the *XeThru Explorer* interface for radar data visualization with the X2M200 module. Amplitude and phase of the radar signals can be visualized if predefined

by the user.

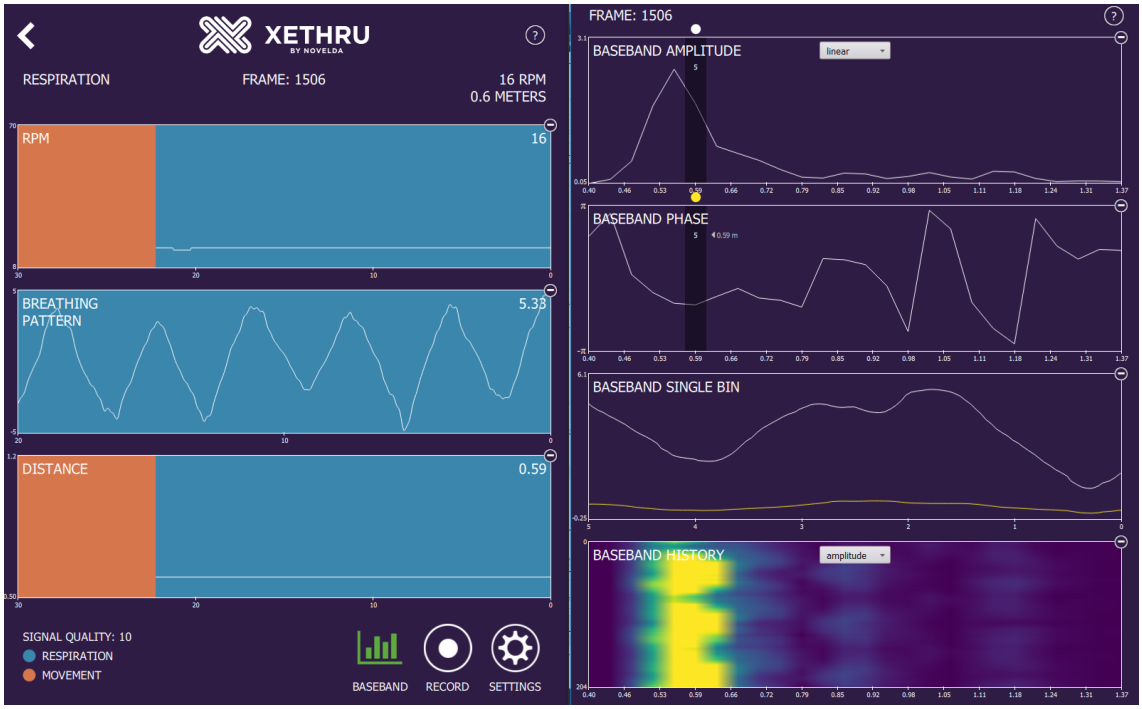


Figure 5.6: XeThru Explorer user interface. Breathing rate in RPM, breathing pattern in mm and distance in meters to breathing target (left), red areas representing non-reading times (artifacts); baseband amplitude and phase data (right). Baseband single bin can be used to visualize values of baseband amplitude (white) and phase (yellow) of the bin selected by the user.

As explained in the previous chapter, the X2M200 module could be set in two different modes: *Sleep monitoring* and *Respiration Detector*. *Sleep monitoring* provides an extra output of movement information, however the respiration detection in this mode is limited to adult users, since the sensor focuses on very low frequency ranges. Since respiration acquisition was the main goal of this research, the X2M200 module was always set in *Respiration Detection* mode. The radar data outputs for the X2M200 module via *XeThru Explorer* are listed in Table 5.4.

Table 5.4: Output data for X2M200 radar module.

Signal	Units	Sampling Frequency
I & Q channels (A & P)	a.u.	20 Hz
Breathing movement	mm	20 Hz
Breathing rate	RPM	20 Hz
Distance to breathing target	m	20 Hz

A: Amplitude, P: Phase, RPM: Respirations per minute. a.u.: arbitrary units.

In the case of the X4M200 module, *Baby Respiration* mode was always set when using this radar module. The radar data outputs of X4M200 module via *XeThru Explorer* are listed in Table 5.5.

Table 5.5: Output data for X4M200 radar module.

Signal	Units	Sampling Frequency
I & Q channels (A & P)	a.u.	17 Hz
Movement in last 6 sec	%	1 Hz
Movement in last 15 sec	%	1 Hz
Breathing rate	RPM	1 Hz
Distance to breathing target	m	1 Hz

A: Amplitude, P: Phase. RPM: Respirations per minute. a.u.: arbitrary units.

5.5. REFERENCE DATA: PSG SIGNALS

PSG equipment used multiple electrodes that monitor the vital signals of the patient over one night of recording. These vital signs include abdominal and thoratic breathing, air flow, ECG, oxygen saturation, HR, chin EMG, partial pressure of carbon dioxide, and EEG. These signals together with their units and sampling frequency are described in Table 5.6.

In the present work, we only focus on breathing signals from thorax and abdomen movements. These signals are extracted by measuring the respiratory effort via *Respiratory Inductance Plethysmography (RIP)* [125]. RIP relies on both *Faraday's Law* and *Lenz's Law*. Firstly, a current applied through a wired loop generates a magnetic field normal to the orientation of the loop (Faraday's Law), whereas a change in the area confined inside the wired loop creates a counter current in opposite direction proportional to this change in area (Lenz's Law). RIP uses two elastic belts surrounding the abdomen and thorax. These elastic belts have a wire sewn inside with a zigzag shape. An AC current is sent through these wires, creating a very small magnetic field that changes with breathing movement and generates another current opposing the first one. Changes in current can be easily translated into changes in area surrounded by the belt. The devices used for the present study were *Embla® XactTrace®* respiratory effort belts. These can be seen on Figure 5.7, a photo of our control patient *Somnus 0*.

The airflow signal was measured by a thermosistor sensor located in the nostrils of the patient. Even though this signal has breathing information as well, the researchers of the present work decided to focus on signals of respiratory effort in the abdomen and thorax due to the fact that both these signals have similar working principle as radar signals, as they all measure respiration by detection of movement.

Table 5.6: PSG signals.

Signal	Units[1]	Sampling Frequency [Hz]
Thorax	μV	25
Abdomen	μV	25
Airflow	μV	25
ECG	μV	250
Sat. O_2	%	10
HR	BPM	5
EMG chin	μV	250
PCO_2	mV	5
EEG channels	μV	250

[1] BPM: Beats Per Minute

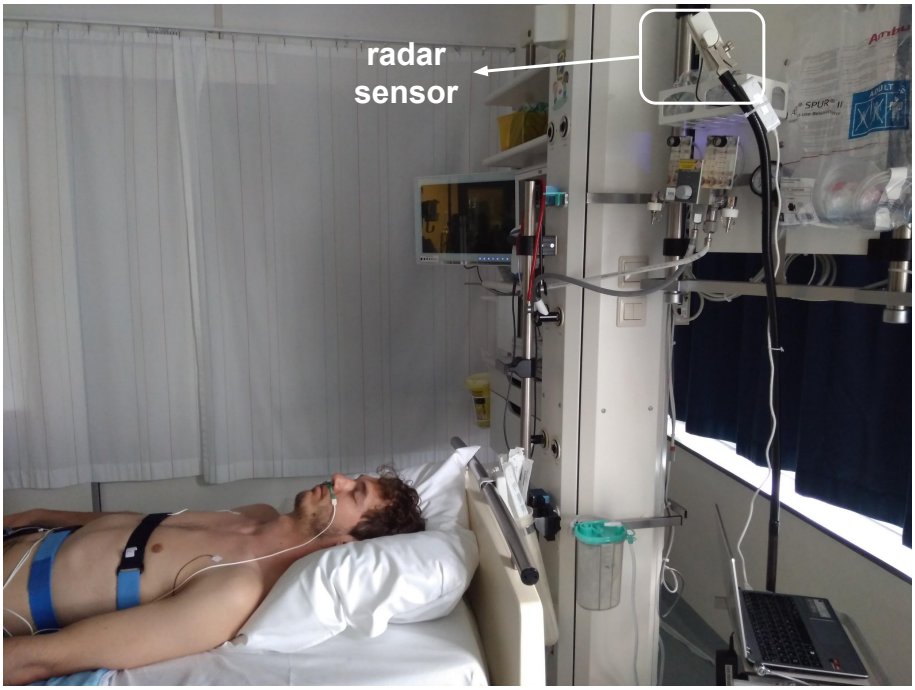


Figure 5.7: SOMNUS 0 set-up showing respiratory effort belts (thorax and abdomen) used for the present study.

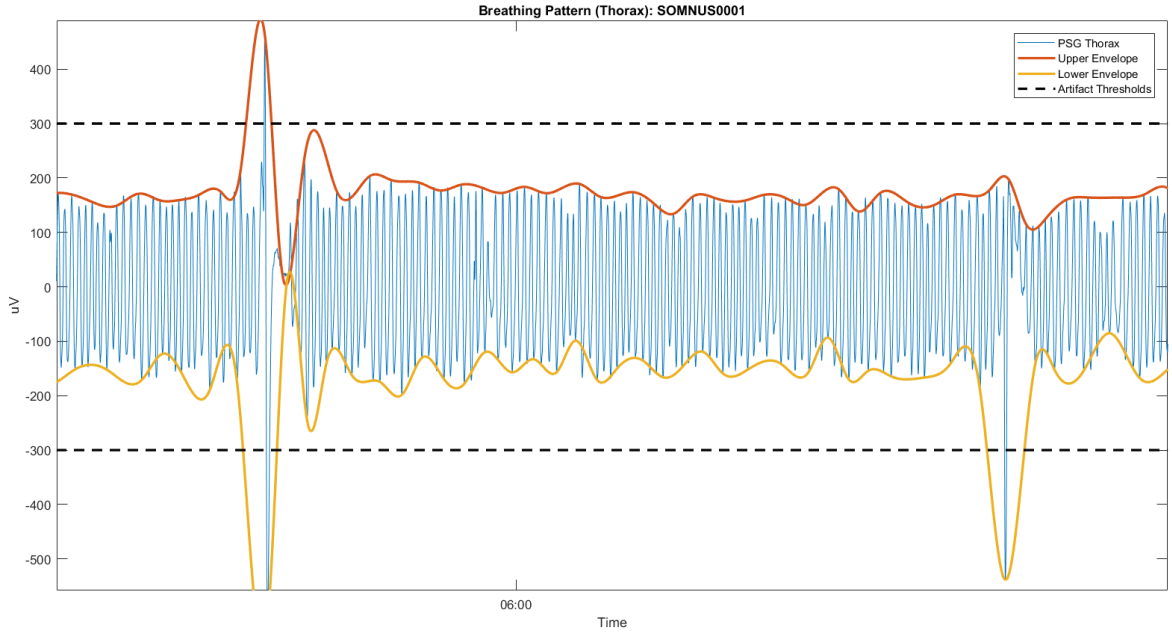
5.6. DATA PREPROCESSING

Since data acquired from the radar sensor was extracted through an external laptop that was not connected to PSG equipment, after the separate acquisition of the PSG and radar data, a synchronization of signals was necessary in order for the time vectors of all signals to coincide. This synchronization was done manually via MATLAB. Since the time difference between signals was not higher than a few seconds, the signals could be easily synchronized.

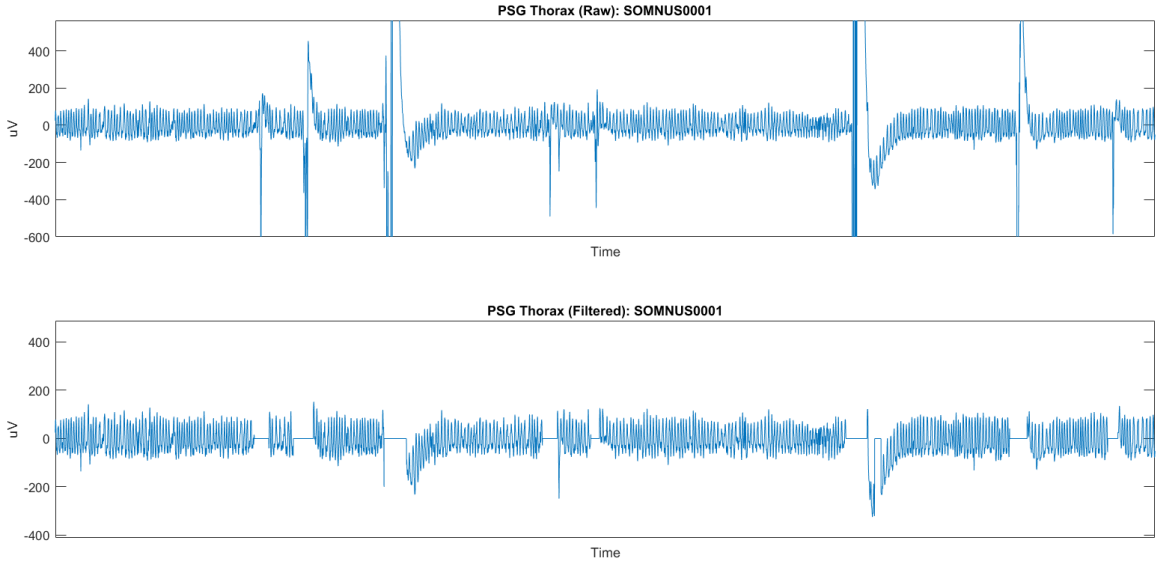
Another data preprocessing step focused on making the sampling frequencies of all signals coincide, as the thorax/abdomen data from the PSG were sampled at 25 Hz and radar data was sampled at either 17 Hz, for the X4M200 module, or 20 Hz, for the X2M200 module. This resampling process was done by interpolation of time signals of thorax and abdomen data from the PSG in order to coincide with the radar sampling frequency.

A filtering step was needed in order to remove the noise from raw abdomen and thorax data from the PSG. This filtering process was done by studying the envelope of the signals and eliminating the part of the signals in which the envelope overshooted a certain threshold. This threshold was manually set for each signal of every patient since each patient presented different μV values in the breathing signals and also patients normally had a more prominent respiration in either abdomen or thorax. This filtering step is represented in Figure 5.8.

Once the abdomen and thorax data were properly filtered, the breathing rate was calculated by finding the peaks of the most prominent signal (thorax or abdomen). Once peaks were detected, the distance between every peak was calculated in order to extract the instantaneous breathing rate, expressed in RPM. Since this signal gave very noisy measurements, a moving average filter with a time window of 15 seconds was employed to obtain the final breathing rate signal that was used as the reference breathing rate signal of our study.



(a) PSG filtering via envelope detection.



(b) Thorax raw (up) and filtered (down) breathing signal from PSG .

Figure 5.8: PSG breathing signal filtering with Subject 1 thorax data.

6

BREATHING DETECTION SYSTEM

Despite the extremely high accuracy of the XeThru radar modules to detect breathing, specially the sub-mm accuracy reported on the X2M200 for its breathing pattern output, when the sensors where initially tested for the present study, overall recordings of an over-night study showed a great amount of noise in the respiration signals due to motion artifacts. This effect motivated us to create a more robust, less noisy breathing detector capable of suppressing motion artifacts *before* actual breathing readings are provided to the user.

Furthermore, another new contribution that the present work provides to the breathing detection system, is the automatic adaptability of the algorithm to every patient's respiration frequency range, since this range may change from 15-20 RPM with old children, to 40-50 RPM with babies. Figure 6.1 illustrates the main differences in block diagrams between the conventional breathing detection in-built algorithm of the XeThru modules and the proposed breathing detection algorithm in the present work.

The outputs of the breathing detection system proposed in this present work will serve as input of the sleep monitoring system. The tasks of the breathing detection system are:

1. To extract baseband raw data from the XeThru sensor (@17-20 Hz), showing reflections along the range bins of the radar (distance bins).
2. To quantify the presence of motion in baseband signals.
3. To identify whether a specific threshold of motion has been reached. If positive, no breathing outputs are provided.
4. As long as no motion is detected, to perform frequency analyses of baseband data to create a range-frequency matrix.
5. To detect Respiration Rate (RR) and distance to respiration by localizing the most prominent peak of the range-frequency matrix.
6. To detect stable values of RR within a time window and adapt the resulting RPM value to the frequency analysis of following iterations.
7. To extract the breathing pattern by focusing on the baseband reading of the distance bin.

An important detail from the conventional breathing detector of the XeThru modules is the fact that the motion detector runs after frequency analyses and not before, thus when motion is present, the sensor reads noisy data during a certain time window before motion is even detected.

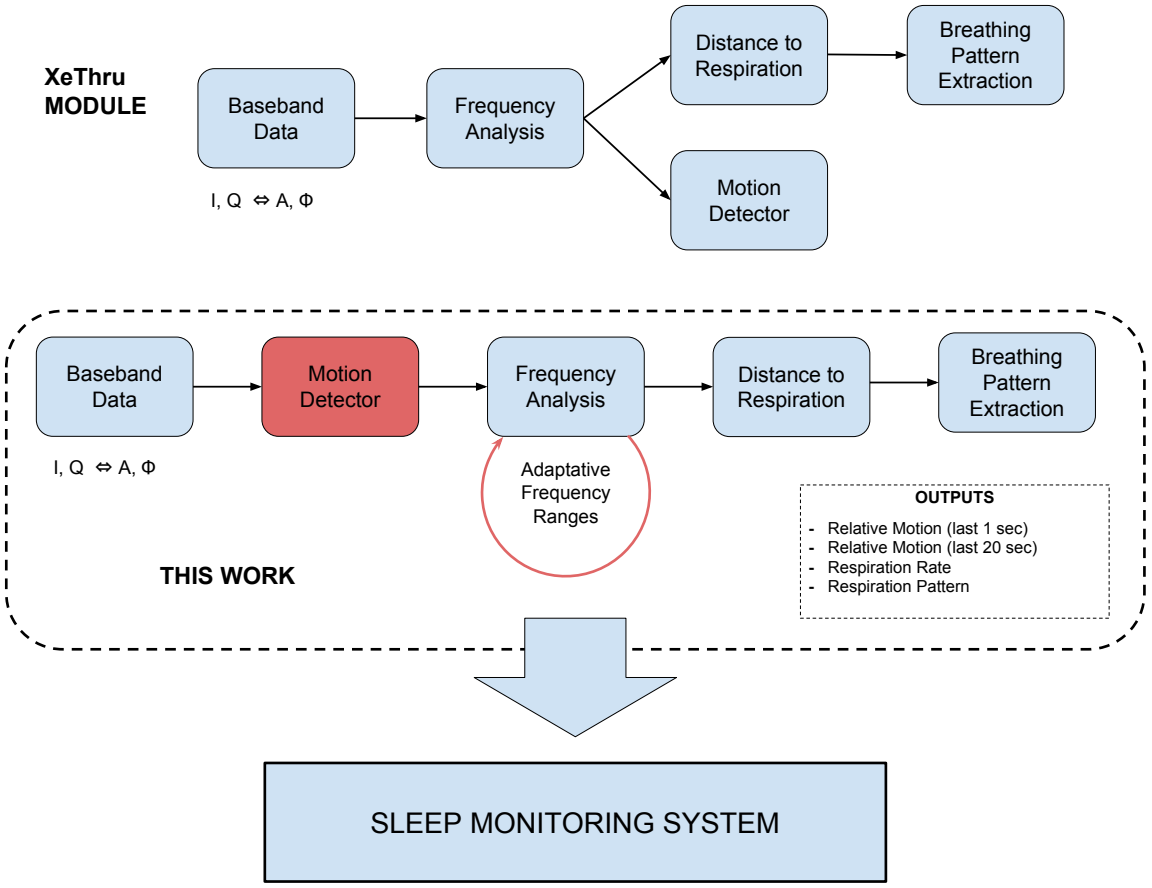


Figure 6.1: Block diagram of the Breathing Detector.

6.1. MOTION DETECTOR

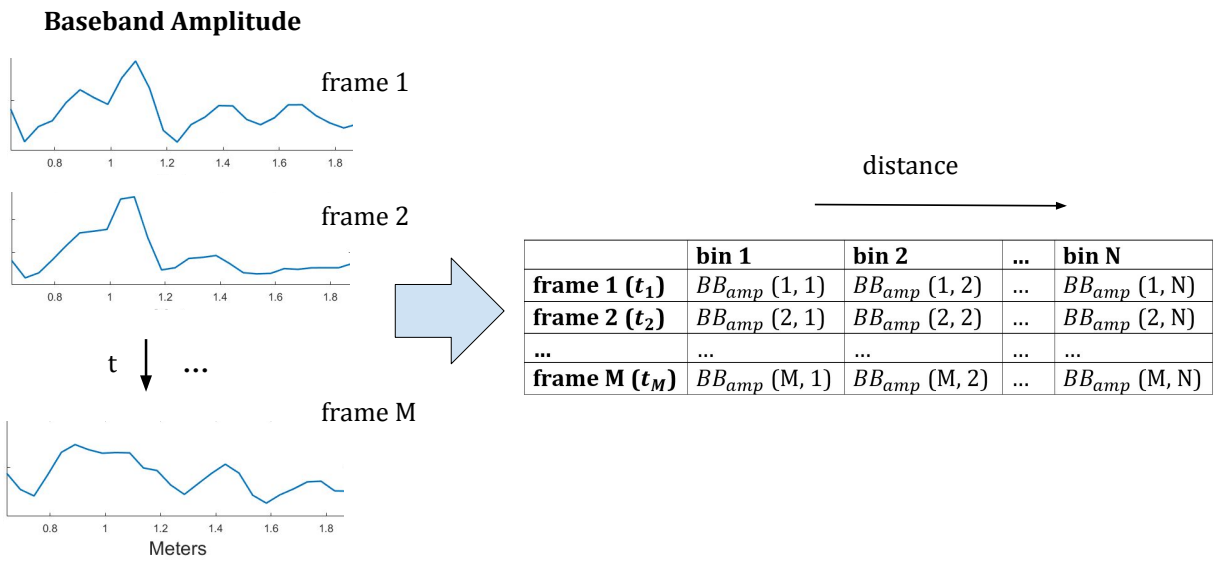
The movement detector algorithm is based on the analysis of raw baseband data. Similarly to Figure 4.1, baseband data can be represented in a matrix BB of M number of frames and N number of distance bins (Figure 6.2). This holds for both amplitude and phase baseband signals. However, for the motion detector algorithm, the amplitude signal will be employed in order to analyze the overall presence of movement in the radar's field of view.

Movement quantification is performed by integrating the difference between two subsequent frames across a specific time window as expressed in Equation 6.1:

$$MVM_{win}(t) = \sum_{t=t_i-win+1}^{t_i} [BB_{amp}(t_i) - BB_{amp}(t_i - 1)], \tag{6.1}$$

$MVM_{win}(t)$ being the movement quantity for a time window win at a specific point in time t.

Similarly as for the XeThru sensors, two movement signals are provided by the motion detector



	bin 1	bin 2	...	bin N
frame 1 (t_1)	BB_{amp} (1, 1)	BB_{amp} (1, 2)	...	BB_{amp} (1, N)
frame 2 (t_2)	BB_{amp} (2, 1)	BB_{amp} (2, 2)	...	BB_{amp} (2, N)
...	...	...	...	...
frame M (t_M)	BB_{amp} (M, 1)	BB_{amp} (M, 2)	...	BB_{amp} (M, N)

Figure 6.2: Schematics of the baseband amplitude signal for N distance bins and M frames.

output: a signal using a window of 1 sec that represents the amount of immediate movement, $MVM_1(t)$, and another one using a time window of 20 sec,$MVM_{20}(t)$, expressing the amount of movement during a not so immediate past.

Afterwards, these two signals are normalized using its maximum value of the recording in order to express a relative measure of movement, as expressed in Equation 6.2:

$$MVM_{win}(t)[\%] = \frac{\sum_{t=t_i-win+1}^{t_i}[BB_{amp}(t_i) - BB_{amp}(t_i-1)]}{MVM_{win,max}} \cdot 100, \tag{6.2}$$

The presence or absence of motion for the next steps of the breathing detector is computed by establishing a movement threshold of 10% within the 20 sec window signal. If this threshold is crossed, motion is considered to be present and no further breathing reading is performed in further steps of the breathing detector system.

This same threshold is also applied to the 1 sec window movement signal. However, since this fast movement might be momentaneous during sleep recordings, the breathing detector does not rely on the presence of movement reported by this signal. If this was the case, breathing readings would be constantly interrupted by spontaneous, fast body movements during sleep. However, both movement signals $MVM_1(t)$ and $MVM_{20}(t)$ are employed in the sleep classification system described in Chapter 7.

6.2. BREATHING DETECTION ALGORITHM

After the motion detector ensures the motion within the last 20 seconds of recording has not exceeded 10% of relative motion, the breathing detector algorithm is enabled.

Just as in the previous step of motion detection, the raw data used for this algorithm is the amplitude baseband data. First, a Fast Fourier Transform (FFT) of the BB_{amp} matrix of F frames within 20 sec and N distance bins is performed every 1 sec, giving rise to a new frequency-range matrix. The

frequency bins of this matrix are expressed in RPM instead of Hz and the number of distance bins depends on the XeThru module. Starting with an initial frequency window of 10-55 RPM, the most prominent peak within this frequency window is found from the FFT. After a peak is localized, the distance bin and frequency bin of the FFT are stored. The distance bin allows for localization of the breathing "target" (the patient). Once this distance bin is found, the system can focus on the exact baseband signal point of future frames that provides the exact periodic movement of the chest due to respiration, i.e. the breathing pattern.

An important contribution of the present work is the ability to adapt the frequency window of the frequency analyses to the RR frequency of the patient in order to improve the quality of the readings and suppress the noise at other frequencies. This frequency range is calculated when a stable reading of RR is given for at least 5 minutes, hence, no big movement has been detected by the sensor in those 5 min. This no movement condition prevents the algorithm from focusing on frequencies of parents or nurses that are next to the patient for a specific period in time. The algorithm will adapt to a new frequency range every time there has been a time window of 5 minutes with no movement.

The complete algorithm of the breathing detection system, including motion detector, is shown below. The algorithm takes into account data points in time at a sampling frequency F_s equal to the one of the radar sensor (20 Hz for X2M200 and 17 Hz for X4M200). However, the FFT is only performed every second, i.e. every F_s frames. Therefore, values of RR for a specific frame are stored for the following frames within a second assuming RR is kept constant for that second. The reason behind this decision is the fact that, in order to extract an accurate breathing pattern, the sampling frequency needs to stay in 17-20 Hz, therefore maintaining movement and RR signals at the same sampling frequency would give homogeneity to the radar data in order to provide easier signal processing.

Algorithm 1 Breathing Detection Algorithm

1: **Input:** Baseband Amplitude matrix.
2: $t = t_0$, Initial frequency window f_{win} = [10 , 50] RPM
3: **if** $MVM_{20}(t) < 10\%$, **then** *continue* , **else**, $RR_t = 0; t = t + 1$, *repeat (3).*
4: Calculate the Fast Fourier Transform of last $Fs \cdot 20$ frames (20 sec frames).
5: Find the most prominent peak within f_{win} window.
6: Retrieve frequency f_t and range value r_t values for Fs frames.
7: Localize baseband amplitude data of r_t and extract breathing pattern of $[t : t + Fs - 1]$ frames.
8: **if** movement not detected in the last 5 min **&** f_{win} not updated in the last 5 min, **then** $f_{win} = [f_t - 10, f_t + 10]RPM$
9: $t = t + F_s$; *continue in (3)*
10: **Output:** Respiration Rate (RR) vector, Breathing Pattern.

An example of the resulting FFT of 20 sec baseband data with a frequency window [29 - 49 RPM] is shown in Figure 6.3. This frequency window was updated from a previous RR value, which was 39 RPM as well.

6.3. RESULTS

The proposed breathing detector system was tested with the raw radar signals of each of the patients decribed in *Section 5.1 Patients*. The results were compared to the ones of the in-built breathing

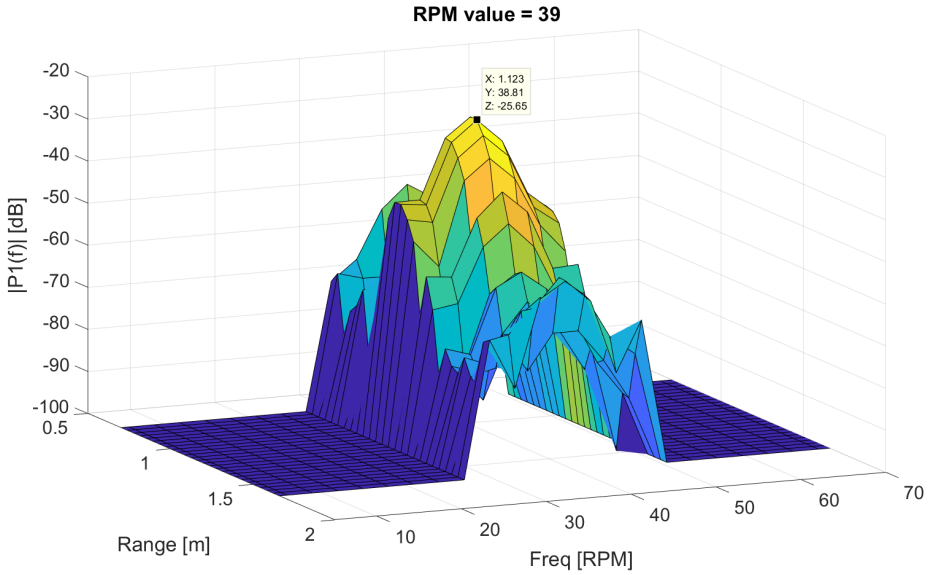


Figure 6.3: Example FFT performed for a window of 20 seconds. A peak at (39 RPM , 1.12 m) can be observed representing the breathing of a sleeping baby.

detector algorithm of XeThru modules. Before comparisons are done, it is important to notice that the signals used in this section have not been filtered to remove motion artifacts of parents/nurses interruptions. The aim of this was to recreate the hospital environment conditions as closely as possible and to study how both algorithms adapt to them.

The adaptive frequency window was not possible for one of the babies (Subject 10) due to the fact that the parents were too close to the sensor during the night of the recording. When a frequency window of 25-40 RPM was fixed, the sensor was able to properly read the RR from the baby.

In the first place, we wanted to compare the results of both the breathing pattern output of the X2M200 module, and the one resulting from the proposed algorithm. However, the researchers of the present work decided to compare RR signals, since the exact breathing waveform extracted from the radar depends on the location of the patient and the exact point of the chest tracked by the sensor. Therefore an overnight comparison of breathing patterns give very low correlation values.

RR was therefore used as the signal for testing the accuracy of both the XeThru RR output, defined from now on as RR_{xet}, and the RR output of the proposed breathing detector, defined as RR_{BD}. The reference RR signal extracted from PSG equipment is defined as RR_{PSG}

Figure 6.4 shows an example of RR_{xet} and RR_{BD} compared to RR_{PSG}. Vertical lines are caused when the RR signal goes to zero due to the presence of motion. From the graph, a great decrease in noise can be observed in RR_{BD} compared to RR_{xet}. This is mainly due to the fact that the proposed algorithm does not provide any reading until no motion is present, hence the expected decrease in noise.

Regarding the resolution of the signals, since the FFT is computed with a 20 sec window, this gives a frequency resolution of 1/20 Hz = 0.05 Hz * 60 sec = 3 RPM for the proposed breathing detector. Hence, the RR_{BD} vector can only return values that are an integer multiple of 3 RPM. In contrast, the XeThru module provides a 1 RPM resolution that may arise from extra processing steps that are unknown to the researchers of the present work. This difference in resolution is shown in Figure 6.5, showing a zoomed area of Figure 6.4.

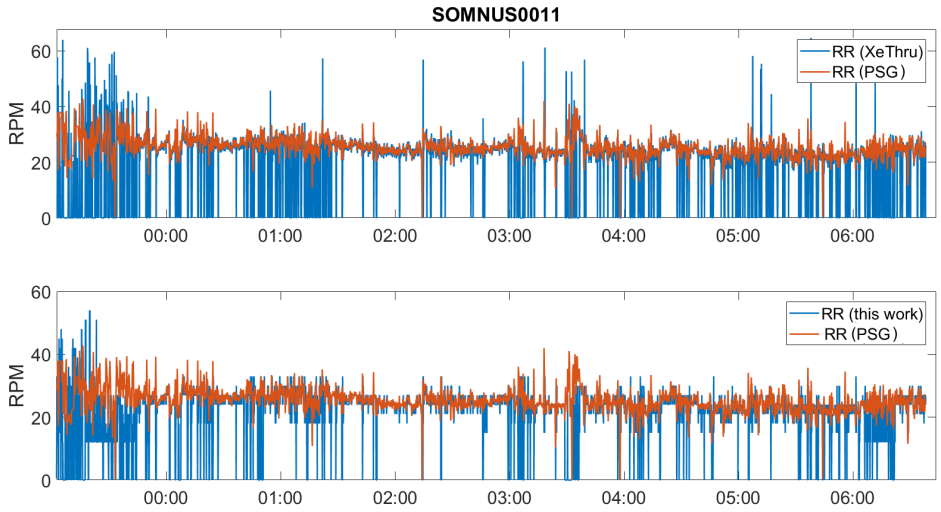


Figure 6.4: RR signals for an over-night sleep study with Subject 11. Up: RR_{xet} compared to RR_{PSG}. Down: RR_{BD} compared to RR_{PSG}.

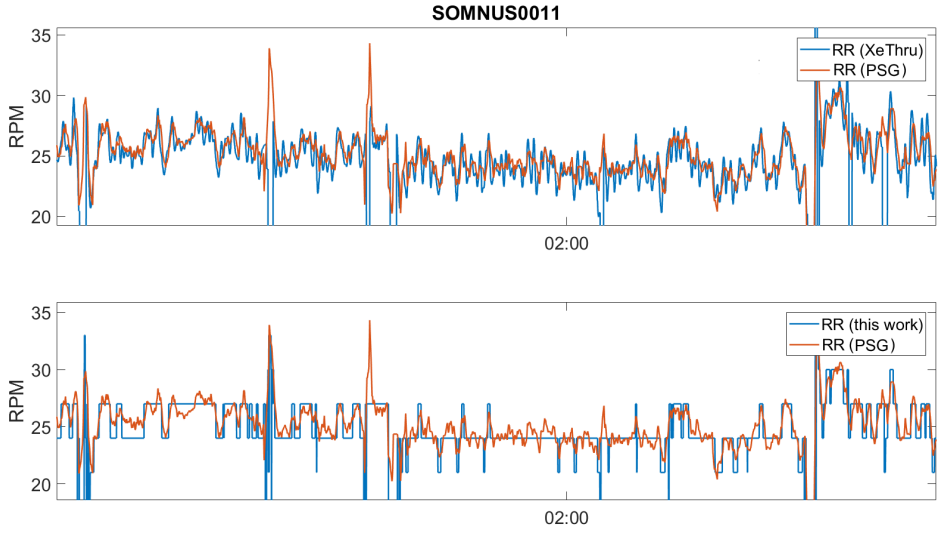


Figure 6.5: Zoomed area of Figure 6.4 showing the difference in resolution between RR_{xet} (up) and RR_{BD} (down).

In order to compare the similarity of both RR_{xet} and RR_{BD} with respect to RR_{PSG}, several parameters were studied. These were: *Mean Error* (ME), *Mean Absolute Error* (MAE), *Root Mean Square Error* (RMSE), *Standard Deviation* (STD) of the error and *Relative Error* (RE) with respect the real value. The formulas for the previously mentioned parameters are:

$$ME = \frac{\sum_{i=1}^{n} |\overline{RR_i}| - RR_i}{n} = \frac{\sum_{i=1}^{n} e_i}{n}, \tag{6.3}$$

$$MAE = \frac{\sum_{i=1}^{n} |\overline{RR_i} - RR_i|}{n} = \frac{\sum_{i=1}^{n} |e_i|}{n}, \tag{6.4}$$

$$RMSE = \sqrt{MSE} = \sqrt{\frac{\sum_{i=1}^{n}(\overline{RR_i} - RR_i)^2}{n}} \tag{6.5}$$

$$STD_{err} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(e_i - \mu_{e_i})^2} \tag{6.6}$$

$$RE = \frac{1}{n}\frac{\sum_{i=1}^{n}|\overline{RR_i} - RR_i|}{RR_i} \cdot 100 = \frac{1}{n}\frac{\sum_{i=1}^{n}|e_i|}{RR_i}, \tag{6.7}$$

$\overline{RR_i}$ being the estimation of RR at time point i, RR_i the real RR value, e_i the error between estimation and real value, STD_{err} the standard deviation of the error, μ_{e_i} the mean of the error and n the total number of points in the vector. All the previous parameters were computed with only the non-zero values of the RR signals.

Since both systems have a great amount of data in time recorded as zero, total recording time[1] is an important parameter to analyze. The expected recorded time for the proposed algorithm was expected to decrease, since the motion detector disables the RR_{BD} output as soon as movement is present in the radar detection area. Thus, we expected a lower recording time in RR_{BD} than for RR_{xet}. However, this was not the case and the recording time was greater for most of the subjects when using the proposed breathing detection algorithm. This can be observed in Figure 6.6. The total recording time of this and XeThru's approach was 172.5 and 146.3 hours respectively (about 15% increase).

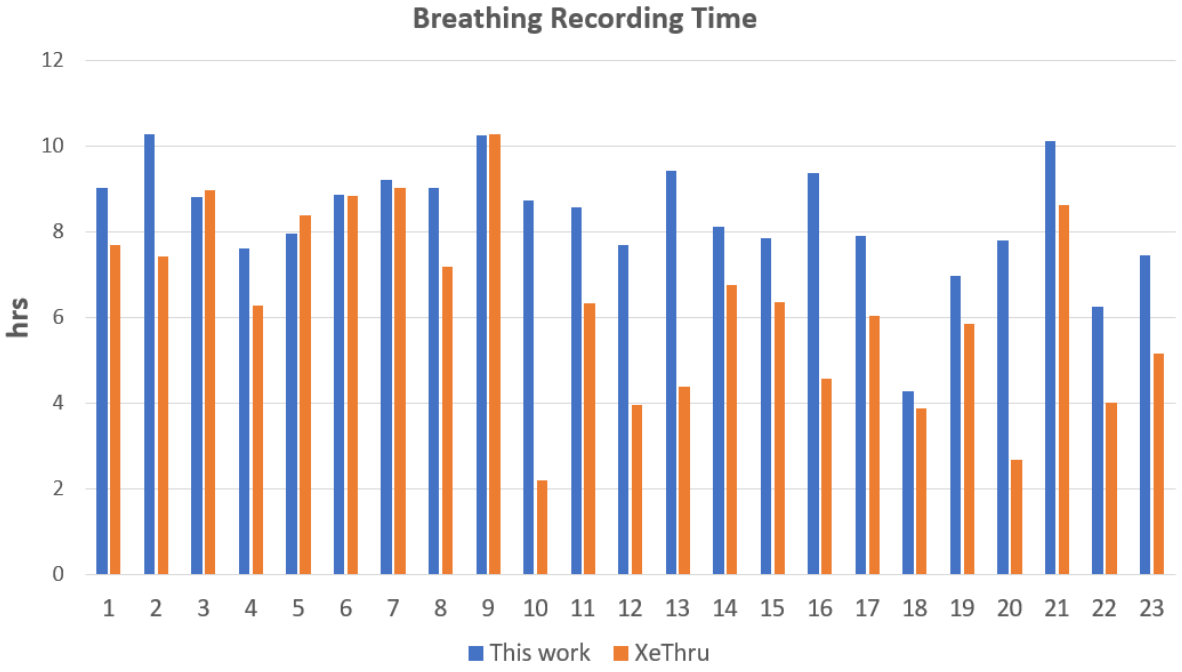


Figure 6.6: Comparison of total recording time (in hours) as extracted from this work and as extracted from the XeThru output results.

[1] *Recording time* understood as the time where RR signal was not 0, thus the time the radar actually *records respiration*. Radar is however providing baseband raw data along the time is switched on by the user.

The accuracy analysis results of RR_{BD} for each subject are presented in Table 6.1. Especially, for subject 20 we got bad results when analyzing RR_{xet}. This may be due to the fact that the Baby mode of module X4M200 was on, thus restricting the detected RPM values to the high ones of babies. However, this patient, being the oldest (14 years old) of them all, had RPM values closer to the adult ones (around 10 RPM). Therefore, the results of this patient are excluded when computing the total average and STD results of Table 6.1 and this patient will be treated as an outlier from now on.

Figure 6.7-11 are provided in order to better visualize the previously mentioned accuracy parameters per subject.

Table 6.1: Breathing Detector error performance compared to XeThru output.

	This work					**XeThru Output**				
	ME (RPM)	**MAE (RPM)**	**RMSE (RPM)**	**STDe (RPM)**	**RE (%)**	**ME (RPM)**	**MAE (RPM)**	**RMSE (RPM)**	**STDe (RPM)**	**RE (%)**
Subj. 1	-0.23	2.98	5.28	5.27	9.66%	-1.32	4.01	8.25	8.14	20.55%
Subj. 2	-1.13	2.53	4.06	3.90	15.69%	0.09	2.26	4.98	4.98	12.37%
Subj. 3	-0.60	2.13	3.51	3.46	12.62%	0.53	2.61	5.76	5.73	12.56%
Subj. 4	-3.46	4.64	7.30	6.43	28.08%	-1.16	5.12	8.87	8.80	28.66%
Subj. 5	-0.55	2.65	5.01	4.98	9.76%	-0.59	3.53	7.08	7.06	15.18%
Subj. 6	-0.77	2.18	3.65	3.57	13.72%	0.51	3.18	7.00	6.98	15.74%
Subj. 7	-1.76	3.15	5.89	5.62	17.68%	-0.99	3.89	7.74	7.67	20.50%
Subj. 8	-4.28	5.51	7.95	6.70	28.78%	-1.96	6.47	10.46	10.28	32.91%
Subj. 9	-0.35	1.34	2.31	2.28	8.02%	0.08	1.26	3.34	3.34	6.49%
Subj. 10	0.66	3.87	5.72	5.68	13.36%	4.31	7.19	11.37	10.53	19.11%
Subj. 11	-2.93	4.82	7.48	6.88	28.78%	0.42	3.02	5.47	5.45	10.87%
Subj. 12	-6.41	7.46	10.64	8.49	43.46%	-0.24	3.19	5.93	5.92	10.13%
Subj. 13	-4.29	4.86	6.78	5.25	34.47%	0.82	2.76	6.10	6.05	9.92%
Subj. 14	-0.16	1.96	2.90	2.90	12.71%	0.75	2.41	5.18	5.12	11.96%
Subj. 15	-1.99	3.22	5.43	5.05	18.52%	-0.26	2.25	4.58	4.57	9.03%
Subj. 16	-2.80	4.36	6.11	5.43	27.88%	0.21	3.19	6.22	6.22	13.16%
Subj. 17	-0.27	1.64	2.56	2.54	10.53%	0.53	1.50	3.38	3.34	7.45%
Subj. 18	-0.79	1.96	3.50	3.41	11.71%	0.07	1.14	2.78	2.78	5.39%
Subj. 19	-0.57	1.34	2.29	2.22	8.86%	0.09	0.78	2.46	2.46	3.72%
Subj. 20	0.24	2.41	3.93	3.92	16.32%	18.10	18.33	23.57	15.10	46.82%
Subj. 21	-0.30	1.93	2.91	2.89	9.30%	0.22	1.57	3.40	3.39	6.37%
Subj. 22	0.06	3.22	5.13	5.13	13.27%	1.41	2.69	5.85	5.68	8.42%
Subj. 23	-1.17	3.77	6.35	6.25	11.55%	-0.84	2.92	5.58	5.51	8.43%
Avg	**-1.55**	**3.25**	**5.12**	**4.74**	**17.66%**	**0.12**	**3.04**	**5.99**	**5.91**	**13.13%**
STD	**1.77**	**1.54**	**2.14**	**1.69**	**9.78%**	**1.23**	**1.60**	**2.34**	**2.23**	**7.39%**

ME: Mean Error. MAE: Mean Absolute Error. RMSE: Root Mean Square Error. STDe: Standard Deviation of the error. RE: Relative Error.
* Total average (Avg) and Standard Deviation (STD) are performed removing Subj. 20.

In terms of ME (Figure 6.7), also called *bias*, it seems there is a small tendency of our breathing detector to underestimate RR values. However, in most of the patients error values tend to cancel out, hence no final conclusion can be made from this bias value. RR_{xet} shows very high accuracy since bias is almost zero for every patient on which the sensor was tested, with the exception of Subject 20. However this recording was considered an outlier, as mentioned before. On average, ME values of RR_{xet} were considerably better than RR_{BD}, showing an average of almost zero bias in RR_{xet}.

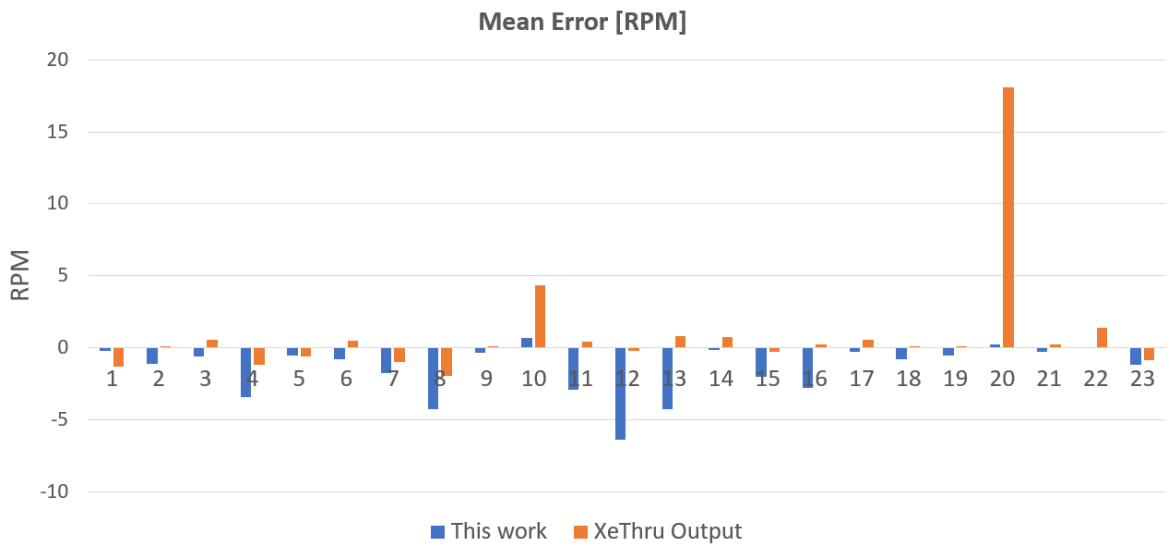


Figure 6.7: Mean Error (estimate vs. real PSG data) comparison between this work results and XeThru output results tested on 23 patients.

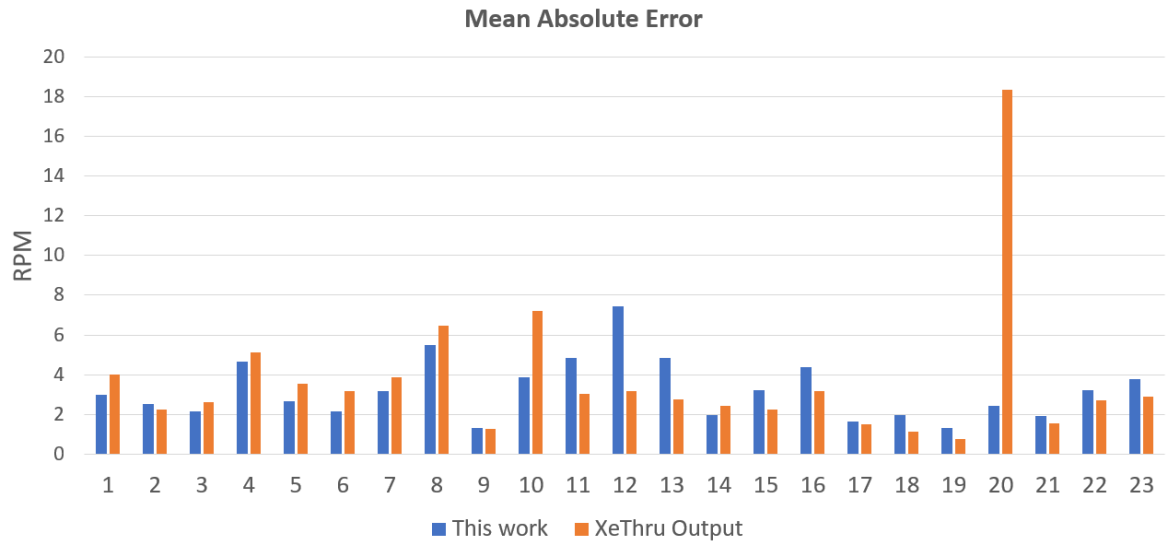


Figure 6.8: Comparison of Mean Absolute Errors (estimate vs. real PSG data) between this work results and XeThru output results tested on 23 patients.

Most publications use either MAE or RMSE in order to study the accuracy of a signal. RMSE is known for penalizing bigger errors. Therefore, sometimes the use of RMSE is preferred instead of MAE. In this work, we present both. Lower MAE errors in RR_{xet} were expected, since the resolution of RR_{xet} is higher, however this was not always the case. On the one hand, a comparison of the MAE errors (Figure 6.8) suggests worse results of RR_{xet} compared to RR_{BD} in studies where X2M200 was used (Subj. 1-9) than in studies using the new generation X4M200 sensor. However, very similar MAE values are reached for most of the studies with very similar MAE averages of 3.25 and 3.04 RPM for RR_{BD} and RR_{xet}, respectively.

In contrast, RMSE values (Figure 6.9) are considerably higher for RR_{xet} than RR_{BD} in the first half of the patients using X2M200. For the rest of patients, with the exception of 11-13, RMSE values

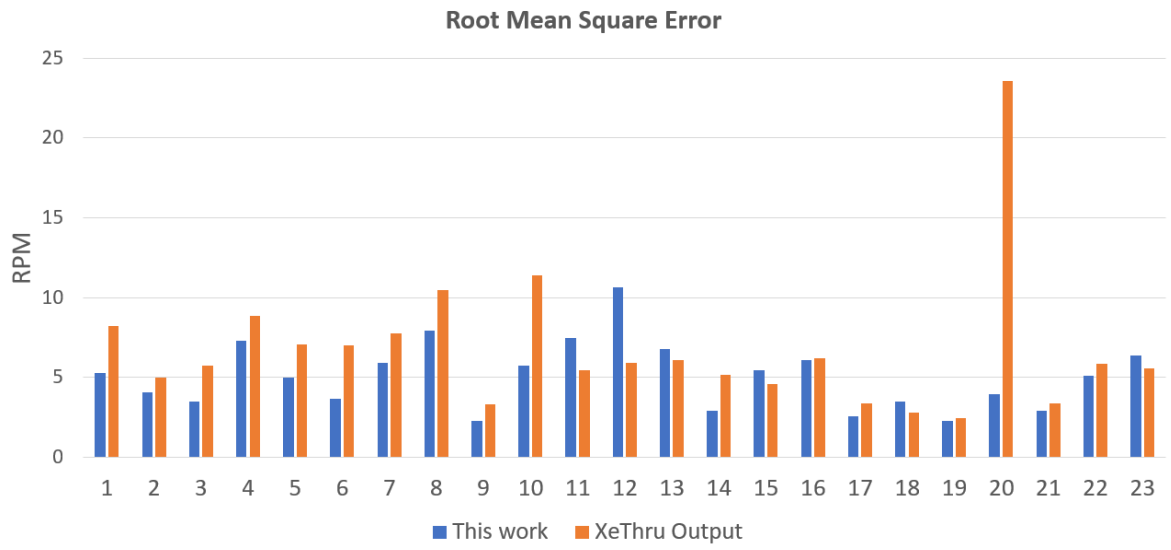


Figure 6.9: Comparison of root mean square errors (estimate vs. real PSG data) between this work results and XeThru output results tested on 23 patients.

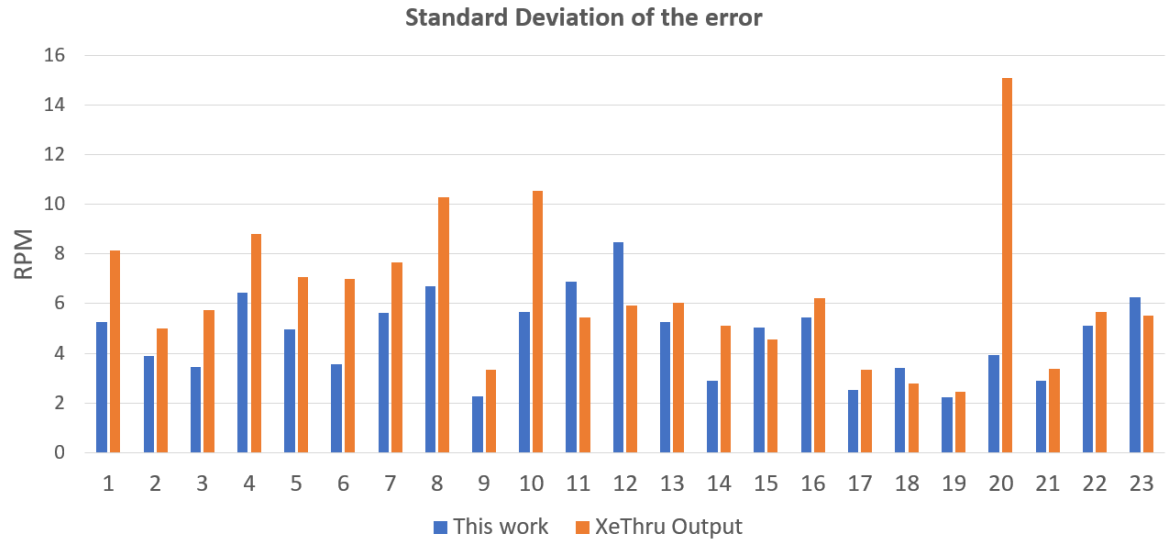


Figure 6.10: Comparison of standard deviation of the error (estimate vs. real PSG data) between this work results and XeThru output results tested on 23 patients.

look quite similar, without taking into account subject 20 (outlier). Some of the patients having higher MAE errors in RR_{BD} have now lower errors of RMSE (subj. 16,17,19,21,22). This may be due to the fact that even though values of RR_{xet} were in general closer to RR_{PSG} than with RR_{BD} with X4M200 module, big errors in RR_{xet} were slightly more common, since RMSE tends to penalize big errors within a dataset. This assumption can be supported by analyzing the STD of the error results (Figure 6.10), which shows that the errors were more dispersed with RR_{xet} in most of the subjects. On average, better values of RMSE and STD of error are shown for RR_{BD} in Table 6.1.

Finally, studies of relative errors (RE) are important due to the fact that the population has quite some differences regarding their mean RR values, which were around 10-15 RPM in old children and around 40 RPM with babies. Figure 6.11 shows considerably bigger RE values in subjects 11,12,13

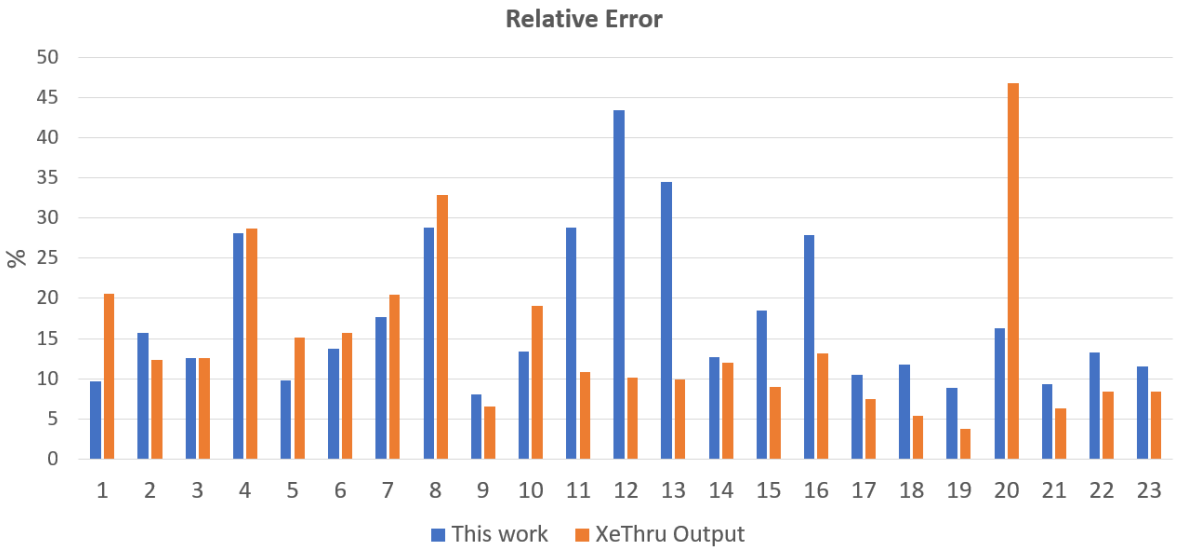


Figure 6.11: Comparison of relative error (estimate vs. real PSG data) comparison between this work results and XeThru output results.

and 16, and similar RE values for the rest of the patients. Again, we can see a slightly better performance of our breathing detector system when compared to X2M200 studies. The final average of RE values for RR_{BD} and RR_{xet} was 17.6% and 13.1% respectively.

When taking a closer look to the patients with the highest relative errors, subjects 4, 8, 11, 12, 13 and 16 were studied in order to detect whether they were recorded in the same room, suggesting high errors may come from experimental settings differences. However, these patients were recorded in different rooms of the two previously mentioned centers, thus, suggesting there is not one specific room presenting especially bad results. No final conclusion can be made regarding experimental settings influencing the quality of recordings.

The previous results were divided into children of less than 1 year old age (babies) and children of more than 1 year old (older children) in order to study whether the fact of monitoring a baby gave worse results than with any other patient. These groups were comprised of 10 patients (younger than 1 year old) and 13 patients (older than 1 year old). The results are shown in Table 6.2. From this table we can observe a decrease in the error in the older population for both algorithms, with a more significant decrease in XeThru's algorithm. This might be related to the fact that babies are specially smaller targets than normal patients, thus their chest movement is slightly more difficult to detect. Nevertheless, this decrease in error is more drastic in XeThru signals, which may suggest the proposed algorithm in this work is less sensitive to age differences. Meanwhile, recording time remains more or less the same in both populations with our proposed algorithm, therefore robusticity is mantained regardless of patient age. However, this recording time significantly decreases in the oldest population with XeThru's algorithm.

A more detailed study of algorithm performance with respect to age could have been done by studying this error in the 7 populations of patients described in Table 5.2. However, the low number of patients per group makes this performance analyses unreliable. Therefore, we decided to focus on comparing the performance between babies and older patients.

In conclusion, error analyses of both RR_{BD} and RR_{xet} show the high accuracy of both algorithms, the in-built one of XeThru modules and the one proposed in this work. Even though RR_{BD} has con-

Table 6.2: Breathing detector error performance compared to XeThru Output with two study populations.

	This work					XeThru Output				
	ME (RPM)	**MAE (RPM)**	**RMSE (RPM)**	**STDe (RPM)**	**RE (%)**	**ME (RPM)**	**MAE (RPM)**	**RMSE (RPM)**	**STDe (RPM)**	**RE (%)**
Avg (less 1 y)	-1.65	3.77	5.84	5.39	18.38%	0.15	3.90	7.25	7.10	16.29%
STD	2.28	1.73	2.35	1.72	11.27%	1.78	1.81	2.46	2.29	9.00%
Avg (older 1 y)	-1.47	2.82	4.53	4.21	17.05%	0.10	2.33	4.94	4.92	10.50%
STD	1.31	1.29	1.83	1.54	8.81%	0.53	1.00	1.70	1.68	4.62%

ME: Mean Error. MAE: Mean Absolute Error. RMSE: Root Mean Square Error. STDe: Standard Deviation of the error. RE: Relative Error.
* Total average (Avg) and Standard Deviation (STD) are performed removing Subj. 20.

siderably higher relative errors, it has shown to be less sensitive to age differences. This proposed algorithm also has the advantage of giving more robustness to the signal at the expense of RR resolution, however the recording time is increased drastically. Therefore, it can be concluded that RR_{BD} can be a very good estimate of RR that will serve as the input of the sleep monitoring system described in the next section.

6.3.1. RESULTS BREATHING PATTERN

Figure 6.12 shows an example of a zoomed in image comparing the breathing pattern extracted from the breathing detector algorithm and the reference breathing pattern from thorax data. An strict analysis of this patterns throughout the night was not possible since the amplitude of the waveform from the radar changes constantly depending on the exact location of the chest the radar is focusing on. Therefore, correlation or error analyses are not useful to compare the similarity of these waveforms. Instead the researchers decided to focus on the frequency of this waveforms (RR) in order to study the accuracy of the signals.

Breathing pattern was not considered as one of the features for the sleep classification system due to the variability of the waveform due to noise and radar characteristics.

When inspecting the breathing pattern signal provided by the breathing detector, it was very interesting to see frequent apneas in one of the subjects of the study. Figure 6.13 shows an example of one of these apneas detected by the radar sensor. This interesting finding demonstrates the capabilities of XeThru modules to detect apnea events, which opens a new door in the application of apnea diagnostics.

In the coming chapter, the complete sleep monitoring system based on signals from the radar modules will be described. This system uses the signals extracted from the breathing detection system (motion + breathing) as input variables of the sleep-stage classification.

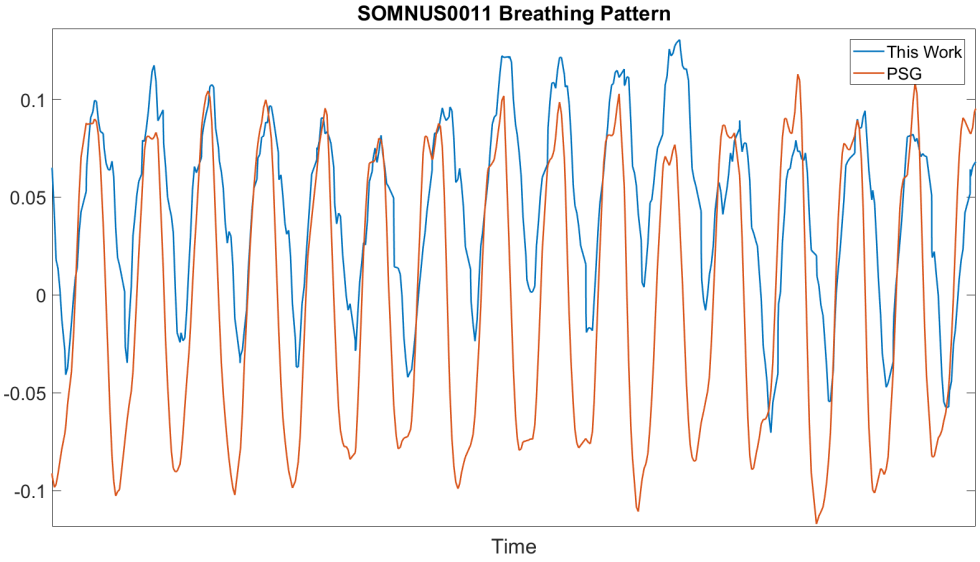


Figure 6.12: Breathing pattern of Subject 1 extracted by breathing detector (blue) compared to PSG referece signal of thorax (orange).

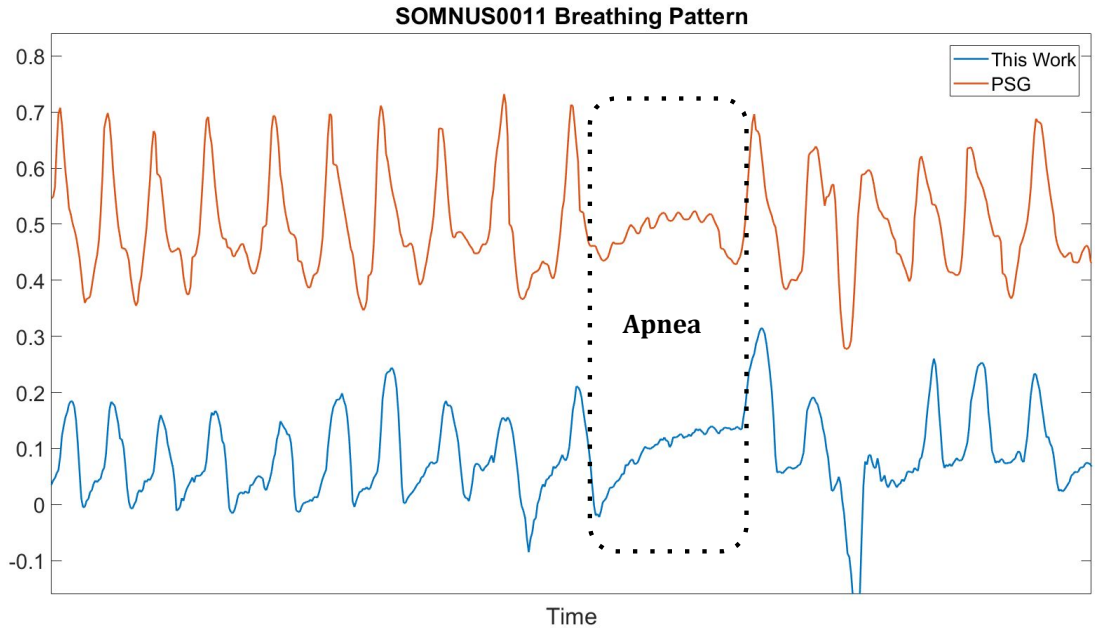


Figure 6.13: Breathing pattern of Subject 1 showing an apnea event.

7

SLEEP MONITORING SYSTEM

Once the raw radar data has been processed in order to acquire movement and respiration information from the motion detector and the breathing detector, these outputs signals are used as the inputs for the sleep monitoring system.

The sleep monitoring system has two main tasks:

1. To classify between *Sleep* and *Awake* events.
2. To classify *Sleep* events into *REM* and *NREM* events.

Movement information was the key feature to classify sleep/awake for this study. A challenging aspect of classifying *Sleep*/*Awake* measurements based on motion is the fact that during the time it takes to fall asleep, patients do not move much. This effect makes nowadays sleep/awake classification systems overestimate the time of sleep [57, 58]. Another big challenge when addressing this problem is that nurses and parents interruptions are misguiding the radar system to classify these motions as patient motions during awake. The details and performance testing of the sleep/awake classification algorithm presented in this work are described in this section.

Regarding REM/NREM classification, the key main difference between these two phases at which the researchers focused their attention, was the breathing irregularity characterizing REM phases, along with the greater amount of movements, as discussed in previous chapters. Therefore both breathing and movements information was used for the REM/NREM classification. Due to the complexity of this classification problem, different analyses were performed in order to find the best approach for this classification. One of these analyses focused on studying the classification performance when using both overall respiration and movement information in contrast to only respiration variability, which is known to be the key characteristic difference between REM/NREM in young patients [26].

7.1. SLEEP/AWAKE CLASSIFICATION

Classification into *Sleep* and *Awake* events was done through the motion signal that described if the amount of relative movement within the last 20 seconds, $MVM_{20}(t)$, crossed a 10 % threshold.

The reason why $MVM_1(t)$ was not used to address this task was because of the assumption that fast movements do not really give a lot of information about whether the subject is sleep or awake. Instead, an accumulation of these fast movements due to awakenings would give a high signal in $MVM_{20}(t)$, and therefore provide information on sleep or awake states.

The algorithm to detect sleep or awake events was performed at every point of movement signal, therefore, at the same baseband sampling frequency (17-20 Hz). This was done by detecting clusters of high $MVM_{20}(t)$ signal within a time window of 30 sec. Meanwhile, when certain data points were scored as awake less than one minute after the last awake event happened, these points were re-scored as awake again. This condition was set in order to avoid results of apparent fast alternation between awake and sleep events. This algorithm, which we will refer to as *S/A algorithm*, is described below.

Algorithm 2 Sleep/Awake (S/A) classification algorithm

1: **Input:** $MVM_{20}(t)$: Relative amount of movement for the last 20 seconds.
2: $t = t_0$.
3: **if** $MVM_{20}(t-30:t) < 10\%$, **then** $sleep(t) = 0$, **else**, $sleep(t) = 1$.
4: **if** $sleep(t) = 1$ & $t_{lastawake} - t > 60$, **then**, $sleep(t_{lastawake} : t) = 0$.
5: $t = t + 1$, *go to (4)*.
6: **Output:** $sleep(t)$ vector showing sleep/awake state.$sleep = 1$ when sleep, $sleep = 0$ when awake.

The S/A classification was performed with all the 23 patients and sensitivity, specificity and precision of scoring sleep were studied along with the overall accuracy of the classification for each of these subjects. Furthermore, a final estimate of Total Sleep Time (TST) was compared to the real TST of the patients according to the PSG data. Table 7.1 shows the results of these analyses.

When analyzing these classification parameters, is important to notice that sleep covers the great majority of time of these sleep studies. This fact has a great influence on the results of sensitivity (the *True Positive Rate* (TPR) of *Sleep* events), precision (the *Positive Predictive Value*) and accuracy. In this type of unbalanced classification problem, these three parameters tend to show a very high, misguiding percentage value. Therefore, specificity of sleep, also called *False Positive Rate* (FPR), is an important parameter to analyze, since it is what provides the information about the percentage of awake events that are being correctly classified as awake. The formulas of the previously mentioned parameters are:

$$Sensitivity_{sleep} = TPR = \frac{TP}{P} = \frac{TP}{TP+FN} \tag{7.1}$$

$$Specificity_{sleep} = FPR = \frac{TN}{N} = \frac{TN}{TN+FP} \tag{7.2}$$

$$Precision_{sleep} = PPV = \frac{TP}{TP+FP} \tag{7.3}$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{7.4}$$

TP being the number of true positives, TN the number of true negatives, FP the number of false positives and FN the number of false negatives. In this case, positives are understood as sleep events and negatives as awake. It is important to notice that specificity of *sleep* is the same as sensitivity of *awake*, and viceversa.

Table 7.1: Sleep/Awake Classification performance

	Sens.(sleep)	Spec.(sleep)	Prec.(sleep)	Acc.	TST(real)	TST(est)	Error(TST)
Subj. 1	94.77%	60.94%	91.76%	88.72%	82.11%	84.81%	2.69%
Subj. 2	99.84%	29.87%	87.35%	87.88%	82.91%	94.76%	11.86%
Subj. 3	89.41%	98.35%	99.87%	90.00%	93.40%	83.62%	9.78%
Subj. 4	96.00%	65.78%	87.06%	87.10%	70.57%	77.82%	7.25%
Subj. 5	91.88%	83.35%	96.95%	90.62%	85.19%	80.74%	4.45%
Subj. 6	94.28%	89.93%	97.75%	93.51%	82.27%	79.35%	2.92%
Subj. 7	97.69%	85.43%	94.73%	94.35%	72.82%	75.10%	2.28%
Subj. 8	90.71%	87.48%	95.94%	89.95%	76.56%	72.38%	4.18%
Subj. 9	98.94%	37.44%	89.09%	88.97%	83.78%	93.04%	9.26%
Subj. 10	95.82%	46.16%	78.19%	79.34%	66.82%	81.88%	15.07%
Subj. 11	97.80%	21.15%	76.19%	76.39%	72.07%	92.51%	20.44%
Subj. 12	96.07%	84.91%	96.09%	93.78%	79.44%	79.42%	0.02%
Subj. 13	98.50%	65.83%	93.16%	92.80%	82.54%	87.27%	4.73%
Subj. 14	95.32%	69.87%	96.38%	92.62%	89.38%	88.40%	0.98%
Subj. 15	98.95%	66.29%	86.02%	88.40%	67.70%	77.88%	10.18%
Subj. 16	94.92%	50.39%	89.03%	86.43%	80.93%	86.28%	5.35%
Subj. 17	98.97%	36.45%	95.10%	94.32%	92.57%	96.34%	3.76%
Subj. 18	97.76%	57.51%	90.76%	90.12%	81.02%	87.27%	6.25%
Subj. 19	99.76%	32.65%	90.30%	90.55%	86.27%	95.31%	9.04%
Subj. 20	98.58%	53.17%	72.86%	78.62%	56.05%	75.84%	19.79%
Subj. 21	98.70%	58.05%	94.21%	93.56%	87.37%	91.53%	4.16%
Subj. 22	86.95%	95.67%	99.28%	87.98%	87.24%	76.33%	10.92%
Subj. 23	90.50%	81.50%	97.83%	89.62%	90.22%	83.46%	6.76%
Avg	**95.74%**	**63.40%**	**91.13%**	**88.94%**	**80.40%**	**84.41%**	**7.48%**
STD	**3.63%**	**22.54%**	**7.31%**	**4.91%**	**9.18%**	**7.15%**	**5.45%**

In terms of TST analysis, as expected, the S/A classification algorithm tends to overestimate the sleep time in most of the subjects of this study. This can be explained by the fact that the lack of movement does not really ensures the patient is sleeping, since awake-sleep transition does not happen exactly at the moment the patient stops moving. This represents one of the main challenges when classifying sleep against awake based on motion signals.

Another challenge to overcome is the fact that motion artifacts produced by nurses and parents were accounted as awake by our S/A algorithm. Figure 7.1 shows the sleep/awake classification diagram estimated by the S/A algorithm compared to the real sleep/awake diagram for subject 18. Here, the similarities between the diagrams and the effect of interruptions in the S/A algorithm can be observed. Appendix A shows each of these diagrams for the rest of the patients, where we can see the similarities in waveforms of the sleep/awake diagrams.

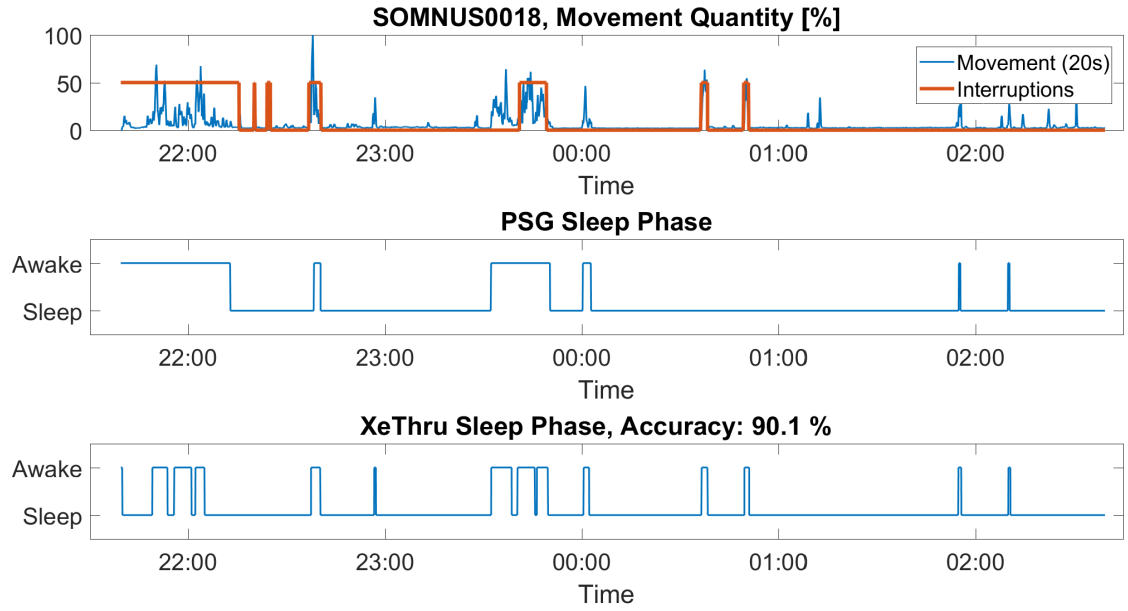


Figure 7.1: Sleep/Awake diagram of subject 18 showing presence/absence of interruptions and relative movement (up), real sleep/awake events (center) and estimated sleep/awake events (below).

7.2. SLEEP CLASSIFICATION

Unlike sleep/awake classification, in which our own algorithm was executed in order to score sleep against awake measurement, REM/NREM classification, defined here as R/N, is addressed as a *Statistical Classification* in terms of *Machine Learning*. Statistical classifications are processes in which an observation, defined by one or multiple features, is assigned to a *class*. In order to do so, a previous *training* step is required for the classifier to study a dataset, called the *training set*, comprised by labeled measurements. In our particular classification problem, the observations are time windows of 2 min, called *epochs*. Each of the observations' features describe information of respiration and movement data of the patient within those two minutes, while classes are labeled as either *REM* or *NREM*. The features of REM and NREM epochs will serve as the training set that the classifier needs to score new epochs as REM or NREM.

There are many different types of classifiers that are already well explored, each of them having different classification algorithms. Some of the most popular ones will be described and studied

for our N/R classification in this section. The main challenge of statistical classification is defining what is considered a "good" training set of data points with different features that provides as much information as possible about a certain class, giving rise to a high performance classification when being tested.

Features are normally expressed in terms of pattern recognition as *dimensions*. This is due to the fact that classifiers work in a *feature space* where the observations are defined as their number of features. Classifiers divide the feature space according to their own classification (or *discrimination*) algorithms. Figure 7.2 shows the representation of a very basic classification problem; two classes in a two-feature space [126]. However, most of classification problems work with observations with many different features that increase the complexity of the classification. Hence, pattern recognition researchers share a common aim of finding approaches that can decrease the dimensionality of the data before the actual classification takes place.

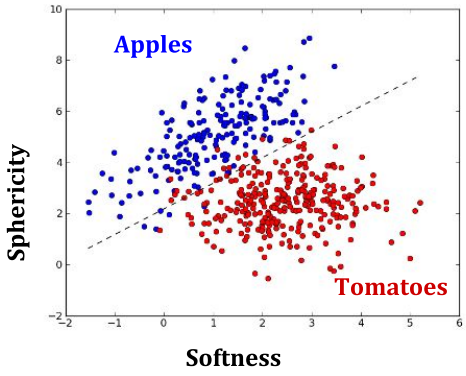


Figure 7.2: Basic example of a two-dimensional feature space with two classes (apples/tomatoes) and two features (sphericity/softness).

In contrast to S/A classification, which was performed for each of the data points of the radar output signals, R/N classification was performed by dividing the studies into 2 min epochs, as previously mentioned. Conventional PSG sleep classification through EEG and other PSG features is performed by using 30 sec epochs. However, the challenge of performing sleep classification with breathing parameters is the low frequency of this signal, which makes it difficult to reliably detect differences between sleep states in such a short time. For instance, a patient presenting an RR of 20 RPM would only breath 10 times in 30 sec, which is not enough to observe the characteristic differences in breathing patterns of REM state.

This R/N classification was performed only taking into account the epochs with a maximum limit of 15 sec where high movement ($MVM(t) > 10\%$) is present. When too much movement was detected the epoch was considered as an artifact.

In order to avoid the accumulation of classification errors of the previous Sleep/Awake classification, R/N classification was firstly studied with only *real* REM and NREM epochs as training sets of our classifiers. This way, we could really see their real classification performance. Classifiers were trained using either the strongest feature (the one who varies the most between REM and NREM epochs) or all five different features coming from movement and breathing signals (further explained in this section).

7.2.1. FEATURE GENERATION

As explained before, the key characteristic of R/N classification of the present work is detecting the amount of movement and the irregularity of respiration. Five different features regarding this two signals were investigated in order to serve as input of our classifiers. These are listed in Table 7.2.

Table 7.2: List of input features for sleep classification analyses.

	Name	Description
1.	$Var_{RR,Norm}$	Normalize variance of the RR_{BD} signal within epoch
2.	$\mu_{RR,Norm}$	Normalize mean of RR_{BD} signal within epoch
3.	$\mu_{MVM_{20}}$	Average of MVM_{20} signal within epoch
4.	μ_{MVM_1}	Average of MVM_1 signal within epoch
5.	FMr_{10}	(Fast Movement ratio) Ratio of time within last 10 min governed by fast movements

Regarding respiration features, two main factors were to be studied in the classification: (i) whether RR variance within the epoch is able to determine the R/N phase, and (ii) whether the mean RR within the epoch increased when REM was present. Since both these features depend on the age of the baby, as they increase when age decreases, a normalization was required. In the first place, Var_{RR} of epoch number i is normalized using the RR mean $\mu_{RR,i}$ registered in epoch i. In contrast, $\mu_{RR,i}$ was normalized using the overall RR mean μ_{RR} along the over-night study. Equation 7.5 and Equation 7.6 describe the computation of these two features.

$$Var_{RR,Norm}(i) = \frac{Var(RR(t:t+2min))}{\mu(RR(t:t+2min))} \tag{7.5}$$

$$\mu_{RR,Norm}(i) = \frac{\mu(RR(t:t+2min))}{\mu(RR(t))} \tag{7.6}$$

Taking into account that mean and variance of a signal x are defined as:

$$\mu(x(t)) = \frac{1}{n}\sum_{t=1}^{n} x(t) \tag{7.7}$$

$$Var(x(t)) = \frac{1}{n}\sum_{t=1}^{n} (x_t - \mu)^2 \tag{7.8}$$

Three other features were used in our sleep classification problem in order to study whether or not there was an increase in motion of the patient during REM phase. To do so, three features were added to our list of classification features: the average amount of relative movement present during the last 20 sec signal over a 2 min epoch, $\mu_{MVM_{20}}$, the average amount of relative movement present during the last 1 sec signal, μ_{MVM_1} and the relative time within the last 10 min epoch in which fast movement defined as $MVM_1(t) > 10\%$ has been present. This latter parameter is the way to account for movement in the previous epochs. These three movement features are defined for an epoch i as:

$$\mu_{MVM_{20}}(i)) = \mu(MVM_{20}(t:t+2min)) \tag{7.9}$$

$$\mu_{MVM_1}(i)) = \mu(MVM_1(t:t+2min)) \tag{7.10}$$

$$FMr_{10} = \frac{no.points(MVM_1(t+10min) > 10\%)}{no.points(t:t+10min)} \tag{7.11}$$

Even though we used these five features for our classification process, which describe information about overall respiration and movement patterns, further analyses were performed using only respiration variability (feature 1) alone. The aim of those analyses was to find out if this feature alone was strong enough to classify R/N events, as literature suggests [26].

There is one last important point to take into account regarding feature generation. As sleep is a cyclic process, is easily seen that sleep state at a time point i has a great influence in the consecutive sleep state of points $i+1, i+2$,... etc. Therefore, it was considered the possibility of adding *previous scored state* as another feature of the classification. However, this can be considered a relatively "dangerous" feature to take into account, since it has a great statistical power against the rest of the features of our classification problem. For instance, a classifier may detect that around 80% of epochs had the same previous epoch labeled in its training set, therefore when testing a new data set, it is highly probable that the classifier would tend to classify all the epochs as the first one. Consequently, we did not consider previous score sleep phase as a sleep feature for this first version of the sleep classification system.

7.2.2. DIMENSIONALITY REDUCTION

When performing statistical classification with multiple features, it is common practice to perform *dimensionality reduction* in order to decrease the number of dimensions, thus decreasing the complexity of the classification. *Feature extraction* is a common technique used for dimensionality reduction. It relies on the principle of transforming the current feature space into a space with fewer dimensions. Two common approaches of feature extraction were used for our sleep classification problem.

On the one hand, *Principal Component Analysis* (PCA) performs a linear transformation of the feature space in the direction of the highest variance. To do so, the process starts with building the covariance matrix of the dataset and then computing its corresponding eigenvectors, also called the *principal components*. The largest eigenvectors are therefore the new orthogonal representation of the feature space with those variables retaining the highest variance [127].

On the other hand, *Linear Discriminant Analysis* (LDA), also called Fisher's discriminant analysis, is another popular technique for dimensionality reduction. In this case, the ratio of *between-class variance* against *within-class variance* is maximized in order to achieve the direction of the biggest separability of data with the lowest data variation within the same class [11, 128].

The graphs shown in Figure 7.3, adapted from [11], show the representation of two classes in a two-dimensional space. When creating a 1-D projection of two classes by joining their means (Figure 7.3a), the classes cannot be fully separated as some data points overlap on the projection. When LDA is performed in Figure 7.3b the projection greatly improves the separability. LDA can be used both as dimensionality reduction technique and as a classifier.

Even though both techniques, PCA and LDA, transform the given data into a new feature space, the great difference between both techniques is that PCA does not take into account the classes of

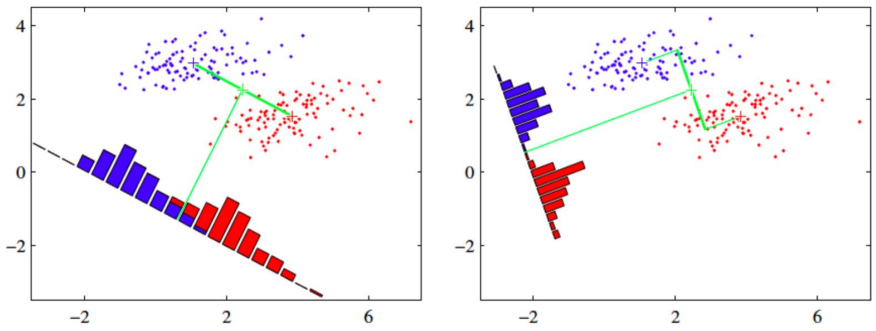


Figure 7.3: Example of LDA analysis in (b) showing a higher separability of classes [11].

the datasets in order to model the difference between them, thus performing a so-called *unsupervised* transformation, and LDA does take into account these classes, therefore being a *supervised* approach. Even though it may look logical to choose LDA over PCA approaches, PCA is still one of the most popular techniques in feature extraction. Some researchers claim that PCA can outperform LDA when a small training set is used since it is less senstive to different training sets [129].

In this work, both PCA and LDA were tested in order to compare the performance of the sleep classification when using each approach. Both approaches reduced the dimensions of the classification from 5 to 2.

7.2.3. POPULATIONS

Classification performance was studied with three different populations: (i) babies younger than 1 year old, (ii) older children and (iii) all patients together. We will refer to these three populations as population Y (young), with 9 patients in total, population O (old), with 11 patients, and population A (all), with 20 patients.

As explained in Section 5.1, subjects 4, 11 and 18 were discarded from the study due to lack of REM events or due to neurological disorders. However, it is important to take into account that the population of this classification study was not as ideal as if it would have been a population of 100% healthy babies. In contrast, since these patients were scheduled by their doctors for PSG studies in order to closely observe sleep behaviors, it is very possible that some of them do not present the typical sleep patterns that healthy children of their age would have presented.

The best possible division of populations would have been made according to Table 5.2 in Section 5.1. However, the low number of patients per group makes it very difficult to perform a reliable classification data set. Therefore, only two groups were formed out of this set of patients. Heterogeneity of age in population O will be taken into account in the analyses. However the main goal of the researchers was to focus on population Y since it is the closest to the ideal study population of preterm babies.

7.2.4. CLASSIFIERS

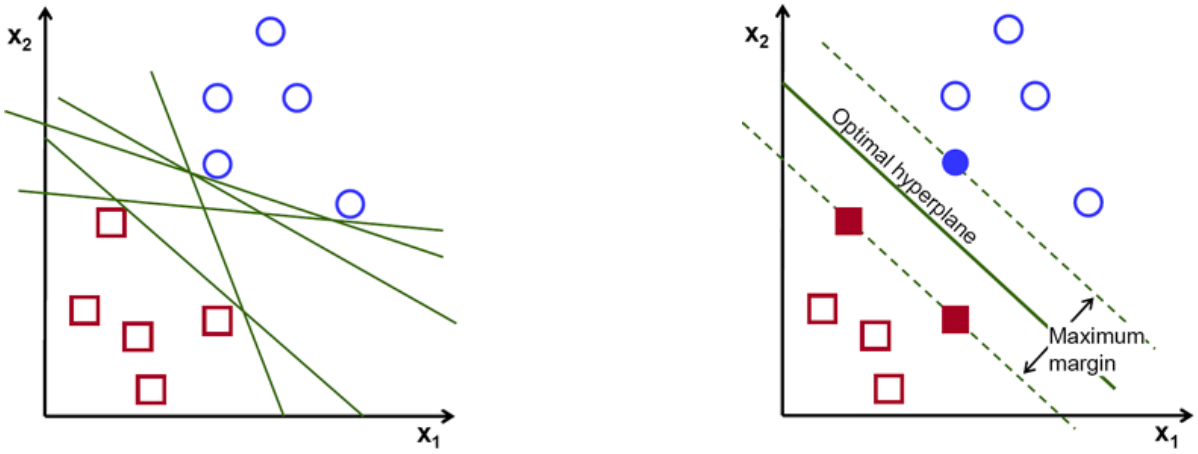
A set of common classifiers used in pattern recognition applications were chosen to test their performance in the R/N classification problem. These were implemented using the MATLAB Toolbox *PRTools 5* (http://prtools.org/), developed by the Pattern Recognition Research Group of the TU

Delft (*DelftPR*).

In this specific problem, we chose three different classifiers: (Fisher) Linear Discriminant Analysis (LDA), Support Vector Machines (SVM) and K-Nearest Neighbor (kNN).

The Fisher LDA classifier, which relies on the previously described principle of maximizing the separability between two classes, can be used as both dimensionality reduction approach and as classifier. Here, we want to test whether its application as classifier has good results in separating REM and NREM events. In order to distinguish between its applications as both classification and dimensionality reduction, we will refer to *Fisher* when using LDA as classification method.

SVM is a supervised technique that looks for a hyperplane that maximizes the distance (or margin) between two different classes. In order to understand it better, Figure 7.4 shows an example of SVM implementation in a 2D (two features) classification. Beginning from a two sets of data points belonging to two different classes, many straight lines can serve to separate them (Figure 7.4a). However, some are better than others, since the further the line is from both classes, the less sensitive it is to error classifications. Thus, a line which is the furthest to both classes is optimized in order to maximize the *margin* between classes, which is defined as double the distance to the first points of each class, defined as *support vectors* (Figure 7.4b). This example can be translated to higher dimensions, where *hyperplanes* are computed, instead of lines [12].



(a) Two classes classification infinite number of possible straight lines acting to classify them.

(b) SVM finds the optimal line with highest margin to each class. Filled colors are the *support vectors*.

Figure 7.4: Implementation of SVM in a 2D feature space [12]. x_1 and x_2 correspond to feature 1 and feature 2.

Fisher and SVM are two very common linear classifiers used in many pattern recognition applications. Fisher is considered a relatively simple classifier, in contrast to SVM, which is considered as a more complex classifier that requires a much higher computational cost (about 600 times longer processing times in this work when being used with five features). Another difference between them is that Fisher assumes a normal distribution of classes with equal covariance, while SVM does not need to assume the distributions of the data sets, what makes it more robust with classes that have completely different distributions. They both were chosen for this classification because (i) linear classifiers have reported good results in sleep classification problems (previously discussed in Section 2.2), and (ii) because of their different level of complexity. The objective behind comparing both performances is to study whether the use of a more complex classifier is actually providing any value to the classification at the expense of a longer processing time.

kNN was chosen as the third classifier used in sleep classification due its high simplicity and its popularity among pattern recognition approaches. kNN classifies a test data point by finding the closest k points in the training set, defined as *neighbors*, studying the predominant class within this labeled training set group and assigning it to the test point. The main concern of this classifier is the choice of k value, since low k values generally increase the over-fitting[1] effect of the classifier. In the kNN algorithm implemented in this work by PRTools, k values are automatically selected by minimizing the classification error by a leave-one-out technique (take out one data point out of the training set and test the algorithm for different values of k, and so on and so forth with every data point of the training set) [130].

Just as SVM, kNN is a *non-parametric* classifier, which means it does not assume any distribution in the classes. The main difference with respect SVM and kNN is that the latter is non-linear, which means the classification margin can adopt any possible shape.

The reason behind adding kNN to our classifier list was to study the performance of the classification using a non-linear classifier. Even though there is a wide range of non-linear classifiers, kNN was chosen due to its simplicity, its high processing speed, low computational cost and non-parametric characteristics that avoids making assumptions about classes distributions.

Even though these three classifiers represent three main classes of classifiers (Fisher: linear, parametric; SVM: linear, non-parametric; and kNN: non-linear, parametric), it is important to take into account that no concluding generalizations can be made out of the performance results of these classifiers regarding the group they belong to, since there is a really long list of classifiers which might perform better or worse than the ones studied in this work.

7.2.5. ANALYSES

Multiple analyses were conducted for the three classifiers previously mentioned using different scenarios defined by their populations, number of features and dimensionality reduction approaches. The aims of these analyses were:

- To compare classification performance when using five features related to both respiration and movement against using respiration variability ($Var_{RR,Norm}$) alone.
- To compare the influence of training and testing the classifiers using the data of all patients together against separate populations according to age.
- To study the influence of dimensionality reduction in the classification.
- To recognize the best dimensionality reduction approach among PCA or LDA.
- To detect the best classifier(s) out of the three of our list for each scenario.

The list of analyses performed to achieve these previously mentioned aims is shown in Table 7.3, with the numbers of features as listed in Table 7.2, and populations names as mentioned in Section 7.2.3. The three classifiers were tested in the twelve analyses listed below.

[1] An over-fitted classifier is one that is too adapted to its training set, leading to a higher chance of classification error.

Table 7.3: Classification analyses in R/N classification

	Features	Population	Dim. Red.
1.	1-5	Y	-
2.	1-5	O	-
3.	1-5	A	-
4.	1-5	Y	PCA
5.	1-5	O	PCA
6.	1-5	A	PCA
7.	1-5	Y	LDA
8.	1-5	O	LDA
9.	1-5	A	LDA
10.	1	Y	-
11.	1	O	-
12.	1	A	-

7.2.6. CLASSIFICATION PERFORMANCE

Classification performance of the previously mentioned analyses was performed by a technique called *Leave-One-Out-Cross-Validation* (LOOCV). This technique is widely used when testing a classifier performance within a specific test population. Given a dataset of n points $x_1, x_2, ...x_n$, the idea behind LOOCV is to leave the first one (x_1) out of the set, so the classifier can be trained with the rest of points ($x_2, x_3, ...x_n$), and afterwards being tested with this left out point. In the next iteration, the next data point (x_2) is left out, and the classifier is trained with the rest data points. This is made in an iterative way until the classifier has been trained and tested one-by-one with each possible combination of data points. Figure 7.5 shows an schematic of the principle behind LOOCV [13].

Figure 7.5: Principle of Leave-One-Out-Cross-Validation [13].

In our case, each of the points being "left out" of the training set were the patients of the study. Afterwards, several parameters of the classification, such as sensitivity, specificity, precision and accuracy, whose formulas are explained in Section 7.1, were averaged for all the subjects of the dataset and these were the results provided in the following subsections.

7.3. RESULTS

Classification performance results were provided by conducting four different analyses for each population. These analyses were:

1. Multiple features (no dimensionality reduction).
2. Multiple features + PCA.
3. Multiple features + LDA.
4. One unique feature ($Var_{RR,Norm}$).

The results were provided in term of sensitivity, specificity and precision of REM phase, along with overall accuracy. Being a two-class problem, sensitivity and specificity of REM directly translate to specificity and sensitivity of NREM. Precision of REM was provided due to the special interest of the researches in REM phase, since it is the most important phase of the sleep cycle of preterm babies, as discussed in Section 2. Equation 7.1-4 already described these performance parameters.

7.3.1. YOUNG POPULATION

Table 7.4 comprises the performance results of the analyses performed for the young population (babies of less than 1 year old), which can be better visualized in Figure 7.6.

Table 7.4: Classification performance results: Population Y. The best results are highlighted in bold.

	Clasf.	Sensitivity (REM)	Specificity (REM)	Precision (REM)	Accuracy
Mult.	Fisher	75.60% (26.29%)	60.69% (32.07%)	59.11% (18.11%)	65.94% (15.59%)
Feats.	kNN	64.37% (10.25%)	72.04% (9.48%)	58.20% (9.59%)	69.21% (5.43%)
	SVM	76.74% (23.19%)	74.13% (19.95%)	66.75% (16.68%)	74.87% (12.35%)
Mult.	Fisher	100.00% (0%)	0.00% (0%)	36.97% (5.40%)	36.97% (5.40%)
Feats.	kNN	46.42% (10.04%)	63.34 % (10.40%)	42.95% (7.72%)	57.50% (7.01%)
(PCA)	SVM	100.00% (0%)	0.00% (0%)	36.97% (5.40%)	36.97% (5.40%)
Mult.	Fisher	75.20% (26.31%)	60.25% (32.46%)	58.86% (18.16%)	65.53% (15.60%)
Feats.	kNN	50.03% (15.24%)	72.27% (19.17%)	55.15% (15.50%)	63.86% (10.41%)
(LDA)	SVM	69.89% (29.40%)	67.90% (26.23%)	60.47% (18.48%)	68.52% (12.36%)
One	Fisher	89.41% (10.47%)	69.08% (21.81%)	65.82% (15.15%)	75.86% (13.82%)
Feat.	kNN	61.94% (6.20%)	76.36% (12.86%)	62.44% (12.81%)	70.81% (7.86%)
	SVM	**77.61% (16.58%)**	**78.70% (21.09%)**	**72.81% (17.81%)**	**77.71% (12.68%)**

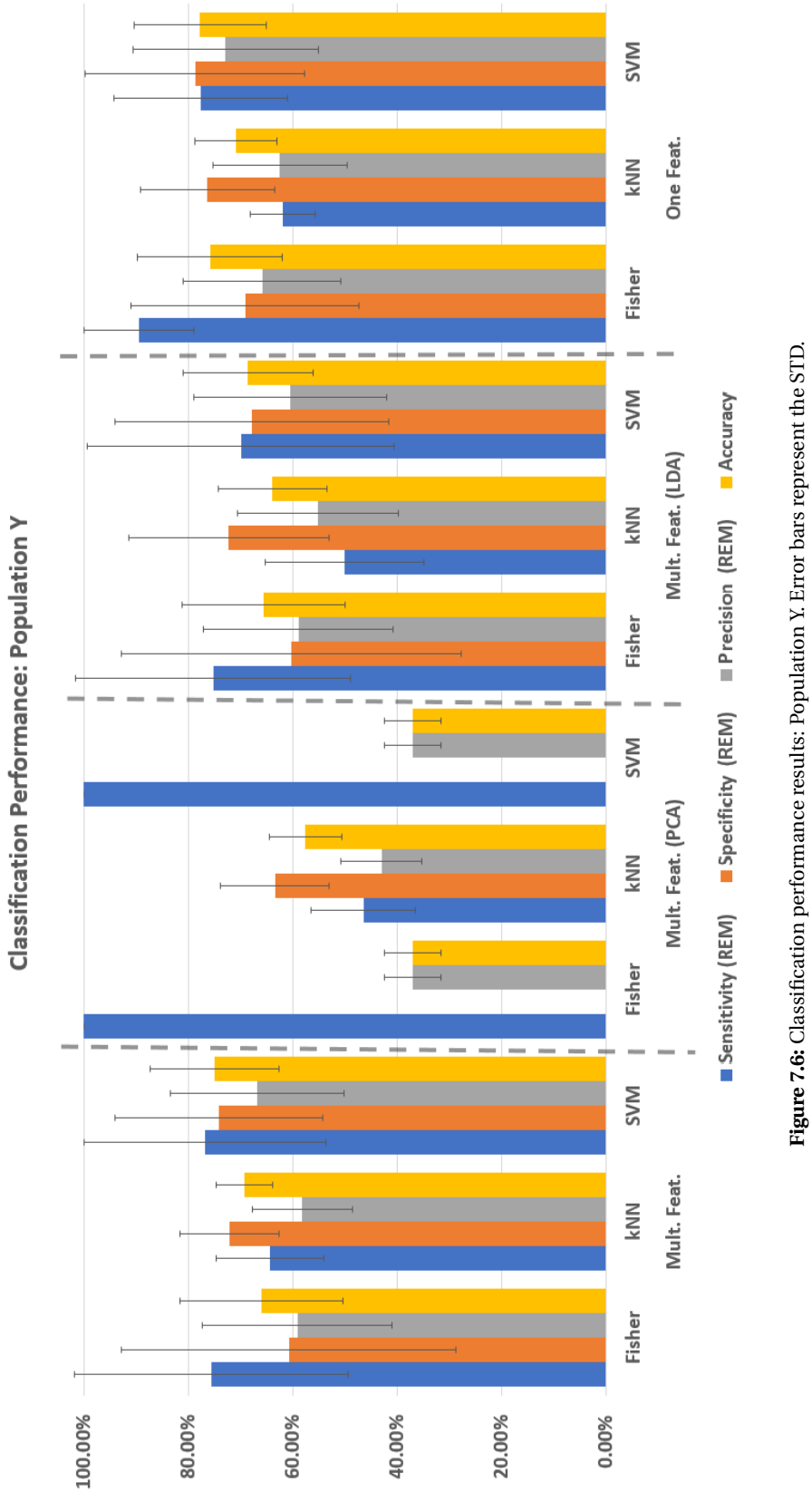


Figure 7.6: Classification performance results: Population Y. Error bars represent the STD.

Firstly, when all features were used for the classification with no dimensionality reduction, all classifiers got performance parameters around or higher than 60%. In this case, SVM outperformed the other two classifiers in all studied parameters.

When PCA was used in this population, there was a drastic decrease in performance in all classifiers, specially with SVM and Fisher, which classified all events (R and N) as REM events. Because of this, specificity of REM had a 0% value while sensitivity a 100% and an overall accuracy equal to the percentage of REM events. kNN did not suffer from this effect, however its overall performance was decreased as well.

In contrast, the LDA approach did not decrease so drastically the performance parameters, whereas it did not increase them either. Only kNN seemed like having a greater specificity for REM but at expense of reduced sensitivity. It is interesting to notice that Fisher performance remains the same after LDA since they are both performing the same discriminant principle. In general, application of dimensionality reduction approaches in this data set did not provide an improvement in classification performance.

In terms of number of features, the use of a unique feature, $Var_{RR,Norm}$, gave an overall increase in classification performance in all the previous classifiers. This may be explained by the fact that $Var_{RR,Norm}$ alone had enough information to provide a clear differentiation between REM and NREM events, and it was very likely that one of the previously used features was adding more noise to the classification, rather than helping to classify the sleep states.

Finally, regarding the best results of classification, both SVM and the Fisher classifier gave much better results when used with $Var_{RR,Norm}$ as the only feature of the classification. If we take into account the number of better parameters, SVM outperformed Fisher in terms of average accuracy, precision and specificity of REM, with similar standard deviations, even though it provided 10% less sensitivity. Therefore, we determined that the best classification results were the ones achieved by SVM, when observations were described by only one feature. However, it is still important to take into account the big spread of the performance parameters.

The detailed results of classification performance per subject are given in Table B.1 for the best classification analysis (SVM, unique feature) with Population Y. Also in Appendix C , the comparison between real and estimated hypnograms from these analysis are shown.

7.3.2. OLDER POPULATION

Classification results for this population showed an overall decrease in classification performance with respect Population Y, which an outstanding higher spread in performance parameters, suggesting that patients classifications provided very inconsistent results with each patient.

When multiple features were inputted to the classifiers with no dimensionality reduction, both SVM and Fisher classifiers showed similar results in specificity and accuracy. Moreover, Fisher tends to outperform SVM in precision , while SVM does regarding sensitivity. Since differences are very small and spread (STD) are specially big, no concluding results can be extracted from this particular comparison.

In terms of dimensionality reduction, PCA has again produced a decrease in performance in every possible parameter, in contrast to LDA, which does neither decrease nor improve the performance that much.

Table 7.5: Classification performance results: Population O. STD in parenthesis. The best results are highlighted in bold.

	Clasf.	Sensitivity (REM)	Specificity (REM)	Precision (REM)	Accuracy
Mult.	Fisher	49.84% (40.84%)	80.48% (24.14%)	62.97% (28.41%)	75.65% (11.94%)
Feats.	kNN	35.75% (17.08%)	80.00% (11.97%)	36.19% (15.93%)	69.53% (9.44%)
	SVM	52.18% (36.86%)	80.89% (23.01%)	56.84% (23.58%)	75.86% (11.68%)
Mult.	Fisher	35.48% (44.34%)	70.14% (39.89%)	24.10% (15.93%)	65.38% (20.94%)
Feats.	kNN	36.36% (9.90%)	77.72% (9.61%)	32.78% (10.24%)	68.05% (8.53%)
(PCA)	SVM	31.13% (43.10%)	76.09% (38.26%)	32.55% (29.00%)	68.21% (20.26%)
Mult.	Fisher	49.84% (40.84%)	80.48% (24.14%)	62.97% (28.41%)	75.65% (11.94%)
Feats.	kNN	35.07% (15.93%)	80.73% (10.39%)	35.87% (12.94%)	70.96% (7.06%)
(LDA)	SVM	51.60% (37.72%)	80.68% (23.73%)	58.03% (25.58%)	75.76% (12.78%)
One	**Fisher**	**47.92% (36.14%)**	**83.03% (23.34%)**	**60.54% (22.67%)**	**77.16% (11.57%)**
Feat.	kNN	33.58% (13.10%)	82.21% (11.88%)	38.79% (17.66%)	72.06% (8.02%)
	SVM	46.62% (33.61%)	82.43% (22.99%)	58.03% (26.81%)	75.87% (10.98%)

In terms of number of features, there is a slight increase of around 2-3% regarding specificity and accuracy, with a similar percentage value getting decreased in sensitivity. Precision slightly increases in kNN and SVM. Therefore, we can conclude that using $Var_{RR,Norm}$ alone gives slightly improved but similar results as using multiple features. Hence, the use of only one feature is preferred as well with this population due to the decrease of classification dimensions, which drastically decreases processing costs and classification complexity.

It is very difficult to choose the best classifier in terms of classification parameters due to the fact that all these parameters are quite similar in every analysis and STD is high. However, due to lower complexity of the classifier as well as the reduced amount of dimensions, we selected the Fisher classifier with datasets of only one feature ($Var_{RR,Norm}$) as the best classification scenario. It is important to notice here that two specific patients (Subj. 9 and 14) gave really bad results as the classifier failed to detect barely or none REM events. When compared to the results using SVM, these two subjects presented very similar results.

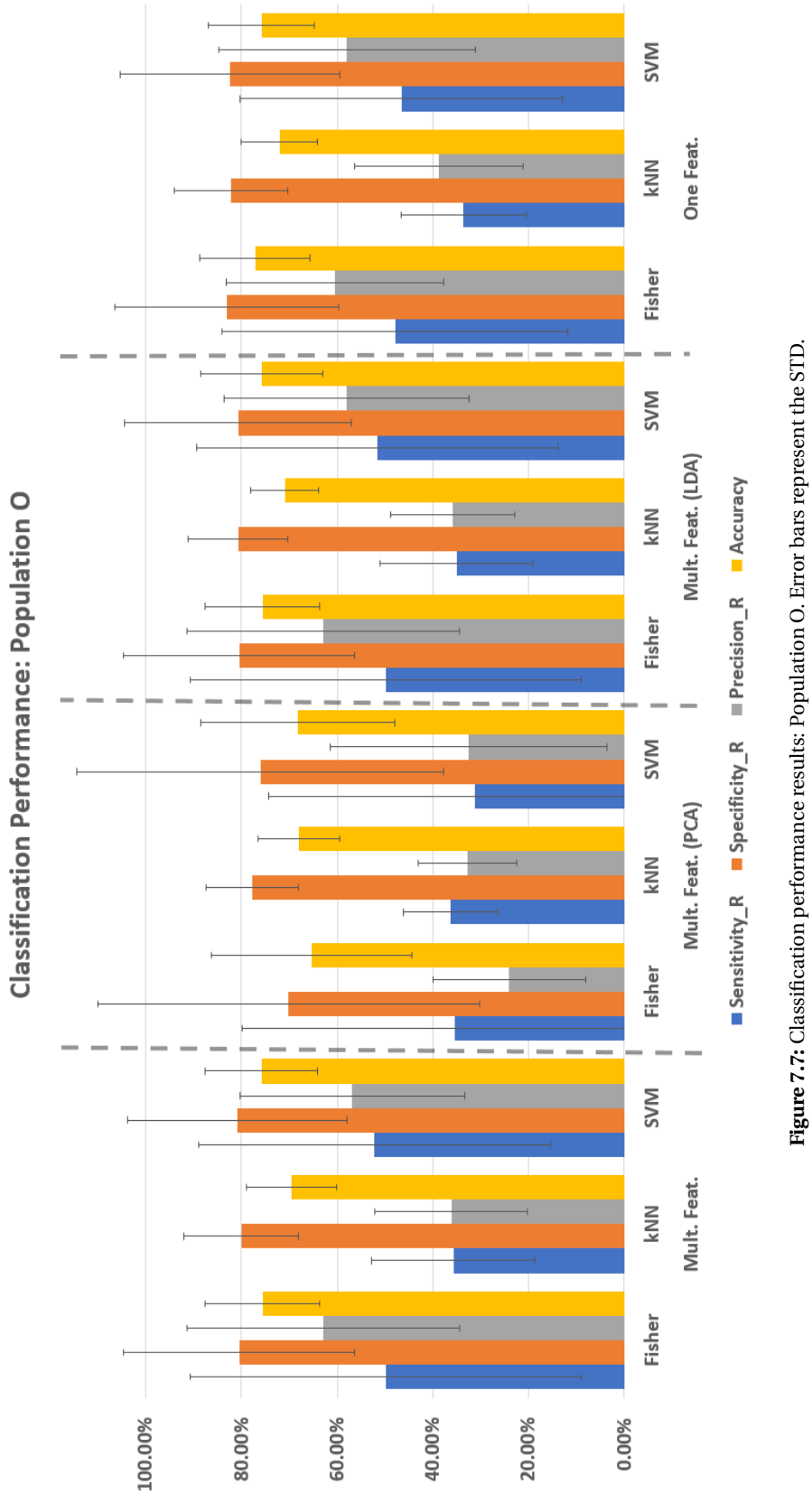


Figure 7.7: Classification performance results: Population O. Error bars represent the STD.

Table 7.6: Classification performance results: Population A. STD in parenthesis. The best results are highlighted in bold.

	Classifiers	Sensitivity (REM)	Specificity (REM)	Precision (REM)	Accuracy
Mult.	Fisher	53.16% (35.15%)	80.38% (22.98%)	64.86% (22.04%)	75.55% (10.95%)
Feats.	kNN	49.67% (17.43%)	76.04% (11.38%)	46.36% (14.66%)	68.26% (9.10%)
	SVM	59.65% (30.18%)	81.00% (20.34%)	61.14% (19.70%)	76.58% (10.93%)
Mult.	Fisher	40.99% (47.14%)	63.00% (46.66%)	31.80% (27.57%)	62.89% (19.14%)
Feats.	kNN	35.95% (9.77%)	72.21% (8.20%)	34.97% (12.27%)	61.55% (6.16%)
(PCA)	SVM	42.27% (46.20%)	61.42% (46.48%)	37.29% (27.59%)	62.19% (19.22%)
Mult.	Fisher	53.10% (35.09%)	80.48% (22.95%)	64.88% (22.06%)	75.60% (11.01%)
Feats.	kNN	45.41% (17.24%)	75.29% (14.84%)	44.78% (17.07%)	67.22% (9.81%)
(LDA)	SVM	60.58% (31.20%)	77.85% (21.91%)	57.45% (21.90%)	74.60% (13.03%)
One	**Fisher**	**52.04% (35.71%)**	**82.51% (20.31%)**	**70.26%** (22.40%)	**76.78% (9.85%)**
Feat.	kNN	43.67% (13.95%)	79.97% (12.95%)	48.53% (18.39%)	70.36% (7.89%)
	SVM	58.71% (31.09%)	81.80% (19.94%)	65.33% (22.50%)	77.19% (10.68%)

The detailed results of classification performance per subject are given in Table B.2 for the best classification scenario (Fisher, unique feature) with Population O.

7.3.3. ALL PATIENTS POPULATION

Finally, analyses were performed using all patients of young and older populations together in order to study whether a good generalization can be made in such a heterogeneous population. Table 7.6 comprises these results, which are better visualized in Figure 7.8.

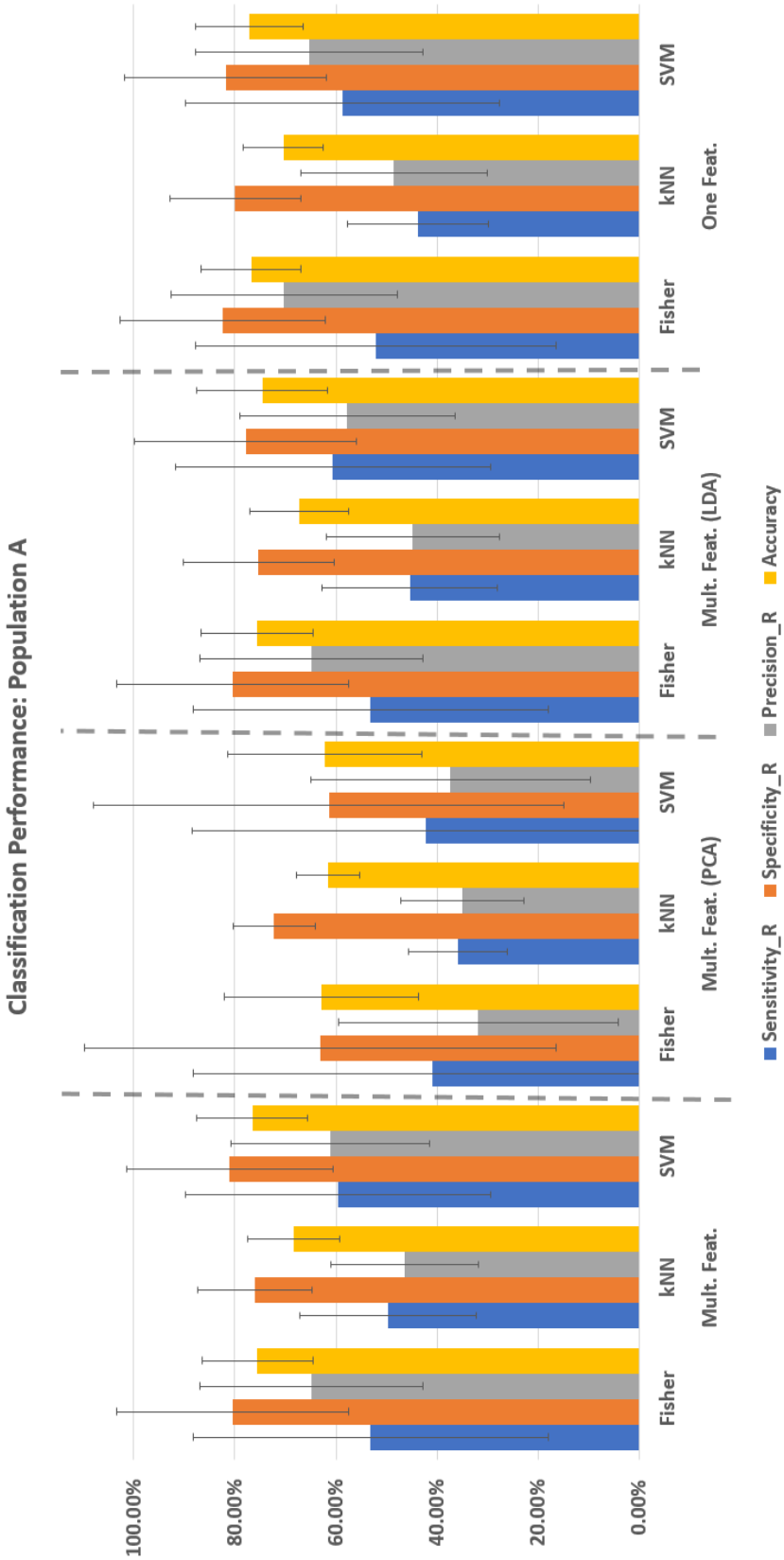


Figure 7.8: Classification performance results: Population A. The best results are highlighted in bold. Standard Deviation in parenthesis.

Regarding dimensionality reduction, there was a considerable decrease in performance when PCA is applied, just as with the rest of the populations. Meanwhile, LDA did not improve the overall performance either.

The overall best classification results of the multi-feature classification were shown with no dimensionality reduction applied, along with the SVM classifier. However both SVM and Fisher show very similar results with very high spread.

In terms of using $Var_{RR,Norm}$ as the only input feature for our classifier, there was an increase in overall performance in kNN and Fisher, and similar results for SVM than when using more features. Just as with the previous population, the best classifier in this specific scenario was quite difficult to calculate, due to the fact that both Fisher and SVM showed relatively good results and similar STD, however due to higher precision to detect REM and lower computation cost, Fisher was selected as the classifier with the best results, also when compared to all the other analyses of this population.

The detailed results of classification performance per subject are given in Table B.3 for the best classification analysis (Fisher, unique feature) with Population A. In this case, subjects 9 and 14 also provide drastically bad classification performance due to the almost non-existent ability of the classifier to detect their REM phases. An independent analysis was done in order to study the average $Var_{RR,Norm}$ values for all the subjects comparing REM and NREM phases. Patients 9 and 14 had considerably low average values of respiration variability during REM, which explains the inability of the classifier to detect these phases for these two patients.

7.3.4. BETWEEN POPULATION COMPARISON

The overall results for the older population of patients are shown to be worse than in the young population, especially regarding sensitivity. This is due to an increase of false negatives (False NREM epochs). Having in general worst classification results for this older population was what we expected because of the age heterogeneity between the patients within this group, which varies from 2 to 14 years old. However, this was the only way we had to test the classification performance in these patients since dividing them into more groups according to Table 5.2 would have leaded to unreliable results due to the low number of patients per group.

When all subjects were included in the classification (Population A analyses), similar results were shown to the ones studying only the older population. However, the young population is the only one whose results improved when left out from the all subjects population. This might be caused by the higher similarity among patients of this young group.

7.4. CONCLUSIONS

The present work has shown the ability to develop a complete radar-based monitoring system based on respiration and movement data acquired from a sleeping patient.

First, an awake/sleep classification was performed based on movement information with an average error in TST estimation of around 7% (STD = 5%), with a general overestimation of TST due to the assumption that time windows with no movement are considered as sleep states.

Regarding REM/NREM classification, previous analyses have reached relatively good results, especially in the analyses performed with the young population alone, which is the one the researchers had more interest in, due to their bigger similarities to a preterm population.

Two specific subjects in the older population showed really bad classification results due to the inability to detect REM sleep. The reason behind this is the low respiration variability of these two subjects during REM phase. This opens a new door for the application of this radar sensor to detect respiration abnormalities. However, this was not thoroughly studied in the present work as it is out of the scope of the project.

The best results within each population were all acquired when using only one feature for the classification: $Var_{RR,Norm}$. This is an important finding since it leads us to the conclusion that respiration variability provides the system with enough information to classify REM/NREM events, and more importantly, that XeThru radar modules were able to correctly detect this respiration variability among sleep phases.

Dimensionality reduction did not provide any improvement to the classification when five features were employed, with PCA showing worse results than LDA. This might be explained by the fact that this technique is an unsupervised approach, as it does not take into account observation labels when the direction of maximum variance is recognized. This may lead to the strong conclusion that $Var_{RR,Norm}$ was indeed the best feature for the classification and any other feature only provided more noise to the feature space.

Even though specific analyses were selected as the best scenario for the classification problem with each population, it is very difficult to conclude which classifier gave the best results since the standard deviations were all really high, showing really spread distributions of classification parameters. This is due to the fact that some patients had really high classification performance parameters and some others did not.

Fisher and SVM normally gave very similar results. That might be explained due to the fact they both are linear classifiers. If a selection must be made between the two of them in case they both show relatively good results, our final choice would lean towards the Fisher classifier due to its low computational cost compared to SVM. However, when only one feature is used for the classification, SVM computational cost drastically decreases. In general kNN never gave very good results, however STDs of its performance parameters were less spread across all populations. This suggests that kNN might not be the best classifier for this specific problem, however it gives more consistent results along all patients.

Another important factor to take into account in our classification results is the variability of technicians scoring sleep states, which were three in total. In an ideal study, only one technician would have scored all patients in order to provide more reliable reference data. Studies such as [64] show a classification accuracy change of around 10% between studies scored by two different technicians, which leads us to the conclusion that having three different technicians might have influenced our classification performance.

No chain classification was performed due to time constraints. An ideal classification performance study needs to be done concatenating results from sleep/awake classification and REM/NREM classification in order to compare the capability of the radar sensor to detect Awake/REM/NREM at the same time.

In conclusion, overall accuracy of classification reached 76-77% in all populations, while sensitivity, specificity and precision of detecting REM sleep reached values of 52-77%, 78-83% and 60-72%. It is difficult to compare these results to previous literature work due to the heterogeneity of performance analyses, populations under study and study protocols as well as the important fact that no much sleep classification studies have been done with only respiration data in such a young pop-

ulation. However, as the first system capable of classifying sleep states via radar-based respiration signals, these results show its potential capability as a sleep-state monitoring system.

8

Discussion, Conclusions and Recommendations for Future Work

The main goal of the project was to be able to monitor sleep in populations of patients as young as premature babies. In order to do so, a previous literature study was performed with the following two main goals:

- To investigate whether it is possible to detect physiological signals that vary the most during the patient's sleep phases (REM/NREM) in young infants.
- To investigate whether it is possible to unobtrusively measure these signals in a population of preterm babies.

After a broad literature study, respiration was selected as the main parameter to take into account in order to differentiate between REM and NREM phases. The main reason was that the AASM reported this parameter as the best physiological signal to score sleep phases in young neonates [26]. In fact, some research has been done in the past for detecting sleep phases through respiration signals in young populations of patients [66–68], with accuracies of classification as high as 93% [66].

Several contactless techniques such as imaging, ballistocardiography and radar are described in this report regarding their potential for sleep monitoring classification. All of them have reported high accuracies when measuring respiration signals. Due to availability of the sensor, XeThru radar modules were employed in the development of the project.

It is important to notice that to conduct a thorough comparative study between radar technologies was not part of the scope of this project, since it was established from the beginning to develop the whole sleep-classification system with a specific UWB module available in the market. As mentioned in Section 1.2, the implementation of the actual radar sensor was not part of the project's scope either, but only the data analysis involved after the signal acquisition.

After the previous literature study, the research question arisen was *to study the feasibility of a radar-based approach for sleep monitoring on the Neonatal Intensive Care Unit that could be eventually applied on premature patients.*

In this chapter, a final discussion will be provided about the two main contributions created in this work: (i) the new approach to perform breathing detection from raw radar signals, and (ii) the sleep monitoring capabilities of the system based on these acquired radar signals.

8.1. BREATHING DETECTION SYSTEM

A detailed algorithm to compute breathing signals has been described in this work. This algorithm is based on the idea of adding a motion detector block before actual breathing processing takes place in order to decrease the noise level induced by environmental movements detected by the radar.

Another contribution of our breathing detector was the ability to adapt to mean patient's Respiration Rate (RR) values by finding time windows with stable RR readings where no motion was detected for a minimum time of 5 minutes.

The radar sensors were tested in a population of 23 patients in order to perform accuracy analyses. Our proposed breathing detection algorithm was compared to the in-built system of XeThru sensors, having as reference RR signals the ones from respiratory effort belts of the PSG equipment. Errors provided very similar average results for both algorithms.

Mean Error (ME) values showed a slight tendency of our algorithm to underestimate RR values, in comparison with a zero average bias in XeThru's in-built algorithm. Mean Absolute Error (MAE) values suggested their algorithm is more accurate due to higher RR resolutions, however their slightly higher values of Root Mean Square Error (RMSE) and Standard Deviation (STD) of the error showed a small increase in their error spread. Higher Relative Errors (RE) values in our algorithm were predictable since this parameter is highly impacted by the frequency resolution of the system. For instance, when we deal with RR values of 11 Respirations per Minute (RPM), our algorithm provides a value of either 9 or 12 RPM (integer multiples of 3 RPM). In the case of providing 9 RPM, the RE would be of around 20%. However, what we were actually interested in was the capability of the system to detect the ups and downs of the patient's respiration, without paying so much attention to the exact values. From the comparison between the errors of the two algorithms (our approach and XeThru), no significant conclusion can be made about which error performance is better, since averages of all parameters looked really similar.

The error analyses performed by dividing the previously described results into two groups of young and older patients showed similar error values in our breathing detector algorithm for both populations. However, the XeThru output signal showed more sensitivity to patients' age, since a higher value of errors was observed in younger patients.

Regarding time of breathing recording, our approach clearly outperformed XeThru's in-built algorithm, showing a significant increase in recording hours in most of our subjects (about 15% increase).

When observing the error analyses of our system, it is important to take into account that no filtering was done in order to remove the moments of the recordings with interruptions of parents or nurses. The aim of this was to really mimic the normal hectic conditions of the NICU environment, which is related to the final goal of the system regarding its applicability to sleep monitoring of premature patients in hospital conditions.

Three papers were found using UWB radar on a neonate population [20–22]. However, a strict comparison of performance of our algorithm with respect to theirs cannot be made due to the the heterogeneity of the accuracy analyses, populations under study and time of recording. Huang et al

[22] also applied frequency windows adapted to patient's age in the frequency analysis of respiration signals. However, the contribution of the present work is that this is done in an automatic way when a stable respiration is detected for at least 5 minutes.

8.2. SLEEP MONITORING SYSTEM

After motion and respiration signals were computed from raw radar signals through our breathing detection system, these were used as input of our sleep monitoring system. This system had two main tasks: (i) to detect Sleep/Awake events based on movements of patient and (ii) to differentiate between REM/NREM phases within sleep based on both movements and respiration.

As expected, Sleep/Awake classification tended to overestimate sleep due to the lack of movement of the patient when falling asleep but not yet sleeping. The system was able to reach an average error in TST of around 7%, however we must take into account the greater amount of sleep time during the over-night PSG, giving rise to already low values of error when only "sleep" events are scored during the whole night. In this case, it is better to take into account the specificity of the classification, which on average was about 63% with quite spread values (STD = 22%). Due to this relatively low specificity, we can conclude that more in depth refinements need to be made in the sleep/awake classification. However diagrams of real against estimated sleep/awake transitions in Appendix A provide a good qualitative study of their similarity in most of the patients. These diagrams also show the fact that the system tends to score as awake states the periods of time with interruptions of parents/nurses, but not always. Thus, some improvements can be made here regarding the sensitivity of the radar system to motion artifacts.

With respect to sleep-state classification, several analyses were performed in order to study the influence of number of features, dimensionality reduction, the classifier used and the training set population in this specific classification problem.

Most analyses showed and improvement in classification performance when only the feature containing information of respiration variability, $Var_{RR,Norm}$, was used as input variable for the classification. This provided better results than the analyses where dimensionality reduction was applied, in which the number of features was reduced from five to two. Regarding which dimensionality reduction technique to choose between Linear Discriminant Analysis (LDA) and Principal Component Analysis (PCA), the latter gave worse results compared to LDA. This might be explained by the fact that LDA is an unsupervised technique that finds the direction with the highest variance without taking into account the labels of the classes. Despite of getting better results with LDA, these did not improve compared to not using any dimensionality reduction technique at all. This, and the fact that $Var_{RR,Norm}$ alone was enough to have good results in the classification, lead us to the conclusion that respiration variability was indeed the best classification feature providing the highest differentiation between REM and NREM sleep states, and that the inclusion of any other variable only adds more noise to the classification.

Regarding classification population, there was an increase in performance when classification was performed on the young population. As expected, this suggests that sleeping patterns within the young population are more similar to each other. This gets us closer to the idea that performing the same analyses with a preterm population would have quite similar results, since they share some major developmental features. On the other hand, the older population had slightly worse classification results, which are very likely due to the heterogeneity of the population age, as expected.

When studying the average classification performance, we must take into account that the STD of

the classification parameters reported in all analyses was quite high, specially in the older population. This is due to the fact that classification in two specific patients within the old population was not possible due to the low respiration variability during REM phase, creating outliers in the analyses. Nevertheless, these results also provide us with the idea that our radar sensor might be able to detect these alterations in respiration patterns that may be affected by a specific syndrome. However we lack more medical information of these two patients in order to truly understand the reason behind their different sleeping patterns in contrast to the rest of patients.

In general, results of the sleep classification when using the population with all the patients were quite similar to the ones of the old population. However, patients within the young population had worse results with classifiers trained on all subjects than on only the young population. This confirms the idea that classifiers provide better results when being trained and tested on populations of patients that are more similar to each other, providing more homogeneity to the sleeping patterns.

This high spread of classification performance parameters, created some difficulty when selecting the classifier with the best results. However, linear classifiers (SVM and Fisher) outperformed kNN. Between SVM and Fisher, when a really similar performance was reported, Fisher was prioritized due to its simplicity and low computational cost. However, when observations had a unique feature $Var_{RR,Norm}$, the computational cost of SVM was not that high. Therefore we concluded that both classifiers reported really similar performances.

Lastly, it is important to take into account that the sleep classification system is highly influenced by the fact that three different technicians were classifying the sleep studies. It has previously been shown in other studies [64] that the alternation between two examinees scoring sleep states changes the classification accuracy up to 10%, showing that sleep scoring is, to a certain extent, a subjective process. Therefore the ideal scenario would have been to have one unique technician scoring all PSG studies. However, the researchers in this present project had no influence on choosing the technicians working in the hospital performing the sleep scoring.

Another big limitation affecting the results of the sleep classification was the fact that the study populations were quite heterogeneous with respect to age and diagnosis. Firstly, a higher number of patients undergoing a PSG would have been ideal in order to be able to perform a proper age aggregation of patients and detailed sleep classification studies. Unfortunately that was impossible due to the small amount of total PSG studies scheduled in the hospitals we were collaborating with along the small duration of this project. On the other hand, the unavailability of healthy patients performing a PSG is a big issue influencing our sleep classification, since now the system was being training with a population that does not present 100% "normal" sleeping patterns.

In conclusion, our sleep monitoring system based on UWB radar has provided both sleep/awake and REM/NREM classifications. Relatively good classification results have been provided so far, however we faced several limitations regarding patients population age heterogeneity, lack of healthy patients and lack of a unique scorer. These limitations had a great influence on the overall performance results, however the final accuracy of the system reached values as high as 90%, with an average accuracy around 77%, sensitivity, specificity and precision of REM of around 47-77%, 78-83% and 60-72%, respectively.

8.3. RECOMMENDATIONS FOR FUTURE WORK

Future work to increase the performance of both respiration detection and sleep classification could potentially enhance the results presented in the present work as well as its reliability. Here, some of

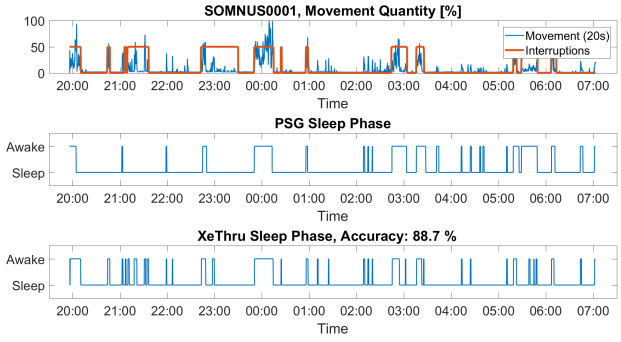
these improvements will be listed:

- To investigate the use of techniques to increase the resolution of the frequency analyses from the radar modules in order to potentially detect HR signal, which could also potentially increase sleep-state classification performance.
- To implement breathing waveform analyses that allow to accurately study the variability of respiration from raw breathing waveforms instead of only RR signal variability. Breathing waveform analyses could potentially be used to detect apnea events as well, a frequent condition present in young patients with extremely high medical significance.
- To improve the sleep/awake classification in order to be less sensitive to nurses/parents interruptions and not score these as awake events by default.
- To add healthy patients to the study population and analyze differences in sleep classification among healthy and non-healthy patients.
- To increase the number of patients' studies per age in order to achieve a more reliable training of the classifiers.
- To study the agreement level among technicians scoring PSG studies in order to have a better idea of the degree of impact they had on the results of the present work.
- To study the effect of adding the previous scored state as a new feature of the sleep-state classification.
- To analyze the results of joining both classifications together, Sleep/Awake and REM/NREM, to study the overall performance of the system.

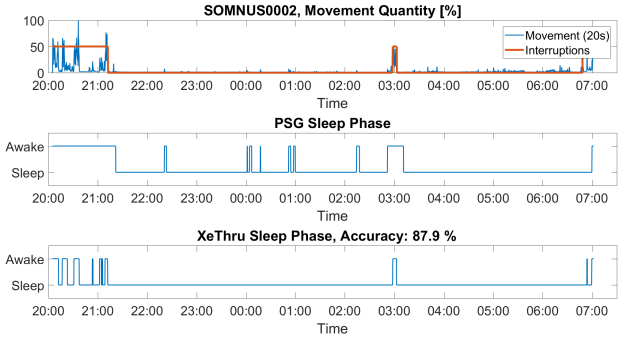
Appendices

A

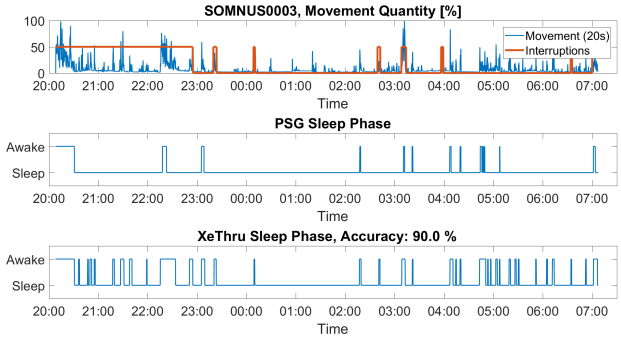
IMAGES SLEEP/AWAKE CLASSIFICATION



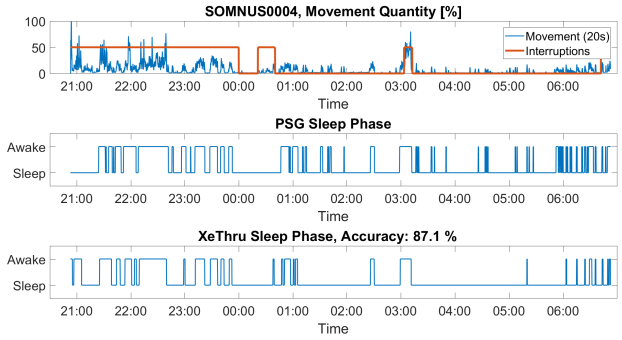
(a) % Movement (up), Real Score A/W (center), Est. Score A/W (down), Subject 1.



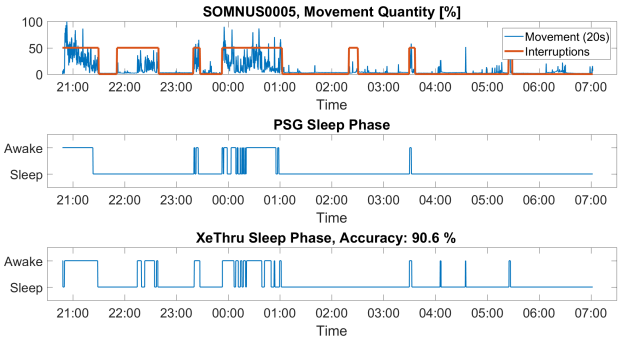
(b) % Movement (up), Real Score A/W (center), Est. Score A/W (down), Subject 2.



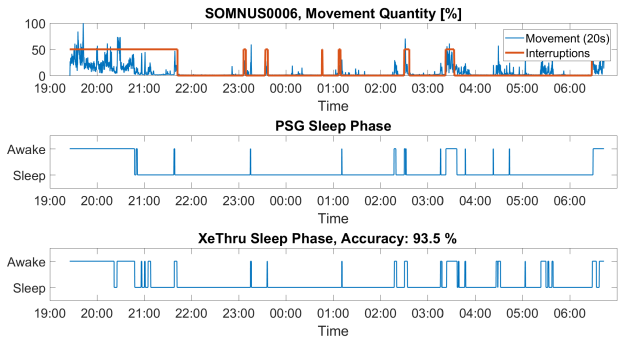
(c) % Movement (up), Real Score A/W (center), Est. Score A/W (down), Subject 3.



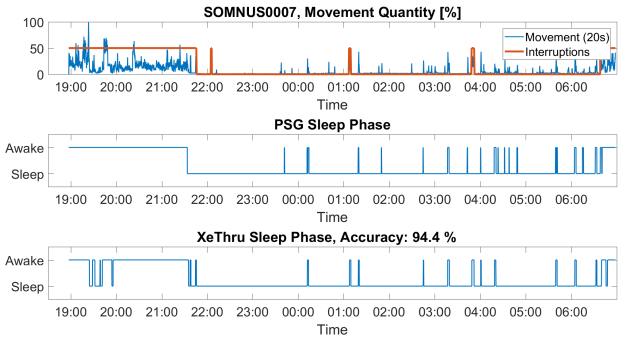
(d) % Movement (up), Real Score A/W (center), Est. Score A/W (down), Subject 4.



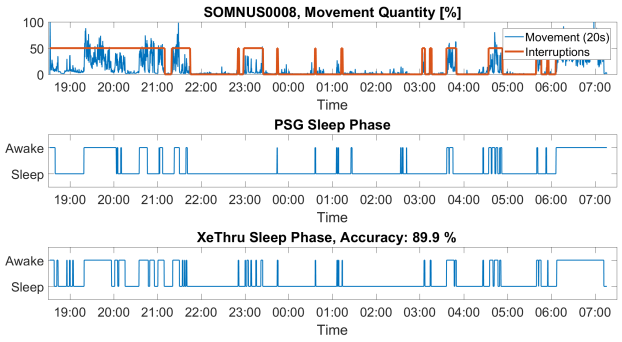
(e) % Movement (up), Real Score A/W (center), Est. Score A/W (down), Subject 5.



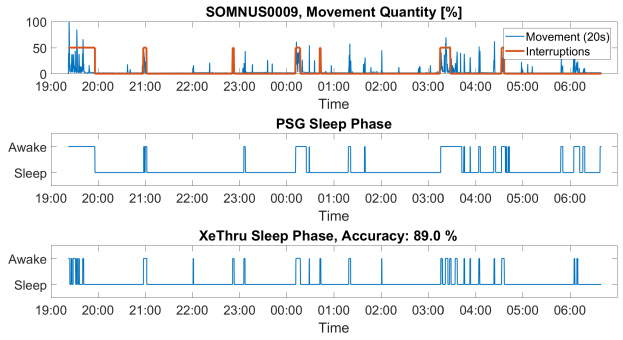
(f) % Movement (up), Real Score A/W (center), Est. Score A/W (down), Subject 6.



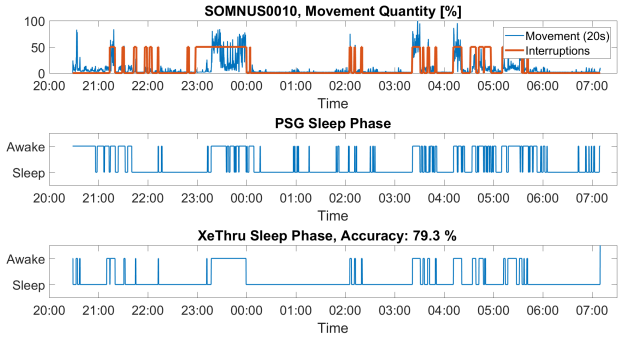
(g) % Movement (up), Real Score A/W (center), Est. Score A/W (down), Subject 7.



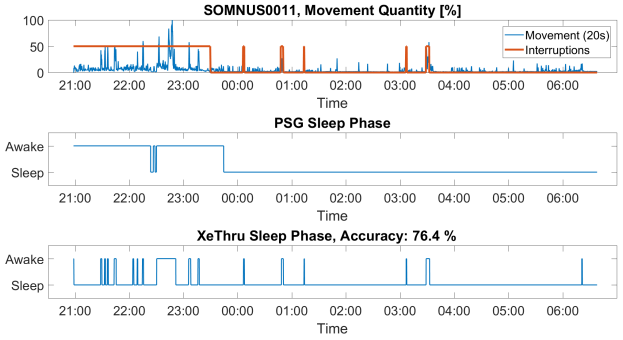
(h) % Movement (up), Real Score A/W (center), Est. Score A/W (down), Subject 8.



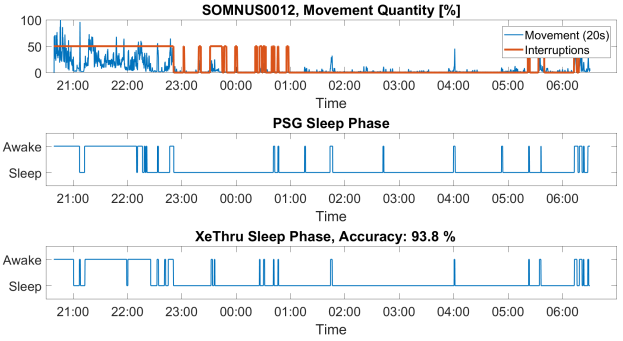
(a) % Movement (up), Real Score A/W (center), Est. Score A/W (down), Subject 9.



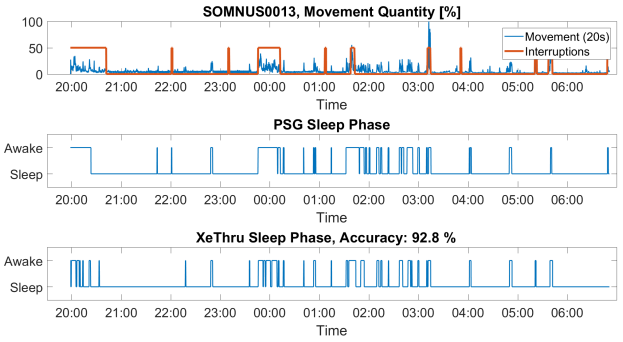
(b) % Movement (up), Real Score A/W (center), Est. Score A/W (down), Subject 10.



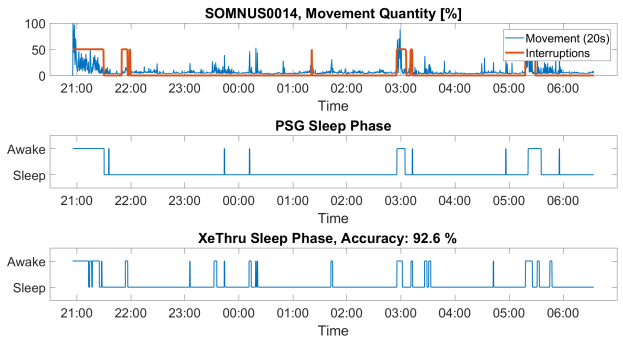
(c) % Movement (up), Real Score A/W (center), Est. Score A/W (down), Subject 11.



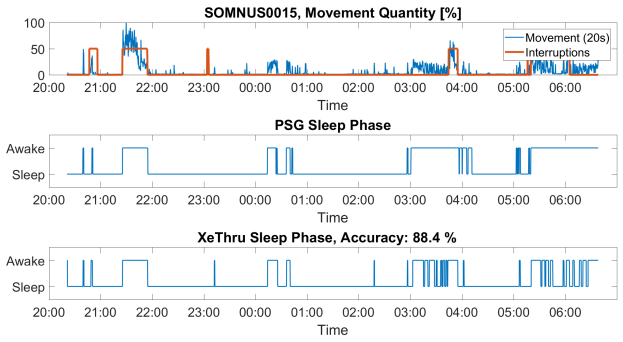
(d) % Movement (up), Real Score A/W (center), Est. Score A/W (down), Subject 12.



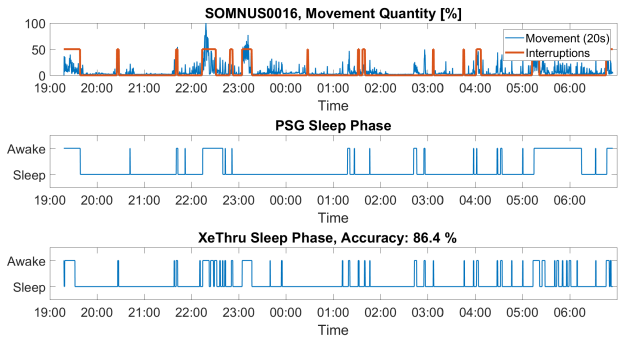
(e) % Movement (up), Real Score A/W (center), Est. Score A/W (down), Subject 13.



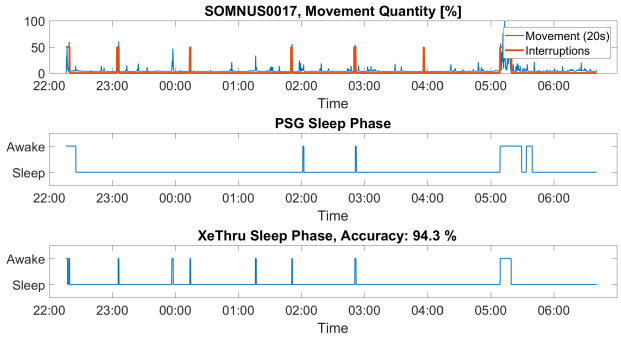
(f) % Movement (up), Real Score A/W (center), Est. Score A/W (down), Subject 14.



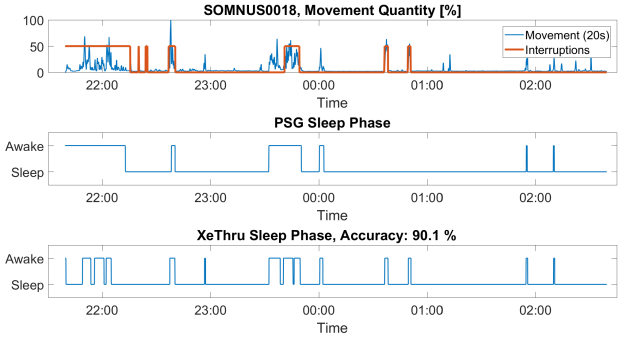
(g) % Movement (up), Real Score A/W (center), Est. Score A/W (down), Subject 15.



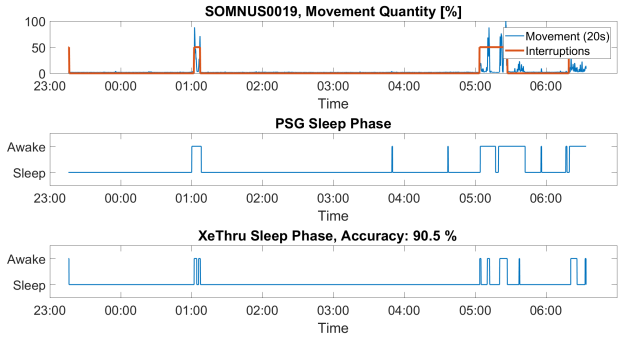
(h) % Movement (up), Real Score A/W (center), Est. Score A/W (down), Subject 16.



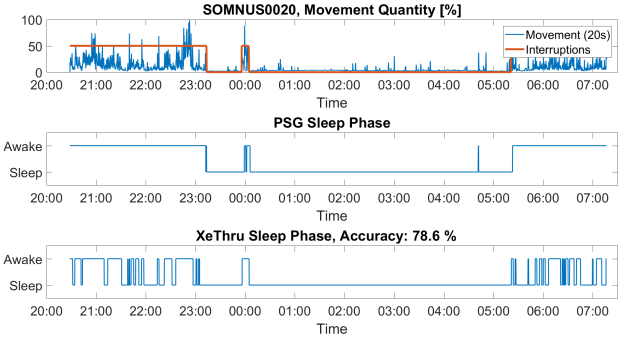
(a) % Movement (up), Real Score A/W (center), Est. Score A/W (down), Subject 17.



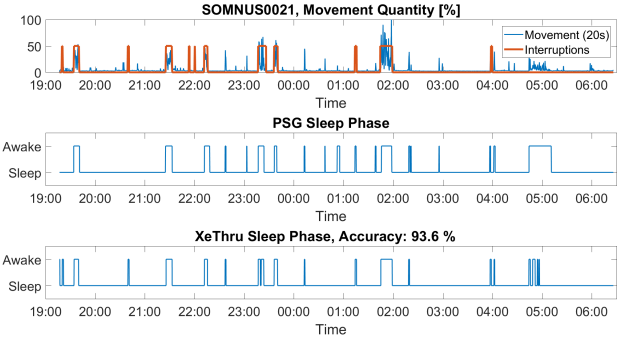
(b) % Movement (up), Real Score A/W (center), Est. Score A/W (down), Subject 18.



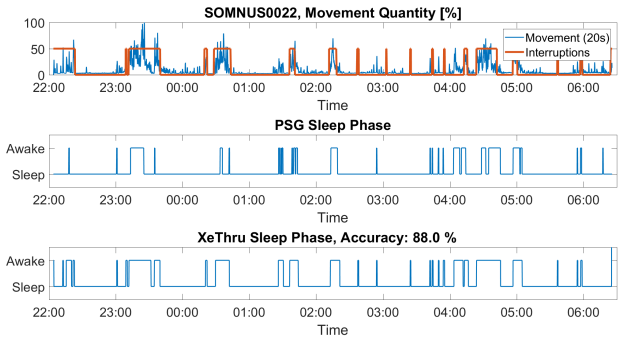
(c) % Movement (up), Real Score A/W (center), Est. Score A/W (down), Subject 19.



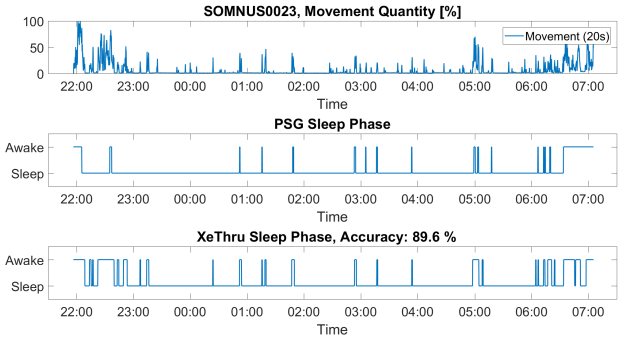
(d) % Movement (up), Real Score A/W (center), Est. Score A/W (down), Subject 20.



(e) % Movement (up), Real Score A/W (center), Est. Score A/W (down), Subject 21.



(f) % Movement (up), Real Score A/W (center), Est. Score A/W (down), Subject 22.



(g) % Movement (up), Real Score A/W (center), Est. Score A/W (down), Subject 23.

B

Classification Results of best analyses per Population

Table B.1: Classification Results: Population Y, SVM Classifier, one feature. Total sums of TP/FN/TN/FP are provided. Positive = REM, Negative = NREM.

	Sens. (REM)	Spec. (REM)	Prec. (REM)	Accuracy	TP	FN	TN	FP
Subject 1	90.20%	90.35%	80.70%	90.30%	46	5	103	11
Subject 3	55.38%	89.77%	80.00%	75.16%	36	29	79	9
Subject 6	80.33%	95.58%	90.74%	90.23%	49	12	108	5
Subject 8	81.82%	46.81%	41.86%	57.97%	36	8	44	50
Subject 10	69.81%	84.44%	72.55%	79.02%	37	16	76	14
Subject 12	100.00%	38.95%	44.23%	58.87%	46	0	37	58
Subject 21	52.50%	94.67%	84.00%	80.00%	42	38	142	8
Subject 22	95.24%	90.00%	86.96%	92.16%	40	2	54	6
Subject 23	73.24%	77.78%	74.29%	75.66%	52	19	63	18
Avg / sum	**77.61%**	**78.70%**	**72.81%**	**77.71%**	**384**	**129**	**706**	**179**
STD	**16.58%**	**21.09%**	**17.81%**	**12.68%**				

Table B.2: Classification Results: Population O, Fisher Classifier, one feature. Total sums of TP/FN/TN/FP are provided. Positive = REM, Negative = NREM.

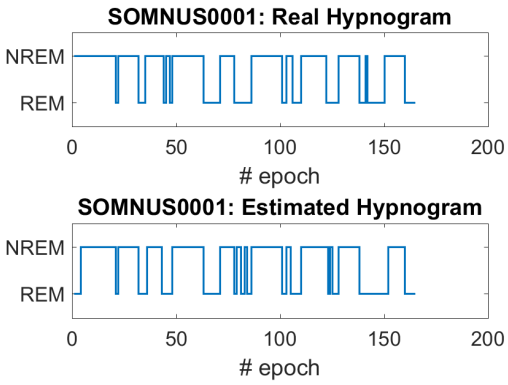
	Sens. (REM)	Spec. (REM)	Prec. (REM)	Accuracy	TP	FN	TN	FP
Subject 2	0.8125	0.76630435	0.47560976	0.77586207	39	9	141	43
Subject 5	0.71428571	0.89423077	0.73170732	0.84246575	30	12	93	11
Subject 7	0.52777778	0.94736842	0.76	0.84666667	19	17	108	6
Subject 9	0	1		0.75586854	0	52	161	0
Subject 13	0.91044776	0.23333333	0.39869281	0.47593583	61	6	28	92
Subject 14	0.06666667	1	1	0.82716049	2	28	132	0
Subject 15	0.95744681	0.74226804	0.64285714	0.8125	45	2	72	25
Subject 16	0.71428571	0.632	0.17857143	0.64028777	10	4	79	46
Subject 17	0.24137931	0.97972973	0.7	0.85875706	7	22	145	3
Subject 19	0.24	0.94782609	0.5	0.82142857	6	19	109	6
Subject 20	0.08695652	0.99065421	0.66666667	0.83076923	2	21	106	1
Avg / sum	**47.92%**	**83.03%**	**60.54%**	**77.16%**	**221**	**192**	**1174**	**233**
STD	**33.61%**	**22.99%**	**26.81%**	**10.98%**				

Table B.3: Classification Results: Population A, Fisher Classifier, one feature. Total sums of TP/FN/TN/FP are provided. Positive = REM, Negative = NREM.

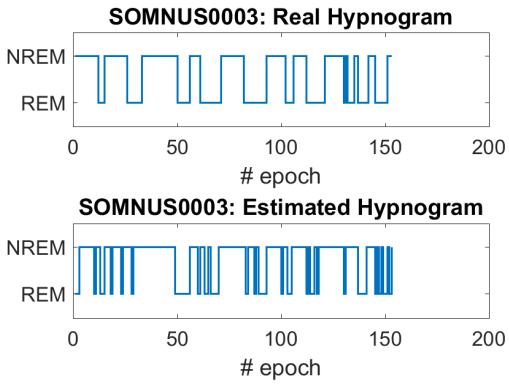
	Sens. (REM)	Spec. (REM)	Prec. (REM)	Accuracy	TP	FN	TN	FP
Subject 1	35.29%	95.61%	78.26%	76.97%	18	33	109	5
Subject 2	68.75%	83.70%	52.38%	80.60%	33	15	154	30
Subject 3	76.92%	81.82%	75.76%	79.74%	50	15	72	16
Subject 5	11.90%	100.00%	100.00%	74.66%	5	37	104	0
Subject 6	96.72%	84.96%	77.63%	89.08%	59	2	96	17
Subject 7	33.33%	97.37%	80.00%	82.00%	12	24	111	3
Subject 8	56.82%	67.02%	44.64%	63.77%	25	19	63	31
Subject 9	0.00%	100.00%		75.59%	0	52	161	0
Subject 10	90.57%	67.78%	62.34%	76.22%	48	5	61	29
Subject 12	76.09%	58.95%	47.30%	64.54%	35	11	56	39
Subject 13	94.03%	20.00%	39.62%	46.52%	63	4	24	96
Subject 14	3.33%	100.00%	100.00%	82.10%	1	29	132	0
Subject 15	27.66%	95.88%	76.47%	73.61%	13	34	93	4
Subject 16	64.29%	71.20%	20.00%	70.50%	9	5	89	36
Subject 17	10.34%	98.65%	60.00%	84.18%	3	26	146	2
Subject 19	16.00%	100.00%	100.00%	85.00%	4	21	115	0
Subject 20	8.70%	100.00%	100.00%	83.85%	2	21	107	0
Subject 21	81.25%	88.00%	78.31%	85.65%	65	15	132	18
Subject 22	100.00%	75.00%	73.68%	85.29%	42	0	45	15
Subject 23	88.73%	64.20%	68.48%	75.66%	63	8	52	29
Avg / sum	**52.04%**	**82.51%**	**70.26%**	**76.78%**	**550**	**376**	**1922**	**370**
STD	**35.71%**	**20.31%**	**22.40%**	**9.85%**				

C

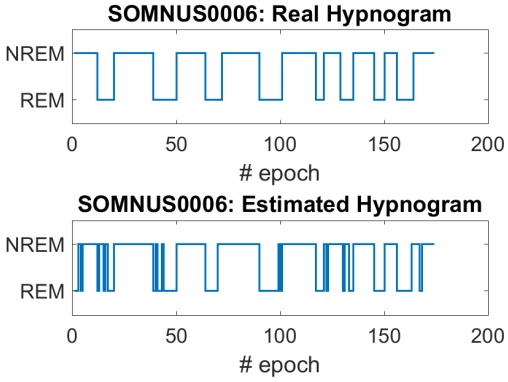
R/N Hypnograms: Population Y, SVM Classifier



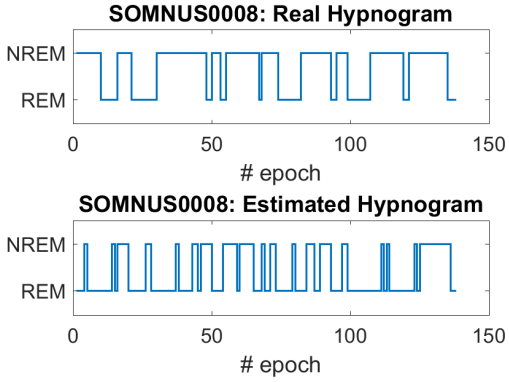
(a) Subject 1.



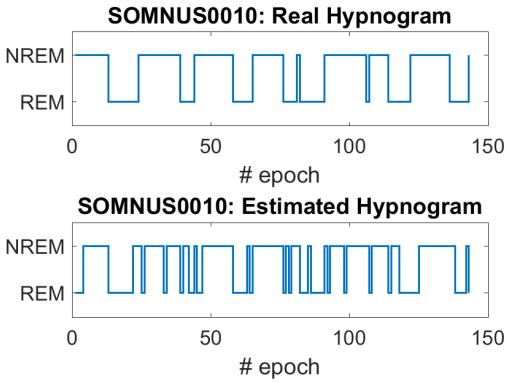
(b) Subject 3.



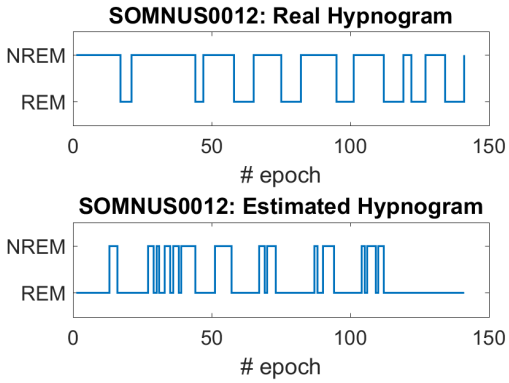
(c) Subject 6.



(d) Subject 8.

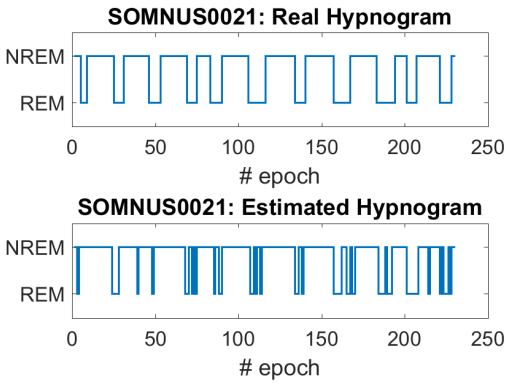


(e) Subject 10.

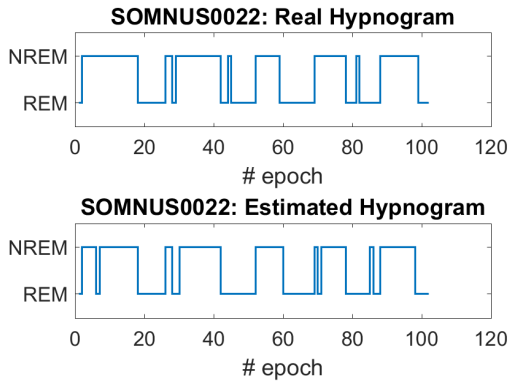


(f) Subject 12.

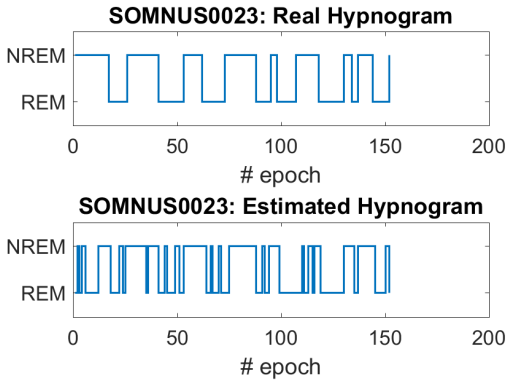
Figure C.1: Classification Results from SVM classifier, one feature, Population Y (part A)



(a) Subject 21.



(b) Subject 22.

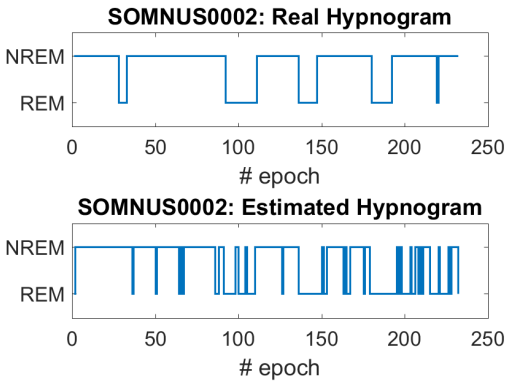


(c) Subject 23.

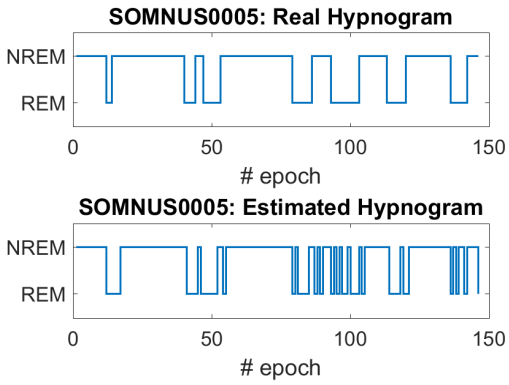
Figure C.2: Classification Results from SVM classifier, one feature, Population Y (part B)

D

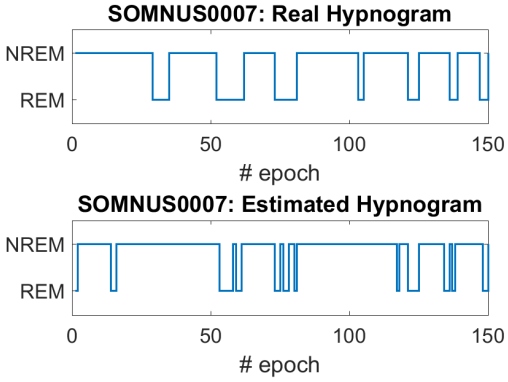
R/N Hypnograms: Population O, Fisher Classifier



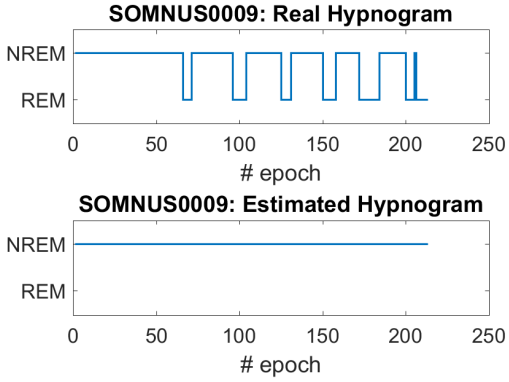
(a) Subject 2.



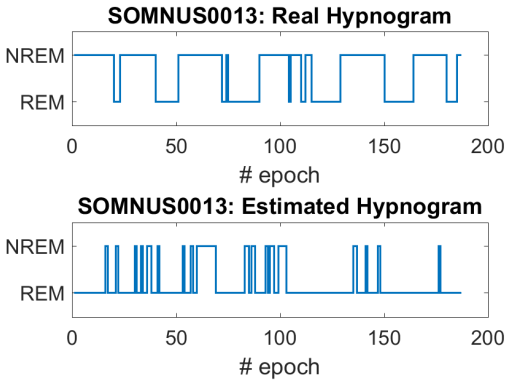
(b) Subject 5.



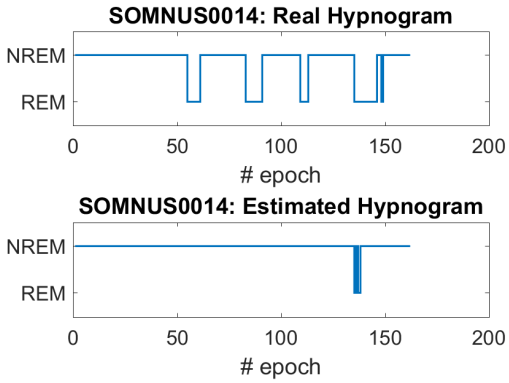
(c) Subject 7.



(d) Subject 9.

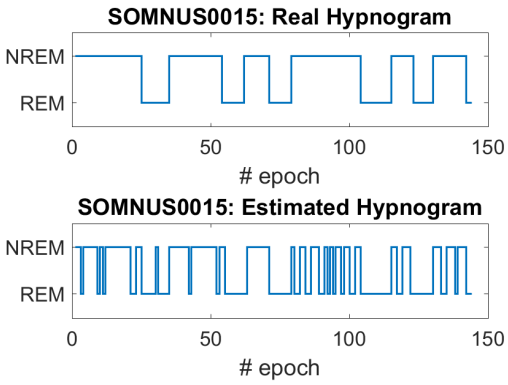


(e) Subject 13.

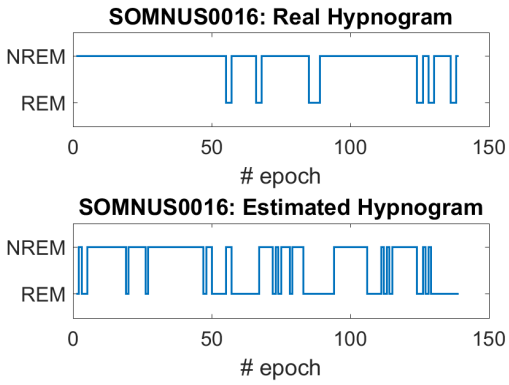


(f) Subject 14.

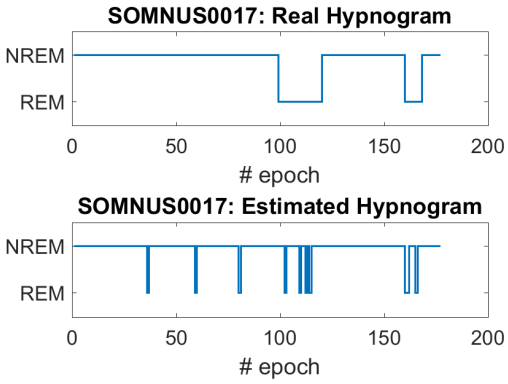
Figure D.1: Classification Results from Fisher classifier, one feature, Population O (part A)



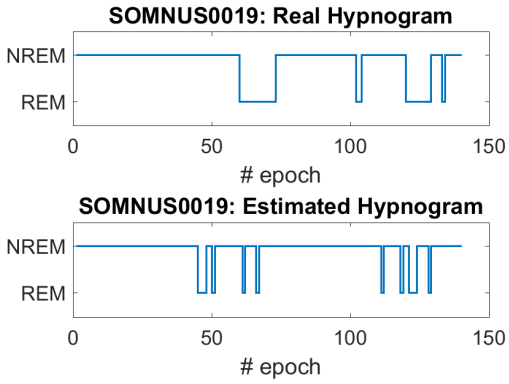
(a) Subject 15.



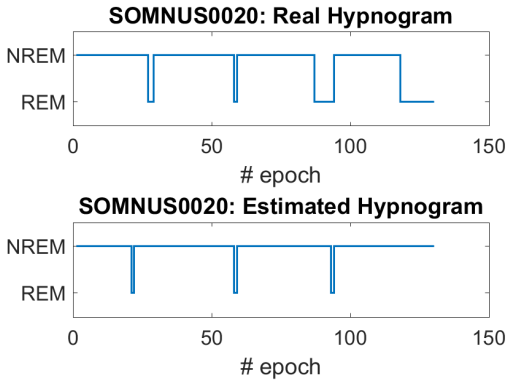
(b) Subject 16.



(c) Subject 17.



(d) Subject 19.

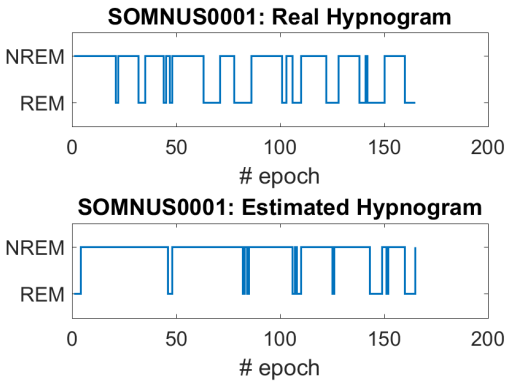


(e) Subject 20.

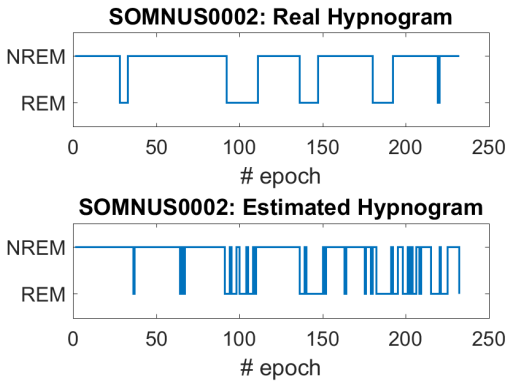
Figure D.2: Classification Results from Fisher classifier, one feature, Population O (part B)

E

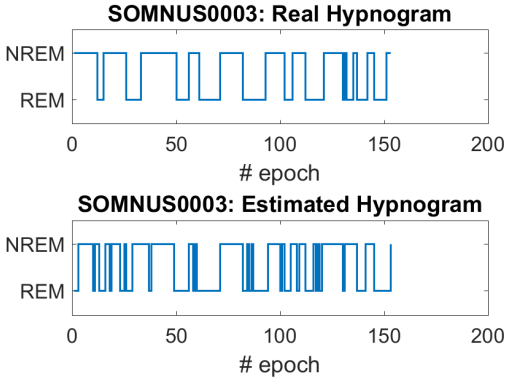
R/N Hypnograms: Population A, Fisher Classifier



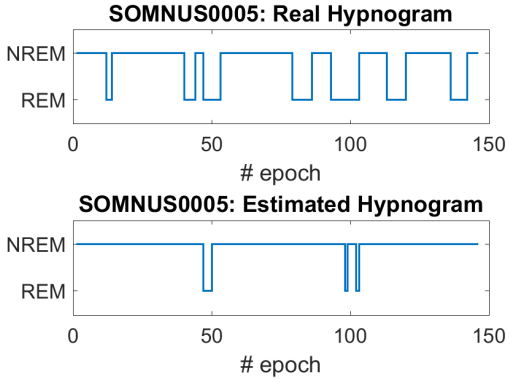
(a) Subject 1.



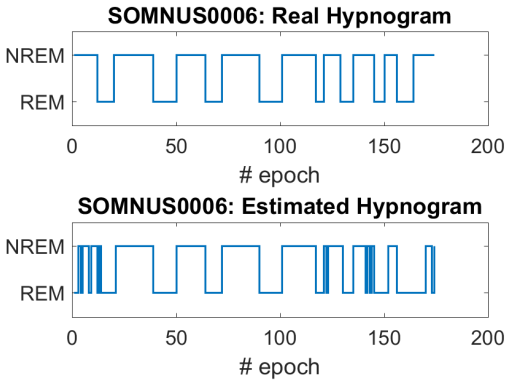
(b) Subject 2.



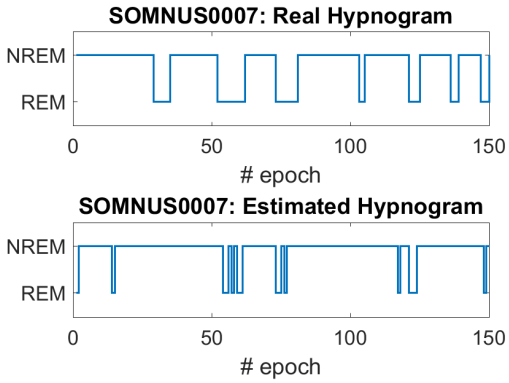
(c) Subject 3.



(d) Subject 5.

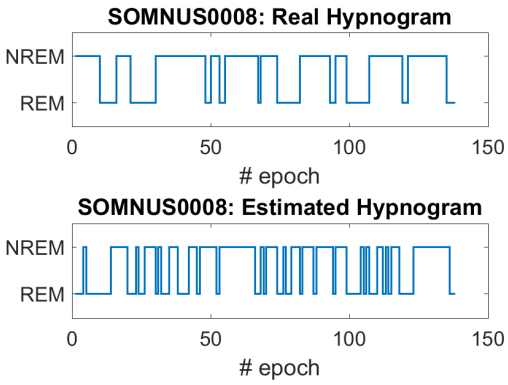


(e) Subject 6.

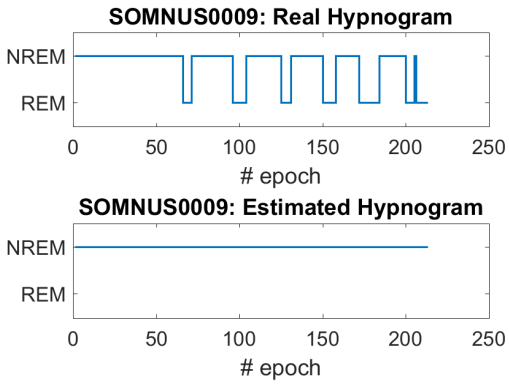


(f) Subject 7.

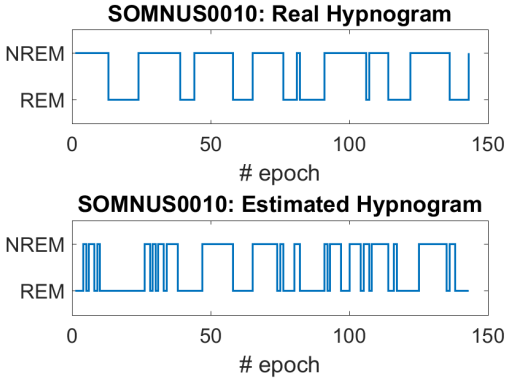
Figure E.1: Classification Results from Fisher classifier, one feature, Population A (part A)



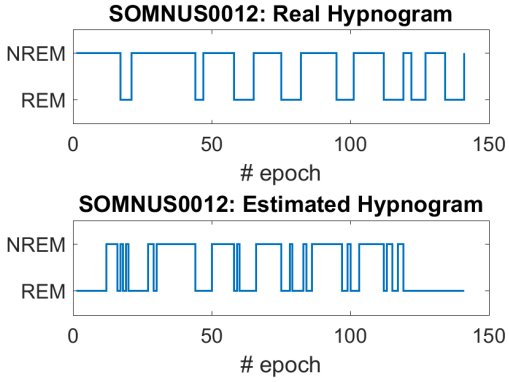
(a) Subject 8.



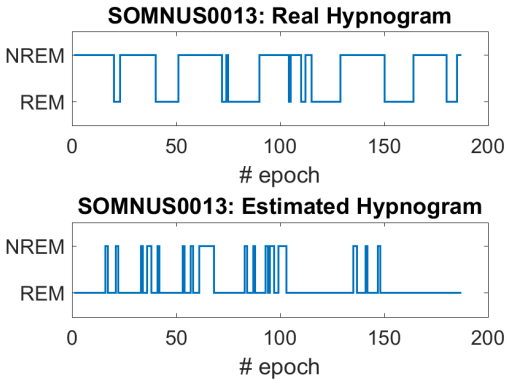
(b) Subject 9.



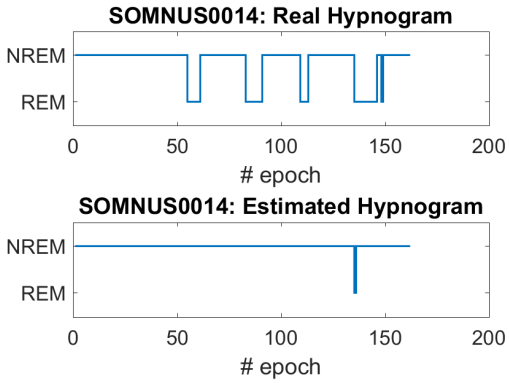
(c) Subject 10.



(d) Subject 12.

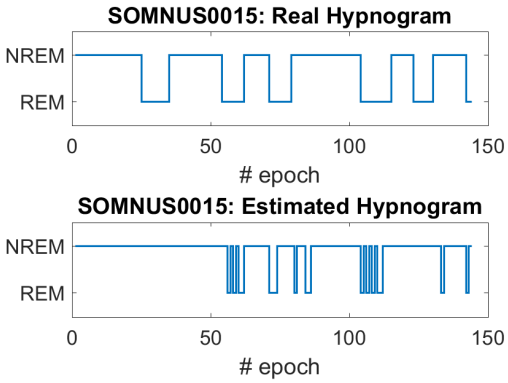


(e) Subject 13.

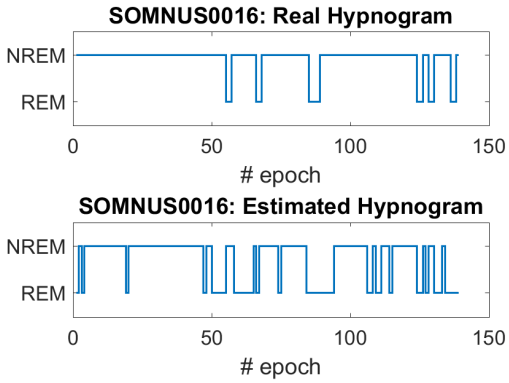


(f) Subject 14.

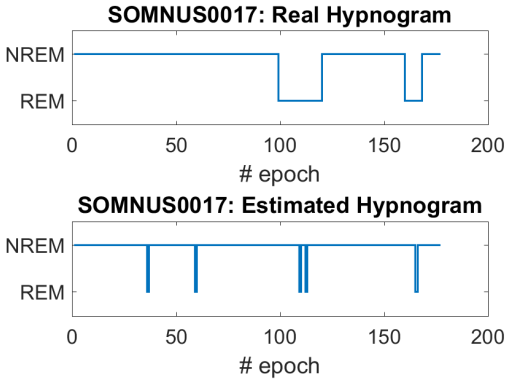
Figure E.2: Classification Results from Fisher classifier, one feature, Population Ama (part B)



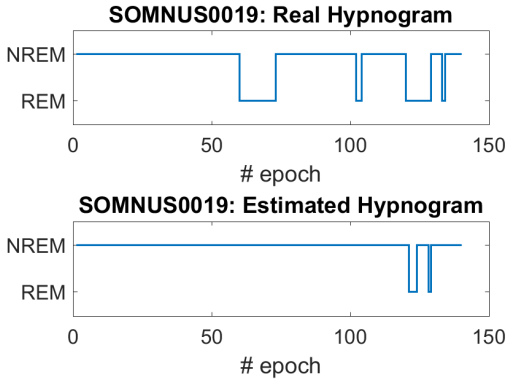
(a) Subject 15.



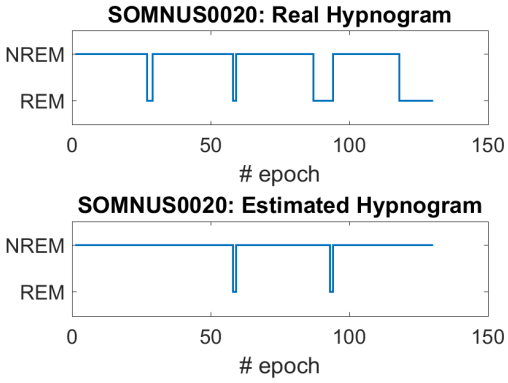
(b) Subject 16.



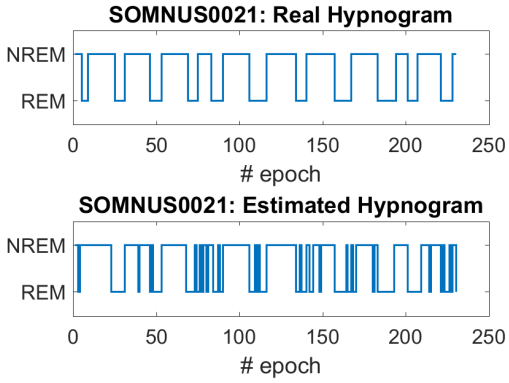
(c) Subject 17.



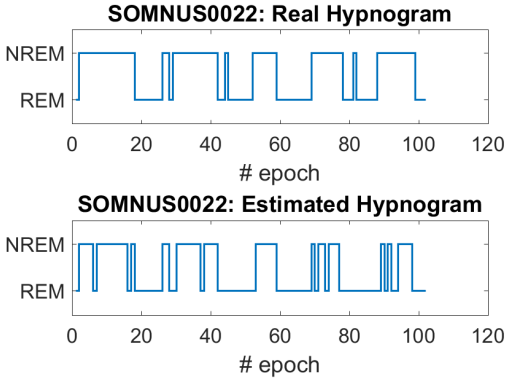
(d) Subject 19.



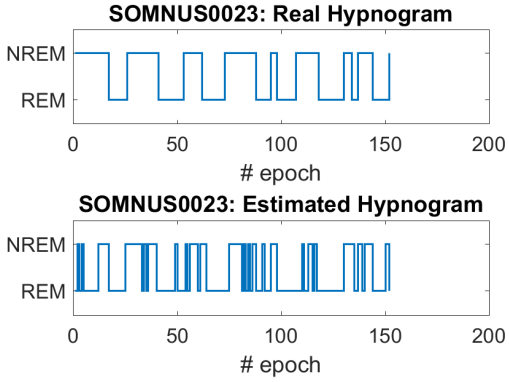
(e) Subject 20.



(f) Subject 21.



(g) Subject 22.



(h) Subject 23.

Figure E.3: Classification Results from Fisher classifier, one feature, Population A (part C)

BIBLIOGRAPHY

[1] N. Bernacchia, L. Scalise, L. Casacanditella, I. Ercoli, P. Marchionni, and E. P. Tomasini, *Non contact measurement of heart and respiration rates based on kinect™,* in *Medical Measurements and Applications (MeMeA), 2014 IEEE International Symposium on* (IEEE, 2014) pp. 1–5.

[2] L. Atallah, A. Serteyn, M. Meftah, M. Schellekens, R. Vullings, J. Bergmans, A. Osagiator, and S. B. Oetomo, *Unobtrusive ECG monitoring in the NICU using a capacitive sensing array,* Physiological measurement **35**, 895 (2014).

[3] A. K. Abbas, K. Heimann, K. Jergus, T. Orlikowsky, and S. Leonhardt, *Neonatal non-contact respiratory monitoring based on real-time infrared thermography,* Biomedical engineering online **10**, 93 (2011).

[4] N. J. Kinzie, *Ultra-wideband pulse Doppler radar for short-range targets*, Ph.D. thesis, University of Colorado at Boulder (2011).

[5] D. P.D. Jenn, *Radar fundamentals,* University Lecture, Naval Postgraduate School, Department of Electrical, & Computer Engineering (2007).

[6] J. Reed, *Introduction to ultra wideband communication systems, an* (Prentice Hall Press, 2005).

[7] S. Huseth, B. Dewberry, and R. McCrosky, *Pulsed-RF ultrawideband ranging for the GLANSER GPS-denied emergency responder navigation system,* in *Proceedings of the 2011 International Technical Meeting of The Institute of Navigation, San Diego, USA* (2001) pp. 389–396.

[8] D. T. Wisland, K. Granhaug, J. R. Pleym, N. Andersen, S. Støa, and H. A. Hjortland, *Remote monitoring of vital signs using a CMOS UWB radar transceiver,* in *14th IEEE International, New Circuits and Systems Conference (NEWCAS) , Vancouver, Canada* (IEEE, 2016) pp. 1–4.

[9] J. R. Pleym, *Xethru pulse doppler radar - the backbone technology for respiration monitoring,* https://www.xethru.com/blog/posts/xethru-pulse-doppler-radar---the-backbone-technology-for-respiration-monitoring, accessed: 2017-10-05.

[10] *X2M200 datasheet,* Novelda AS. Rev. F Preliminary. November 14, 2016.

[11] C. Li and B. Wang, *Fisher linear discriminant analysis,* (2017).

[12] O. CV, *Open cv: Introduction to support vector machines,* https://docs.opencv.org/2.4/doc/tutorials/ml/introduction_to_svm/introduction_to_svm.html, accessed on 20-10-2017.

[13] S. Raschka, *Model evaluation, model selection, and algorithm selection in machine learning. part iii - cross-validation and hyperparameter tuning,* https://sebastianraschka.com/blog/2016/model-evaluation-selection-part3.html, accessed: 2017-10-15.

[14] *X4M200 datasheet,* Novelda AS. Rev. A Preliminary. June 30, 2017.

[15] *Xethru sensor emissions: An exploration of safety regulations,* Novelda AS. XeThru Application Note. February 04, 2016.

[16] S. Graven, *Sleep and brain development,* Clinics in perinatology **33**, 693 (2006).

[17] J. A. Hobson, *Sleep is of the brain, by the brain and for the brain,* Nature **437**, 1254 (2005).

[18] J. Dudink and A. van den Hoogen, *Assessing, improving and utilizing sleep in high risk neonates.* Early human development (2017).

[19] A. K. Abbas and S. Leonhardt, *Intelligent neonatal monitoring based on a virtual thermal sensor,* BMC medical imaging **14**, 9 (2014).

[20] E. Ziganshin, M. Numerov, and S. Vygolov, *UWB baby monitor,* in *5th International Conference on Ultrawideband and Ultrashort Impulse Signals (UWBUSIS), Sevastopol, Ukraine* (IEEE, 2010) pp. 159–161.

[21] I. Immoreev and T.-H. Tao, *UWB radar for patient monitoring,* IEEE Aerospace and Electronic Systems Magazine **23**, 11 (2008).

[22] X. Huang, L. Sun, T. Tian, Z. Huang, and E. Clancy, *Real-time non-contact infant respiratory monitoring using UWB radar,* in *IEEE 16th International Conference on Communication Technology (ICCT), Chengdu, China* (IEEE, 2015) pp. 493–496.

[23] R. B. Berry, R. Budhiraja, D. J. Gottlieb, D. Gozal, C. Iber, V. K. Kapur, C. L. Marcus, R. Mehra, S. Parthasarathy, S. F. Quan, *et al.*, *Rules for scoring respiratory events in sleep: update of the 2007 AASM manual for the scoring of sleep and associated events: deliberations of the sleep apnea definitions task force of the American Academy of Sleep Medicine,* Journal of clinical sleep medicine, JCSM: official publication of the American Academy of Sleep Medicine **8**, 597 (2012).

[24] M. Grigg-Damberger, D. Gozal, C. L. Marcus, S. F. Quan, C. L. Rosen, R. D. Chervin, M. Wise, D. L. Picchietti, S. H. Sheldon, and C. Iber, *The visual scoring of sleep and arousal in infants and children,* J Clin Sleep Med **3**, 201 (2007).

[25] T. F. Anders, R. N. Emde, and A. H. Parmelee, *A manual of standardized terminology, techniques and criteria for scoring of states of sleep and wakefulness in newborn infants* (UCLA Brain Information Service/BRI Publications Office, NINDS Neurological Information Network, 1971).

[26] M. M. Grigg-Damberger, *The visual scoring of sleep in infants 0 to 2 months of age,* Journal of clinical sleep medicine: JCSM: official publication of the American Academy of Sleep Medicine **12**, 429 (2016).

[27] Z. Shinar, S. Akselrod, Y. Dagan, and A. Baharav, *Autonomic changes during wake–sleep transition: A heart rate variability based approach,* Autonomic Neuroscience **130**, 17 (2006).

[28] T. F. Anders and M. Keener, *Developmental course of nighttime sleep-wake patterns in full-term and premature infants during the first year of life.* Sleep **8**, 173 (1985).

[29] T. Leppänen, J. Töyräs, A. Muraja-Murro, S. Kupari, P. Tiihonen, E. Mervaala, and A. Kulkas, *Length of individual apnea events is increased by supine position and modulated by severity of obstructive sleep apnea,* Sleep disorders **2016** (2016).

[30] R. Michaelis, A. H. Parmelee, E. Stern, and A. Haber, *Activity states in premature and term infants,* Developmental psychobiology **6**, 209 (1973).

[31] C. L. Booth, H. L. Leonard, and E. B. Thoman, *Sleep states and behavior patterns in preterm and fullterm infants,* Neuropediatrics **11**, 354 (1980).

[32] M. Hayes, M. Akilesh, M. Fukumizu, A. Gilles, B. Sallinen, M. Troese, and J. Paul, *Apneic preterms and methylxanthines: arousal deficits, sleep fragmentation and suppressed spontaneous movements,* Journal of Perinatology **27**, 782 (2007).

[33] H. Prechtl, *Qualitative changes of spontaneous movements in fetus and preterm infant are a marker of neurological dysfunction,* Early human development **23**, 151 (1990).

[34] J. A. Hobson, *A manual of standardized terminology, techniques and criteria for scoring of states of sleep and wakefulness in newborn infants.* Clinical Neurophysiology **33**, 614 (1972).

[35] A. Sadeh, C. Acebo, R. Seifer, S. Aytur, and M. A. Carskadon, *Activity-based assessment of sleep-wake patterns during the 1st year of life,* Infant Behavior and Development **18**, 329 (1995).

[36] M. André, M.-D. Lamblin, A.-M. d'Allest, L. Curzi-Dascalova, F. Moussalli-Salefranque, S. N. T. Tich, M.-F. Vecchierini-Blineau, F. Wallois, E. Walls-Esquivel, and P. Plouin, *Electroencephalography in premature and full-term infants. developmental features and glossary,* Neurophysiologie clinique/Clinical neurophysiology **40**, 59 (2010).

[37] A. W. de Weerd and R. A. van den Bossche, *The development of sleep during the first months of life,* Sleep Medicine Reviews **7**, 179 (2003).

[38] H. Prechtl, *The organization of behavioral states and their dysfunction,* in *Seminars in Perinatology*, Vol. 16 (Elsevier, 1992) pp. 258–263.

[39] B. Koley and D. Dey, *An ensemble system for automatic sleep stage classification using single channel eeg signal,* Computers in biology and medicine **42**, 1186 (2012).

[40] T. Shimada, T. Shiina, and Y. Saito, *Detection of characteristic waves of sleep eeg by neural network analysis,* IEEE Transactions on Biomedical Engineering **47**, 369 (2000).

[41] A. Kaplan, J. Röschke, B. Darkhovsky, and J. Fell, *Macrostructural EEG characterization based on nonparametric change point segmentation: application to sleep analysis,* Journal of neuroscience methods **106**, 81 (2001).

[42] F. Ebrahimi, M. Mikaeili, E. Estrada, and H. Nazeran, *Automatic sleep stage classification based on eeg signals by using neural networks and wavelet packet coefficients,* in *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE, Vancouver, Canada* (IEEE, 2008) pp. 1151–1154.

[43] S. Gudmundsson, T. P. Runarsson, and S. Sigurdsson, *Automatic sleep staging using support vector machines with posterior probability estimates,* in *International Conference on Computational Intelligence for Modelling, Control and Automation and International Conference on Intelligent Agents, Web Technologies and Internet Commerce, Vienna, Austria*, Vol. 2 (IEEE, 2005) pp. 366–372.

[44] I. Gath and E. Bar-On, *Computerized method for scoring of polygraphic sleep recordings,* Computer programs in biomedicine **11**, 217 (1980).

[45] E. Stanus, B. Lacroix, M. Kerkhofs, and J. Mendlewicz, *Automated sleep scoring: a comparative reliability study of two algorithms,* Electroencephalography and clinical neurophysiology **66**, 448 (1987).

[46] S. Khalighi, T. Sousa, G. Pires, and U. Nunes, *Automatic sleep staging: A computer assisted approach for optimal combination of features and polysomnographic channels,* Expert Systems with Applications **40**, 7046 (2013).

[47] R. Agarwal and J. Gotman, *Computer-assisted sleep staging,* IEEE Transactions on Biomedical Engineering **48**, 1412 (2001).

[48] J. Virkkala, J. Hasan, A. Värri, S.-L. Himanen, and K. Müller, *Automatic sleep stage classification using two-channel electro-oculography,* Journal of neuroscience methods **166**, 109 (2007).

[49] T. Watanabe and K. Watanabe, *Noncontact method for sleep stage estimation,* IEEE Transactions on biomedical engineering **51**, 1735 (2004).

[50] M. Migliorini, A. M. Bianchi, D. Nisticò, J. Kortelainen, E. Arce-Santana, S. Cerutti, and M. O. Mendez, *Automatic sleep staging based on ballistocardiographic signals recorded through bed sensors,* in *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE* (IEEE, 2010) pp. 3273–3276.

[51] J. M. Kortelainen, M. O. Mendez, A. M. Bianchi, M. Matteucci, and S. Cerutti, *Sleep staging based on signals acquired through bed sensor,* IEEE Transactions on Information Technology in Biomedicine **14**, 776 (2010).

[52] F. Ebrahimi, S.-K. Setarehdan, J. Ayala-Moyeda, and H. Nazeran, *Automatic sleep staging using empirical mode decomposition, discrete wavelet transform, time-domain, and nonlinear dynamics features of heart rate variability signals,* Computer methods and programs in biomedicine **112**, 47 (2013).

[53] G. S. Chung, B. H. Choi, J.-S. Lee, J. S. Lee, D.-U. Jeong, and K. S. Park, *REM sleep estimation only using respiratory dynamics,* Physiological measurement **30**, 1327 (2009).

[54] A. Tataraidze, L. Anishchenko, L. Korostovtseva, B. J. Kooij, M. Bochkarev, and Y. Sviryaev, *Sleep stage classification based on respiratory signal,* in *37th Annual International Conference of the IEEE, Engineering in Medicine and Biology Society (EMBC), Milano, Italy* (IEEE, 2015) pp. 358–361.

[55] X. Long, J. Yang, T. Weysen, R. Haakma, J. Foussier, P. Fonseca, and R. M. Aarts, *Measuring dissimilarity between respiratory effort signals based on uniform scaling for sleep staging,* Physiological measurement **35**, 2529 (2014).

[56] C. Robert, C. Guilpin, and A. Limoge, *Review of neural network applications in sleep research,* Journal of Neuroscience methods **79**, 187 (1998).

[57] J. M. Kelly, R. E. Strecker, and M. T. Bianchi, *Recent developments in home sleep-monitoring devices,* ISRN neurology **2012** (2012).

[58] P. De Chazal, N. Fox, E. O'Hare, C. Heneghan, A. Zaffaroni, P. Boyle, S. Smith, C. O'Connell, and W. T. McNicholas, *Sleep/wake measurement using a non-contact biomotion sensor,* Journal of sleep research **20**, 356 (2011).

[59] A. Zaffaroni, B. Kent, E. O'Hare, C. Heneghan, P. Boyle, G. O'Connell, M. Pallin, P. Chazal, and W. T. McNicholas, *Assessment of sleep-disordered breathing using a non-contact bio-motion sensor,* Journal of sleep research **22**, 231 (2013).

[60] M. S. Scher, S. G. Dokianakis, M. Sun, D. A. Steppe, R. D. Guthrie, and R. J. Sclabassi, *Computer classification of sleep in preterm and full-term neonates at similar postconceptional term ages,* Sleep **19**, 18 (1996).

[61] G. Pfurtscheller, D. Flotzinger, and K. Matuschik, *Sleep classification in infants based on artificial neural networks. schlafklassifikation mit hilfe neuronaler netzwerke,* Biomedizinische Technik/Biomedical Engineering **37**, 122 (1992).

[62] P. Estévez, C. Held, C. Holzmann, C. Pérez, J. Pérez, J. Heiss, M. Garrido, and P. Peirano, *Polysomnographic pattern recognition for automated classification of sleep-waking states in infants,* Medical and Biological Engineering and Computing **40**, 105 (2002).

[63] A. Piryatinska, G. Terdik, W. A. Woyczynski, K. A. Loparo, M. S. Scher, and A. Zlotnik, *Automated detection of neonate EEG sleep stages,* Computer methods and programs in biomedicine **95**, 31 (2009).

[64] M. Čić, J. Šoda, and M. Bonković, *Automatic classification of infant sleep based on instantaneous frequencies in a single-channel EEG signal,* Computers in biology and medicine **43**, 2110 (2013).

[65] R. M. Harper, V. L. Schechtman, and K. A. Kluge, *Machine classification of infant sleep state using cardiorespiratory measures,* Electroencephalography and clinical neurophysiology **67**, 379 (1987).

[66] G. G. Haddad, H. J. Jeng, T. L. Lai, and R. B. Mellins, *Determination of sleep state in infants using respiratory variability,* Pediatric research **21**, 556 (1987).

[67] J. R. Isler, T. Thai, M. M. Myers, and W. P. Fifer, *An automated method for coding sleep states in human infants based on respiratory rate variability,* Developmental psychobiology **58**, 1108 (2016).

[68] P. I. Terrill, S. J. Wilson, S. Suresh, D. M. Cooper, and C. Dakin, *Application of recurrence quantification analysis to automatically estimate infant sleep states using a single channel of respiratory data,* Medical & biological engineering & computing **50**, 851 (2012).

[69] D. Samson-Dollfus, B. Nogues, and G. Delapierre, *Automatic analysis of sleep-wake states during the first year of life,* Neurophysiologie clinique= Clinical neurophysiology **22**, 133 (1992).

[70] M. S. Scher, S. G. Dokianakis, D. A. Steppe, D. L. Banks, and R. J. Sclabassi, *Computer classification of state in healthy preterm neonates,* Sleep **20**, 1432 (1997).

[71] J. Löfhede, M. Thordstein, N. Löfgren, A. Flisberg, M. Rosa-Zurera, I. Kjellmer, and K. Lindecrantz, *Automatic classification of background eeg activity in healthy and sick neonates,* Journal of neural engineering **7**, 016007 (2010).

[72] M. Erkinjuntti, P. Kero, J.-P. Halonen, H. Mikola, and K. Sainio, *SCSB method compared to EEG-based polygraphy in sleep state scoring of newborn infants,* Acta Paediatrica **79**, 274 (1990).

[73] T. Kirjavainen, D. Cooper, O. Polo, and C. Sullivan, *Respiratory and body movements as indicators of sleep stage and wakefulness in infants and young children,* Journal of sleep research **5**, 186 (1996).

[74] L. Giovangrandi, O. T. Inan, R. M. Wiard, M. Etemadi, and G. T. Kovacs, *Ballistocardiography—a method worth revisiting,* in *2011 Annual International Conference of the IEEE, Engineering in Medicine and Biology Society, EMBC, Boston, USA* (IEEE, 2011) pp. 4279–4282.

[75] M. Brink, C. H. Müller, and C. Schierz, *Contact-free measurement of heart rate, respiration rate, and body movements during sleep,* Behavior research methods **38**, 511 (2006).

[76] W. Zhao, H. Ni, X. Zhou, Y. Song, and T. Wang, *Identifying sleep apnea syndrome using heart rate and breathing effort variation analysis based on ballistocardiography,* in *37th Annual International Conference of the IEEE, Engineering in Medicine and Biology Society (EMBC), Milano, Italy* (IEEE, 2015) pp. 4536–4539.

[77] D. Austin, Z. T. Beattie, T. Riley, A. M. Adami, C. C. Hagen, and T. L. Hayes, *Unobtrusive classification of sleep and wakefulness using load cells under the bed,* in *2012 Annual International Conference of the IEEE, Engineering in Medicine and Biology Society (EMBC), San Diego, USA* (IEEE, 2012) pp. 5254–5257.

[78] S. Nukaya, M. Sugie, Y. Kurihara, T. Hiroyasu, K. Watanabe, and H. Tanaka, *A noninvasive heartbeat, respiration, and body movement monitoring system for neonates,* Artificial Life and Robotics **19**, 414 (2014).

[79] J. Alihanka and K. Vaahtoranta, *A Static Charge Sensitive Bed: A new method for recording body movements during sleep,* Electroencephalography and clinical neurophysiology **46**, 731 (1979).

[80] B. H. Jansen, B. H. Larson, and K. Shankar, *Monitoring of the ballistocardiogram with the Static Charge Sensitive Bed,* IEEE Transactions on Biomedical Engineering **38**, 748 (1991).

[81] H.-Y. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand, and W. T. Freeman, *Eulerian video magnification for revealing subtle changes in the world,* ACM Transactions on Graphics (Proc. SIGGRAPH 2012) **31** (2012).

[82] M. Kumar, A. Veeraraghavan, and A. Sabharwal, *Distanceppg: Robust non-contact vital signs monitoring using a camera,* Biomedical optics express **6**, 1565 (2015).

[83] L. Tarassenko, M. Villarroel, A. Guazzi, J. Jorge, D. Clifton, and C. Pugh, *Non-contact video-based vital sign monitoring using ambient light and auto-regressive models,* Physiological measurement **35**, 807 (2014).

[84] M. H. Li, A. Yadollahi, and B. Taati, *A non-contact vision-based system for respiratory rate estimation,* in *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE* (IEEE, 2014) pp. 2119–2122.

[85] M. Villarroel, A. Guazzi, J. Jorge, S. Davis, P. Watkinson, G. Green, A. Shenvi, K. McCormick, and L. Tarassenko, *Continuous non-contact vital sign monitoring in neonatal intensive care unit,* Healthcare technology letters **1**, 87 (2014).

[86] N. Koolen, O. Decroupet, A. Dereymaeker, K. Jansen, J. Vervisch, V. Matic, B. Vanrumste, G. Naulaers, S. Van Huffel, and M. De Vos, *Automated respiration detection from neonatal video data.* in *ICPRAM (2)* (2015) pp. 164–169.

[87] M. J. Maisels and A. F. McDonagh, *Phototherapy for neonatal jaundice,* New England Journal of Medicine **358**, 920 (2008).

[88] R. Murthy and I. Pavlidis, *Non-contact monitoring of breathing function using infrared imaging,* Tech. Rep. (Technical Report Number UH-CS-05-09, 9 April. Department of Computer Science, University of Houston, Houston, TX, 2005).

[89] Z. Zhu, J. Fei, and I. Pavlidis, *Tracking human breath in infrared imaging,* in *Bioinformatics and Bioengineering, 2005. BIBE 2005. Fifth IEEE Symposium on* (IEEE, 2005) pp. 227–231.

[90] A. K. Abbas, K. Heiman, T. Orlikowsky, and S. Leonhardt, *Non-contact respiratory monitoring based on real-time ir-thermography,* in *World Congress on Medical Physics and Biomedical Engineering, September 7-12, 2009, Munich, Germany* (Springer, 2009) pp. 1306–1309.

[91] H.-S. Cho, H.-K. Lyu, and Y.-J. Park, *Noninvasive heartbeat extraction from IR UWB radar signals,* in *Information and Communication Technology Convergence (ICTC), 2015 International Conference on* (IEEE, 2015) pp. 977–980.

[92] J. Eaves and E. Reedy, *Principles of modern radar* (Springer Science & Business Media, 2012).

[93] *Radar tutorial: Frequency-modulated continuous-wave radar (FMCW radar),* http://www.radartutorial.eu (), accessed: 2017-09-26.

[94] M. Skolnik, *Introduction to Radar Systems, ch. 3: MTI and Pulse Doppler Radar* (McGraw-Hill, 3rd ed., 2002).

[95] D. Kocur, M. Švecová, and J. Rovňáková, *Through-the-wall localization of a moving target by two independent ultra wideband (UWB) radar systems,* Sensors **13**, 11969 (2013).

[96] V. M. Lubecke, O. Boric-Lubecke, A. Host-Madsen, and A. E. Fathy, *Through-the-wall radar life detection and monitoring,* in *Microwave Symposium, 2007. IEEE/MTT-S International, Honolulu, USA* (IEEE, 2007) pp. 769–772.

[97] I. Immoreev and S. Ivashov, *Remote monitoring of human cardiorespiratory system parameters by radar and its applications,* in *4th International Conference on Ultrawideband and Ultrashort Impulse Signals. UWBUSIS 2008. Sevastopol, Ukraine* (IEEE, 2008) pp. 34–38.

[98] *Revision of part 15 of the commission's rules regarding ultra-wide- band transmission systems,* (2002).

[99] *Radar tutorial: Pulse integration,* http://www.radartutorial.eu (), accessed: 2017-09-28.

[100] N. Hafner, I. Mostafanezhad, V. M. Lubecke, O. Boric-Lubecke, and A. Host-Madsen, *Non-contact cardiopulmonary sensing with a baby monitor,* in *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE, Lyon, France.* (IEEE, 2007) pp. 2300–2302.

[101] C. Li, J. Lin, and Y. Xiao, *Robust overnight monitoring of human vital signs by a non-contact respiration and heartbeat detector,* in *Engineering in Medicine and Biology Society, 2006. EMBS'06. 28th Annual International Conference of the IEEE, New York, USA* (IEEE, 2006) pp. 2235–2238.

[102] J. Lin and C. Li, *Wireless non-contact detection of heartbeat and respiration using low-power microwave radar sensor,* in *Microwave Conference, 2007. APMC 2007. Asia-Pacific* (IEEE, 2007) pp. 1–4.

[103] A. D. Droitcour *et al.*, *Non-contact measurement of heart and respiration rates with a single-chip microwave doppler radar,* Ph.D. thesis, Stanford University (2006).

[104] D. Dei, G. Grazzini, G. Luzi, M. Pieraccini, C. Atzeni, S. Boncinelli, G. Camiciottoli, W. Castellani, M. Marsili, and J. Lo Dico, *Non-contact detection of breathing using a microwave sensor,* Sensors **9**, 2574 (2009).

[105] M. Sekine and K. Maeno, *Non-contact heart rate detection using periodic variation in doppler frequency,* in *Sensors Applications Symposium (SAS), 2011 IEEE, San Antonio, USA* (IEEE, 2011) pp. 318–322.

[106] S. M. Sharpe, J. Seals, A. H. MacDonald, and S. R. Crowgey, *Non-contact vital signs monitor,* (1990), uS Patent 4,958,638.

[107] Y. S. Lee, P. N. Pathirana, C. L. Steinfort, and T. Caelli, *Monitoring and analysis of respiratory patterns using microwave doppler radar,* IEEE journal of translational engineering in health and medicine **2**, 1 (2014).

[108] B.-J. Jang, S.-H. Wi, J.-G. Yook, M.-Q. Lee, and K.-J. Lee, *Wireless bio-radar sensor for heartbeat and respiration detection,* Progress In Electromagnetics Research C **5**, 149 (2008).

[109] H. Avagyan, A. Hakhoumian, H. Hayrapetyan, N. Pogosyan, and T. Zakaryan, *Portable non-contact microwave doppler radar for respiration and heartbeat sensing,* Armenian Journal of Physics **5**, 8 (2012).

[110] S. Wang, A. Pohl, T. Jaeschke, M. Czaplik, M. Köny, S. Leonhardt, and N. Pohl, *A novel ultra-wideband 80 GHz FMCW radar system for contactless monitoring of vital signs,* in *Engineering in Medicine and Biology Society (EMBC), 37th Annual International Conference of the IEEE, Milano, Italy* (IEEE, 2015) pp. 4978–4981.

[111] L. Scalise, P. Marchionni, and I. Ercoli, *A non-contact optical procedure for precise measurement of respiration rate and flow,* Biophotonics: Photonic Solutions for Better Health Care II, Proceedings of SPIE , 77150G (2010).

[112] G. Ossberger, T. Buchegger, E. Schimback, A. Stelzer, and R. Weigel, *Non-invasive respiratory movement detection and monitoring of hidden humans using ultra wideband pulse radar,* in *International Workshop on Ultra Wideband Systems joint with Conference on Ultrawideband Systems and Technologies, Kyoto, Japan.* (IEEE, 2004) pp. 395–399.

[113] K. Higashikaturagi, Y. Nakahata, I. Matsunami, and A. Kajiwara, *Non-invasive respiration monitoring sensor using UWB-IR,* in *IEEE International Conference on Ultra-Wideband. ICUWB . Hannover, Germany.*, Vol. 1 (IEEE, 2008) pp. 101–104.

[114] J. Yan, H. Hong, H. Zhao, Y. Li, C. Gu, and X. Zhu, *Through-wall multiple targets vital signs tracking based on VMD algorithm,* Sensors **16**, 1293 (2016).

[115] F. Khan and S. H. Cho, *A detailed algorithm for vital sign monitoring of a stationary/ non-stationary human through IR-UWB radar,* Sensors **17**, 290 (2017).

[116] A. Lazaro, D. Girbau, and R. Villarino, *Analysis of vital signs monitoring using an IR-UWB radar,* Progress In Electromagnetics Research **100**, 265 (2010).

[117] P. Tworzydlo, *Monitoring Breathing Using a Doppler Radar*, Ph.D. thesis, Carleton University Ottawa (2016).

[118] C. C. Johnson and A. W. Guy, *Nonionizing electromagnetic wave effects in biological materials and systems,* Proceedings of the IEEE **60**, 692 (1972).

[119] J. C. Lin, *Microwave sensing of physiological movement and volume change: A review,* Bioelectromagnetics **13**, 557 (1992).

[120] P. Marchionni, L. Scalise, I. Ercoli, and E. Tomasini, *An optical measurement method for the simultaneous assessment of respiration and heart rates in preterm infants,* Review of Scientific Instruments **84**, 121705 (2013).

[121] *Radar tutorial: In-phase & quadrature procedure,* http://www.radartutorial.eu (), accessed: 2017-10-01.

[122] *Signal processing stack exchange: What is the connection between analog signal to noise ratio and signal to noise ratio in the IQ plane in a quadrature demodulation system?* https://dsp.stackexchange.com/questions/24372/what-is-the-connection-between-analog-signal-to-noise-ratio-and-signal-to-noise, accessed: 2017-10-05.

[123] *Electromagnetic compatibility and radio spectrum matters (ERM); short range devices (SRD) using ultra wide band technology (UWB); harmonized encovering the essential requirements of article 3.2 of the R&TTE directive; part 1: Requirements for generic UWB applications,* ETSI. April , 2014.

[124] A. Ahlbom, U. Bergqvist, J. Bernhardt, J. Cesarini, M. Grandolfo, M. Hietanen, A. Mckinlay, M. Repacholi, D. Sliney, J. A. Stolwijk, *et al., Guidelines for limiting exposure to time-varying electric, magnetic, and electromagnetic fields (up to 300 GHz),* Health physics **74**, 494 (1998).

[125] G. G. Mazeika and R. Swanson, *Respiratory inductance plethysmography an introduction,* Pro-Tech Services Inc., Mukiteo, WA, USA, Tech. Rep (2007).

[126] *Linear discriminant analysis classifier (LDAC),* http://mlpy.sourceforge.net/docs/3.4/lin_class.html, mlpy v3.4.0 documentation.

[127] S. Wold, K. Esbensen, and P. Geladi, *Principal component analysis,* Chemometrics and intelligent laboratory systems **2**, 37 (1987).

[128] M. Welling, *Fisher linear discriminant analysis,* Department of Computer Science, University of Toronto **3** (2005).

[129] A. M. Martínez and A. C. Kak, *PCA versus LDA,* IEEE transactions on pattern analysis and machine intelligence **23**, 228 (2001).

[130] Wikipedia, *K-nearest neighbors algorithm — Wikipedia, the free encyclopedia,* https://en.wikipedia.org/wiki/K-nearest_neighbors_algorithm, accessed on 20-10-2017.