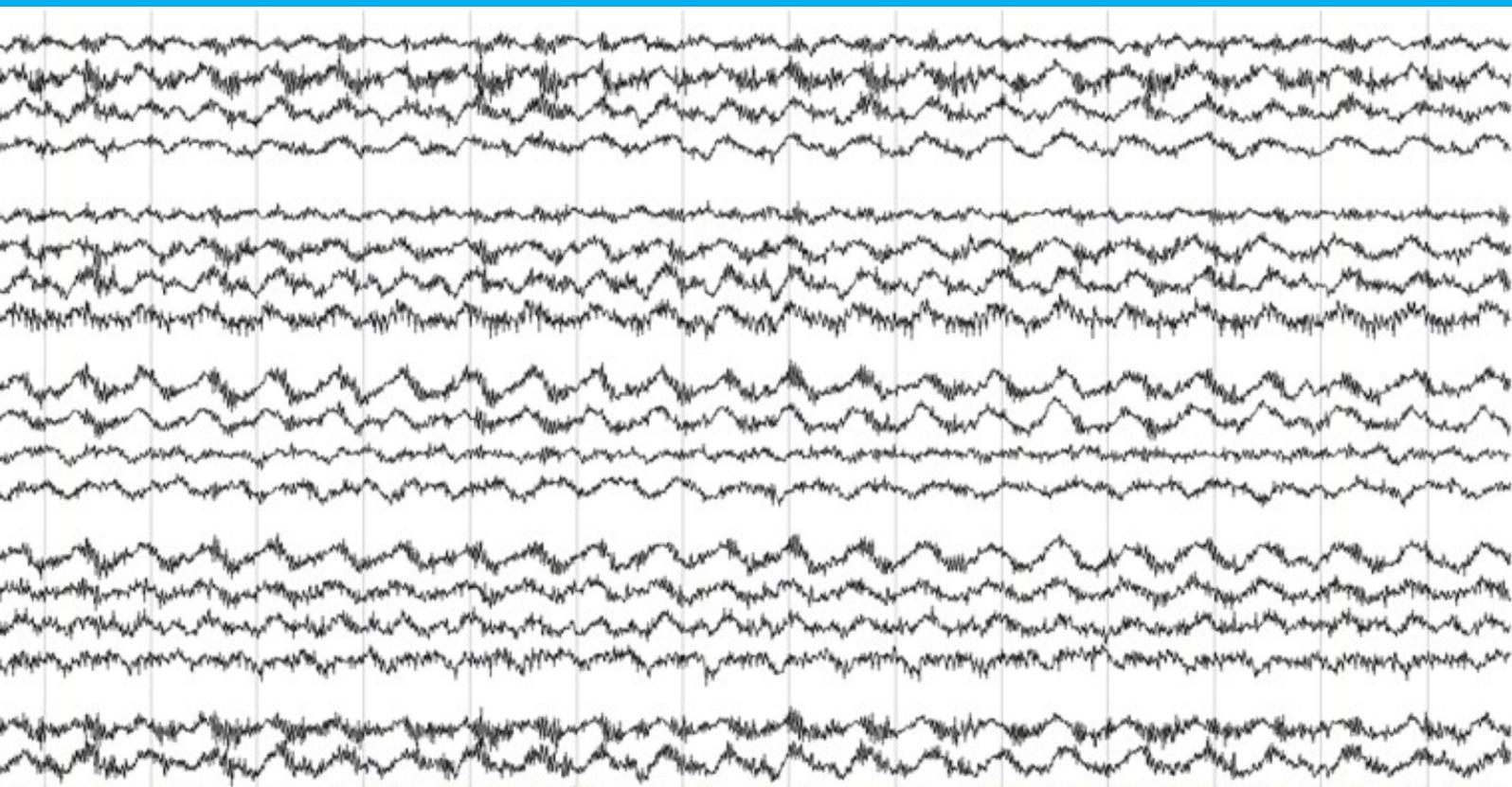


MSc Thesis

Quantitative Analysis in Neonatal Healthcare

T. Beintema

Detecting Delta Brushes
with the Wavelet Transform



MSc Thesis

Quantitative
Analysis in
Neonatal
Healthcare

by

T. Beintema

to obtain the degree of Master of Science
in Biomedical Engineering
at the Delft University of Technology,
to be defended publicly on Thursday October 24, 2019 at 13:00.

Student number: 4094379
Graduation committee: Prof. dr. ir. W. A. Serdijn, TU Delft, Bio-electronics
Dr. B. Hunyadi, TU Delft, Biomedical Signal Processing
Ir. T. G. Goos, TU Delft, Medical Instruments & Bio-Inspired Technology

An electronic version of this thesis is available at <http://repository.tudelft.nl/>.

Abstract

In the Netherlands the preterm birth rate is around 7%. The problematic part is that these preterm infants make up for 72% of all deaths during or shortly after birth. To provide a way to aid these infants the following question was stated by the doctor: "turn on a light in the Neonatal Intensive Care Unit when the preterm infant's development is not as expected". This question was interpreted as following: first of all, find an informative signal present for preterm infants which can provide knowledge on the state of development. Secondly, find a way to quantify this signal through measurement and provide a way to automate the detection.

Based on literature the delta brush was selected as a promising waveform in the neonatal EEG, which can be used to follow the maturation of the brain. The spectral behavior of the EEG is used to identify the specific waveforms, but this analysis of the EEG is still mainly based on visual analysis, which is a subjective and time-consuming task. A way to quantify and automate this process was found in the Continuous Wavelet Transform (CWT). Using the Morlet mother wavelet, a time-frequency distribution of the EEG was calculated. The CWT's redundant behavior was found to provide a finer sampling which could be used to reinforce specific waveform behavior.

A third goal was to investigate the troubles that new innovations face in being accepted by the medical field. This was found to be related to the building of trust. The key elements for building trust in this work were identified as: interpretability of the results, transparency of the process, and ease of operation. These elements will be taken into account in the proposed solution.

A set of high-density EEG measurements of preterm infants was acquired through collaboration with the Academic Hospital Antwerp and the Erasmus Medical Center. This high-density consists of 19 electrode channels instead of the more commonly used 11 electrode channels. The measurement was done at 30 weeks postmenstrual age (PMA). These measurements were accompanied with annotation for the waveform of interest. An interactive application for visualization of the EEG data and wavelet coefficients was designed in MATLAB. This application can run standalone and can be used to visualize the detector's results for the end user. Based on the inherent characteristics of the target waveform two detection features were chosen. The first feature is a measure of the ratio between peak in the high-frequency region (3.3-40 Hz) and the low-frequency region (0.1-3.3 Hz). The second feature is an energy calculation based on a windowed Squared Energy Operator (SEO). A triple-threshold multi-channel detector, based on these features, was initialized and performance was tested over the full threshold range.

A Precision-Recall (PRC) curve was provided showing the total solution space for both feature thresholds. Based on the point closest to (100,100) in the training set PRC, a single threshold was selected for validation with the validation datasets. This has resulted in a detector with a sensitivity of 73.66% and a precision of 21.71%. This performance is expected to increase with cross-validation, due to lower quality EEG datasets in the validation set. The differences in performance can be explained by the uncertainties in the exact delta brush behavior and the incomplete annotation set, which causes possible detected delta brush activity to be labeled as false positives.

In the end an informative waveform was found and a suitable quantification method has been proposed. A delta brush detector has been designed based on the inherent characteristics of the waveform. A standalone interactive application has been developed to visualize the EEG signal in combination with its time-frequency behavior. The application can also be used in combination with the detector to provide insight into the decision-making process and to visualize the detector's output.

“It is good to have an end to journey toward;
but it is the journey that matters, in the end.”
— *Ursula K. Le Guin, The Left Hand of Darkness*

Preface

Dear reader,

This thesis features a look into the effort spent on answering the seemingly easy question: "Turn on a light in the NICU when the infant's development is not as expected". In the process of finding the answer, I did not only learn a lot about the broad range of the scientific fields featured in this thesis. But in the end, I learned at least as much about myself. The precise "magic" needed to make the light glow in the NICU remains a mystery for know, but looking back at it all, I can say that it was surely a rewarding journey.

I would like to thank Wouter Serdijn for his guidance through this expedition. It was a real pleasure discussing all facets that we encountered along the way. It's amazing how you and the bio-electronics department have this continuous positive mindset and interest in researching the unknown. The availability for problem-solving and philosophical coffee-breaks was also very appreciated. I believe this atmosphere is one of the many secrets to why such amazing results come from this group.

Leen de wispelaere, Jeroen Dudink, Bori Hunyadi, Tom Goos, and all other people I met along the way, thank you for sharing your own personal passion into helping me get that one step further. It was due to your aid that I was able to get into depth in all of the fields that I encountered.

A big thanks to my lovely friends and family, through your continuous love and support it was possible to keep going in any way necessary. You where there when I needed you and gave me the space when I needed some time for myself. In the end, it wouldn't have been possible without you!

Last but surely not least, I would like to thank Heleen for her continuous effort to understand what I was working on and your intuitive feeling for when it was time to talk about something else. I know that it wasn't always easy for you to see how to help me, but I want you to know that I do appreciate everything you did to help me along the way.

Back to you, dear reader. Thank you for taking the time to read this thesis. I hope that by sharing my own personal passion through this writing, I can help you get that one step further into answering seemingly easy questions.

*T. Beintema
Delft, October 2019*

Contents

| | | |
|----------|--|-----------|
| 1 | Introduction | 1 |
| 1.1 | Problem Statement | 1 |
| 1.2 | Neonatal Development | 2 |
| 1.3 | Possible Analysis Methods | 3 |
| 1.4 | Importance of Trust for New Innovations | 3 |
| 1.5 | Thesis Objectives | 4 |
| 1.6 | Thesis Outline | 4 |
| 2 | Neonatal Electroencephalography | 5 |
| 2.1 | Basic Principle of EEG. | 5 |
| 2.2 | Maturation of Neonatal EEG | 7 |
| 2.3 | Neonatal EEG Analysis | 10 |
| 3 | Quantitative EEG Analysis | 11 |
| 3.1 | Analysis of the Neonatal EEG | 11 |
| 3.2 | Spectral Analysis | 12 |
| 3.3 | Wavelet Analysis | 14 |
| 3.4 | Numerical Implementation of the Wavelet Transform | 15 |
| 3.5 | Quantitative EEG Analysis with the Wavelet Transform | 17 |
| 4 | Tools & Materials | 19 |
| 4.1 | Choice for Maturation Waveform | 19 |
| 4.2 | Choice for Quantitative Analysis Method | 19 |
| 4.3 | Available Dataset | 20 |
| 5 | Preprocessing and Visualization | 23 |
| 5.1 | Preprocessing for the EEG Dataset | 23 |
| 5.2 | Wavelet Transform for the EEG Dataset | 25 |
| 5.3 | <i>Intermezzo: Importance of Building Trust</i> | 28 |
| 5.4 | Interactive Application Environment | 28 |
| 5.5 | Validation of Trust in Interactive Application | 30 |
| 5.6 | Conclusion | 32 |
| 6 | Data Acquisition | 33 |
| 6.1 | Target Sample Collection | 33 |
| 6.2 | Non-Target Sample Collection | 39 |
| 6.3 | Conclusion | 40 |
| 7 | Waveform Detector | 41 |
| 7.1 | Detection Principle | 41 |
| 7.2 | Detection Features | 42 |
| 7.3 | <i>Intermezzo: Performance Measures</i> | 46 |
| 7.4 | Threshold Selection | 48 |
| 7.5 | Detector Principle | 50 |
| 7.6 | Comparison Referenced to Non-Referenced Features | 55 |
| 7.7 | Single-Threshold Detector | 56 |
| 7.8 | Conclusion | 58 |
| 8 | Conclusions & Recommendations | 59 |
| 8.1 | Conclusions | 59 |
| 8.2 | Scientific Contributions | 60 |
| 8.3 | Recommendations | 61 |

| | | |
|----------|---|-----------|
| A | Application's Assignment & Questionnaire | 65 |
| B | Examples of Collected Samples | 69 |
| C | Comparison of Energy Calculation Principles Feature II | 75 |
| D | Comparison of PRC Solution Space For Referenced and Non-Referenced Detection | 77 |
| | Bibliography | 83 |

Introduction

1.1. Problem Statement

We have all experienced the wonders of human reproduction at the beginning of our very own life. We have learned in biology that during conception the male sperm cell fuses with the female egg cell and 40 weeks later we can expect a baby to be born. But what we don't all realize is that a lot can go wrong during the development of a baby, which may result in the infant being born preterm. Even though the techniques used in the neonatal care unit have advanced, there are still a lot of preterm infants born with disabilities [1].

An infant is born preterm if its birth is before 38 weeks of development. In the Netherlands, the minimum age for a baby to be considered viable is 24 weeks [2]. A baby born between 22 weeks of development and 28 weeks is classified as extremely preterm [3]. In the Netherlands the preterm birth rate is around 7%, which includes 0.4% that is born extremely preterm. The problematic part is that these preterm infants make up for 72% of all deaths during birth or shortly after, where the extreme preterms make up for a third of all deceased infants [4]. As illustrated in Fig. 1.1. Even if the infant survives it still has a high chance of acute disability or later in life, like cognitive, motor, visual or hearing impairment, behavioral problems or chronic injury. For extreme preterm infants that survive there is a 20-50% chance of disability [2].

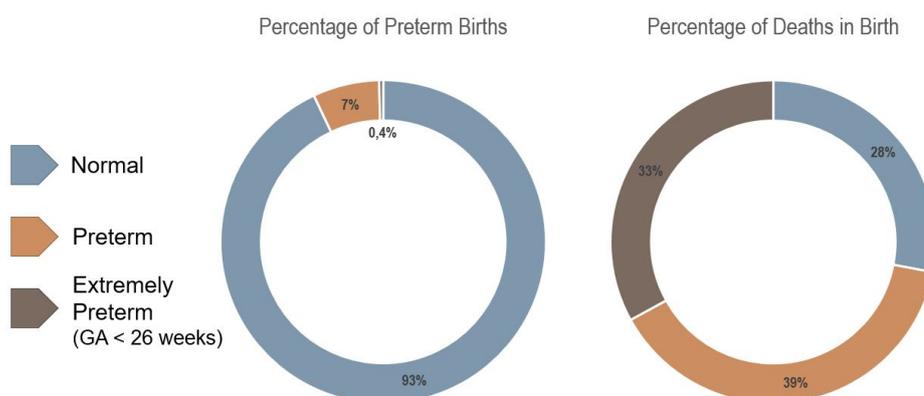


Figure 1.1: Percentages of preterm birth and deaths related to preterm birth. Data from [4].

To be able to assess the development of a preterm infant it is important to know at what age the baby was born and how much time has passed since the conception. Because the exact moment of conception is not readily available, another moment is chosen to depict the start of pregnancy, the first day of the last menstrual period of the mother (LMP). The pregnancy's normal due date follows 40 weeks from this moment [5].

The number of weeks passed between the LMP and birth is labeled as the *gestational age* (GA). The elapsed time after birth is called the *chronological age*. Because it is also of interest to know how long the infant has been developing, the summation of the gestational age and the chronological age is recorded as the *postmenstrual age* (PMA) [5]. A graphic overview of these terms is provided in Fig. 1.2.

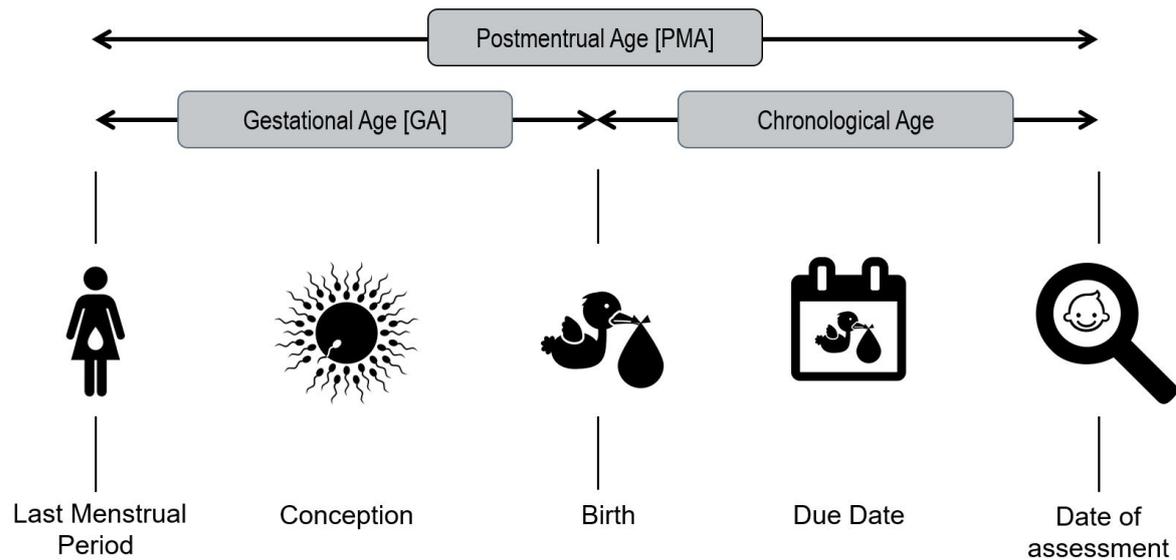


Figure 1.2: Neonatal age terminology.

1.2. Neonatal Development

The first eight weeks of development are part of the embryonic period. During this period the general structure of the central nervous system, which consists of the brain and the spinal cord, is formed. The neurons will have been generated and migrated to form the essential processing networks. By the end of the embryonic period, the basic formation of the sensory and motor region will be completed [6].

From the ninth week of gestation, human development enters the fetal period. The brain starts with a totally smooth surface and the gyral (ridges) and sulcal (grooves) foldings will develop gradually [6]. The maturation and gyration of the cortex in the occipital regions is the first to occur, followed by the frontal and temporal regions¹ [1]. Around 17-23 weeks of neonatal development the brain activity starts [7]. From 28 weeks to 40 weeks PMA the brain goes through a rapid enlargement with its weight increasing fourfold [8]. See Fig. 1.3 for an illustration of the growth and gyration in the fetal period.

¹The brain's outer layer (cortex) can be subdivided into four major regions. From front to back: the frontal region, the parietal region and the occipital region. The temporal regions can be found at both the left and right side of the brain.

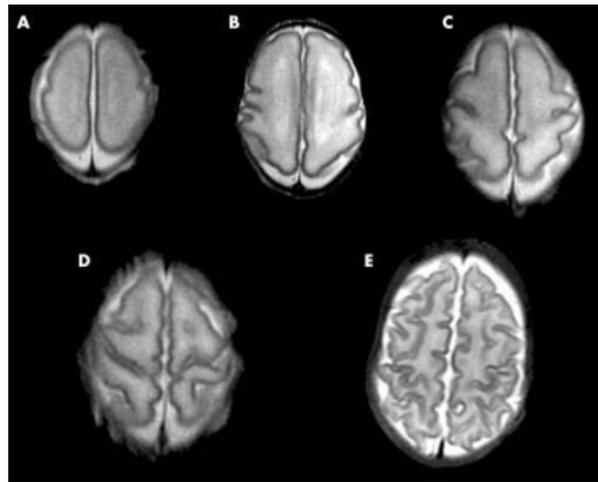


Figure 1.3: Growth and gyration of the brain from (A) 25 weeks GA to (E) 39 weeks GA. Adapted from [9].

After 40 weeks of development (*term age*), the brain will consist of 10^{11} neurons [7]. This is in the same order of magnitude as the number of planets in the Milky Way galaxy [10]. These neurons will make up the different cortical regions, each with its own distinct function.

A form of feedback on the development of these regions could be used as a diagnostic tool in the neonatal care unit. Luckily, the neuron's activity is based on the change in potential in its cell body and along the axon connecting to the subsequent neurons, the so-called action potential. The activation of subsequent neurons follows from the change in postsynaptic potential, which is triggered by the action potential [7]. These potentials can be measured with *electroencephalography* (EEG). This is also one of the technologies used to get clinical insight into the infant's well-being.

1.3. Possible Analysis Methods

Through EEG the bio-electric activity of the brain can be measured. By placing electrodes uniformly over the head it is possible to get feedback on the development of the different brain regions. The analysis of an EEG is still based on visual analysis, which depends mostly on the appearance and duration of specific waveforms. This makes the analysis highly dependent on the expertise of the electroencephalographer and difficult to compare on a broader scale [1], [11]. The use of a quantitative and automatic approach could be a valuable addition to the toolbox of the doctor to be able to receive continuous diagnostic feedback on the development of the preterm infant.

1.4. Importance of Trust for New Innovations

If a new analysis method is found to perform well and is suggested for use in the medical field, it is not certain that it will indeed be used. In the field of healthcare, 21% of employees is afraid to lose job due to artificial intelligence and automation [12]. One of the issues concerns the problems in the interpretation of the outputs of these so-called "black box models". Meaning that it is not clear why a certain output is reached, for not only the physician but also the engineer who created them [13]. To counter this, more energy is spent on providing rich visualizations and automatic explanations in combination with these outputs [13]. One of the key factors in the acceptance of new innovations is the building of trust [14], some of the examples on how this could be done are: the innovation should operate smooth and intuitively, the innovations should be able to explain and justify its decisions, and the innovation should have positive reviews from other users.

1.5. Thesis Objectives

This need for monitoring and analysis in the NICU was also emphasized during a visit to the Neonatal Intensive Care Unit of the Wilhelmina Children's Hospital in Utrecht. The wish of the doctor was to have more insight in the development of the preterm infants and the ability to react immediately when development when something is wrong. In response to asking what he needed to most, he asked for: "a light to turn on in the Neonatal Intensive Care Unit when the preterm infant's development was not as expected". This question was the starting point of this thesis work. It was interpreted into two goals which should be answered in this graduation work.

Find an informative signal which is present for preterm infants which can provide knowledge on the state of development. There are a lot of different signals a doctor can see and measure to see that a patient is not healthy, but is it possible to find a signal which can predict this outcome beforehand? This could open up the possibility for preventive care and a possible better outcome for the patient.

Find a way of quantifying this signal through measurement and provide a way to automate the detection. There are many different ways for a doctor to see that something is wrong with a patient, but only by writing his observations down and measuring these phenomena in precise values it becomes possible to see trends and compare findings among patients and even hospitals. This opens up the possibilities for automatic measurement and analysis.

Develop the proposed solution in a way that it would be accepted by the medical field. The difficulties in acceptance of current innovations and ways of increasing this acceptance will be studied. The outcome of this will be used in the design of the proposed solution.

1.6. Thesis Outline

This thesis will act as a guide through the necessary steps to answer these goals. Chapter 2 and 3 provides the background knowledge needed to answer these questions. Chapter 2 will focus on the medical-oriented literature around the electrical activity of the brain and ways of measuring it. It will conclude with a range of possible waveforms which serve as candidates for detection. Chapter 3 focuses on the technical side of measurements. It will provide different ways of interpreting signals and will compare the advantages and disadvantages of different techniques when applied to biosignals.

In Chapter 4 the focus will be on the tools and materials available for this thesis work. It will provide the choices for the waveform to detect and the analysis method to use. It will also provide details on the EEG dataset that was available for this study.

Chapter 5 and 6 will provide all the preparations needed to work with the EEG data. Chapter 5 focuses on the preprocessing steps needed to perform on the EEG data and the initialization of the quantitative analysis method. It will also focus on the development of visualization possibilities, which were not yet available for these types of data. Chapter 6 will contain all necessary preparation steps needed before detection of the waveform is possible. Based on information provided by the doctor it will collect EEG signal examples which should be detect and examples of activity that should be avoided.

Chapter 7 will focus on the development of the detector. Detector features based on the inherent characteristics of the waveform will be provided and the difficulties in threshold selection will be discussed. In the end the detector will be validated by applying it on validation data which wasn't used for the feature and threshold selection.

Chapter 8 will provide conclusion based on the initial goals. It will also provide the contributions of this work and recommendations based on the limitations of the current work.

All the chapters follow roughly the same structure. First background knowledge and general principles will be explained. This will continue into the methods or analysis steps based on the knowledge and principles. The chapter will end with results and conclusions. In some occasions an *intermezzo* section is inserted, providing additional information needed before further steps can be taken. Along the lines text in different markup can be found, providing an additional meaning. *Emphasize* is used for key definitions and software package used. The Typewriter font is used to specify the use of functions available in MATLAB. All functions and code has been written and tested in MATLAB version R2019A.

2

Neonatal Electroencephalography

As mentioned in the previous chapter, the brain's activity can be measured with a procedure called *electroencephalography* (EEG). This chapter will provide an overview on the basic principle behind neonatal EEG and insights into the possibilities to follow maturation of the infant through the EEG's characteristic features.

2.1. Basic Principle of EEG

The majority of the activity that is measured with EEG comes from the combined inhibitory (suppressing) and excitatory (stimulating) postsynaptic potentials from thousands of pyramidal cells close to the recording electrode. Only large pools of active cells can generate enough current to be measured on the head surface since the current needs to travel through the skin, bone, and other layers [7], [15]. Because EEG is essentially a two-dimensional representation of a three-dimensional situation it is not possible to pinpoint the exact neuronal source of a specific EEG signal. This is the so-called *inverse problem* [8].

Changes in the electrical activity of the neonatal brain and the maturation of the neuronal systems are closely related. The rule of ontogenic scheduling tells us that the bioelectric maturation occurs in a way that is predictable and time-linked [8]. This process follows the age of the brain itself and is independent of time after birth (*extrauterine life*). Though some minor differences are visible between full-term and preterm infants, they have not been regarded as clinically significant [8]. This means that it is possible to compare EEG based on the PMA regardless of the GA at which the infant was born.

The basic principle of EEG measurement consists of measuring the potential differences between a recording electrode placed on the skull and a reference electrode [7]. A third electrode, the ground electrode, is used to be able to distinguish a common voltage that is present for the recording and reference electrode and is used to reduce the common mode interference [16]. The measured analog signal will be filtered from high and low frequencies to reduce artifacts and converted into a digital signal [7].

EEG electrode placement follows the 10-20 system. In this system the precise location of each electrode is standardized in a proportional way, making it independent of head size. The head is divided into regions of 10% or 20% where starting points are marked by prominent skull landmarks. Each electrode is designated a letter with respect to its underlying brain region (Frontal, Central, Temporal, Posterior and Occipital) and a number (odd for the left side of the head and even for the right side of the head) [7], [17]. As illustrated in Fig. 2.1. When placing the electrodes it is recommended to keep the electrode-skin impedance below $5k\Omega$ for high quality recordings [15].

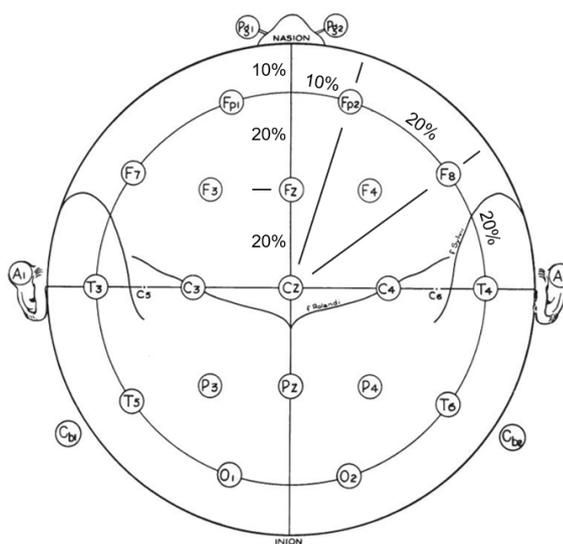


Figure 2.1: The 10-20 system. Adapted from [17].

The brain typically produces oscillating activity with an amplitude in the range of 0.5 to $100\ \mu V$. Compared to ECG signals, this is about 100 times smaller [7]. Like most biosignals, EEG-signals are noisy, nonlinear and nonstationary, meaning that its statistical characteristics are constantly changing [18]. The oscillating activity that is measured is thought to be related to the communication loops between cortex and other deeper brain structures when in rest. When activated by a task the EEG desynchronizes, the amplitude lowers and higher frequency rhythms dominate [15].

The clinical EEG is usually measured between 0.5 Hz and 40 Hz [16]. This range follows from conventional principles to reduce noise and artifacts. The low-frequency filtering follows from problems in recording low-frequency signals, the drifts experienced by the electrodes and amplifier caused saturation of the dynamic range and pushed the polygraph's pens out of range [19]. In the high-frequency spectrum, filtering was provided to reduce the influence of 50/60-Hz interference and muscle artifacts [19]. The EEG frequency band has been further classified into four activity regions: *delta activity* (<4 Hz), *theta activity* (4-7 Hz), *alpha activity* (8-13 Hz) and *beta activity* (>13 Hz) [8]. For preterm infants, an increase is found in the lower frequency regions [8]. Especially the delta frequencies are very abundant making up for more than 40% of the total EEG energy [20], as illustrated in Fig. 2.2.

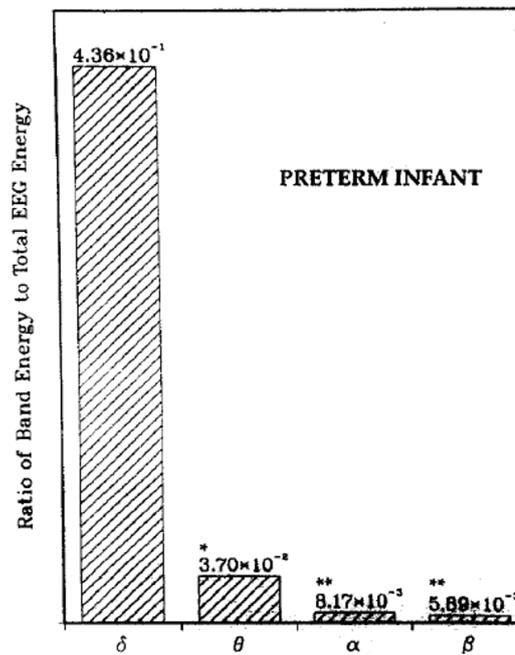


Figure 2.2: Ratio of band energy to total EEG energy in preterm infants.
Adapted from [20]

2.2. Maturation of Neonatal EEG

As the neonatal brain matures, the EEG signal develops according to four distinct trends: (1) increase in continuity, (2) appearance of sleep cycles, (3) changes in synchrony, and (4) appearance of specific spontaneous activity waveforms [21]. Each of these features is linked to specific moments in the maturation of the brain and can be used to provide diagnostic feedback on its development. Examples of the features mentioned can be found below and in Fig. 2.3.

Continuity

A distinct characteristic of the preterm neonatal EEG is that it is not continuous. The electrical activity of preterm infants is interrupted by periods of inactivity, this phenomenon is called *tracé discontinu*. The periods of inactivity are described by an amplitude below $30 \mu\text{V}$ and may even appear to be flat [8]. The active period contains low-frequency burst activity in the delta and theta regions. As the infant matures this degree of discontinuity will decrease until disappearing around term age [15].

The periods of inactivity are described by their duration, the *interburst interval* (IBI). The duration of the IBI decreases from around two minutes around 21-22 weeks PMA to zero as the brain matures to term age [8]. At 30 to 34 weeks PMA the IBI will have a duration of 4-5 seconds [8]. Maturation of the interburst interval reflects the development of the cortical folding [21]. For diagnostic evaluation, a rule proposed by Scher et al. states that multiple IBIs should not be longer than 30 seconds for preterm infants younger than 30 weeks PMA. Above the age of 30 weeks PMA the IBIs should remain below a duration of 20 seconds [22].

Sleep-Wake Cycling

The behavioral states of the preterm infant influence the EEG pattern that is measured. The behavioral states that a preterm infant experiences are: *awake* (AW), *active sleep* (AS/REM) and *quiet sleep* (QS). Active sleep will show increased continuity when compared to quiet sleep [1]. Differentiation between these states is possible from 31 weeks PMA, but differences can be found as early as 27 weeks PMA [8], [15]. It is important to capture all states when measuring a neonatal EEG, as the feature of interest could be state-dependent. It is advised to record for at least 2-3 hours to overcome this problem [15].

Synchrony

Synchrony captures the degree of bursts occurring at relatively the same time. In everyday practice this means that bursts appearing within 1.5-2 s from each other are still considered synchronous [1]. Synchronous bursts that are measured in both hemispheres are a sign of interhemispheric synchrony [1]. In extremely preterm neonates the cortical activity in the *tracé discontinu* is hypersynchronous with an interhemispheric synchrony of 90-100%. This will dip to 60-80% from 30 to 35 weeks PMA and will gradually increase to 100% as the infant approaches term age [8]. The degree of synchrony is said to be connected to the maturation of the *corpus callosum* (major pathway between the hemispheres) and can be used for diagnostic feedback on the interconnection of the different brain regions [21].

Spontaneous Activity Waveforms

Spontaneous activity in the neonatal EEG is believed to be responsible for the construction of the first neuronal connections before the sensory-activity-dependent construction takes over [23], [24]. These spontaneous bursts of activity are strongly related to the developmental age of the brain since distinct bursts can be found for different PMA. They differ from each other with respect to amplitude, frequency, and spatial location. Their migration follows the gyral development of the brain regions [8].

Delta Activity

- **Delta Waves**
Waves inside the typical delta frequency range (0-4 Hz) with an amplitude of 50-250 μV . Following maturation, the waves become more prominent in the occipital part of the brain, with duration longer than one minute until disappearing after 34 weeks of PMA [1], [21].
- **Delta Brushes**
In preterm infants, the most characteristic pattern is the *delta brush* pattern. A delta brush is a combination of short bursts of high-amplitude 8-20 Hz activity superimposed on low-amplitude waves with a frequency of 0.8-1.5 Hz, mostly found on the ascending part. The range in amplitude of the delta wave is found to vary from 50-250 μV [25]. The delta brush is found from 26 weeks PMA in the central region and will gradually migrate to the temporal and occipital regions where they are found after 32 weeks PMA [1], [8]. As the brain matures, the delta brush will show a decrease in amplitude and an increase in the frequency of superimposed activity [21].
- **Slow Anterior Dysrhythmia (SAD)**
In the last weeks of neonatal development, around 36 weeks PMA, delta patterns of short duration and an amplitude of 50-100 μV appear in the frontal area of the brain.

Theta Activity

- **Premature Temporal Theta (PTT)**
Also known as the *temporal sawtooth*. A sharp rhythmic sawtooth-like activity in the theta frequency range (4-7 Hz) found in the temporal region. Visible between 26-32 weeks PMA with an amplitude of 100-250 μV [1], [8].
- **STOPs**
Sharp rhythmic activity in the theta frequency range (4-7 Hz) found in the occipital part of the brain is called STOP (Sharp Theta rhythms on the Occipital region of the Premature). Only present for a short time (24-26 weeks PMA). They can be differentiated from the temporal sawtooth by their slightly higher frequency and lower amplitude [1] [8].

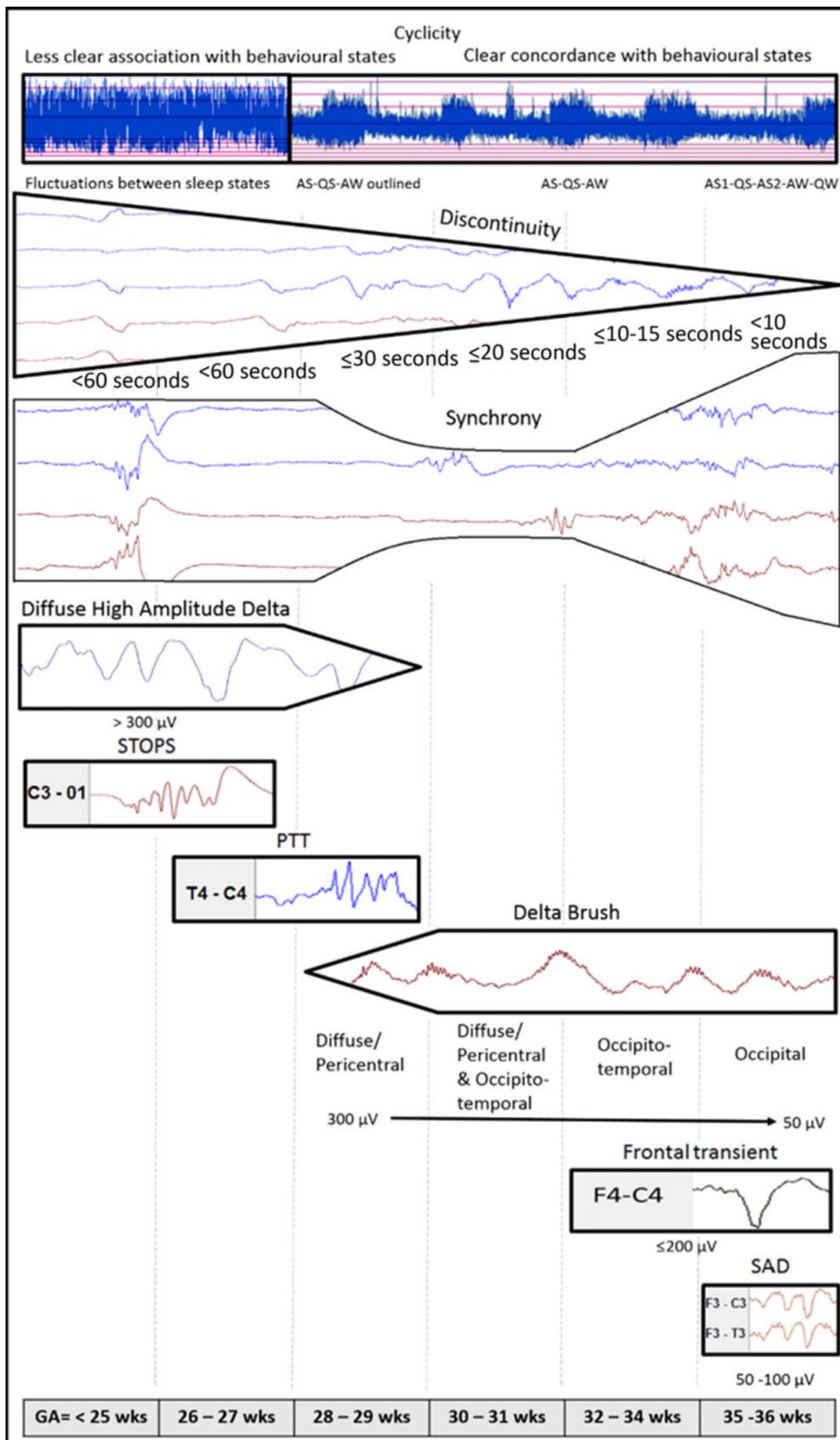


Figure 2.3: Maturation of Features in Neonatal EEG. Adapted from [21]

Artifacts

When measuring and analyzing the EEG it is important to be familiar with the possible sources of artifacts and to be able to recognize them. Artifacts that can be found in the EEG are either patient- or technology-related and can cause fluctuations in the amplitude of the original signal or the appearance of new activity in combination with the original signal. The most common patient-related artifacts are: eye-movement artifacts, ECG/EMG artifacts, movement artifacts or medicinal artifacts. The most common technology-related artifacts are: 50/60Hz interference or electrode-impedance fluctuations [7], [15].

2.3. Neonatal EEG Analysis

The above mentioned maturational features are important when the neonatal EEG is analyzed for diagnostic purposes. They can be subdivided into three categories based on their characteristics in the spatial, temporal, and spectral domain. The spatial characteristics, following from the electrode's location, can be used to get feedback on the development of specific regions. The temporal characteristics are considered when analyzing the duration of features, e.g. the IBI or specific waveforms, but also to identify synchronous behavior. This provides diagnostic feedback on the interconnections inside and between brain regions. The spectral characteristics of the features are considered when looking at the composition of the different waveforms and the background activity. As illustrated in Fig. 2.3, different frequencies in the EEG will be dominant and the composition will change as the brain matures. By looking at these different characteristics it is possible to compare the current state of development to the expected state with regards to the PMA of the infant.

The traditional clinical analysis of neonatal EEG is based on an orderly visual analysis of the EEG dataset. Taken into account are the following features: continuity, symmetry, synchrony, amplitude, reactivity, and the presence and composition of the spontaneous activity waveforms [15]. The goal of the clinician is the identification of seizures and the grading of the brain's maturation and injury [26].

The downside of such an approach is that it is purely dependent on a time-consuming visual analysis by an experienced EEG expert. This not only makes this traditional approach expensive and subjective (because definitions are not uniform)[1], but also limits the possibilities of analyzing the spectral characteristics, which can be easily visualized with analytic methods. The temporal characteristics are also easily measured by an analytic approach and could give more accurate results than a visually-based approach.

By looking at other diagnostic analysis methods, which could be used as additions to the traditional methods, it could be possible to achieve a faster, cheaper, and objective analysis approach. This also creates the possibility of getting continuous, bed-side feedback on the infant's well-being. We will look into the different possibilities for analysis in the upcoming chapter.

3

Quantitative EEG Analysis

The previous chapter concluded that a lot of interesting features are present in the neonatal EEG, each with the possibility of providing diagnostic feedback on a specific state of development of the infant. The downside was that the analysis procedures were mainly based on the visual analysis of the EEG, making it subjective, time-consuming and expensive. What is needed is a quantitative approach, to be able to reliably compare results between different measurements and patients. This chapter focuses on the quantitative analysis methods available to aid the clinician in forming his diagnosis on the neonatal EEG.

3.1. Analysis of the Neonatal EEG

In Section 2.3, a division of the maturational features was made based on their spatial, temporal, and spectral characteristics. The analysis will take a different approach for each of these different feature classes. Spatial information can be derived from the specific location at which an electrode is placed. The analysis of the spatial information can be based on the detection and comparison of specific activity or waveforms in a certain region. Temporal information is mainly based on the duration of activity and waveforms or the specific moment at which multiple bursts are measured simultaneously. The spectral characteristics will focus on the frequency composition of the different types of activity and waveforms. Based on the spectral content it will be possible to identify and detect the different kinds of activity and waveforms.

Since people working in the field of EEG are already familiar with the frequency representation of EEG in the different frequency bands and the representation of the maturational features in these different bands is well documented, an analysis methods based on the spectral domain looks promising. The main focus of this chapter will therefore be on the analysis of the spectral features

3.2. Spectral Analysis

A well-known spectral analysis tool in signal processing is the Fourier transform. With the Fourier transform it is possible to convert a signal from the time domain into the frequency domain. This allows for the construction of a so-called spectrogram of the signal showing the energy density per frequency.

$$\mathcal{F}\{x(t)\} = X(\omega) = \int_{-\infty}^{+\infty} x(t)e^{-j\omega t} dt \quad (3.1)$$

Where $x(t)$ is the original signal, which is transformed to its Fourier representation $X(\omega)$, based on the frequency denoted by ω . The analyzing function $e^{-j\omega t}$ is a complex function of time, which can be seen as a sinusoid using Euler's formula [27], [28]. The problem with the Fourier transform is that all temporal information is lost in the process; there is no information about the exact moment a feature was present. This follows from the fact that the Fourier transform integrates over the whole signal using the infinitely oscillating sinusoidal analyzing function [29].

Using other spectral analysis techniques could overcome this problem. A nonlinear energy operator (NEO) like the Teager energy operator computes the instantaneous energy per time step [30]. Using the Wigner-Vile distribution it's possible to generate a time-frequency representation of a signal based on the auto-correlation function [31]. The problem is that both these functions suffer from cross-interference terms when the signal consists of multiple frequency components occurring simultaneously [30], [32], which is undesired when (visually) analyzing a frequency-rich signal like EEG.

Short-time Fourier Transform

There are ways of retrieving temporal information using the Fourier transform. One can cut the total signal into different segments and perform the Fourier transform in a piece-wise fashion or use a windowing function, which is non-zero for the segment of interest, to shift along the target signal selecting the areas of interest [29]. These methods are known as the short-time Fourier transform (STFT) or the windowed Fourier transform [33]

$$STFT\{x(t)\} = X(\tau, \omega) = \int_{-\infty}^{+\infty} x(t)g(t - \tau)e^{-j\omega t} dt. \quad (3.2)$$

Where the standard Fourier transform (3.1) is complemented with the window function $g(t - \tau)$. It is now possible to visualize a spectrogram per time step τ . To be able to plot this three dimensional data in a two-dimensional graph, it is common to use a color map representing the intensity of the energy density. A remaining problem is that the Fourier-based methods make use of periodic oscillating sinusoidal analyzing functions which cannot be used for the precise localized spectral analysis of the non-periodic EEG signal. Since the main principle of the STFT is based on a matching approach between the windowed analyzing function and the target signal, a trade-off exists between the window size and the time-frequency resolution.

Heisenberg's Inequality

To be able to measure the precise frequency ω_t of a signal, it is necessary that the signal is observed for at least as long as the time it would need to propagate [34], the period of the frequency $T_t = 1/\omega_t$. Analysis methods using a fixed window design the window for the lowest frequency present, but this same window is applied for the whole frequency range. The resulting problem is that it is impossible to say with certainty where the frequency is measured exactly within the window. This trade-off is explained through the Heisenberg uncertainty principle. By calculating the variance of time (σ_t^2) and frequency (σ_ω^2) around a certain point in the time-frequency space a measure can be given to the uncertainty. Here a trade-off is found between the time and frequency resolution, resulting in the Heisenberg Inequality [33]

$$\begin{aligned}\sigma_t^2 \sigma_\omega^2 &\geq 1/4 \\ \sigma_t \sigma_\omega &\geq 1/2.\end{aligned}\tag{3.3}$$

In the time-frequency space, a so-called *Heisenberg-box* can be drawn based on these standard deviations, visualizing the uncertainty in the time-frequency decomposition [33], as shown in Fig. 3.1. A consequence of the Heisenberg inequality is that the area of the Heisenberg box cannot be smaller than 1/2.

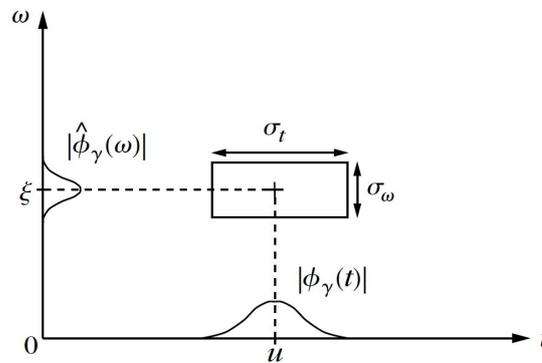


Figure 3.1: Illustrative representation of a Heisenberg-box in the time-frequency space. Adopted from [33].

This inequality is true for all functions, including window functions. When a Gaussian window is used the inequality becomes an equality [34], thus providing the optimal time-frequency area. In summary, a short window length is needed for a high temporal resolution but will result in a decrease in the spectral resolution [29], and vice versa. An analysis approach using an adaptable window size, tailored to the local characteristics of the target signal, would be preferable.

3.3. Wavelet Analysis

Instead of a rectangular window function, used in the short-time Fourier transform, a different window function can be used. The combination of a Gaussian window and the sinusoidal analyzing function used in the Fourier transform will create a localized pattern of finite length, a so-called *wavelet* (little wave), as illustrated in Fig. 3.2.

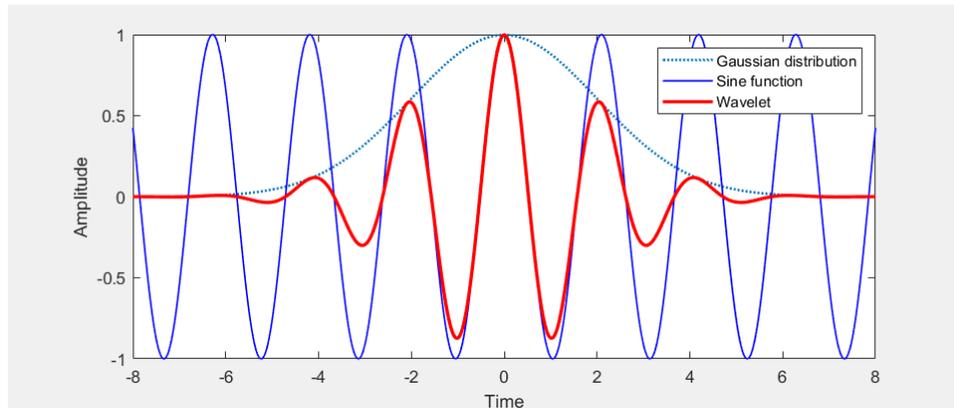


Figure 3.2: Construction of a wavelet from a sine and Gaussian function.

This wavelet can then be used in a correlation-like process with the target signal, through the calculation of the inner product [35]. This results in a maximal output where the wavelet most resembles the target signal [36]. By scaling and shifting this wavelet, a multi-frequency analysis is performed per time-step. This approach is known as the *continuous wavelet transform*

$$W(s, \tau) = \frac{1}{\sqrt{s}} \int_{-\infty}^{+\infty} f(t) \psi_0^* \left(\frac{t - \tau}{s} \right) dt. \quad (3.4)$$

The ψ_0 in (3.4) denotes the general form of the analyzing function, known as the *mother wavelet*. Under the effect of the scaling parameter s and the translation parameter τ the specific wavelets are formed from the mother wavelet. The factor $1/\sqrt{s}$ in (3.4) has a scaling effect in the vertical direction. It functions as a normalization factor needed to make sure that the energy is preserved in the area of the wavelet spectrum [29].

All mother wavelets should satisfy the admissibility condition for reconstruction to be possible. The admissibility condition is satisfied if the mother wavelet has a zero mean, $\int \psi_0(t) dt = 0$, and a band-pass like Fourier spectrum, $\int \hat{\psi}_0(\omega) = 0$ for $\omega = 0$ [35]. One of the major advantages in wavelet analysis is the possibility of tailoring the analyzing function to match the target signal. Since the wavelet transform's main principle is correlation-based, an analyzing function with similarities to the activity of interest is preferred.

Heisenberg's Inequality for Wavelets

Since the continuous wavelet transform typically uses Gaussian-based window function, an optimum is found in the time-frequency resolution. The area of the Heisenberg boxes for the wavelet transform will be as small as possible. When the Heisenberg boxes are drawn for the wavelet transform, another interesting feature shows. Through the use of the scaling parameter s , the borders of the Heisenberg-box are manipulated. The time and frequency width will be equal to $s\sigma_t$ and σ_ω/s respectively, but the area remains unchanged [33], as seen in Fig. 3.3.

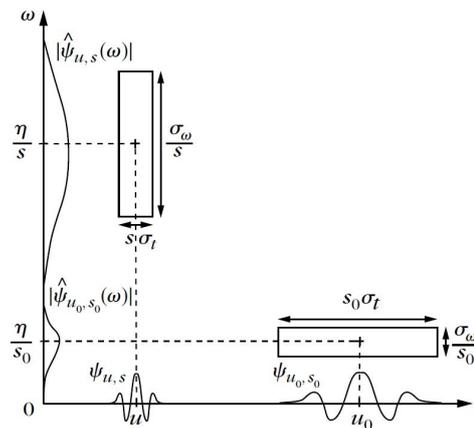


Figure 3.3: Heisenberg-boxes for different wavelet scales. Adopted from [33].

This results in a transform which is better adapted for both the analysis of high and low-frequency activity, because of this adaptable window size. An increased temporal resolution is found for high-frequency signals with a short time duration when compared to the Fourier transform. The reduced temporal resolution for low-frequency signals is traded for an increase in spectral resolution. This property to be able to "zoom in" on short duration activity makes the wavelet transform a promising choice for the analysis of non-stationary biosignals, like EEG.

3.4. Numerical Implementation of the Wavelet Transform

To be able to use the wavelet analysis in a computational practical manner, it is necessary to provide a discretized version of the wavelet transform. The continuous wavelet transform makes use of a set of continuous-valued scaling and translation parameters. It is not possible to work with continuous values in the digital domain, instead discrete steps are used. In general, the numerical implementation of the wavelet transform can be done in two different ways, each with its own advantages.

The Sparse Way

Instead of using a continuous set of values for the scaling parameter s , a discrete set can be used, like the fractional power of a base value. Typically a base value of two is taken, $s = 2^j$, providing an octave scaling of the frequency domain [37]. Since the dimensions of the Heisenberg box are controlled by the scale, it is possible to create a perfect tiling of the Heisenberg-boxes in the time-frequency plane by making the translation parameter dependent on the scale as well $\tau = 2^j n$ [33]. This so-called dyadic tiling of the time-frequency plane is illustrated in Fig. 3.4.

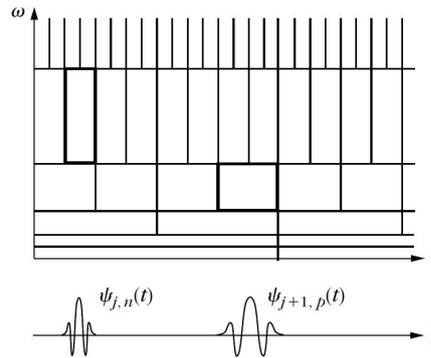


Figure 3.4: Optimal tiling of Heisenberg-boxes in the time-frequency plane with dyadic scales . Adapted from [33].

Through the use of this dyadic tiling, it is possible to represent a signal in the sparsest way with the wavelet coefficients, in other words, representing the signal in as little coefficients as possible. One condition for dyadic tiling is that an orthonormal mother wavelet is used because the wavelet coefficients need to be independent of each other [35]. This approach is commonly referred to as *Multiresolution Analysis* and can, because of its filtering in octave bands, be compared to subband coding [34].

The Redundant Way

When using the continuous wavelet transform there is significant overlap in the calculation of adjacent wavelet coefficients, providing a lot of redundant information, redundancy referring to the less compact form used to represent the signal [38]. This redundancy can be exploited to strengthen transient signal characteristics [37] or to increase the robustness when high precision is needed for reconstructing the signal [35].

This redundancy is lost with the dyadic discretization steps in the sparse approach but can be recovered by forcing overlap in the coefficients. This can simply be done by adding intermediate discretization steps. In MATLAB, this is done by defining a certain number of *voices per octave*, the number of steps needed to reach the next octave [37]. In doing so, a different set of mother wavelets becomes available, since there is no need for them to be orthonormal. The redundancy also emphasizes the transient behavior in the signal, producing ridges of local maxima in the time-frequency decomposition [38].

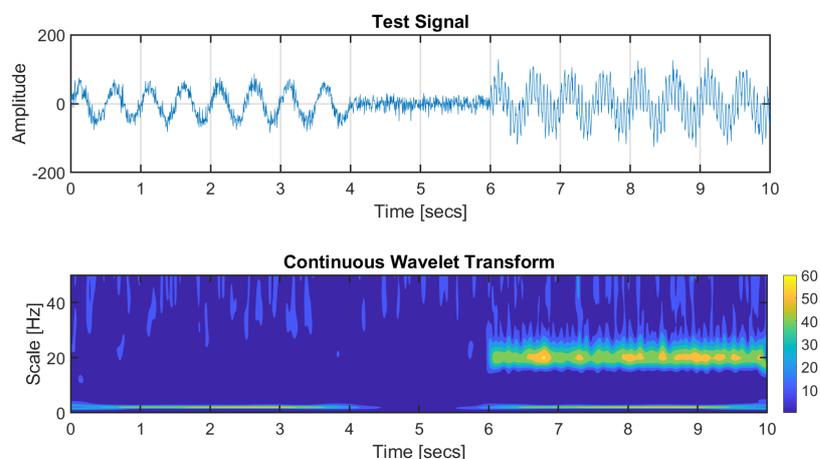


Figure 3.5: Formation of local maxima ridges in the CWT of a test signal.

3.5. Quantitative EEG Analysis with the Wavelet Transform

By using the redundant numerical implementation of the continuous wavelet transform in the field of medicine, it will be possible to visualize the time-frequency behavior of the EEG with the ridges of local maxima. These ridges can make the transient behavior of the signal visible and provide a new quantitative way for the clinician to interpret the data. This can help explain why certain detector decisions are made through a visualization. This opens up the possibility for a new approach on general EEG analysis. This approach will not only be based on the visual analysis of the EEG but will also provide a quantitative representation of the EEG dataset which can be used to compare and discuss results in an objective way.

Combined with the insight of Chapter 2 on the maturation of the EEG. It becomes possible to construct a detector based on the time-frequency characteristics of the specific maturational features, each linked to specific moments in the maturation of the neonatal brain. This will aid in the development of faster, cheaper, and objective analysis approaches for neonatal healthcare.

4

Tools & Materials

Now that the general background around quantitative analysis in the field of neonatal EEG is clear, the next step will be focused on getting the tools and materials together. Before focusing on the design of the quantitative method for EEG diagnostics, a couple of steps have to be completed: (1) selection of a suitable waveform for detection, (2) selection of the quantification method, and (3) acquisition of a suitable EEG dataset and waveform annotations. The outcome of these preparations are presented in this chapter.

4.1. Choice for Maturational Waveform

In Chapter 2, multiple EEG waveforms were presented, each linked to specific a maturational age of the infant. Of these waveforms, the delta brush provides a number of promising properties which could support its detection and provide diagnostic value. First of all, the delta brush is present from 28-29 weeks PMA till term age, making it the waveform with the longest presence during preterm development. Secondly, during the time span in which the delta brush is present, it migrates from the central region to the temporal and occipital regions in the brain. This means that the location where the delta brush is detected, can provide information about the maturation of the brain. Lastly, since the delta brush consists of delta frequency activity in combination with bursts of higher frequency activity, it is easily recognized and well-known in the medical field. This makes the delta brush a promising waveform for quantification and further analysis.

4.2. Choice for Quantitative Analysis Method

In Chapter 3, an overview was provided of the different spectral analysis methods which can be used to quantify analysis in the medical domain. One method stood out in terms of adaptability and visual interpretation. The redundant continuous wavelet transform (CWT) provides an analysis method with analysis window which can be adapted to the signal of interest. It also provides means to increase the temporal resolution for higher frequency activity, this made it possible to "zoom in" on this short duration activity. The possibility of the wavelet transform to analyze the data for different frequency scales makes it especially promising in combination with the delta brush, since the delta brush is characterized by its multicomponent-frequency content.

4.3. Available Dataset

To be able to perform the research on the quantitative analysis of the delta brush waveform, two important aspects need to be obtained through a collaboration with a hospital. First, a neonatal EEG dataset is needed, which contains delta brush activity. Meaning that the EEG is acquired from preterm infants with a minimum age of 28 weeks PMA and that the measurement locations included the central, temporal and occipital regions. Secondly, this dataset needs to be reviewed for delta brush activity by a specialist in neonatal EEG to provide a set of delta brush time annotations.

Provided EEG Dataset

The neonatal EEG dataset was obtained through a collaboration with the Academic Hospital Antwerp and the Erasmus Medical Center. This dataset was collected in a study concerning the added value of high-density EEG measurement, where 19 electrode channels are used instead of the 11 electrode channels used in the traditional neonatal EEG montage [15], as compared in Fig. 4.1. In this study 24 neonatal infants received an high-density EEG analysis at an age of 30-32 weeks PMA. This high-density EEG analysis was provided through the use of electrode cap (WaveguardTM, ANT Neuro, Enschede, The Netherlands) containing 21 electrodes. The recording was done with a sampling frequency of 200 Hz and all measurements had a minimum duration of 4 hours. Each measurement was provided with a *time till placement caps* (TPC) and a *time of good quality* (TGC). During the measurements the following artifacts were reported in one or more datasets: incubator, mains, tube feeding, movement, ECG, bridging between electrodes. These were provided as additional commentary, indicating presence of these artifacts in one or more electrodes. The dataset were provided in the 'European Data Format' (EDF), a format designed for the storage and transfer of biological signals [39].

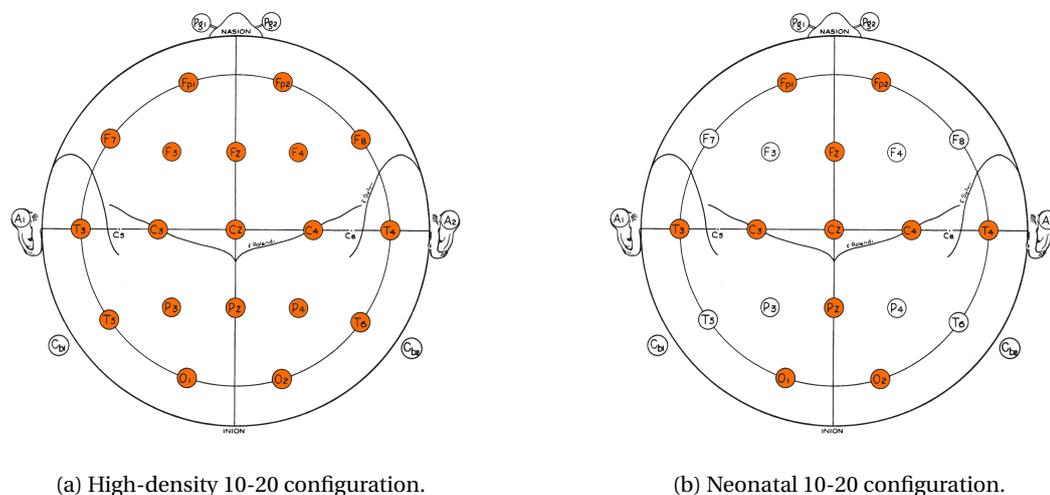


Figure 4.1: Different 10-20 electrode placement configurations, active electrodes depicted in orange. Adapted from [17].

Since the setup of this initial study had a general goal, the neonatal infants which were recruited for the study had varying health conditions. For the current study, it is first necessary to find a healthy baseline for the delta brush before taking into account the possible alterations to the delta brush activity due to low health conditions. Of the 24 available datasets, 8 datasets were selected based on their measurement quality and positive outcome in follow-up study at term age. Of the 21 electrodes used in the high-density EEG measurement, 2 electrodes were used for referencing purposes (A1 and A2), these were excluded from the analysis. The remaining 19 electrodes were used as the EEG channels to be analyzed.

Provided Delta Brush Annotations

The doctor in charge of the initial high-density EEG study, L. De Wispelaere, was asked to review the selected EEG datasets for delta brush activity. This was done by visual inspection of all channels simultaneously. No software for visualization was available in the hospital, so a suitable visualization software was selected based on the recommendations available on the EDF-format website [39]. The visualization of the edf-files was eventually done through the use of the *Online EDF Viewer* software, which is a web browser-based software running locally and offline [40]. See Fig. 4.2 for an example of the EDF visualization software with the provided dataset.

One time stamp was provided per delta brush, rounded to the nearest second. Per dataset a minimum of 50 annotations was provided. In correspondence, the doctor mentioned that she focused on the best examples of the delta brush, which may result in the presence of non-annotated delta brushes within the reviewed time range. For many of the annotations, additional comments have been provided on the channels where delta brush activity was most significant.

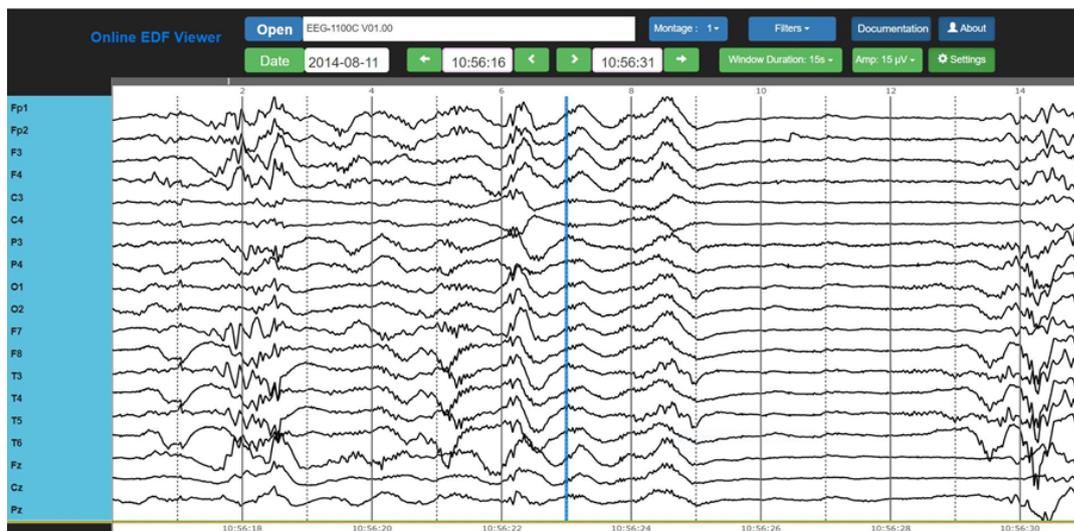


Figure 4.2: Example of the software used to visualize the dataset for annotation

5

Preprocessing and Visualization

With the analysis method, target waveform, and available materials ready, the next step is to preprocess the EEG data and to implement the wavelet transform. The technical environment used for this is provided by MATLAB. This section will explain the steps necessary to implement the wavelet transform for an EEG dataset and ways of visualizing the output.

5.1. Preprocessing for the EEG Dataset

The raw unfiltered EEG data is rarely used in the clinical setting, since it often contains a lot of artifacts and interference [7]. To aid the visual inspection of the EEG data, it is often filtered. To be able to compare the outcome of the detector with the provided annotations, the EEG data will be filtered in the same way as in the clinical setting. This section will explain the necessary preprocessing steps needed before the wavelet transform can be implemented.

Converting the EEG Dataset

The EEG datasets were provided in the European Data Format (EDF), containing all patient, technology, and signal specific information in a single file. To be able to work with the single-channel EEG signal data, this information had to be extracted from the EDF-file. This is done through the use of the freely available EEGLAB MATLAB package [41]. EEGLAB provided the EDF-file information in a MATLAB structure array. This structure was then saved locally as a MAT-file, ready for further use.

Digital Filtering

It is common practice in the medical field to filter the EEG signal before further use. High-pass filtering is used to reduce the influence of low frequency artifacts like breathing and drifting of the signal. The cut-off frequency is conventionally around 0.1-0.7 Hz [7],[19]. To reduce high frequency artifacts, e.g. due to movement or the 50/60 Hz influence of the mains, it is possible to use a low-pass filter with a cut-off frequency below 50 Hz. The mains interference can also be suppressed with a local notch filter [7]. The commonly used filter range runs from 0.5 Hz till 50 Hz [19]. To match further visualization of the EEG to the common standard, these conventional filter ranges were used, with one small adjustment. Since the frequency range of interest for the delta brush is said to run until 20 Hz, the low-pass filter value is set to 40 Hz. This eliminates the necessity of an additional notch filter, while maintaining most of the conventional filter range for an EEG signal.

To implement a frequency filter in the digital domain, two general filter structures are available: Finite Impulse Response (FIR) filters and Infinite Impulse Response (IIR) filters. The basic difference between the two is that the FIR filter's output depends only on the input data points, while an IIR filter's output also takes into account its previous outputs [42]. The general disadvantage for FIR filters over IIR filters is that a FIR filters require a higher filter order to reach the same level of performance [43]. Since the delay caused by the filtering is directly linked to the filter order, FIR filters will impose a greater delay for the same level of performance. There is however one major advantage for FIR filters in combination with biosignals. FIR filters have a constant group delay, where the delay for IIR filters is frequency dependent. For biosignals, consisting of multi-frequency components, IIR filtering would result in the distortion of the general shape of the biosignal waveforms [44]. Since preserving the shape of the waveform was of major importance in the current work, the FIR filter structure was further investigated.

The FIR filter used in this work has been designed using the `designfilt` function available in the MATLAB *Signal Processing Toolbox* [45]. A bandpass FIR filter running from 0.5 Hz till 40 Hz for a signal with a sample rate of 200 Hz of adequate performance was found for a filter order of 300 or more. Since the delay for a FIR filter is equal to half the filter order [43], a filter order of 400 was chosen. This results in a delay of 200 samples, which equals to a 1 second delay. The magnitude and phase response of the designed filter are presented in Fig. 5.1.

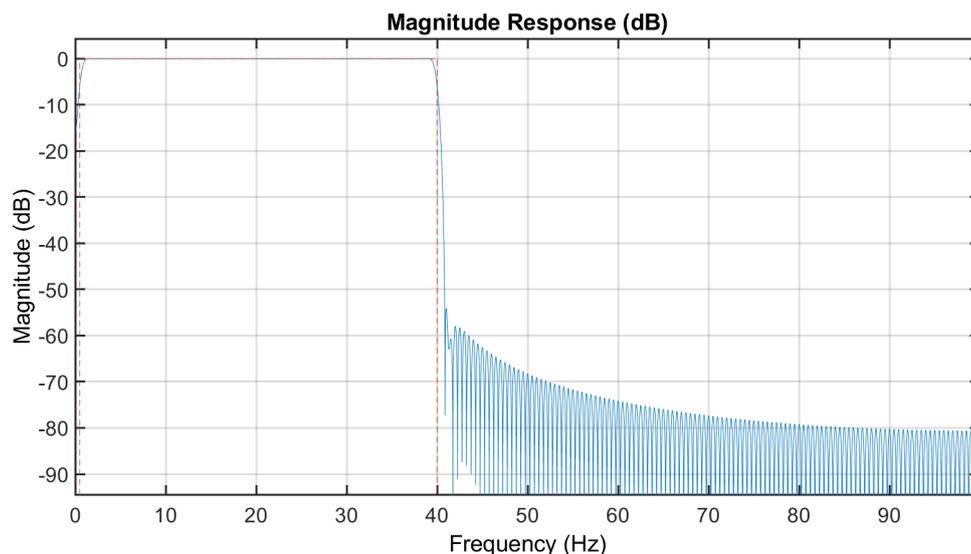


Figure 5.1: Magnitude and Phase Response of Designed FIR Filter

5.2. Wavelet Transform for the EEG Dataset

Once the signal was filtered, the next step was to calculate the wavelet transform for the EEG dataset. As explained in Chapter 3, the continuous wavelet transform calculates the output of a correlation-like calculation between the target signal and the scaled and translated version of the mother wavelet. The continuous wavelet transform was implemented using MATLAB's `cwt` function. This section will explain the steps and choices needed to implement the wavelet transform for the EEG Dataset.

Mother Wavelet Selection

By matching the mother wavelet to the waveform of interest it is possible to increase the performance of the correlation-like process. The mother wavelets most commonly used by the continuous wavelet transform for biosignals are mainly based on the Gaussian function and its derivatives, e.g. Morlet or Mexican hat [29],[38]. There are many more options available, like the orthonormal mother wavelets, e.g. Daubechies, symlets or coiflets. These are more commonly used in the non-redundant discrete wavelet transform for their energy-preserving properties [37].

The wavelet transform was implemented numerically using the `cwt` function available in the MATLAB *Wavelet Toolbox*. In 2016 MATLAB released a new version of the `cwt` function and labeled the old version as "no longer recommended". In this new version only a small selection of mother wavelets is available:

- **Morse Wavelet**
The Morse wavelet is presented as a "super family" of most common wavelets. Its shape and behavior is controlled by two parameters, symmetry and time-bandwidth product, allowing it to match the behavior of many other wavelet, e.g. Cauchy, Paul or Mexican hat [46].
- **Morlet Wavelet**
The Morlet wavelet is a well-known complex mother wavelet. Created through the combination of a sinusoidal plane wave and a Gaussian window function. Since it is based on smooth harmonic function it is well-suited for the detection of oscillatory EEG patterns and sleep spindles [29]. It is described as having a balanced variance in time and frequency resolution [37].
- **Bump Wavelet**
The Bump wavelet is only described as having a wide variance in time resolution and narrow frequency resolution [37]. No further information was available on its analytic representation and shape.

Because the precise time-frequency behavior of the delta brush waveform is not well-defined at the moment, it is difficult to predict beforehand which wavelet will match its characteristics best. The Morse wavelet might provide a high performance, because of the freedom to tune the wavelet to the characteristics of the target waveform. But for now, too little information and validation of this behavior is found in literature. Since no further characteristics of the Bump wavelet are mentioned in MATLAB's documentation or in literature, it is not a suitable candidate for this study. The Morlet wavelet, however, is well established in literature and has already been found to perform well in combination with oscillatory EEG patterns. The Morlet wavelet will be used as the mother wavelet in the further analysis using the wavelet transform in this study.

Wavelet Transform Implementation

MATLAB's `cwt` function allows for redundancy in the output coefficients by specifying the number of *voices per octave*, i.e. the number of data points between each doubling in frequency [37]. The possible range runs from 4 till 48 and has to be an even integer value. Per default 10 voices per octave are used. Since the voices per octave are defined for every doubling of the frequency, more data points will be available for the lower frequencies when compared to the higher frequencies. Meaning that the frequency resolution will not be evenly distributed along the frequency range.

For the default value of 10 voices per octave, a frequency resolution of 0.014 Hz was found for frequencies around 0.1 Hz and a frequency resolution of 2.68 was found for frequencies around 40 Hz. Since the goal of this study is partly exploratory, having a higher resolution would be advantageous. A downside would be a higher load on performance since more data points have to be calculated and saved. For 20 voices per octave, a frequency resolution of 0.004 Hz was found for frequencies around 0.1 Hz and a frequency resolution of 1.36 Hz was found for frequencies around 40 Hz, while no major performance issues were found. Unless specified differently, 20 voices per octave will be used. Since the EEG signal was filtered between 0.5 Hz and 40 Hz, the wavelet transform calculation frequency range was also limited. The high frequency boundary was set to 40 Hz, equal to the FIR filter's cutoff frequency. However, the low frequency boundary was set to 0.1 Hz instead of 0.5 Hz, because a lot of activity was still found below the cutoff frequency.

Wavelet Transform Normalization

The analytical wavelet transform commonly uses a $L2$ normalization, needed to preserve the energy [37]. As a consequence of the Heisenberg inequality, the frequency resolution differs for signals of different frequency. Resulting in a broader frequency footprint for a higher frequency signal. Since the energy is represented as the area under the power spectrum, this will result in a frequency-dependent behavior. The power spectrum's amplitude for a higher frequency signal will be lower, to maintain an equal area under the curve [29]. When the cwt is implemented numerically, however, the preservation of energy is lost, so there is no need for a $L2$ normalization to be used [37]. The cwt function uses $L1$ normalization, in practice this means that waveforms of equal amplitude but different frequency, will result in wavelet coefficients of equal magnitude [37]. Fig. 5.2 demonstrates the effect of $L1$ normalization (5.2a) and $L2$ normalization (5.2b) for a signal containing two different frequencies of equal amplitude (Fig 5.2c).

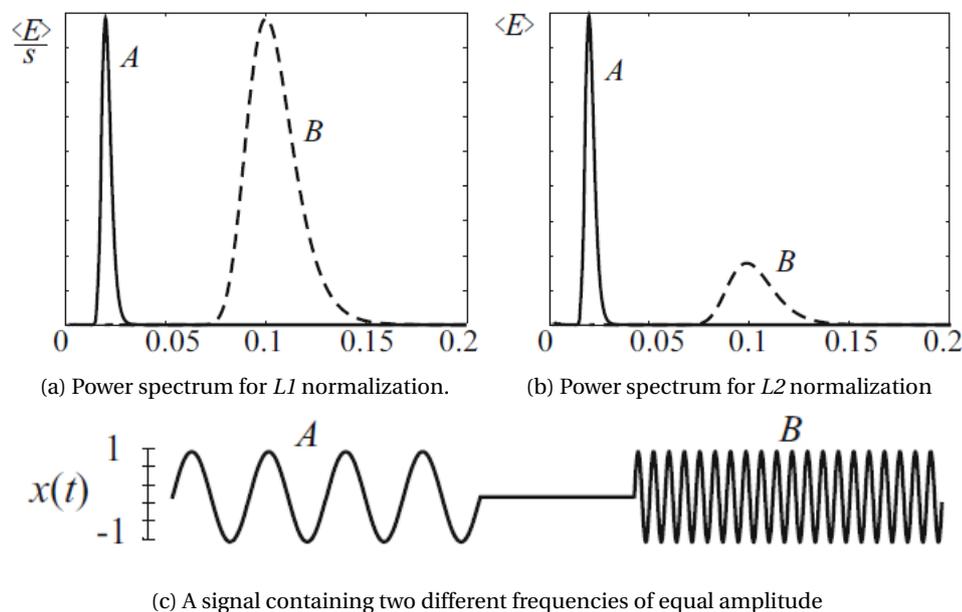


Figure 5.2: The effect of normalization on the visualization of wavelet coefficients. Adapted from [29]

Wavelet Transform Visualization

With all settings regarding the wavelet transform clear, it is time to perform the required calculations. The output of the `cwt` function consists of the wavelet coefficient's covering the whole frequency range for every signal sample. When using 20 voices per octave limited between 0.1 Hz and 40 Hz, the `cwt` produces 173 wavelet coefficients per sample. Every wavelet coefficient represents the magnitude after performing the correlation-like calculation. The output contains data specified three dimensions, i.e. sample number, frequency, and magnitude. To be able to plot this three dimensional data in a two-dimensional graph, it is common to use a color map representing the magnitude. This kind of visualization are commonly known as *spectrograms*, although for the wavelet transform the name *scalogram* is more appropriate. This comes from the fact that the precise frequencies don't necessarily match with *scales* used by the wavelet transform, although, when using the Morlet mother wavelet the scales and frequencies are matched [29].

The results of all the processing steps on the EEG data is visualized in Fig. 5.3. The 'Signal' windows show a 15 second section of the original and filtered single-channel EEG data in blue. The red lines indicate the provided delta brush annotations. The 'Fourier Transform' shows the frequency content of the total single-channel EEG signal, indicating the presence of 50Hz-interference for the non-filtered signal. The 'Continuous Wavelet Transform' window shows the calculated wavelet coefficients for the filtered signal.

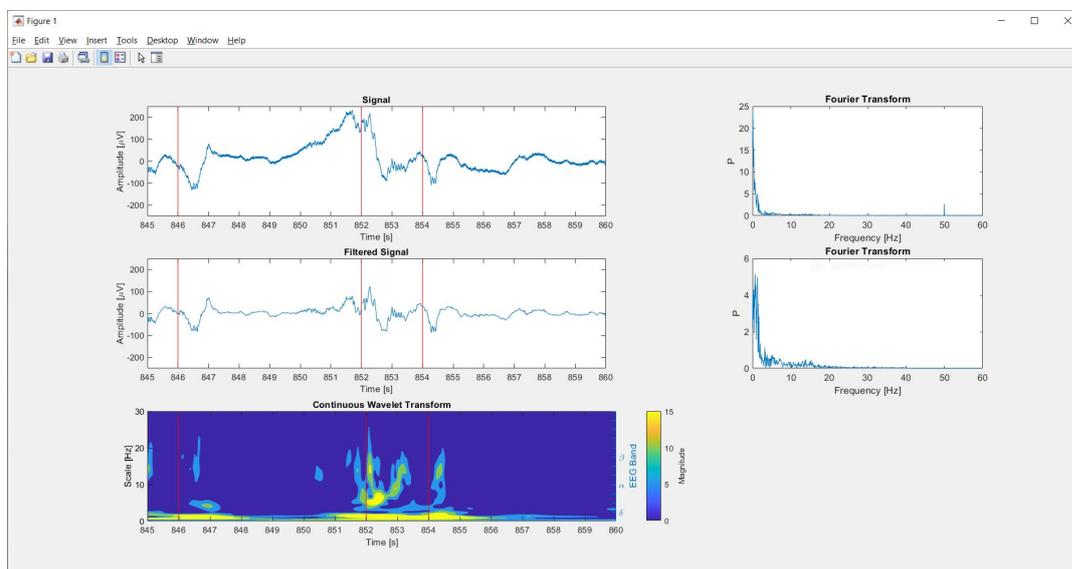


Figure 5.3: Screenshot taken from the output of the MATLAB script

5.3. *Intermezzo: Importance of Building Trust*

As discussed in Section 1.4, one of the main reasons new innovations have a hard time being accepted in the medical field is a lack of trust. Especially with innovations based on artificial intelligence it seems difficult to explain on what precise basis a certain outcome is reached [13]. Siau et al. indicated the factors involved in trust building for artificial intelligence, machine learning, and robotics. These factors could be arranged according to three measures: performance, process, and purpose [14]. An overview of these different factors is provided in Table 5.1.

Table 5.1: Factors influencing trust building in technology

| Initial Trust | | Continuous Trust | |
|----------------|--------------|-----------------------|-----------------|
| Performance | Process | Performance | Purpose |
| Representation | Transparency | Usability/Reliability | Job replacement |
| Perception | Trialability | Communication | Goal Congruence |
| User Reviews | | Socialbility | |
| | | Security/Privacy | |
| | | Interpretability | |

The key elements for building trust in the development phase of the delta brush detector are based on the interpretability of the results, transparency of the process, and the usability. A user needs to understand what is happening and have justification of why a certain outcome is reached. At the same time, the operation of the innovation should run smoothly and work in an intuitive manner. Keeping this in mind will aid in the trust building process. The proposed detector model will therefore be accompanied by an interactive application environment to visualize the spectral information for the user, give feedback on the decision-making process, and present the detector results.

5.4. Interactive Application Environment

In Fig. 5.3, only 15 seconds of the total dataset is visualized per iteration. Since the provided EEG datasets all consist of a minimum of four hours of data, one can understand that it will be difficult to oversee the complete EEG dataset in a single window. To be able to get a grasp on the total dataset, some kind of interaction with the data has to be provided. This was done through MATLAB's App Designer environment. Within App Designer, it is possible to create interactive graphical user interfaces. It provides predefined graphical components, which behavior can be further specified. The programming language used in the app environment is quite similar as the script-based language, the major difference is that App Designer is based on an Object-Oriented Programming (OOP) style. The applications can be exported as web-access applications or complete standalone applications.

All of the processing steps mentioned in this chapter until now have been implemented in one single application. Fig. 5.4 shows a screenshot of the application.

1. Import provided datasets

The provided EEG dataset and delta brush annotation files can be imported with the click of a button. A pop-up window will appear, where the time range of interest can be specified. For now the maximum time-range has been limited to 2000 seconds, to increase performance and prevent memory issues. After importing the EEG dataset, the single-channel EEG signal will be plotted in blue in the 'EEG Signal' window. The delta brush annotations are indicated by the vertical red lines. The active channel can be specified using the drop-down button on the left side of the window. The window has a fixed size of 15 seconds, as is also standard in the medical domain. By using the navigational tools below the window it is possible to navigate through the total signal in three different ways: (1) shift the window by one second, (2) shift the window by one window length, or (3) jump to a specified time.

2. Filter signal

The desired frequency range for filtering can be specified using the button for the High-Pass Frequency (HPF) and Low-Pass Frequency (LPF). After the button 'Filter Signal' is pressed, the required FIR filter is designed and the filtering operation will be performed on the original signal. The filtered signal's frequency content can now be reviewed in the 'Frequency Content (FFT)' window, showing the frequency content calculated with the FFT function. The 'EEG Signal' window will now show both the original signal, in grey, and the filtered signal, in blue.

3. Calculate wavelet transform

The wavelet coefficients for the complete time range of the active channel can be calculated by pressing the 'Calculate CWT' button. The wavelet transform will be calculated for the active signal, which is plotted in blue in the 'EEG Signal' window. In other words, when no filtered signal is calculated, the original signal will be used. The scalogram will be plotted in the 'Time-frequency Spectrum (CWT)' window for the current time window after pressing the 'Plot CWT' button, this is because the required memory to plot the whole time-frequency spectrum at once is too high. The time range is directly linked to the time range shown in the 'EEG Signal' window and will follow the use of the navigational tools. When the 'Plot local CWT' button is pressed a cross section of the scalogram, for the time specified in the 'Jump To' window, will be plotted in the 'Local T-F spectrum' window.

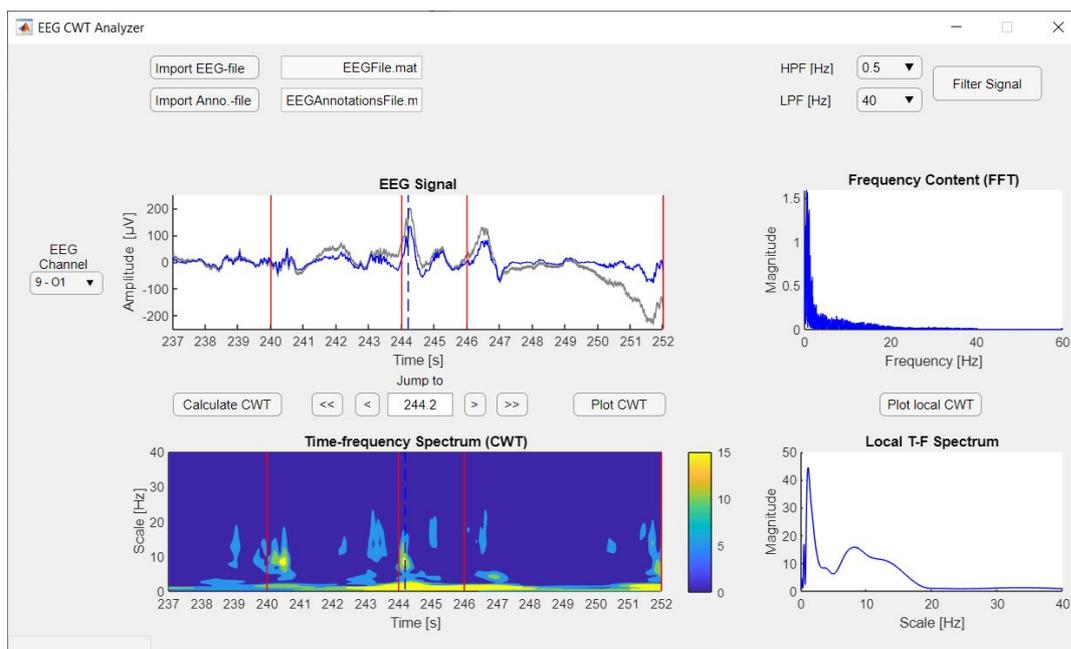


Figure 5.4: Screenshot taken from the interactive application

5.5. Validation of Trust in Interactive Application

For the application to generate trust, it is important that the user understand what is seen and that the application runs smooth and intuitively. To test this a short assignment and questionnaire has been developed. The main function of the questionnaire is getting initial insight in the interpretability and operation of the interactive application environment. Before further roll-out, a extensive large-scale survey is recommended. The target group for this validation was health care professionals with experience in the conventional ways of analyzing neonatal EEG. In the following sections the content and purpose of the assignment and questionnaire are explained. An example of the assignment and questionnaire is provided in Appendix A.

Assignment

A small assignment has been designed to walk the user through all the main functionalities of the interactive application in 5 to 10 minutes. This assignment differs from the eventual use-case of the application to limit the duration of the assignment and to be able to receive feedback on the personal interpretation of the time-frequency information. The instruction 1 till 8 walk the user through the setup procedure, featuring the import of the signal and annotations, the filtering of the signal, the calculation of the wavelet coefficients, and the visualizations. A small section of 400 seconds from the dataset 107G and the corresponding annotations was provided as test material.

Instructions 9 and 10 were used to test if the user had an understanding of the available information. Since the provided annotations were all rounded to the nearest second, a mismatch is present between the annotation time and the exact time instance the delta brush was present. The user was asked to specify a better matching annotation time, based on the visualized data available in the application.

Questionnaire

To be able to compare the different outcomes and feedback on the functionality of the application, a standardized structure was provided in the form of a questionnaire. The goal of the questionnaire is to gain insights on the following factors: does the application work intuitively, does the application work smoothly, and does the user understand the information that is visualized. As a reference for some of the questions, an image of a general EEG viewer visualizing the same dataset is provided. The questionnaire consists of 12 questions, the following points will explain the goal of each of these questions.

- Question 1 till 3 asked for some personal information, regarding their name, function and employment. This information can be used to make sure that the user is indeed familiar with the conventional evaluation of EEG.
- Questions 4 and 5 check if the assignment's instructions were clear and intuitive to follow. The instruction provided in the assignment were deliberately written in a non-specific way. Instead of giving the command to press a specific button, the user was asked to perform a certain action. If the user was able to do so, it would mean that the naming and location of these buttons was done in an intuitive manner. If the instruction would not be clear, then this could also influence all the following questions in a negative way.
- Questions 6 till 8 were focused on how well the user was able to interpret the data visualized in the following three windows: "EEG signal", "Frequency Content (FFT)", "Time-frequency Spectrum (CWT)". Three multiple choice options, featuring increasing insight and usability, were presented to standardize the provided outcome. The first option could be selected if there was no understanding of the visualized data. The second option could be selected when there was understanding and a clear usability of the visualized data. The third options featured and increased understanding and greater usability, when compared to a general EEG viewer.
- Question 9 featured the same multiple choice structure, but was focused on the application as a whole. The main goal was to retrieve feedback on how the application compares to a conventional EEG viewer. The same ordering of understanding and usability has been used. The third option specifically states that an increased insight into the features present in the EEG signal is witnessed.

- Question 10 continues with the focus of the previous question. It asks if the application could serve as a replacement for the conventional EEG viewer and is followed by an open question to what is missed in the application, in case of a negative answer. The main goal is to identify the essential elements of the conventional EEG viewer that are neglected in the proposed application.
- Question 11 returns the focus on the assignment itself. The goal was to have the user reevaluate a possible advantage of the proposed application.
- Question 12 is an open question asking the user for further comments.

Results

The questionnaire was sent out to the Academic Hospital in Antwerp and the Erasmus Medical Center in Rotterdam. The questionnaire was eventually only filled in by a small group of people, only two responses were analyzed. One response from each hospital. This low amount of responses can be explained due to the already small target group of health care professionals with experience in evaluating neonatal EEG and the fact that the questionnaire was but a minor element of the total thesis project, limiting the time spent on retrieving more responses. Since the main goal was to test if the elements for building trust had the right impact, the provided data can still be used to get an initial insight into how the application is received. The results will be discussed based on the three key element to building trust, featured in the questionnaire.

In terms of interpretability, equal responses were recorded. The "EEG signal" window was found to give a comparable insight as a conventional EEG viewer. The "Frequency Content (FFT)" window was understood, but did not give an improved insight as to the conventional EEG viewer. The "Time-frequency (CWT)" window was interpreted as to how strong certain frequencies were present. When looking at the application as a whole, the proposed application gave at least comparable insight as the general EEG viewer, it was also noted once that the application increased the insight into the waveforms present in the EEG signal.

All respondents answered that the application worked in an intuitive manner and the instructions were clear and easy to follow. There were some downsides to the functionality of the application itself. The proposed application featured only limited interaction possibilities, which were named as complicated. The respondents would have liked to more see possibilities, like e.g. mouse interaction, data-point comparison, zoom functionalities, and scaling of the window. The start-up time of the application was also perceived as too long. These are all steps that would need to be regarded to make the proposed application run smoothly.

The respondents were clear that the proposed application could not be used as a replacement for the conventional EEG viewer. They missed an overview covering all EEG channels at the same time, instead of just single-channel information. There was also no information available on the heart rate and respiratory behavior. The respondents also stated elements of added value in the proposed application. It was possible to see trends in the waveforms and pinpoint the exact time of presence more accurately. This way it could work as an additional in combination with the conventional EEG viewer. One of the respondents saw possibilities of using the application to provide annotations for research or use it for educational purposes to help explain waveform characteristics. If the application is to be developed further, it is advised to implement the application in a software environment dedicated for application development. MATLAB's technical environment features only limited options for interaction, which could never match the current expectations in an efficient manner. This would also reduce the long start-up time, as this is related to the start-up of the MATLAB Compiler Runtime (MCR) software package, containing all basic MATLAB libraries, which is required for standalone operation [47].

5.6. Conclusion

This chapter explained all steps necessary to preprocess the provided EEG data, to calculate the wavelet coefficients, and to provide a tool to interact with the data. All these steps are integrated in the proposed interactive application and this can be implemented easily for the different EEG datasets. The interpretability of the time-frequency information by health care professionals was validated with a small-scale study and the application was found to operate in an intuitive manner. The interactive application doesn't have some limitations regarding the functionalities and performance. The interactive application will not be used for further processing of the data, as mentioned in the following chapters. This is due to the already high processing load required for basic application operation. The interactive application will be used for the visual validation of the results gathered in the following chapters.

6

Data Acquisition

Before focusing on the precise detector principle, it is important to know what to detect and what to avoid during detection. This can be done by defining two sample sets, one with the target samples and one with the non-target samples. This chapter will focus on the collection of the samples for these sample sets. When this is done, the collected sample sets will be used in the development of the detector. Data acquisition will be done on the EEG datasets marked for training: 105K, 106S, 106V, and 107G. The target training set will contain samples of the delta brush and can be used to visualize the characteristic behavior of the waveform. The non-target training set will contain the samples that are to be avoided during detection.

6.1. Target Sample Collection

For each of the datasets a minimum of 50 delta brush annotations is available. These delta brush annotations will be the guide for the collection of the target samples. The provided time stamp per annotation were rounded to the nearest integer value, meaning that the indicated delta brush sample has an error when compared to the exact moment the delta brush was present. The delta brush sample should be found in a window of ± 0.5 second around the annotation time. Another error follows from the duration of the delta brush. While the annotation time is based on the initial onset of the delta brush, the spectral behavior of the delta brush shows a peak in activity halfway the delta brush's presence. Since the average duration of a delta brush is roughly equal to 1 second [25], the additional error between the provided annotation and the spectral peak activity is expected to be equal to half this duration, thus equal to 0.5 seconds. This adds up to an expected error of ± 1 second around the annotation time.

In a standard EEG viewer the delta brush is often visible in multiple channels simultaneously. Since the provided annotation times were based on a visual analysis of all 19 channels simultaneously, it is possible that for one annotation time a delta brush can be found in multiple channels. It is not clear if this presence is based on different regions being active simultaneously or if it's the same active region being picked up by multiple electrodes. Because this is unclear, only one delta brush sample will be selected for each annotation time.

To decide which sample to pick from the 19 channels, the known characteristics of the delta brush are consulted. Based on literature, the distinctive feature of the delta brush, when compared to regular delta activity and background activity, lies in the expected activity in the frequency range 8-20 Hz [25]. In the same paper, however, inconsistencies in these description are mentioned. In other literature it is also described that the precise boundaries between the EEG frequency bands are used inconsistently [48]. There is a chance that, due to these inconsistencies, good delta brush samples are excluded from the target sample set.

Pretest Target Sample Collection

A small-scale exploratory pretest was conducted to retrieve insights in the frequency region containing the delta brush's higher frequency activity. For dataset 107G, a small set of 10 delta brush samples was collected manually based on the provide annotations and annotation-related comments. The wavelet coefficients of these samples where combined through averaging and the mean and standard deviation was calculated. This provided insight in the regions where an increase in spectral activity was found for these delta brush samples.

Results Pretest Target Sample Collection

Fig. 6.1 shows the result of the exploratory pretest. The blue line shows the average value of the wavelet transform coefficients, the area around the blue line indicates the spread of the standard deviation.

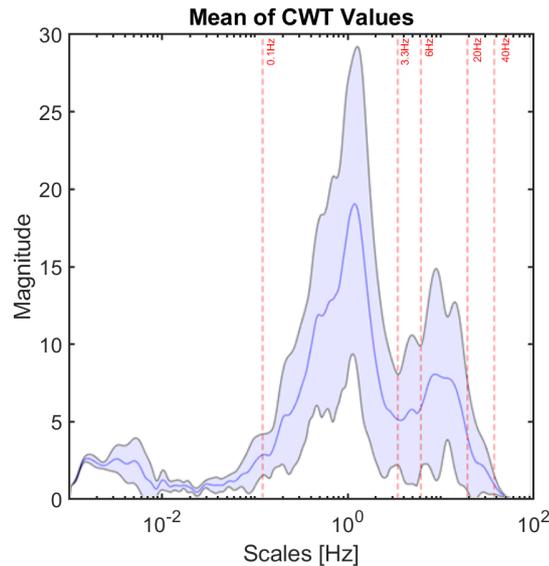


Figure 6.1: Mean and Standard Deviation of Wavelet Coefficients in Pretest Target Sample Set.

Evidence was found for non-delta peak activity upto and even below the 4 Hz delta-band boundary. A nadir in mean and standard deviation values was found for 3.3 Hz. The value of 3.3Hz has since been used as the border to cut the frequency range in a low and high frequency part. All higher frequency activity above this 3.3 Hz border was taken into account as a possible candidate for delta brush activity. Since the signal is filtered from all activity higher than 40 Hz, due to the filter setting discusses in Chapter 5, there is a closed region in which the delta brush activity will be measured. The region running from 3.3 Hz until 40 Hz will be addressed in this thesis as the delta brush's *high-frequency (HF) region*. The validity of this border will be assessed when the complete target sample set is available.

Target Sample Collection

Based on the information and experience obtained from the exploratory pretest, it is now possible to design a sample collection procedure for all training datasets. The design will be based on the following insights. Firstly, The manual collection of the samples was a labor intensive procedure. Secondly, since the dataset provides 200 EEG samples per second, there was a high subjectivity in selecting the precise time instance with maximal activity. The proposed solution is to automate the collection process with a MATLAB script. This takes care of the downsides of manual collection, whilst providing means for reproducibility. The same script can be used for future work on datasets and annotations of the same form.

The sample collector is developed in a MATLAB script. The script can be divided into two separate parts. First of all, the EEG dataset and the annotations are imported. The data preprocessing follows the same steps as explained in Section 5.4. The signal is filtered from all frequencies below 0.5 Hz and above 40 Hz with a FIR filter. The wavelet coefficients are then calculated for this filtered signal.

The second step makes use of the provided annotations to search for the samples with delta brush activity. Since the delta brush distinguishes itself from the other EEG activity with its high-frequency activity, the collection script searches for a maximum peak value in the HF-region to detect possible delta brush activity. This activity is searched for within in a window of 400 samples (2 seconds) around the annotation time, as explained in Section 6.1. The pseudocode for the sample collector is provided in Algorithm 1.

Algorithm 1: SAMPLE COLLECTOR

Input: EEG Dataset, EEG Annotations

```

1 %Preprocess EEG Signal
2 for every EEG Channel do
3   %Calculate Filtered Signal
4    $Signal_{Filtered} \leftarrow \text{filter}(Signal) \{bandpassfir; HPF = 0.5; LPF = 40; Order = 400\}$ 
5   %Calculate CWT
6    $wt \leftarrow \text{cwt}(Signal_{Filtered}) \{MW = amor; V/O = 20; F_{range} = [0.1 \ 40]\}$ 
7 %Collect Delta Brush Samples
8 for every Annotation Time do
9   for every EEG Channel do
10    %Find max value of cwt coefficients in HF-Range (3.3-40 Hz) around Annotation Time (+/-1 s)
11     $wt_{section} \leftarrow wt(3.3 : 40, t_{Anno} - 1 : t_{anno} + 1)$ 
12     $Sample_{Max,Channel} \leftarrow \max(\text{abs}(wt_{section}))$ 
13     $Sample_{Collected} \leftarrow \max(Sample_{Max,Channel})$ 
14     $Struct \leftarrow sample_{Collected}$ 

```

Output: Struct containing collected samples

Results Target Sample Collection

Fig. 6.2 shows the results of the initial target sample collection. The blue line shows the average value of the wavelet transform coefficients, the area around the blue line indicates the spread of the standard deviation. Based on the characteristic of the delta brush, a double peaked activity is expected. One peak should represent the delta activity in the lower frequency range, the other peak should represent the superposition of higher frequency activity.

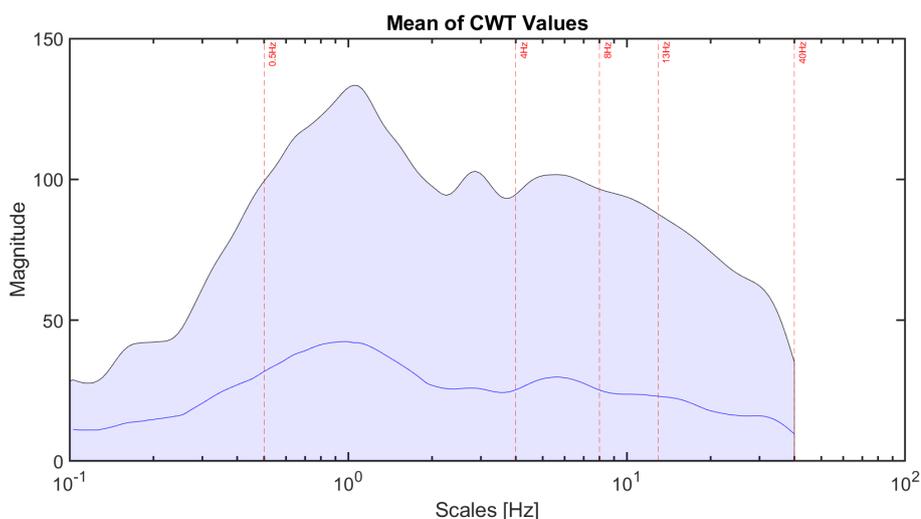


Figure 6.2: Mean and Standard Deviation of Wavelet Coefficients in Initial Target Sample Set.

The mean value does represent some of these expected characteristics, but from the standard deviation it becomes clear that there is a large variation in the target sample set. This could mean that there is a large variation within the delta brush samples themselves, or that non-delta brush activity has been selected by the automated collection script. The exact cause of this variation has to be examined before the sample set is used to train the detector, since this could mean that the sample set is contaminated with non-target activity.

Visual Analysis

To examine the target sample set, the collected samples are visually inspected using the interactive application, as described in Chapter 5. Every collected sample has been visualized and classified based on its behavior. For this manual classification task, four classes were selected based on the activity that is to be expected in the provided EEG dataset, i.e. for preterm infants aged 30 weeks PMA. These classes have been further subdivided into subclasses, based on the variation of the specific encountered cases.

The first class contains all delta brush activity, where the subclasses specify all variations of this activity. Of special interest is the subclass 'Christmas Tree', shown for of broad spectrum delta brush activity. It is named after the specific visual appearance in the time-frequency distribution. The shape is related to the change in time-frequency resolution for different scales. This behavior can be used as a visual identification marker for the detector's end user. Based on the overview provided in Chapter 2, there are no other waveforms expected besides the delta brush activity. Other EEG activity is expected to be pure delta activity or background activity found during the periods of quiescence. All these sample were added to the 'Other Activity' class. The third class contains all possible artifactual activity that is collected as a sample. The fourth class is defined for all cases of doubt, that aren't strictly confined to one group, like combinations of delta brush and artifactual activity or sub-optimal samples. An overview of the classes, subclasses, and a further description is provided in Table 6.1. Examples for all classes are provided in Appendix B.

Table 6.1: Classification of sample classes in collected sample set

| Class | Subclass | Description |
|----------------|-------------------------|---|
| Delta Brush | High Magnitude | High magnitude HF-region activity |
| | Low Magnitude | Low magnitude HF-region activity |
| | Broad Spectrum | Full-range (0.1-40Hz) frequency activity |
| | Narrow Spectrum | HF-range-peak value between 3.3 and 4Hz |
| | Christmas Tree | Broad spectrum activity generating a Christmas-tree-like appearance |
| Other Activity | Background Activity | Broad-spectrum low-magnitude activity |
| | Delta Activity | Delta-range activity |
| Artifact | Discontinuity | Sharp discontinuous signal causing strong HF-region activity |
| | High-frequency Activity | Bursts of high magnitude HF-region activity |
| | Heart-rate Activity | Heart activity causing periodic HF-region activity (>20 Hz) |
| Doubt | Combined Activity | Combination of delta brush and artifact activity |
| | Narrow Spectrum | Second peak's frequency below 3.3 Hz |
| | Duplicate | Duplicate sample caused by overlapping windows |
| | Local Max | Sub-optimal activity due to max outside window |

Following the classification of all collected samples, the occurrence of these classes can be calculated for each of the training datasets. Table 6.2 provides an overview of these occurrences. It becomes indeed clear that a lot of non-delta brush samples were collected, causing the high variation in the target sample set. When only the collected delta brush samples are considered, two of the EEG datasets stand out in a positive way. Dataset 106S and 107G both have 49 delta brush samples collected. When the doubt and artifact sample are removed from the sample set, there should be enough delta brush samples for analysis. This isn't the case for the other two datasets. Dataset 105K suffers from a high ratio of cases of doubt and dataset 106V suffers from

artifacts. These two datasets are evaluated further, with the goal to discover a way to convert the acquired doubt/artifact samples to delta brush samples.

Table 6.2: Occurrence of sample classes in initial training sample set

| Dataset | 105K | 106S | 106V | 107G |
|----------------------|------|------|------|------|
| Provided Annotations | 50 | 73 | 50 | 57 |
| Delta brush | 25 | 49 | 11 | 49 |
| Doubt | 20 | 8 | 5 | 4 |
| Artifact | 5 | 16 | 34 | 4 |

Rerun

For the further analysis of the datasets 105K and 106V, the additional information provided with the EEG datasets is consulted. For dataset 106V comments have been provided about the heart rate artifact being present in three of the channels, the affected channels are not specified however. When the 34 collected samples of the artifact class are regarded, all samples were collected from channel '10-O2' or '16-T6/P8'. The third channel affected with heart rate activity was found to be channel '8-P4'. Also channel '9-O1' seems to be affected with a smaller magnitude of the same periodical activity. Based on this knowledge, the sample collection script was rerun for dataset 106V, with the four affected excluded from analysis.

For dataset 105K, the additional comments state no channel-specific artifacts. It is only stated that the EEG signal occasionally suffer from an artifact caused by the feeding probe. When the 20 cases of doubt class are regarded, it is revealed that all are caused by the combination of a high-frequency activity. This activity is probably the result of the proposed feeding tube interference. The collected samples regarded as cases of doubt are spread over five different channels. Since there is no clear link between the feeding tube artifact and some specific channels, a rerun of the collection script is not recommended. Instead, new samples are collected manually with the aid of the interactive application. For every annotation time where a sample was collected with the feeding tube activity, another channel was selected for sample collection. With the provided annotations, additional comments where provided on the channels with strong delta brush activity. These comments where used to aid the selection of the new delta brush samples. For one of the annotations it was not possible to find activity in the HF-region in the commented channel

Results Rerun Target Sample Collection

The resulting class occurrence for the final sample set are presented in Table 6.3. As a result of the rerun procedure, it was possible to convert a large deal of artifact and cases of doubt to delta brush samples. All newly collected samples in the rerun have been visually validated. The remaining artifact and cases of doubt for all datasets showed a large variation and could not be linked to a single artifactual cause or channel.

Table 6.3: Occurrence of sample classes in rerun training sample set

| Dataset | 105K | 106S | 106V | 107G |
|----------------------|------|------|------|------|
| Provided Annotations | 50 | 73 | 50 | 57 |
| Delta brush | 44 | 49 | 41 | 49 |
| Doubt | 1 | 8 | 7 | 4 |
| Artifact | 5 | 16 | 2 | 4 |

Every training dataset has provided at least 41 delta brush samples. If more delta brush samples should be needed for these datasets further on, it is advised to ask for more annotations to be used for sample collection. All delta brush samples are used in the final target sample collection, while all remaining samples containing artifacts or cases of doubt are excluded. Fig. 6.3 shows the results of the final target sample collection. Again, the blue line shows the average value of the wavelet transform coefficients, the area around the blue line indicates the spread of the standard deviation.

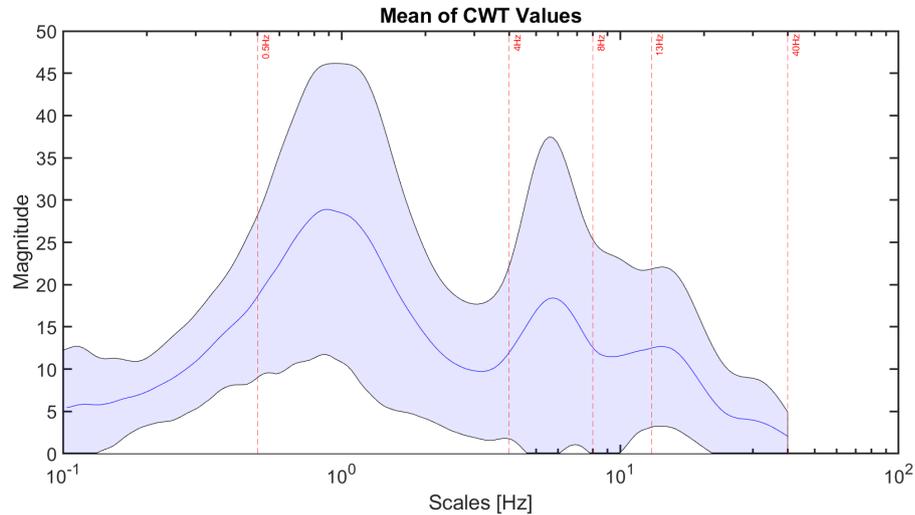


Figure 6.3: Mean and Standard Deviation of Wavelet Coefficients After Rerun of Target Sample Set.

As expected, a clear twin-peak characteristic is seen, indicating the superposition of low-frequency delta activity and activity in the HF-region and the variation between the samples has decreased to reasonable levels. In line with the hypothesis from the pretest, a nadir in activity between the peaks is seen below the conventional 4-Hz delta boundary. Figure 6.4, represents the distribution of the HF-region-peak frequency. This shows a distribution of the high-frequency-region's activity running from 3.42 Hz until 24.62 Hz, with a median value of 6.60 Hz. The interquartile range, representing 50% of the samples, runs from 5.27 Hz until 13.66 Hz. This is in line with the expected inconsistencies around the 8-20 Hz frequency range mentioned in literature [25].

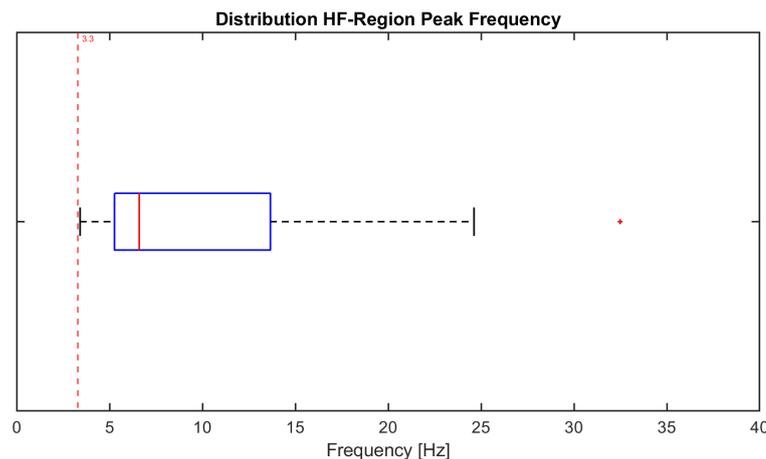


Figure 6.4: Boxplot showing Distribution HF-Region Peak Frequency of Target Sample Set. The red dotted line depicts the minimum value set by the 3.3 Hz boundary.

6.2. Non-Target Sample Collection

Now that the target sample set collection has been finished it is time to focus on all activity that shouldn't be detected. The collection of non-target samples provided difficulties with regard to an automated collection procedure. As discussed in Section 2.2, the neonatal EEG is characterized by its dominant delta activity and its periods of inactivity, known as the background activity. What is also known, from correspondence described in Section 4.3, is that not all delta brush activity had been annotated. If an automated collector script would be designed to take random samples of non-annotated EEG sections, a couple of hazards could occur. Firstly, the detector would contain an increased amount of regular delta and background activity. Since the characteristics of this activity already differ a lot from the delta brush activity, this activity doesn't have to be present abundantly in the non-target sample set. Secondly, it would probably miss artifactual activity which is only active for a small percentage of the total EEG range. Since this activity shows resemblance with the delta brush characteristics it should be included in the non-target sample set. Lastly, there is a chance that non-annotated delta brush activity could be included in the non-target sample set.

To take these hazards into account, a manual selection routine was used. The goal was to create a balanced dataset, which has the same order of samples as the target sample set and consists of a wide range of non-delta brush activity. The non-target samples were collected during the visual analysis of the initial target sample set. For every delta brush sample or case of doubt, a non-target sample was taken from the same channel. The non-target samples contain samples from the 'Other Activity' class and 'Artifact' class, as explained in Table 6.1.

Non-Target Sample Collection Results

Fig. 6.5 shows the result of the non-target sample collection. The blue line shows the average value of the wavelet transform coefficients, the area around the blue line indicates the spread of the standard deviation. Since the non-target sample set contain a wide range samples from different classes, the large variation, seen in the figure, is expected.

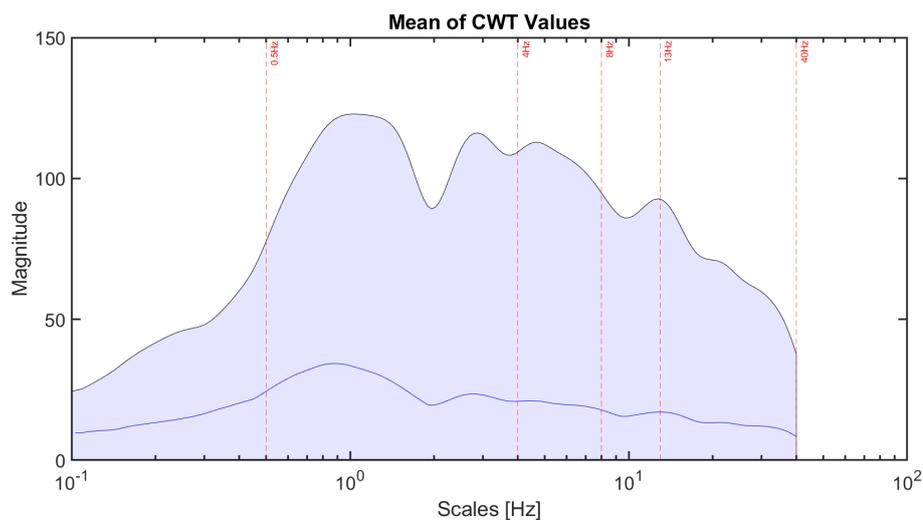


Figure 6.5: Mean and Standard Deviation of Wavelet Coefficients in Non-Target Sample Set.

6.3. Conclusion

The samples containing target and non-target activity have been collected for all datasets marked for training. The final composition for the target sample set and the non-target sample set is provided in Table 6.4. The slight increase in the total amount of non-target samples follows from the collection routine based on delta brush samples and cases of doubt in the initial target sample set. The target and non-target sample sets are now ready to be used as an input for the detector.

Table 6.4: Composition of target and non-target sample set.

| | Target Set | Non-Target Set |
|----------------------|------------|----------------|
| Total Samples | 183 | 202 |
| 105K | 44 | 40 |
| 106S | 49 | 63 |
| 106V | 41 | 46 |
| 107G | 49 | 53 |

7

Waveform Detector

With everything in place to process and visualize the provided EEG data, it is time to focus on the detection of the delta brush waveform. This chapter will focus on the applied detector principle, the choice for the detection features, and the validation of the detector model. In combination with the visualization provided by the application, the eventual detector should provide an understandable approach to why a certain outcome was reached.

7.1. Detection Principle

A wide range of possible detection methods is available. At the moment an increasing interest goes out to the use of machine learning, like e.g. neural networks and deep learning [13]. Although these methods provide promising results, the medical world doesn't welcome them with open arms [12]. One key element is still missing before general acceptance in the medical field can be expected, namely trust [14]. As described in Section 1.4 and 5.3, trust has a major impact on the acceptance of new innovations. There are multiple ways of building trust for new innovations [14]. Often missing are ways of interpreting the data and an explanation why a certain outcome is reached [13]. In this study, the goal will be to provide insight into the detector principle in a way the doctor is able to understand. This way it is possible to provide a new tool to aid the doctor in the decision-making process.

The detector will be based on the principles used in statistical pattern recognition. Since the provided delta brush annotations contain knowledge on the EEG samples with delta brush activity, the design of the detector will be in line with *supervised* pattern classification approaches [49]. Since the classification is based on detecting delta brush EEG samples from all other EEG samples, the delta brush is regarded as a single *class* containing the target samples.

There are different approaches to the design of the detector. In a one-class classification problem there is only information on one class available. The goal is to define a detection boundary around this class, to be able to reject all non-class activity [50]. In a two-class detector approach, a second class is used to optimize the classification boundary. By using multiple classes, overlap between the different classes and the resulting detection boundary's trade-off can be analysed. When more than two classes are considered, the approach is known as a multi-class classification problem. In practice, it doesn't really differ from the two-class approach, since a multi-class problem can be considered as a combination of two-class problems [50].

From the visual analysis of the target and non-target samples collected in Chapter 6, a different set of classes besides the delta brush class was identified. This information can be of added value in finding a fitting classification boundary. If all this different non-delta brush activity is regarded as another single class, the non-target class, the classification problem can be approached as a two-class classification problem.

7.2. Detection Features

The available data that can be used as an input for the detector consists of the collected target and non-target sample sets. At this moment these datasets consist of all the calculated wavelet coefficients for the selected samples. The goal is to select key elements in these datasets that will generate the most distinctiveness between the target sample set and the non-target sample set, these elements are known as the detection *features*. Since a feature can be found in all possible combinations and calculations of the available dataset, there is a large freedom in the selection of the specific detection feature to use.

To aid the interpretability of the selected feature for the doctor, the feature selection will be based on the inherent characteristics of the target signal, i.e. the delta brush. Based on the findings in Section 2.2, it became clear that the delta brush could be easily recognized because of the combination of delta frequency activity in combination with higher frequency bursts. Since the other activity consists mainly of plain delta activity, the difference is in the higher frequency activity only seen with the delta brush. A direct measurement of peak or average values in this high-frequency region could be used as a feature. The disadvantage of this lies in the fact that it would be directly dependent on the amplitude of the signal itself. This is problematic because the recorded signal amplitude is dependent on both the used equipment and the quality of the measuring procedure.

A more elegant approach would be to make the feature independent of the signal amplitude. This could be done by comparing the high-frequency part of the signal with a reference, like e.g. the total average or part of the low-frequency region. A possibility, in line with the inherent characteristics, would be the ratio between the low-frequency activity and the high-frequency activity. There are different measures that can be used to compare the low-frequency region to the high-frequency region, like e.g. the sum, the average or the max of the wavelet coefficients. Since there is logarithmic spread in these coefficients, the sum and average measures would put an emphasis on the lower frequency activity, because of the increased amount of coefficients in this region. An advantage for the comparison of single values instead of regions follows from the *L1* normalization. As explained in Section 5.2, there exists a direct relation between magnitude of the scalogram and the signal's amplitude for *L1* normalization. Because of this, the comparison of the ratio between the peak values of the two frequency regions is proposed as a detector feature, where the ratio is defined as the maximum value in the high-frequency region as a part of the maximum in the low-frequency region.

Initial Feature Selection Results

For all samples in the target and non-target sample sets the proposed feature has been calculated. The distribution of the samples for each of sample sets is provided in a histogram and a boxplot in Fig. 7.1.

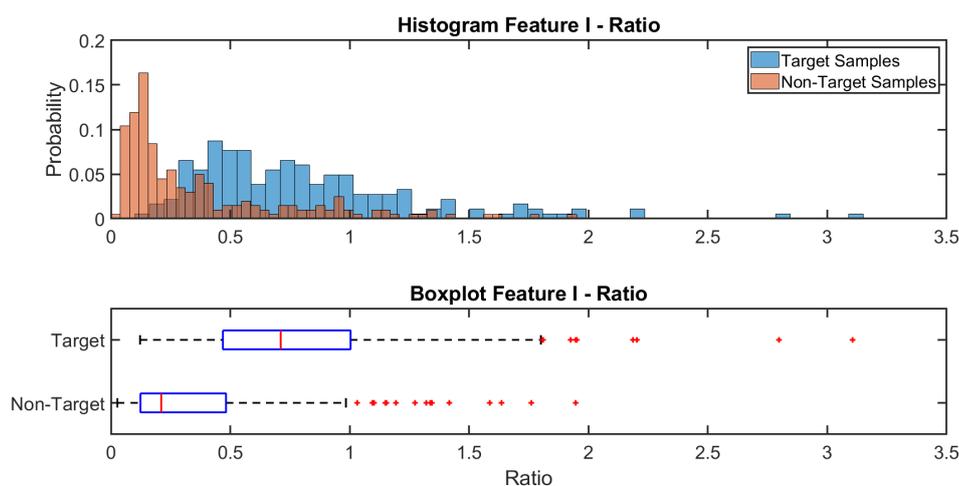


Figure 7.1: Distribution of the Ratio as a Detection Feature

The distribution of the ratio as a feature shows a clear distinction between the two sample sets, there is still an overlapping region which would decrease the detector's effectiveness. From the visual analysis of the collected sample set, described in Chapter 6, it is known that activity besides the pure delta and delta brush activity is also present. If this activity would also agree with the proposed ratio detection feature, it would decrease the performance of the detector. The background activity during quiescence is known to be of very low amplitude, which may even appear flat. A flat signal will probably also be accepted by the ratio detection feature, as the max in the low-frequency region and the high-frequency region will be close to equal. Another problem follows from broad-spectrum artifacts produced by discontinuities in the signal. There is a chance that these artifacts will create samples that will be accepted by the ratio detection feature. It could be wise to try and protect against these measures, to increase the detector performance.

Secondary Feature Selection

A second detection feature could be used as a protective measure against these scenarios. Both the background and artifactual activity can be characterized by the (in-)activity of the signal, a feature using this characteristic could provide a high distinctiveness. The measure of activity can be quantified by calculating the energy. One way to do this with the available data is to calculate the sum over all wavelet coefficients. Again, a direct disadvantage follows from the logarithmic spread in the wavelet coefficients, this would put an emphasis on the lower frequency activity. Another way to calculate the energy is to use a separate calculation method on the original EEG signal to calculate the energy present.

Two commonly used approaches to represent the instantaneous energy in a signal are the Squared Energy Operator (SEO) and the Teager Energy Operator (TEO) [51]. The SEO is equal to the squared amplitude of the signal, $SEOx(t) = x^2(t)$ [51]. It provides a simple measure to approximate the signal's energy. The TEO is a frequency-weighted energy measure, that takes into account the amplitude and the frequency of the signal [30]. The TEO is calculated by taking the squared-value of the signal's first derivative and subtracting the original signal times the second derivative, $TEOx(t) = \dot{x}^2(t) - x(t)\ddot{x}(t)$ [52]. There are some disadvantages in using the TEO. First of all, when the TEO is applied to a broad-band signal it will be sensitive to interference due to cross terms [30]. Secondly, since the TEO is an energy-weighted function, it will assign a higher energy value to higher frequency activity when compared to lower frequency activity [30].

To introduce stability in the outputs of the energy operator, it is common to calculate the mean over an analysis window [51]. The width of the analysis window has been set to one second, which is equal to the interval between the provided annotations and the detector's eventual analysis window. Since it is not straightforward which energy measure will generate the greatest distinctiveness, both the SEO, TEO, and their windowed versions will be calculated and their feature distributions will be compared.

Secondary Feature Selection Results

For all samples in the target and non-target sample sets the proposed energy features have been calculated. The distribution of all the energy measurement methods is provided in a histogram and a boxplot in Appendix C. Of all the energy measurements, the windowed squared energy operator displayed the least overlap and thus the highest distinctiveness. This windowed SEO is further used as the secondary detector feature, the distribution is provided in Fig. 7.2.

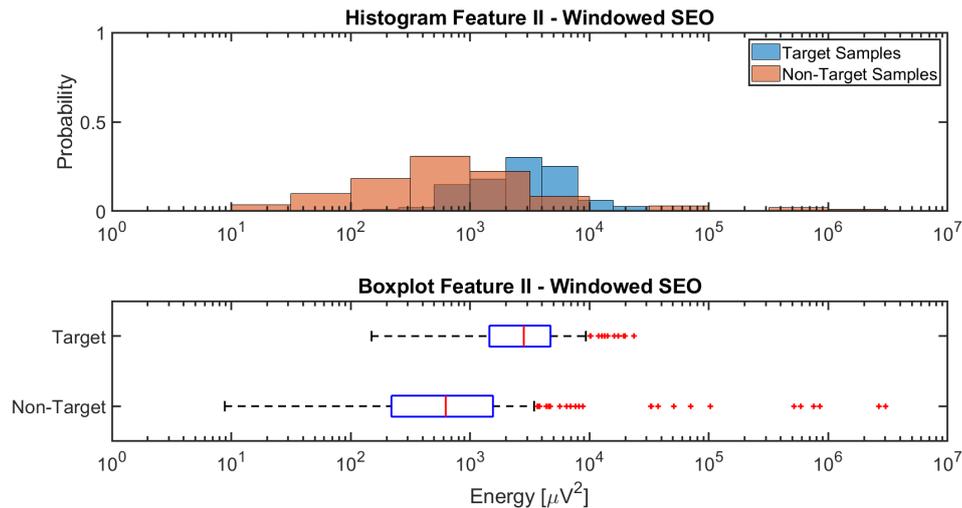


Figure 7.2: Distribution of the Windowed Squared Energy Operator as a Detector Feature

Although the windowed SEO performed the best in comparison with the other proposed energy features, there is still significant overlap between the sample sets. This will result in an eventual trade-off in the selection of the detection threshold depending on the goal of the end user. Another important thing to keep in mind is the composition of the non-target class. Since this class exists of multiple sub-classes, it will show a wide variation. The boxplot, as provided in 7.2, tries to fit a single mean and variation to this class, resulting in all high-energy sub-classes to be handled as outliers. Since the frequency axis is plotted on a logarithmic scale, there is actually a significant gap between the minimum value of the high-valued energy-class outliers and the maximum value of the target-class distribution. This behavior suggests the use a second energy detection boundary between the maximum value of the target-samples and the high-frequency non-target samples.

Another thing to consider is the fact that this feature is directly dependent on the value of the amplitude. As explained in Section 7.2, the amplitude level shows variation between different EEG datasets. To increase the independence of the feature, its value should be adjusted according to a reference value. In the provided EEG datasets a value for the maximum amplitude measured in the total recording is provided. This could be used as a reference value to adjust the specific energy boundaries before being used for detection. This way the energy-boundaries will be made independent of the differences in amplitude levels.

Visualization of Feature Space

Since neither of the two detection features provided enough distinctiveness on its own, it is worth it to look at combinations of the features to increase the detector's performance. Fig. 7.3 shows the total range of the combined feature space. In the figure, the target samples and non-target samples form clear clusters. A wide range of thresholds can be used to try and divide the clusters. The goal is to find thresholds for the two detection features that divide the clusters as good as possible, but since there is overlap between the clusters there will always be a trade-off in the performance.

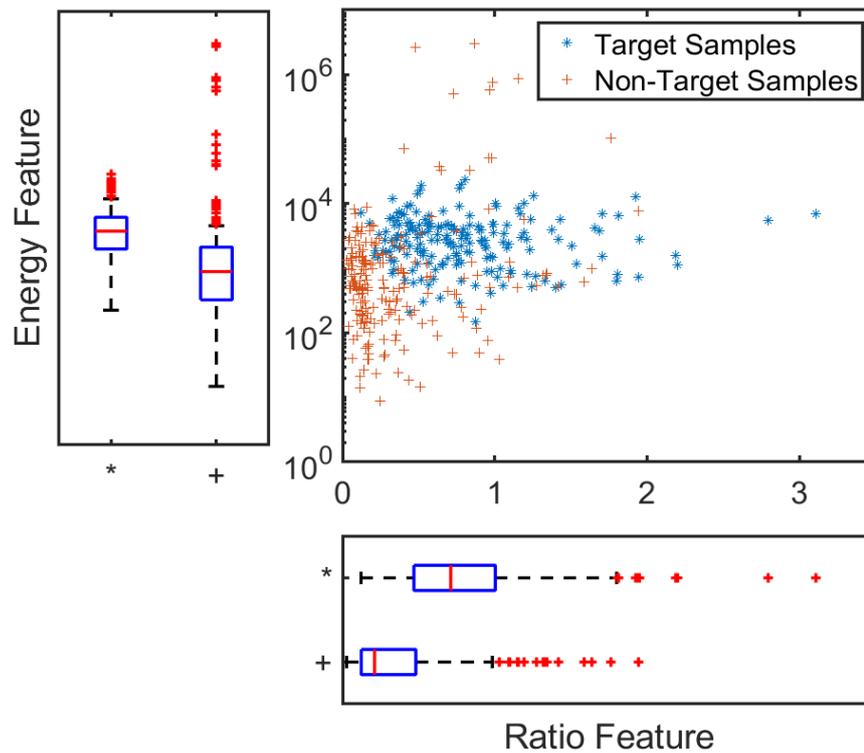


Figure 7.3: Combined Distribution of the Two Detector Features

7.3. *Intermezzo: Performance Measures*

To find the best fit for the feature thresholds, it is first important to produce a measure of performance. This way it is possible to quantify why a certain threshold should be used or neglected. This section provides an overview of the different performance measures that were used in this work.

Confusion Matrix

A common performance measure used in classification is the *confusion matrix* [53]. The confusion matrix consists of four different measures needed to evaluate the performance: True Positive (TP), False Positive (FP), True Negative (TN), False Negative (FN). Each of these measures indicates if the detector's behavior is matching the actual case. Table 7.1 provides the confusion matrix for the current situation.

Table 7.1: Confusion Matrix

| | | Waveform Present | |
|----------|----------|------------------|----|
| | | Yes | No |
| Detector | Positive | TP | FP |
| | Negative | FN | TN |

To evaluate the final performance of a single threshold, the measures of the confusion matrix are combined to produce single values of performance [49]. These performance measures can tell how good one combination of thresholds is and can be used to compare different threshold combinations. A lot of different measures are possible, the performance measures addressed in this study are:

- Sensitivity

How many of the waveforms that are present have been detected. Also known as *recall*.

$$Sensitivity = \frac{TP}{TP + FN} \quad (7.1)$$

- Specificity

How many of the samples without a waveform have been successfully ignored.

$$Specificity = \frac{FP}{TN + FP} \quad (7.2)$$

- Precision

How many of the detections are actually positive. Also known as *Positive Predictive Value (PPV)*.

$$Precision = \frac{TP}{TP + FP} \quad (7.3)$$

ROC Curve

When multiple threshold levels of the same feature are to be compared, a common approach is to calculate the performance measures over the whole range of threshold combinations and plot the outcome in a curve. A Receiver Operating Characteristics (ROC) curve displays the full-range of outcomes for different threshold combinations by plotting the sensitivity versus 1-specificity [54]. An example of a ROC curve for overlapping sample sets is provided in Fig. 7.4. A single value to compare different ROC curves is the Area Under the Curve (AUC) [54]. The higher the AUC of a ROC curve, the better its performance. The ROC's AUC can be compared to a fixed baseline running from (0,0) to (1,1) which indicates an AUC of 0.5, equal to the outcome of a totally random classifier [53].

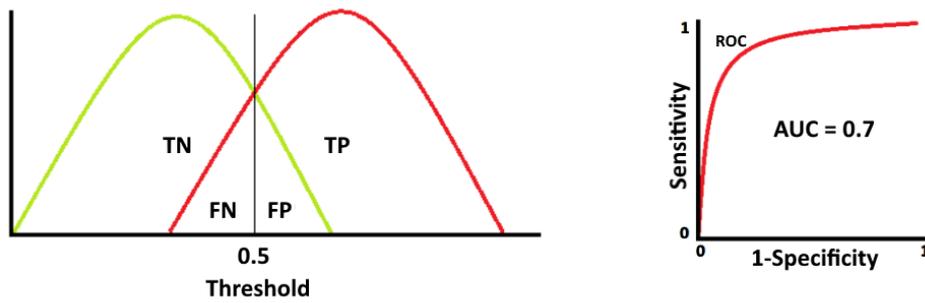


Figure 7.4: Example of a ROC Curve. Adopted from [55]

PRC Curve

When the dataset used in the classification contains far more samples without the waveform than samples with the waveform, the dataset is known to be imbalanced. When evaluated the performance of unbalanced datasets using the ROC curve, the outcome could be misleading [53]. A performance comparison based on the samples containing a waveform is suggested in such cases, resulting in the Precision-Recall (PRC) curves [53]. Since both the ROC and PRC curve are based on similar performance measures, both the ROC and PRC curve can be evaluated in similar manners. The major difference follows from the variable baseline featured in the PRC curve, which varies with the balance of the dataset. The baseline for the PRC curve is equal to the ratio of waveform samples present versus the total number of samples available [53]

$$y_{baseline} = \frac{P}{P + N}. \quad (7.4)$$

When evaluating the area under the PRC curve, an approximation of this value is calculated, i.e. the average precision (AP) [56]. This average precision is calculated by multiplying the precision vs the change in recall over all threshold combinations

$$AP = \sum_{n=1}^N P(n) \Delta R(n). \quad (7.5)$$

7.4. Threshold Selection

Based on the feature space, as presented in Fig. 7.3 it became clear that a wide range of possible threshold is available, where each combination will provide a different performance. One way to find the threshold is by using automated procedure like the Linear Discriminant Analysis (LDA). The LDA tries to find an one-dimensional combination of the two detection features [49]. The LDA optimizes the threshold based on maximizing the distance between the means and minimizing the variance within the class [49]. An example of LDA threshold selection is provided in Fig. 7.5.

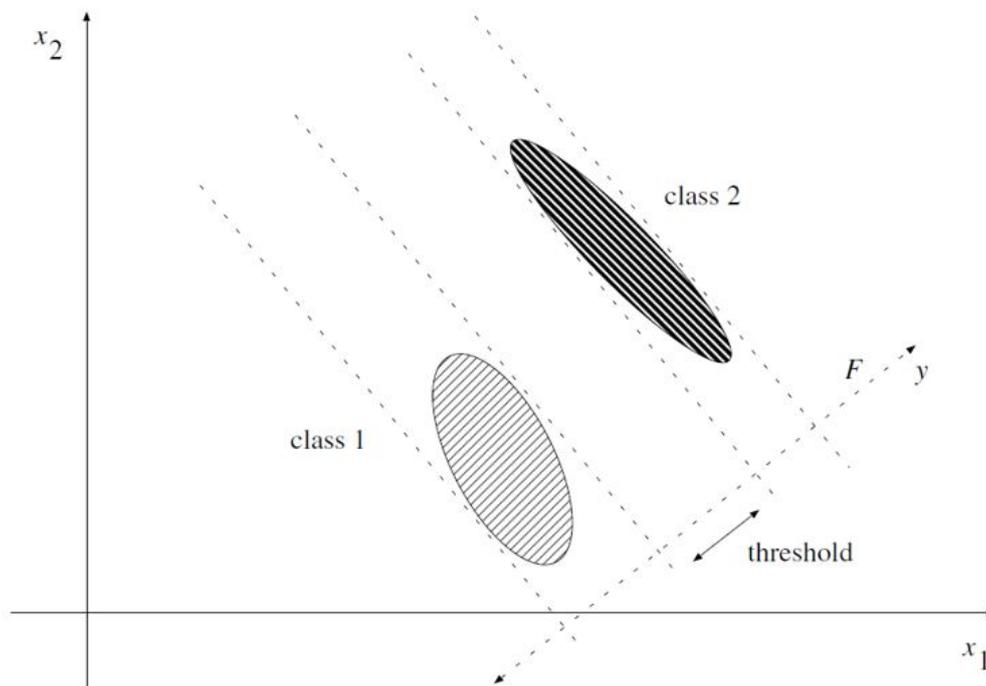


Figure 7.5: Example of threshold selection using LDA. Adopted from [49].

The downside of using the LDA for threshold selection with current feature set, is that the LDA assumes the features to be normally distributed [49]. It assumes that a clear mean and variance are present for the features of the target and non-target samples sets. In Section 7.2, it was already discussed that the non-target sample set consists of multiple dissimilar classes, which don't agree to a normal distribution. Secondly, the distribution of the energy feature is plotted on a logarithmic scale, this gives the impression of a log-normal distribution of the energy. The LDA was applied to the feature set to find a threshold. As seen in Fig. 7.6a, the LDA threshold depends heavily on the high-frequency activity in the non-target sample set. For the sake of the simulation, it is possible to see all non-target samples with an energy higher than the highest target sample as artifacts and exclude them from the simulation. This has been done for the results presented in Fig. 7.6b, where a better matching LDA threshold has been found. Although it seems that the optimization still suffers from the non-normal distribution of the non-target samples. Based on the theory and these results, the LDA is rejected for threshold selection.

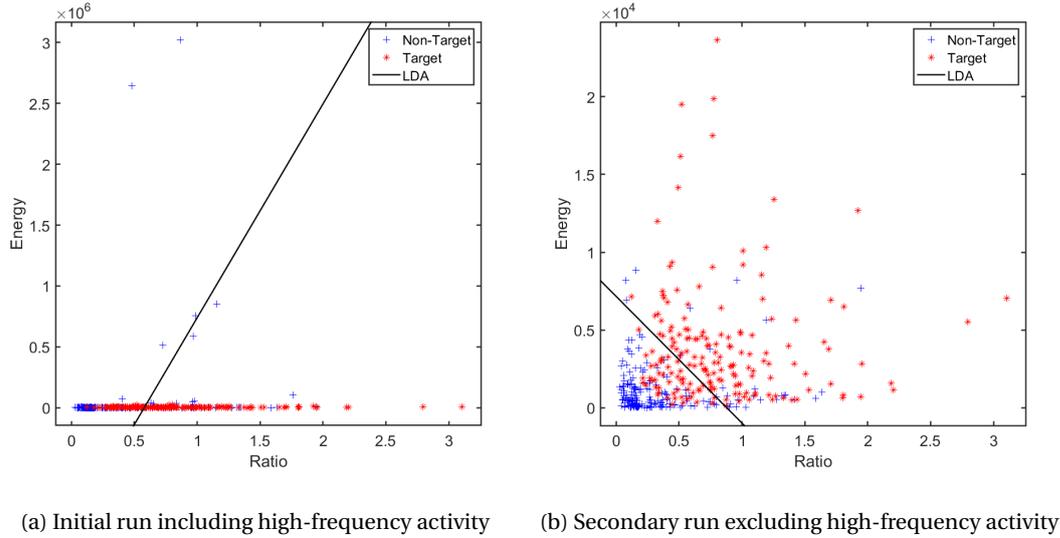


Figure 7.6: LDA threshold selection for the target and non-target sample sets.

Until now, the threshold selection was seen as an optimization problem. However, it is difficult to speak of an "optimal" threshold in the medical field, since the performance will always be dependent on the goal of the classification. For a doctor, for example, it can be more important to identify all sick patients (true positives), with the drawback of classifying some healthy people (false positives) as sick as well. For the hospital itself it might be more important to be as cost efficient as possible, taking into account the costs of follow-up treatment. All these financial, ethical, and comfort criteria are difficult to assess in advance and will be ignored for now. But instead of providing just a single threshold to be used for detection, the detector will be run over the full range of threshold combination. This will provide the total solution space for the two detection features in the form of a PRC curve. This can then be used to select the best-fitting threshold based on the requirements of the specific task at hand.

The full-range for both threshold was specified as the region starting at 0 until the end of overlap between the target and non-target sample sets. This range was then divided in 10 evenly spread intervals. Because 0 is included in the range, it is possible to analyze the detector's behavior when no threshold is applied. Fig. 7.1 shows the distribution for the ratio feature with an overlapping region running until a value of 2. The ratio threshold levels run from 0 to 2 with a step size of 0.2 ([0:0.2:2]). This resulted in 11 ratio threshold to be used. Fig. 7.2 shows the distribution for the energy feature with an overlapping region running until a value of $10^4 \mu V^2$. Since this distribution was plotted on a logarithmic scale, the intervals will be evenly spaced in the logarithmic domain. This was done through the evenly spread of the exponents used for exponentiation. The energy threshold levels run from 10^1 to $10^4 \mu V^2$, where every decade is subdivided in 3 intervals, and includes 0 ([0 10¹(1:(1/3):4)]). This resulted in 11 energy threshold to be used and a total of 121 possible threshold combinations. Since this energy threshold is focused on the low energy samples, a second energy threshold is used for all high-energy samples. Since there was no overlap but a gap in between the high-energy target samples and the high-energy non-target samples, a single static threshold was used. This threshold was chosen to be the average value of the maximum energy of the target samples and the minimum value of the high-energy non-target samples. This makes the third threshold equal to $28,060.5 \mu V^2$.

To be able to provide a single performance value for comparison, a single threshold will be chosen from the total solution space and its performance will be tested on the validation datasets. One of the threshold-selection techniques available for ROC and PRC curves is finding the point closest to the point of perfect performance [54]. For the ROC curve this is the top-left corner (0,100) for the PRC curve this is the top-right corner (100,100). This can be done by calculating the euclidean distance to (100,100) for every PRC point

$$D = \sqrt{(100 - \text{Sensitivity})^2 + (100 - \text{Precision})^2}. \quad (7.6)$$

7.5. Detector Principle

The waveform detector is developed in a MATLAB script. Many of its functions are comparable to the sample collector script, explained in Section 6.1. The script can be divided into two separate parts.

First of all, the EEG dataset is imported and preprocessed. The preprocessing steps for the detector are: filtering the signal, calculating the wavelet coefficients, calculating the ratio feature for every sample, and calculating the energy feature for every sample. To save future processing time, it is possible to save the calculated feature set, which can be used as an input in future runs on the same dataset.

The second part had the goal to detect delta brush samples based on the feature set. To reduce computational load the waveform detection is applied on a window of 200 samples (1 second) instead of sample-by-sample. It first searches for the sample with the maximum wavelet coefficient in the HF-range for the current window. If this sample agrees to the three threshold levels, it is stored in the output struct. The pseudocode for the waveform detector is provided in Algorithm 2.

Algorithm 2: WAVEFORM DETECTOR

Input: EEG Dataset, EEG Feature Set

```

1 %Preprocess EEG Signal
2 for every EEG Channel do
3   %Calculate Filtered Signal
4    $Signal_{Filtered} \leftarrow \text{filter}(Signal) \{bandpassfir; HPF = 0.5; LPF = 40; Order = 400\}$ 
5   %Calculate CWT
6    $wt \leftarrow \text{cwt}(Signal_{Filtered}) \{MW = amor; V/O = 20; F_{range} = [0.1 \ 40]\}$ 
7   %Calculate Feature I: Ratio
8    $Sample_{Max, HF} \leftarrow \max(\text{abs}(wt(3.3 : 40, Sample)))$ 
9    $Sample_{Max, LF} \leftarrow \max(\text{abs}(wt(0.1 : 3.3, Sample)))$ 
10   $Sample_{Ratio} \leftarrow \frac{Sample_{Max, HF}}{Sample_{Max, LF}}$ 
11  %Calculate Feature II: Windowed SEO
12   $Sample_{SEO} \leftarrow \text{signal}(Sample)^2$ 
13   $Sample_{SEOAvg} \leftarrow \text{mean}(Sample_{SEO}(Sample - 0.5 : Sample + 0.5))$ 
14 Optional: Save EEG Feature Set to be used as input
15 %Detect Delta Brush Samples
16 Set Threshold Values:  $Th_{Ratio}, Th_{Energy, Low}, Th_{Energy, High}$ 
17 for every window-step over Signal do
18   for every EEG Channel do
19     %Find max value of cwt coefficients in HF-Range (3.3-40 Hz)
20      $Sample_{Max, Channel} \leftarrow \max(\text{abs}(wt(3.3 : 40)))$ 
21      $Sample_{Max} \leftarrow \max(Sample_{Max, Channel})$ 
22     %Logic test Feature Data of  $Sample_{Max}$  against Thresholds
23      $Test1 \leftarrow (Sample_{Max, Ratio} > Th_{Ratio})$ 
24      $Test2 \leftarrow (Sample_{Max, SEOAvg} < Th_{Th_{Energy, High}})$ 
25      $Test2 \leftarrow (Sample_{Max, SEOAvg} > Th_{Th_{Energy, Low}})$ 
26     if  $Test1 \ \& \ Test2 \ \& \ Test3$  then
27        $Struct \leftarrow Sample_{Max}$ 

```

Output: Struct containing Detected Samples

For the calculation of the total solution space the script will be run for all training datasets and all validation datasets. All results for the training datasets will be combined in one single file and all results for the validation datasets will be combined in one single file. This way the solution space between the training and validation sets can be compared. For the single threshold performance, the solution space will be calculated for the training datasets. This will result in one point being closest to the point (100,100). The threshold levels accompanying this point will be used as input for the detector run on all four validation datasets.

Validation of Collected Samples

The waveform detector provides detected samples of the datasets without knowledge of the annotations that are present. Before the detected samples can be presented as results they need to be validated against these annotations. This comparison should provide the number of true positives, false positives, true negatives, and false negatives. These values can then be used as input for the performance calculations. First the validation procedure will be explained.

A true positive was counted for every detected sample accompanied by an annotation. As discussed in Section 6.1, there was a mismatch between the annotations and the actual waveform present. Because of this, a window of 400 samples (2 seconds) was used around the annotation times in the sample collection script. The same mismatch is found in the validation of the true positives. When an equal window of 400 samples was used for validation, visual analysis of the calculation found detected samples just outside of this window. This can be explained by literature, since the window used in Section 6.1 was based on the average duration of the delta brush. This would mean that delta brushes with an above average duration would fall outside the window. Because of this, the validation window was chosen to be 600 samples (3 seconds). Because of overlapping windows, it was possible to detect more than one delta brush per annotation. Since this could create a situation where more true positives could be counted than the number of annotations, all additional detected samples around an annotation were seen as one group and further neglected. A false negative was counted for every time an annotation was not accompanied by a detected sample. All detected samples that didn't find a match with an annotation were counted as false positives. The measure for true negatives was not used in this study, since the doctor did not provide annotations for signal segments without delta brush activity.

Because the true negative measure is not included, it is not possible to calculate the specificity. This means that it won't be possible to calculate a ROC curve. But maybe the ROC curve wasn't the best measure to visualize this dataset to begin with. As explained in Section 7.3, the ROC curve can be misleading when evaluating imbalanced datasets and the PRC curve would be a better visualization method. The imbalance in the dataset is represented by the baseline which can be calculated for the PRC curve. The training dataset has 230 annotations present in a signal segment with a total duration of 1704 seconds. The baseline for the training dataset will thus be equal to 13.50. The validation dataset has 224 annotations present in a signal segment with a total duration of 2046 seconds. The baseline for the training dataset will thus be equal to 10.95. Since the datasets are indeed imbalanced, the PRC curve will be further used to represent to solution space.

Solution Space for Detected Samples

The detector and validation measures have been run for all 121 combinations of threshold values. Every combination has provided one set of performance measures. Each of these points can be drawn as a single point in the solution space. Since a PRC curve is drawn for all variation of a single threshold, it is not possible to combine all these 121 points into a single PRC curve. Instead 22 PRC curves can be drawn through all the 121 points. The 11 combinations of the ratio feature provide 11 PRC curves and the same goes for the energy feature. The following results show the solution space for non-referenced energy features. The reason for doing this will be explained in Section 7.6.

Fig 7.7 presents the total solution space for the training set in two plots. In the left plot the effect of variation of the ratio feature can be observed per static energy thresholds. In the right plot, the effect of variation of the energy feature can be observed per static ratio threshold. Keep in mind that the same 121 points are present in both plots, the points are just connected in a different manner. The magnitude of the threshold is also represented in the specific color of the PRC curves. Dark blue indicates the effect of no threshold and the warmer the colors get, the higher the threshold's magnitude. The red dotted line presents the baseline for the training dataset.

From the figure it becomes clear that when the threshold values are set to zero, the performance is roughly equal to the baseline. The performance for PRC curves is related to the area under the curve. For an initial increase in threshold magnitudes for both features, an increase in area is visible. The detector is able to keep a stable sensitivity while increasing its precision. After a certain point the threshold values get too high and start rejecting samples which are accompanied by an annotation, this results in a rapid decrease in the sensitivity. Interestingly, while the sensitivity decreases there is still an increase of precision. This can be explained by the fact that the precision is equal to the number of true positive against all detected samples. Due to the increase of the threshold's magnitude, the number of total detected samples is decreasing more rapidly than the decrease in number of true positives.

Another note has to be made on the overall performance. There is a high chance that the performance presents in the solution space presents a pessimistic view on the actual case. As discussed in Section 4.3, the annotations were focused on the best examples and not all delta brushes were provided with an annotation. This will result in the waveform detector detecting samples which are indeed delta brushes, but are classified as false positive since no annotations are available. If the detected samples would be reevaluated, some of the current false negative samples could become true positives. This would increase performance for both the sensitivity and the precision.

Fig 7.8 presents the total solution space for the validation set in the same two plots. What becomes clear is that the overall performance is far less for the validation sets when compared to the training sets. It was already expected that the validation datasets were of lesser quality when compared to the training datasets. The datasets for infants with the best medical outcome were used for the training datasets. This was also seen in the maximum signal amplitude for the validation set when compared to the training set. The average of the maximum signal amplitude for the training sets was equal to $9,230\mu V$, for the validation dataset this average was $2,000\mu V$ which is five times as low. Since the non-referenced energy features are presented here, this could mean that the threshold values are too high for the validation datasets. This would explain the decrease in performance.

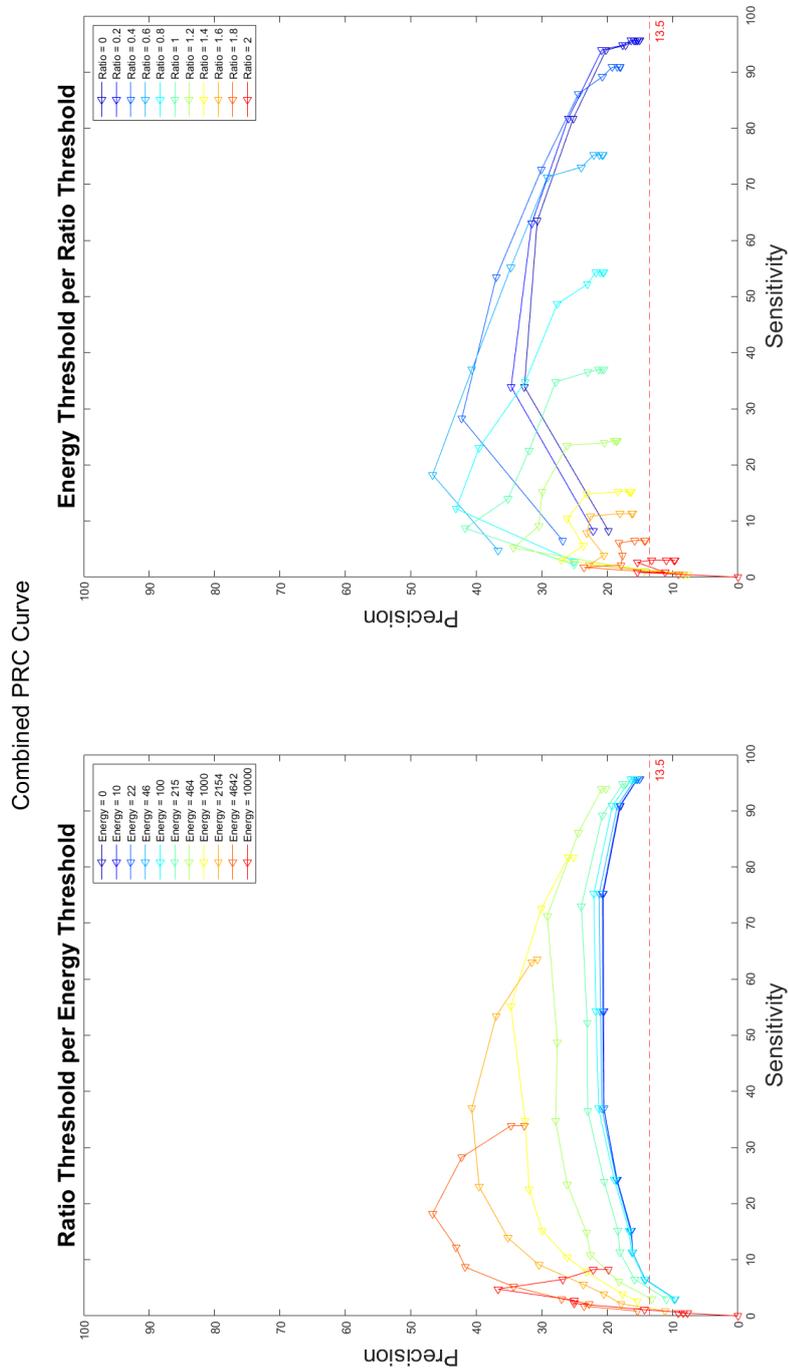


Figure 7.7: PRC plot showing solution space for all non-referenced feature combinations on training data.

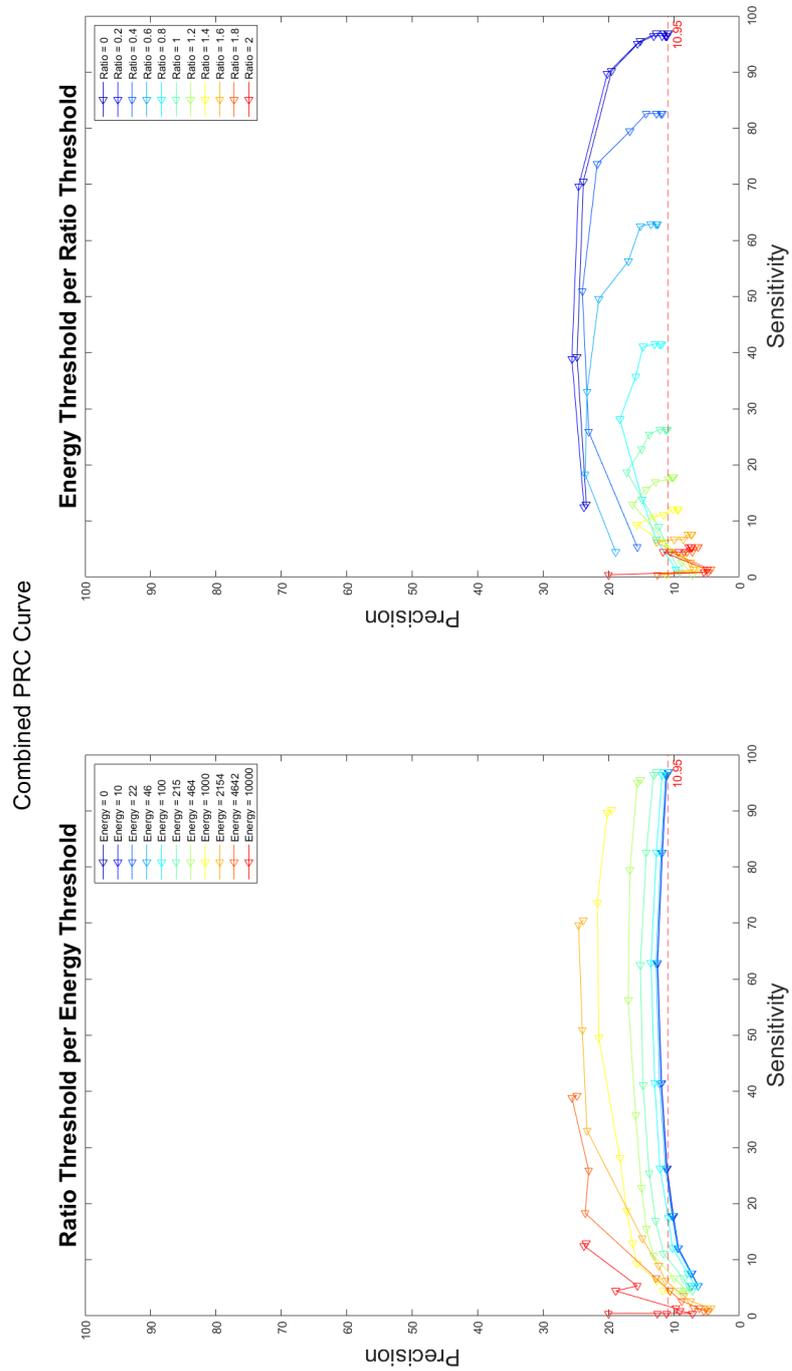


Figure 7.8: PRC plot showing solution space for all non-referenced feature combinations on validation data.

7.6. Comparison Referenced to Non-Referenced Features

The performance limitations in the solution space could be solved by using a referenced energy threshold. In Section 7.2, the need for referencing the energy feature was already discussed to decrease independence of the signal's amplitude. To test if this did increase performance, the detector script has also been applied to the training and validation datasets using referenced energy thresholds. The referencing of the was done by multiplying each energy threshold with the ratio between the maximum signal amplitude, as provided with the dataset, and the average maximum signal amplitude for all datasets in the training dataset

$$Th_{referenced} = Th_{original} * \frac{xmax_{dataset}}{xmax_{average,trainingset}}. \quad (7.7)$$

Through the use of this calculation, the energy threshold will be slightly higher for datasets with an above average maximum signal amplitude and vice versa. Appendix D holds the solution spaces for the referenced and non-referenced energy features for both the training and validation sets. Since its hard to visually compare on performance, the average precision has been calculated for every PRC curve with a static energy value. The results of these calculation are presented in Table 7.2. From the difference between the referenced and non-referenced sets, it becomes clear that the referencing does more harm than good. The only drastic difference in average precision is seen for the high-magnitude thresholds for the validation set, but this is mainly due to the sharp decrease in performance for the non-referenced set.

There is one interesting phenomenon which can be detected by visual inspection. The sensitivity for the referenced validation sets shows a sharp decrease and seems limited to a value of 88%. This is even the case for a low-energy threshold of 0. There is however still a threshold active at this moment, namely the third threshold serving as a high-energy cut-off. This threshold is also referenced according to (7.7). From this finding it seems that this threshold has been decrease to the point where it also affect the detection of samples which are accompanied by an annotation. Based on these findings, the referencing seems too strong, causing a decrease in performance when used. A different way of making the energy feature independent of the signal's amplitude should be investigated.

Table 7.2: Comparison of average precision for referenced and non-referenced energy levels.

| Energy Threshold [V^2] | Average Precision | | | | | |
|----------------------------|-------------------|------------|------------|----------------|------------|------------|
| | Training Set | | | Validation Set | | |
| | Non-referenced | Referenced | Difference | Non-referenced | Referenced | Difference |
| 0 | 1738.37 | 1738.45 | +0.07 | 1020.20 | 946.00 | -74.20 |
| 10 | 1738.45 | 1738.52 | +0.07 | 1021.38 | 946.00 | -75.38 |
| 22 | 1745.59 | 1739.76 | -5.84 | 1025.33 | 946.00 | -79.33 |
| 46 | 1768.30 | 1757.93 | -10.37 | 1041.52 | 947.23 | -94.29 |
| 100 | 1816.57 | 1822.83 | +6.26 | 1108.37 | 951.37 | -157.01 |
| 215 | 1943.53 | 1936.27 | -7.26 | 1266.79 | 968.26 | -298.53 |
| 464 | 2344.26 | 2314.28 | -29.98 | 1405.05 | 1041.56 | -363.495 |
| 1,000 | 2425.82 | 2495.38 | +69.56 | 1636.91 | 1213.70 | -423.22 |
| 2,154 | 2219.21 | 2123.45 | -95.98 | 1298.39 | 1337.04 | +38.65 |
| 4,642 | 1333.52 | 1129.45 | -204.08 | 685.60 | 1541.72 | +856.11 |
| 100,000 | 194.51 | 192.02 | -2.49 | 176.25 | 1055.15 | +878.91 |

7.7. Single-Threshold Detector

The solution spaces for the training and validation datasets already give the total performance for the designed detector. The goal of this chapter is to provide an example on how a threshold can be chosen based on the values presented in the total solution space. As explained in Section 7.3, there exist a threshold selection procure for PRC curves based on the distance between the threshold combination and the point (100,100). Fig. 7.9 shows the possible candidates for threshold selection in the solution space of the training set. The available data and solution of the distance calculation is provided in Table 7.3. A minimum distance to point (100,100) is found for a energy threshold of 1000 and a ratio threshold of 0.4.

Table 7.3: Comparison of distance to point (100,100) for different threshold combinations.

| Threshold Energy Low | Threshold Ratio | Sensitivity | Precision | Distance |
|-------------------------|--------------------|-------------|-----------|----------|
| 464 | 0.4 | 86.09 | 24.38 | 76.89 |
| 1000 | 0.2 | 81.74 | 25.93 | 76.29 |
| 1000 | 0.4 | 72.61 | 30.04 | 75.13 |
| 2154 | 0.4 | 53.48 | 36.94 | 78.37 |
| 2154 | 0.6 | 36.96 | 40.67 | 86.57 |
| 4642 | 0.4 | 28.26 | 42.21 | 92.12 |
| 4642 | 0.6 | 18.26 | 46.67 | 97.60 |

Single-Threshold Performance

This threshold combination is provided to the detector script and been applied to all four validation datasets. The detector detected a total of 2,600 samples over the four validation sets. The detector performance was a sensitivity of 76.66% and a precision of 21.71%. Finding a comparison for these performances measures is difficult since there are no other automatic detectors found in literature which specifically focus on the delta brush.

A detector designed by O'Toole and all, achieved an area under the ROC curve of 0.989 with a sensitivity-specificity of 95.8-94.4% [57]. Their approach focuses on detecting regions with general bursting activity from quiescent regions. They did this through the brute-force calculation of 26 features and try all possible combinations to get maximal performance. The input data they used was also cleaned from artifacts, meaning that a decrease in performance is to be expected when applied to data with artifacts.

Possible Performance Optimization

There was already a major difference of total performance between the solution spaces for the training and validation sets. Because of the difference in quality between the training and validation datasets, comparing the performance between these datasets is not fair, a fairer performance measure would be to use a cross-validation method [49]. In cross-validation, all-but-one of the datasets are used for threshold selection, the performance is then tested on the leftover dataset. This procedure is then repeated until every dataset is used once for validation. The average performance can then be calculated as the final performance measure. This would exclude difference in datasets between training and validation.

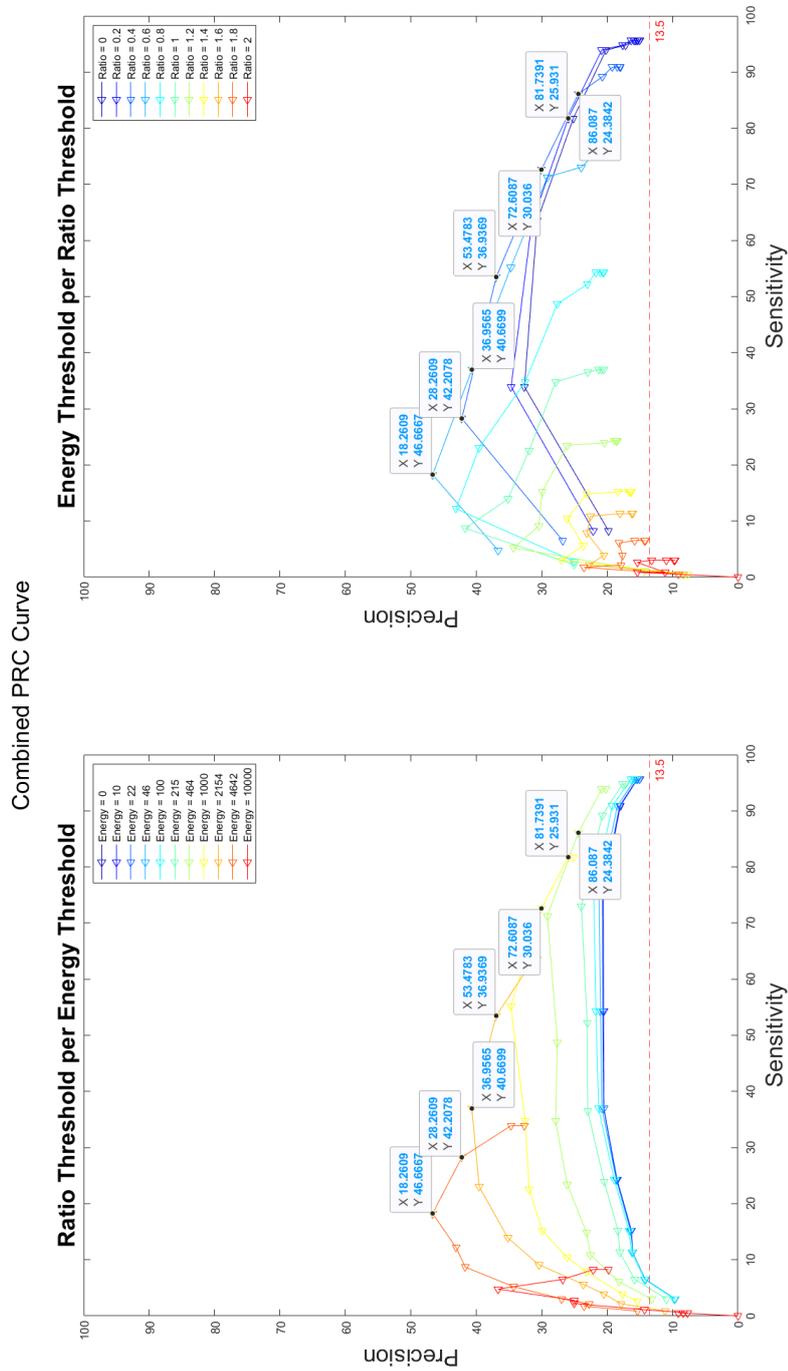
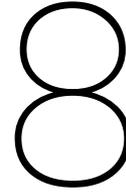


Figure 7.9: PRC plot with candidates for threshold selection based on distance to point (100,100).

7.8. Conclusion

In this chapter a detector has been designed based on the inherent characteristics of the delta brush. A triple-threshold detector based on two detector features, ratio in the time-frequency domain and energy, has been designed and its performance has been tested on both the training and validation dataset. The overall performance was still limited, but limitations related to the provided annotations and energy referencing technique were discussed. An examples was provided for threshold selection based on the provided solution space for the training set and a detector based on this threshold was evaluated with the validation datasets.



Conclusions & Recommendations

This chapter will focus on the conclusions related to the initial goals. The end results will be discussed together with the limitations. The scientific contributions of this thesis will be discussed. The chapter will end in recommendations for the current and future work.

8.1. Conclusions

The initial goals of this study were to find an informative signal which can provide knowledge on the state of development for preterm infants, to find a way of quantifying this signal, to provide means to automate the detection of this signal, and to develop the solution in such a way that it would be accepted by the medical field. In the course of this thesis all goals have been covered.

An informative signal was found in the delta brush which can be measured in the neonatal EEG from 28-29 weeks PMA till term age. There is already evidence in literature for its relevance related to developmental outcome. This evidence can be further investigated if quantitative analysis and automatic detection of the delta brush are available. This would also provide means to decrease the inconsistencies concerning the precise characteristics of the delta brush.

A way of quantifying the delta brush was found in the continuous wavelet transform. Spectral information is widely used in the medical field to characterize the waveforms present in the (neonatal) EEG. Though the spectral analysis is still mainly based on visual analysis. A way of quantifying the spectral content would provide valuable information concerning the signal's content. The continuous wavelet transform feature an adaptable analysis window which can be adjusted to the signal of interest. The possibility to increase the spectral resolution for short-duration activity makes it a suitable analysis method for the analysis of non-stationary biosignals, like EEG.

An automatic detection method was found by defining two detection features based on the inherent characteristics of the delta brush. The first feature is a measure of the ratio between peak in the high-frequency region (3.3-40 Hz) and the low-frequency region (0.1-3.3 Hz). The second feature is an energy calculation based on a windowed Squared Energy Operator (SEO). A triple-threshold multi-channel detector, based on these features, was initialized and performance was tested over the full threshold range. A Precision-Recall (PRC) curve was provided showing the total solution space for both feature thresholds. Based on the point closest to (100,100) in the training set PRC, a single threshold was selected for validation with the validation datasets. This has resulted in a detector with a sensitivity of 73.66% and a precision of 21.71%

The troubles in the acceptance of new innovations in the medical field were found to be related to a lack of trust. The elements for building trust in technology were discussed and the key factors for building trust for the proposed solution were identified. Through the use of the inherent characteristics of the delta brush the interpretability of the detector is increased. The transparency has been increased by using the redundant behavior of the continuous wavelet transform to visualize the features used for detection. An interactive application for visualization of the EEG data and wavelet coefficients was designed in MATLAB. This application can run standalone and can be used to visualize the detector's results for the end user. The ease of operation is inherently limited by the development in the technical environment provided by MATLAB.

Limitations

The limitations to the overall performance can be explained by the uncertainties in the exact delta brush behavior and the incomplete annotation set, which causes possible detected delta brush activity to be labeled as false positives.

The EEG datasets were collected from infants with an age of 30-32 weeks PMA. This limits the validation of the results to this age as well. Since the delta brush is active from 28-29 weeks PMA till term age, the detector is expected to be able to detect the samples here as well, though this has to be further investigated.

The delta brush detector is still partially dependent on the EEG signal's amplitude. The comparisons between the training and validation set showed that deviations from the average signal amplitude can cause performance issues.

8.2. Scientific Contributions

The scientific contributions made in this thesis work are as follows:

Interactive Application

An interactive application has been developed, which can be used to visualize single-channel EEG signals in combination with its time-frequency behavior. The visualization of the spectral characteristics has been validated to be of added advantages if combined with traditional EEG visualization measures. Different purposes for the application in educational and scientific settings has been suggested: it could be used to learn new medical professional to learn about the spectral characteristics of the EEG signal and specific waveforms, it could also be used to provide high-accuracy annotations for research purposes. The application can be distributed in a standalone software package, which provides the possibility for it to be used in a clinical setting.

Sample Collection

The quantitative analysis of the collected samples provided new insights into the spectral characteristics of the delta brush. A larger spread in frequency activity was observed in this study than was suggested in literature. Since a high-density EEG measurement dataset was used for this study, the sample analysis provides spectral information and delta brush activity on these additional channel locations. The information on the incidence of delta brush activity per HD-EEG channel will be used in a follow-up study by L. de Wispelaere.

Waveform Detector

Since there is no other automatic detector specifically focused on the delta brush activity mentioned in literature, this thesis provides a novel detection method to be used on neonatal EEG. It features multi-channel analysis on uncleaned EEG signal segments, i.e. featuring artifact activity. The detection features are based on the intrinsic characteristics of the waveform itself, providing an understandable analysis method for the medical field. In combination with the interactive application, the detector's decision-making process can be made transparent. These elements should increase the acceptance for this method by the medical field.

8.3. Recommendations

The results and performance of this work could be improved by the following recommendations:

Interactive Application

The ease of operation for the interactive application should be further increased. The wide-range of interaction possibilities available for application nowadays sets high expectations to the interactivity of the interactive application. The technical environment, provided by MATLAB, features only limited options for interaction. If the interactive application is to be developed further, it is advised to implement it in a software environment dedicated to the development of these kind of applications. The performance related to computational and storage load can be further increased by subsampling the data and limiting the number of wavelet coefficients to be calculated. This could, for example, be done by specifying a specific number of voices per octave per frequency region.

Sample Collection

In the analysis of the samples and detector results, a large variation in EEG quality was observed between the different subjects. It is recommended to include EEG data from more perterm infants to find a representative sample set. It is also recommended to include EEG datasets measured at different ages. This way the detector's performance won't be limited to a narrow age range. It will also open up the possibility to follow the change of the delta brush characteristic as the brain matures.

More waveform annotations should be collected, spanning the whole signal segment. Instead of providing the specific time instances a waveform is present, it could be of added advantage to ask the annotator to provide annotations on a window. By reviewing each window with a positive or negative annotation for waveform activity, the validation of the detected samples should be less complex. By also providing a negative annotation for regions without activity, the true negative can be included in the performance measures. This will provide far more ways of validating the performance. Because of the inconsistencies in the precise waveform characteristics, it is advised to use a second annotator to decrease subjectivity of just singular annotations.

Waveform Detector

The current proposed detector features multi-channel detection, because the annotations were provided for all channels at the same time. By first focusing on single-channel detection a more robust detector can be developed. A more accurate sample set can be collected if just one single channel is used, since there is no ambiguity on the samples which best represents the annotation. This will also provide a robust validation for single-channel detectors. The insights and validations generated in the development of this single-channel detector can then be used for a multi-channel solution.

To open up the possibility for other threshold optimization methods, the non-target samples set should be split up into its separate subclasses. This should provide samples sets which are normally distributed. This multi-class classification problem can then be approached as a combination of multiple two-class problems.

The detector is still partially dependent on the signal's amplitude. A different way of making the energy feature independent should be investigated.

To investigate the idea that a large portion of the detected samples labeled as false positive are indeed true positives, it is advised to ask an independent neonatal EEG expert to validate the detected samples.

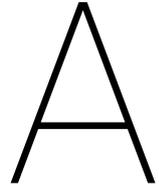
Future work

The current work has provided evidence for the added value for the use of time-frequency information in the medical field. The current work is limited to the characteristics of the delta brush. A more general scientific base for use of the continuous wavelet transform can be established if the same evidence can be found for the other waveforms present in neonatal EEG. Since all of the waveforms present in the neonatal EEG can be described in terms of their spectral behavior, it could be possible to find similar spectral detection features for these waveform as well.

The current work focuses on using the continuous wavelet transform with the Morlet mother wavelet. It is very well possible that a different mother wavelet can be found which increases the performance for delta brush detection even further. It could also be interesting to test the effect of using different mother wavelets for different maturational waveform in the neonatal EEG.

This work already provides quantitative behavior on the characteristics of the delta brush. This information can be further used in the medical field to research the link between the development of the brain and delta brush activity. The outcome of the detector can be used as a guideline for delta brush activity, but should still be validated by a neonatal EEG expert. The validated delta brush detection could then be used for different experiment. The number of delta brushes per time instance could be tested versus the infant's developmental level. To provide a measure of connectivity, the synchrony in delta brush measurements between the left and right hemisphere can be measured.

Appendix



Application's Assignment & Questionnaire

Assignment EEG CWT Analyzer

The EET CWT Analyzer application makes it possible to view single EEG channels of a measured EEG dataset. It is possible to filter these signals and calculate a time-frequency representation of the signal using the continuous wavelet transform (CWT). This small assignment will guide you through the set-up and use of the EEG CWT Analyzer.

1. Open "EEG_CWT_Analyer.exe".
2. Import the "EEGFile.mat" EEG-file.
3. Change the active EEG channel to "9 – O1".
4. Jump to the time-window around 245 seconds.
5. Filter the signal with a high-pass frequency of 0.5 Hz and a low-pass frequency of 40 Hz.
6. Calculate the CWT of the active channel.
7. Import the "EEGAnnotationsFile.mat" Annotation-file.

The red lines in the signal window represent the annotations of an EEG feature called the Delta Brush. The annotations times are rounded to the nearest second.

8. Plot the CWT of the current window to view the time-frequency representation.
9. Jump to the time-window around 244 seconds and press "Plot local CWT". The "Local T-F spectrum now shows a cross section of the time-frequency representation for the annotation placed on 244 seconds (indicated by the blue dashed line).

*Since there is a mismatch between the annotation time and the exact moment the Delta Brush is present, we will try to find a better match by specifying a new annotation time with a precision of 0.1 seconds. By filling in this new time value with a higher precision in the "jump to" field it is possible to find a better matching new annotation time.
(Note that the application uses a dot as a decimal point).*

10. Based on your intuition, try finding a better matching annotation time by filling in new time value with a higher precision in the "jump to" field and press "Plot local CWT". Do this for the annotation set in the window around 245 seconds (running from 238 seconds until 253).

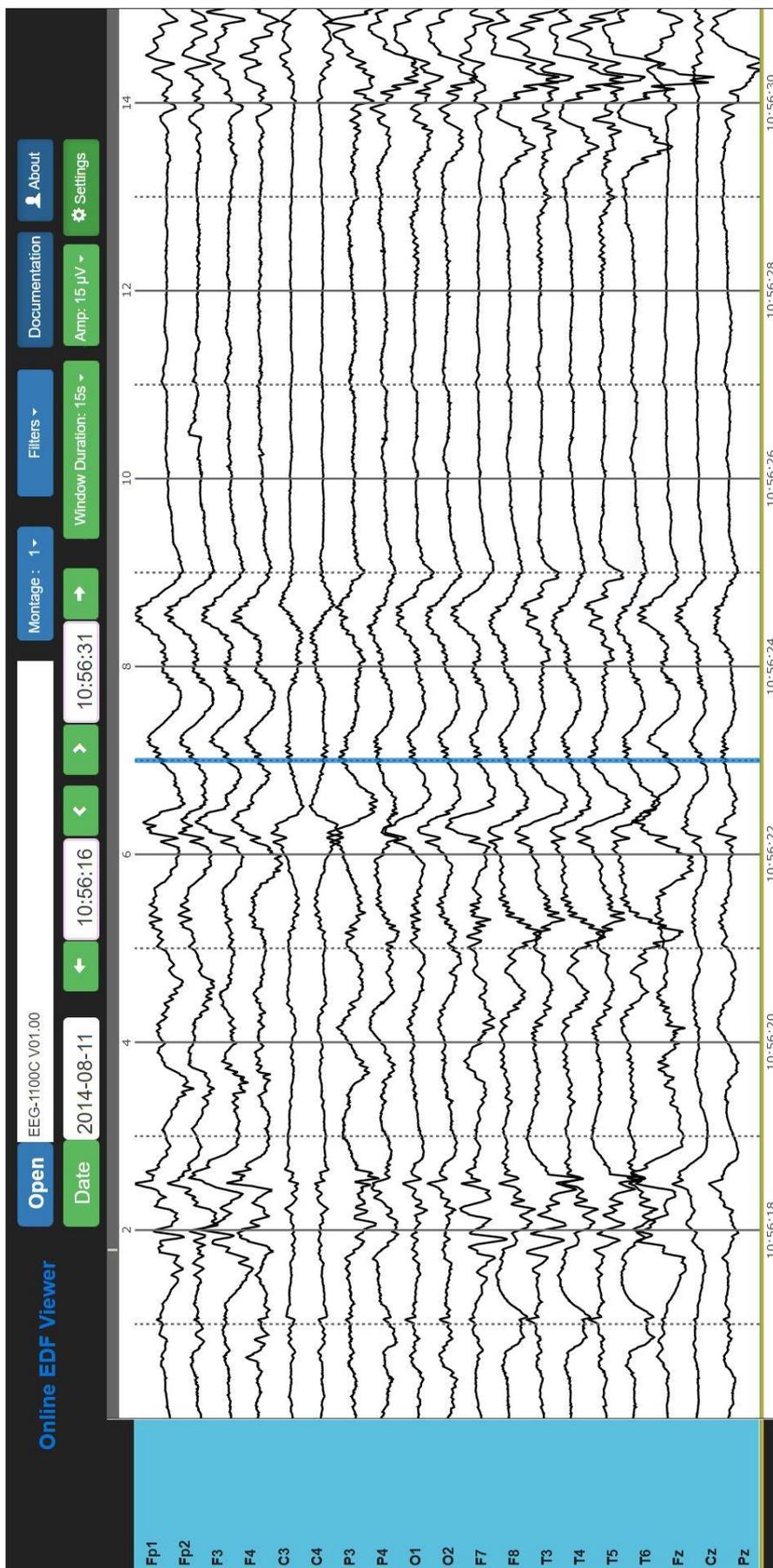
| | | | | |
|--------------------------|-----|-----|-----|-----|
| Previous annotation time | 240 | 244 | 246 | 252 |
| New annotation time | — | — | — | — |

Questionnaire EEG CWT Analyzer

Thank you for completing the small assignment. Hopefully it helped you get a feeling the EEG CWT Analyzer. Please fill in the following questionnaire to aid in the further improvement of the application. See page 3 for an example of a general EEG Viewer for the same dataset.

1. Name: _____
2. Employment: _____
3. Function: _____
4. The assignment's instructions were clear and easy to follow: Yes / No
5. The application works in an intuitive manner: Yes / No
6. When looking at the "EEG Signal" window...:
 - ... I had no idea what I was looking at.
 - ... it gave me comparable insight as a general EEG viewer.
 - ... it improved the insight as to a general EEG viewer.
7. When looking at the "Frequency Content (FFT)" window...:
 - ... I had no idea what I was looking at.
 - ... I felt that is had something to do with the signal's frequency content
 - ... it improved my insight as to a general EEG viewer.
8. When looking at the "Time-frequency Spectrum (CWT)" window...:
 - ... I had no idea what I was looking at.
 - ... it was of added value combined with the "signal" window.
 - ... I understand it represents how strong certain frequencies are present.
9. When looking at the application as a whole...
 - ... I had no idea what I was looking at.
 - ... it gave me comparable insight as to a general EEG viewer.
 - ... it increased my insight into the features present in the EEG signal.
10. I could use the application as a replacement for a general EEG viewer: Yes / No
If no, what did you miss when compared to a general EEG viewer?
11. It was easy to find a better matching annotation time using the application: Yes / No
12. Further Comments:

Example General EEG Viewer



B

Examples of Collected Samples

Delta Brush Class

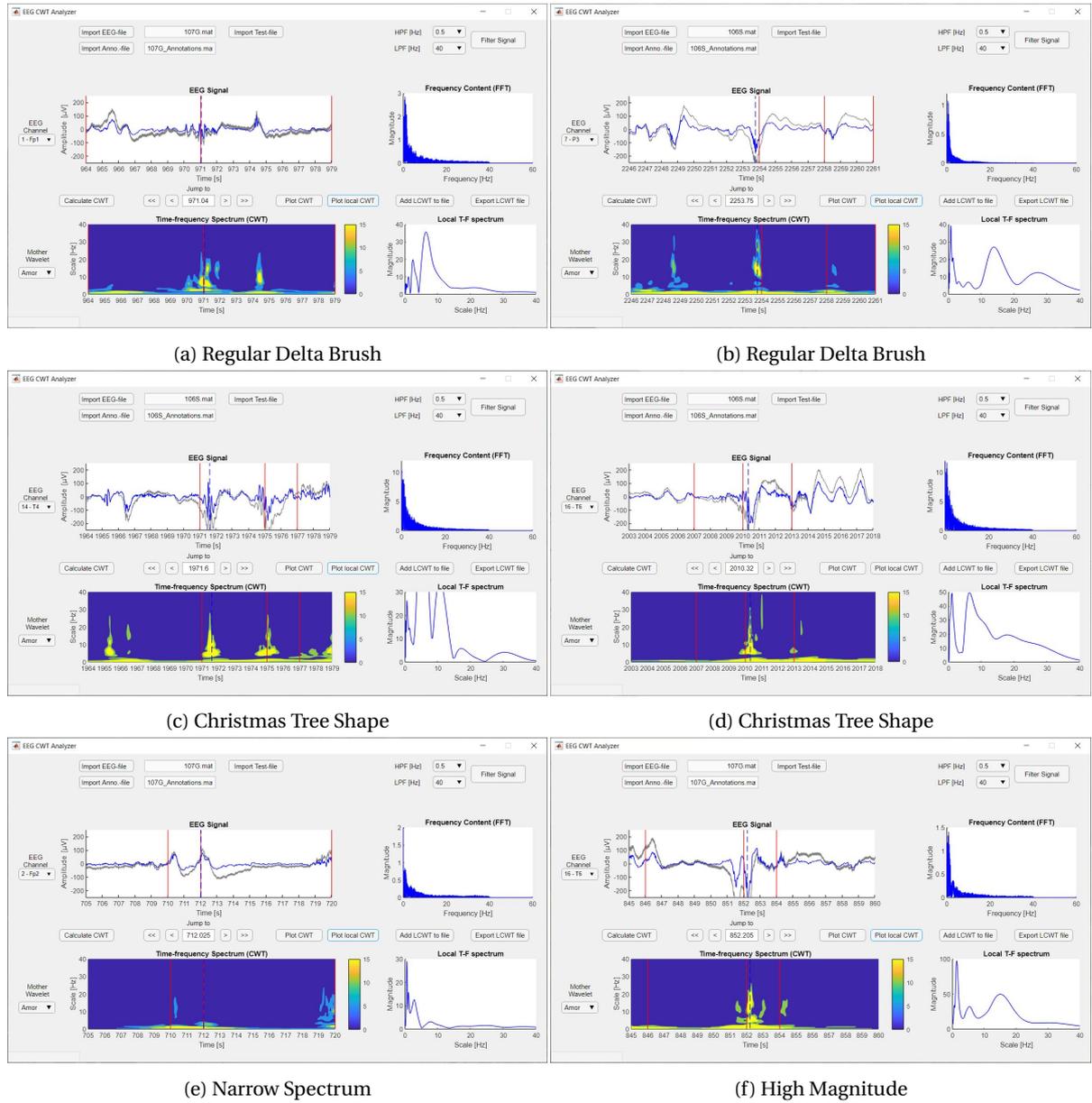


Figure B.1: Examples 'Delta Brush' Class. Dotted line and 'Local T-F Spectrum' window indicate collected sample.

Other Activity Class

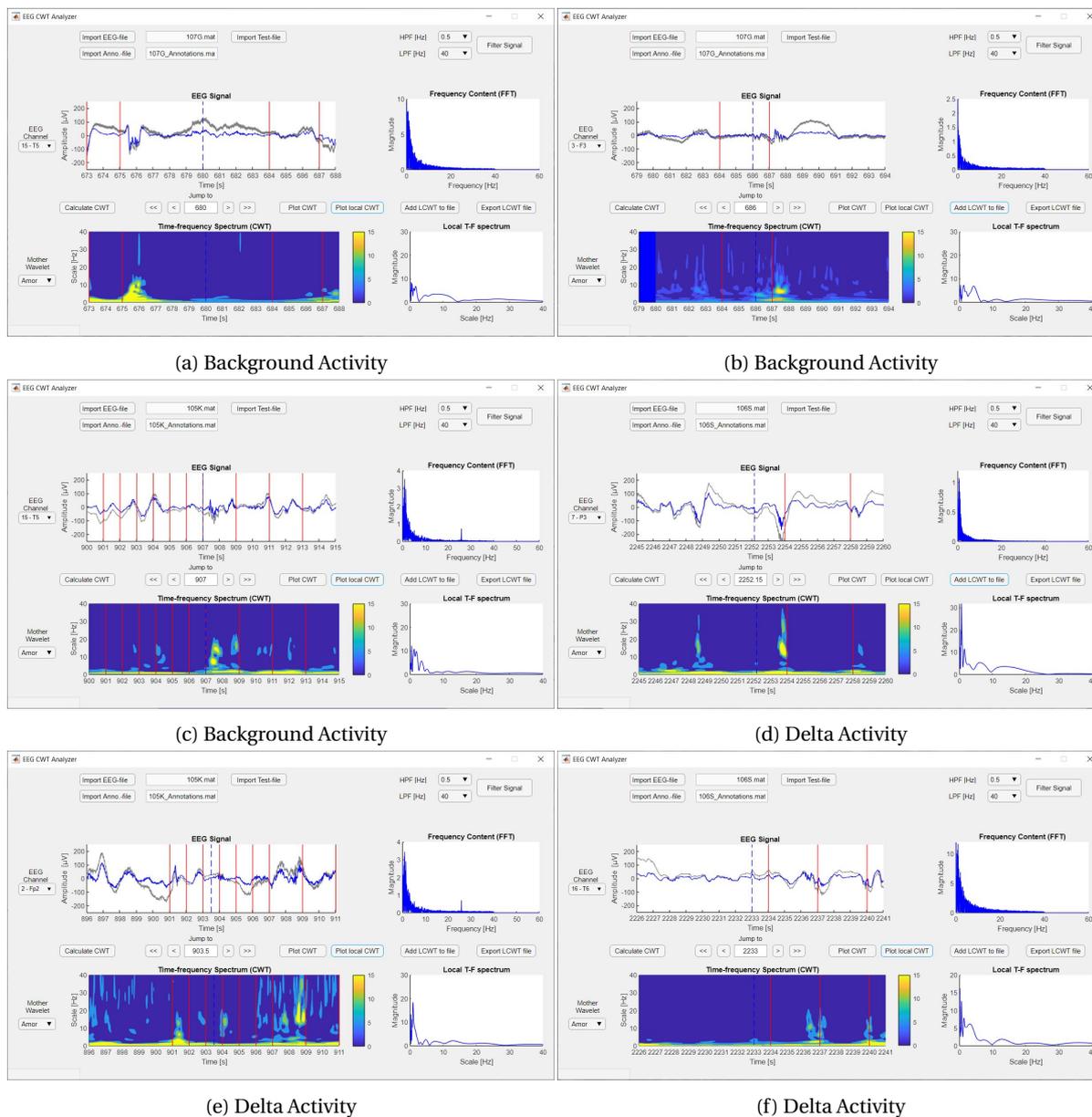


Figure B.2: Examples 'Other Activity' Class. Dotted line and 'Local T-F Spectrum' window indicate collected sample.

Artifact Class

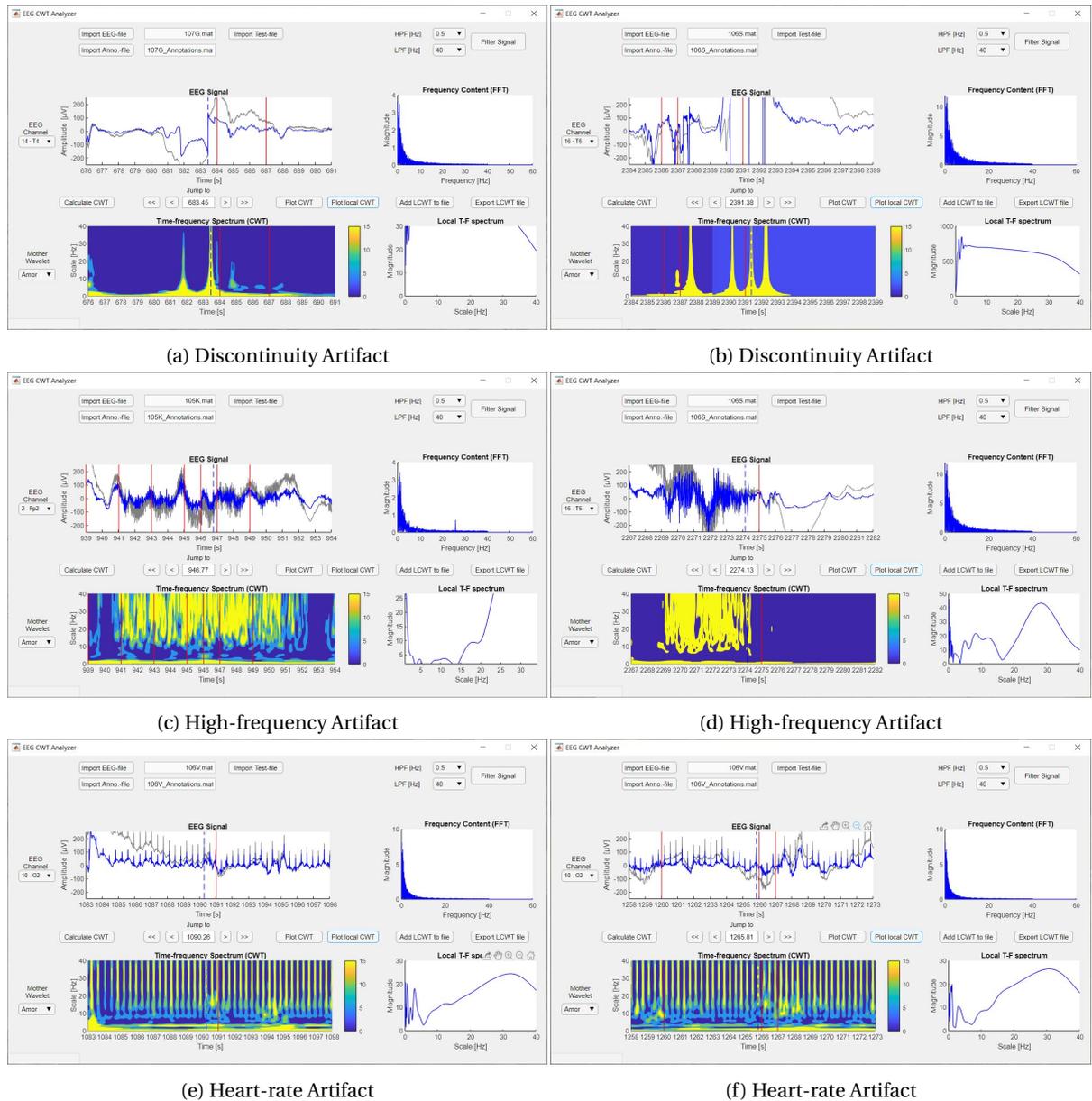
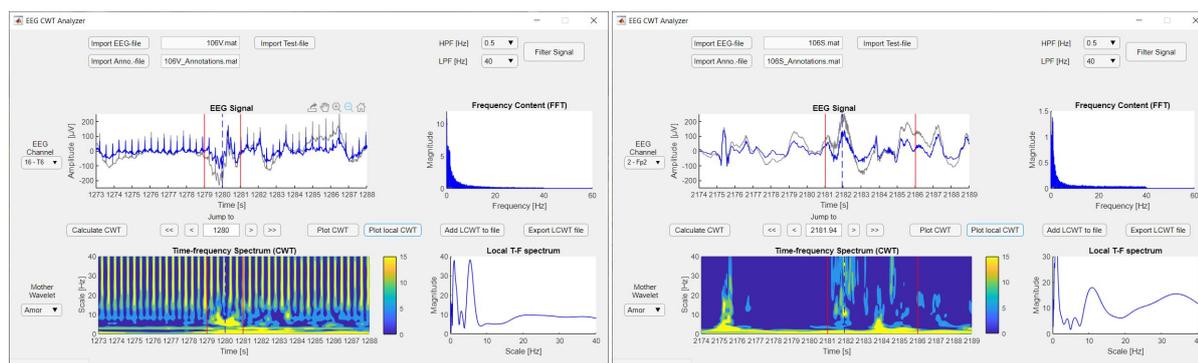


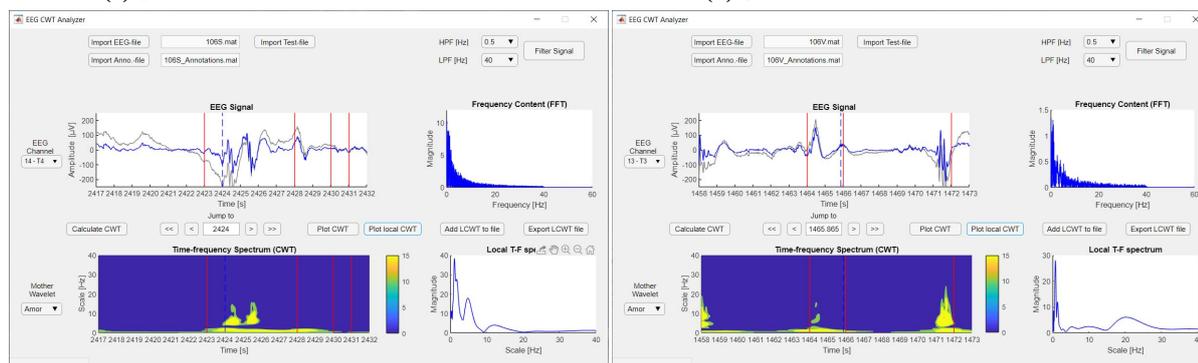
Figure B.3: Examples 'Artifact' Class. Dotted line and 'Local T-F Spectrum' window indicate collected sample.

Doubt Class



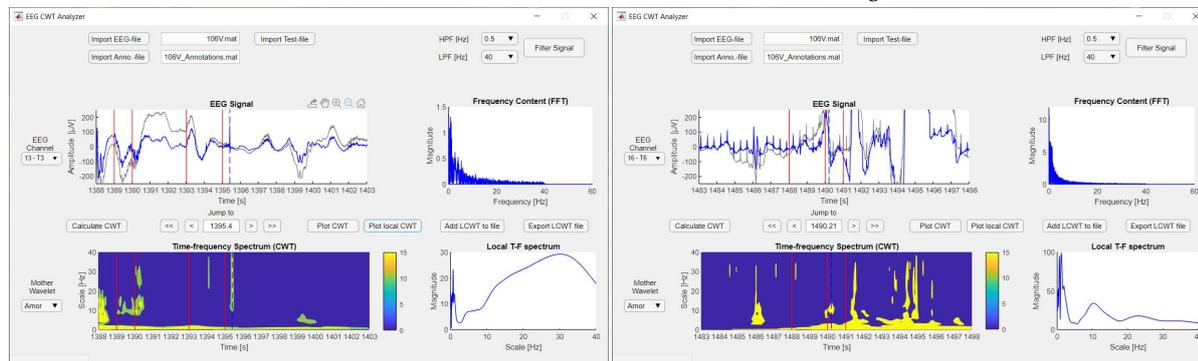
(a) Combination Delta Brush and Artifact

(b) Combination Delta Brush and Artifact



(c) Local Maximum

(d) Low Magnitude Delta Brush



(e) Possible Artifact

(f) Possible Artifact

Figure B.4: Examples 'Doubt' Class. Dotted line and 'Local T-F Spectrum' window indicate collected sample.

C

Comparison of Energy Calculation Principles Feature II

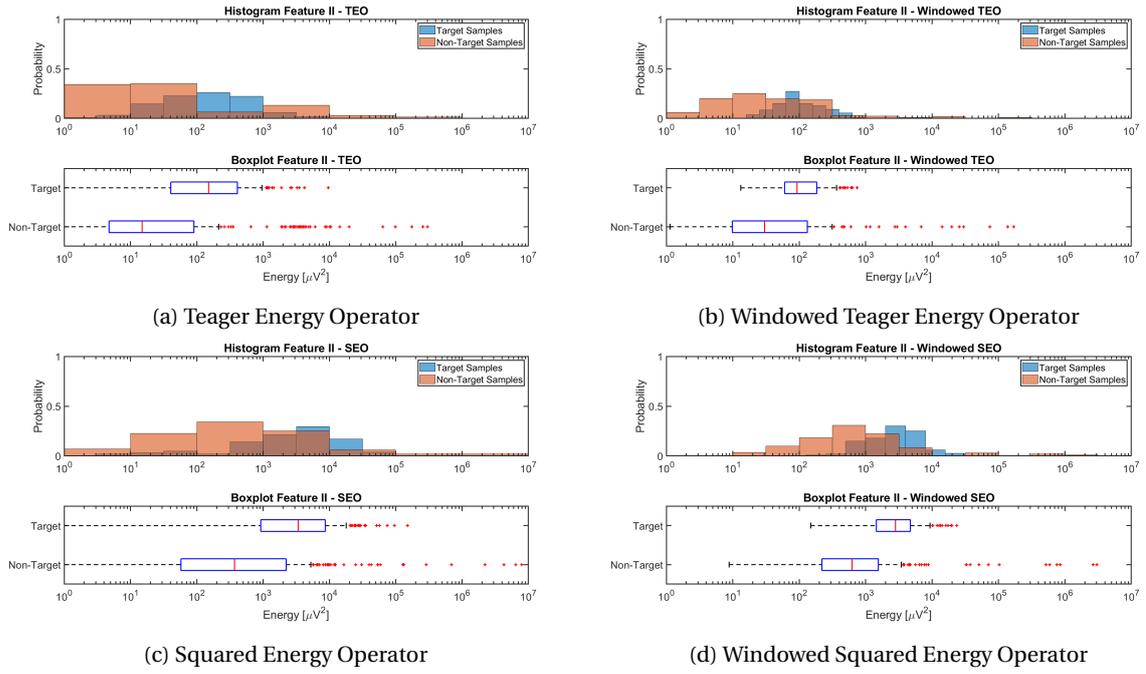


Figure C.1: Feature II - Energy Operator



Comparison of PRC Solution Space For Referenced and Non-Referenced Detection

Training Set - Non-Referenced

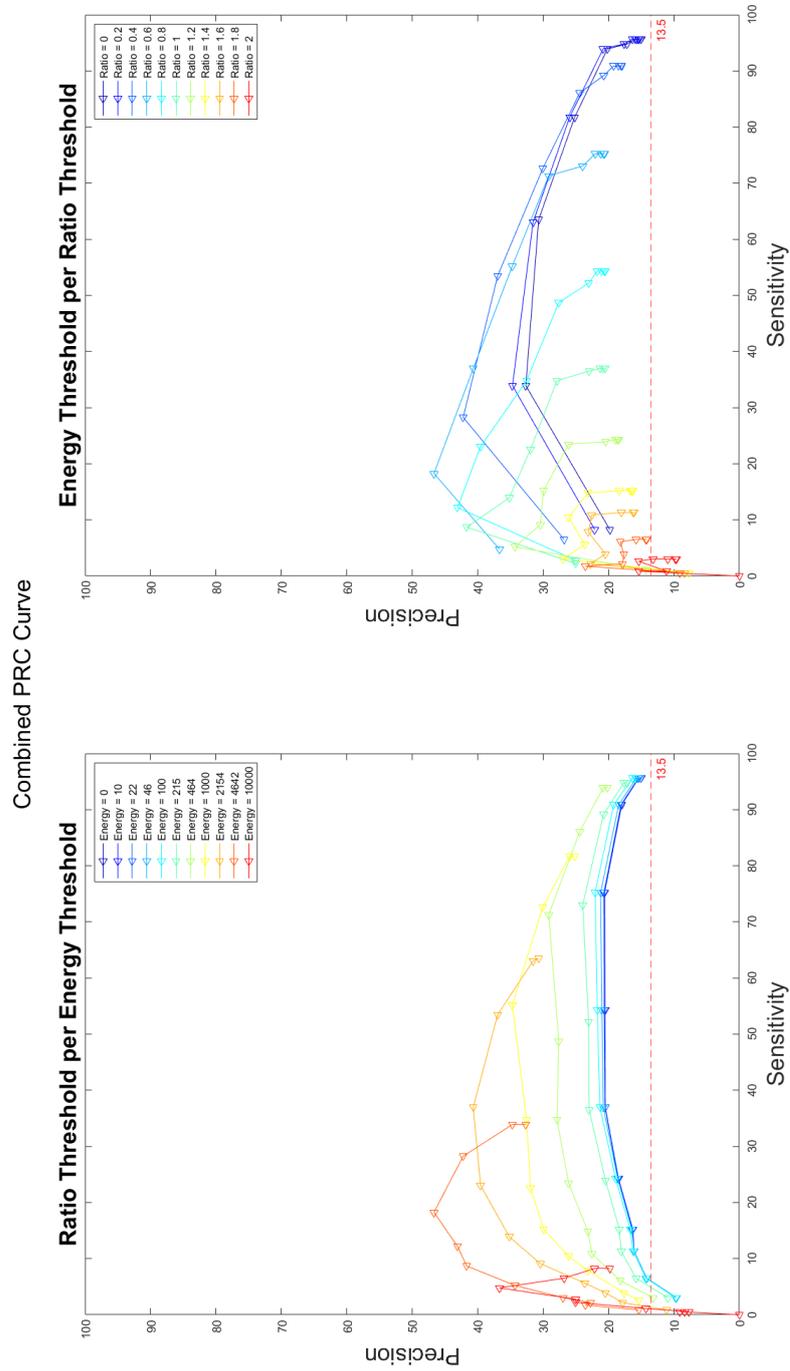


Figure D.1: Solution Space for Non-Referenced Energy Feature of Training Set

Training Set - Referenced

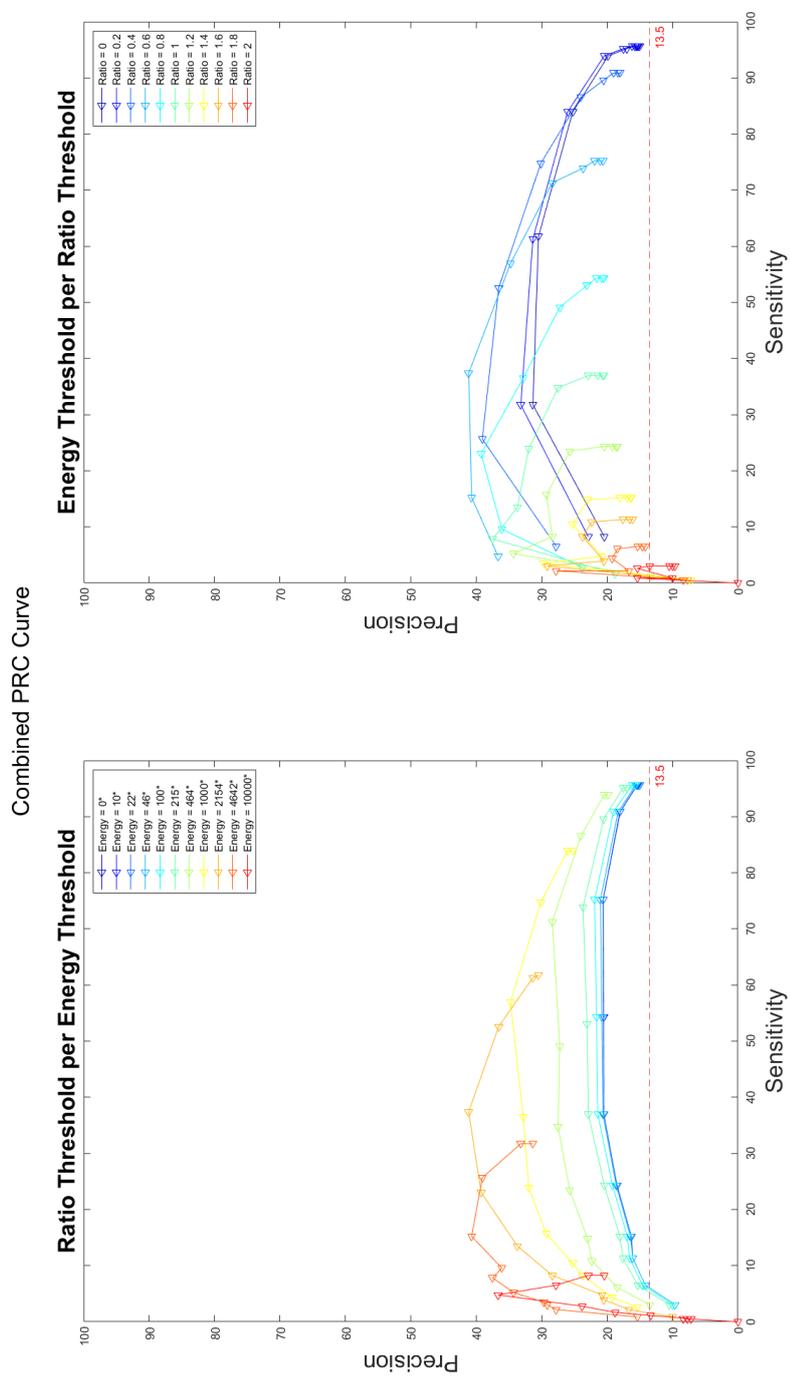


Figure D.2: Solution Space for Referenced Energy Feature of Training Set

Validation Set - Non-Referenced

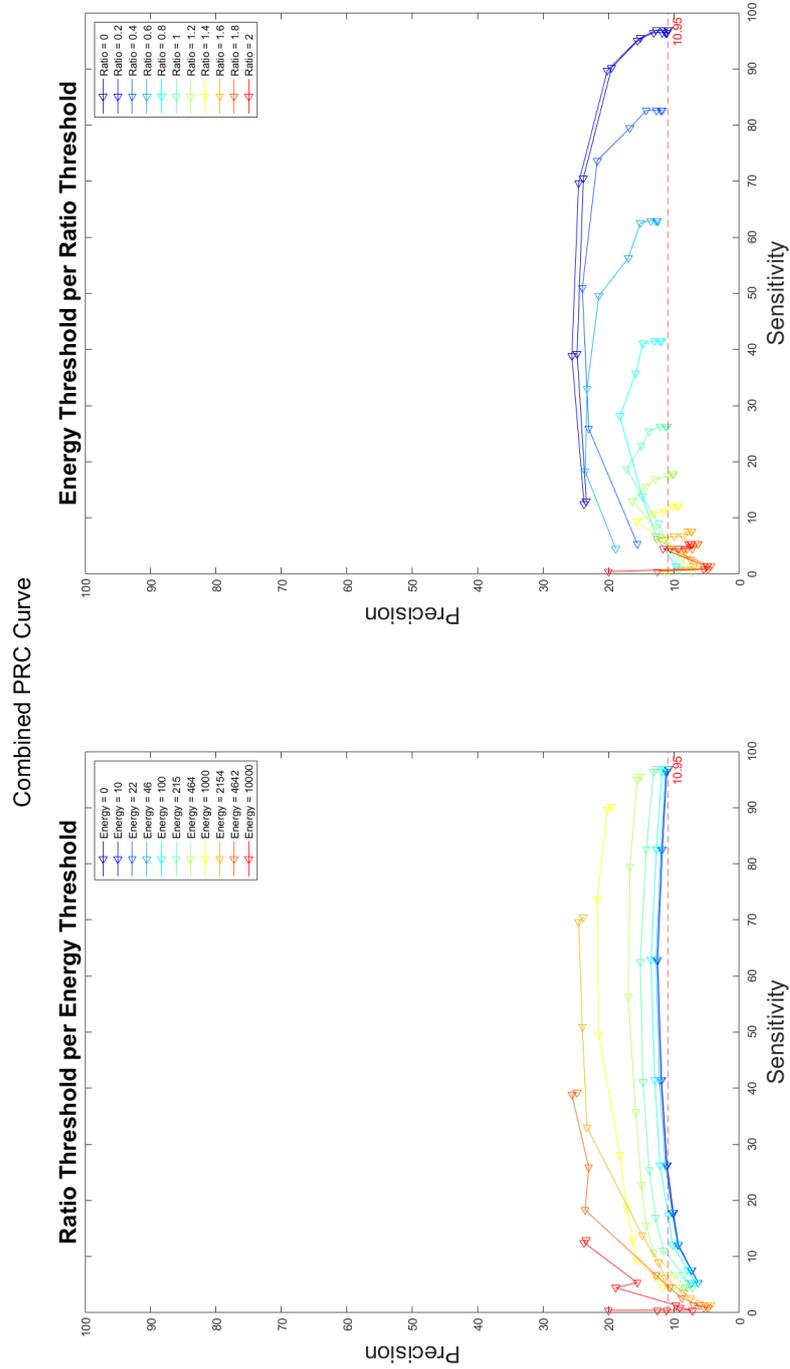


Figure D.3: Solution Space for Non-Referenced Energy Feature of Validation Set

Validation Set - Referenced

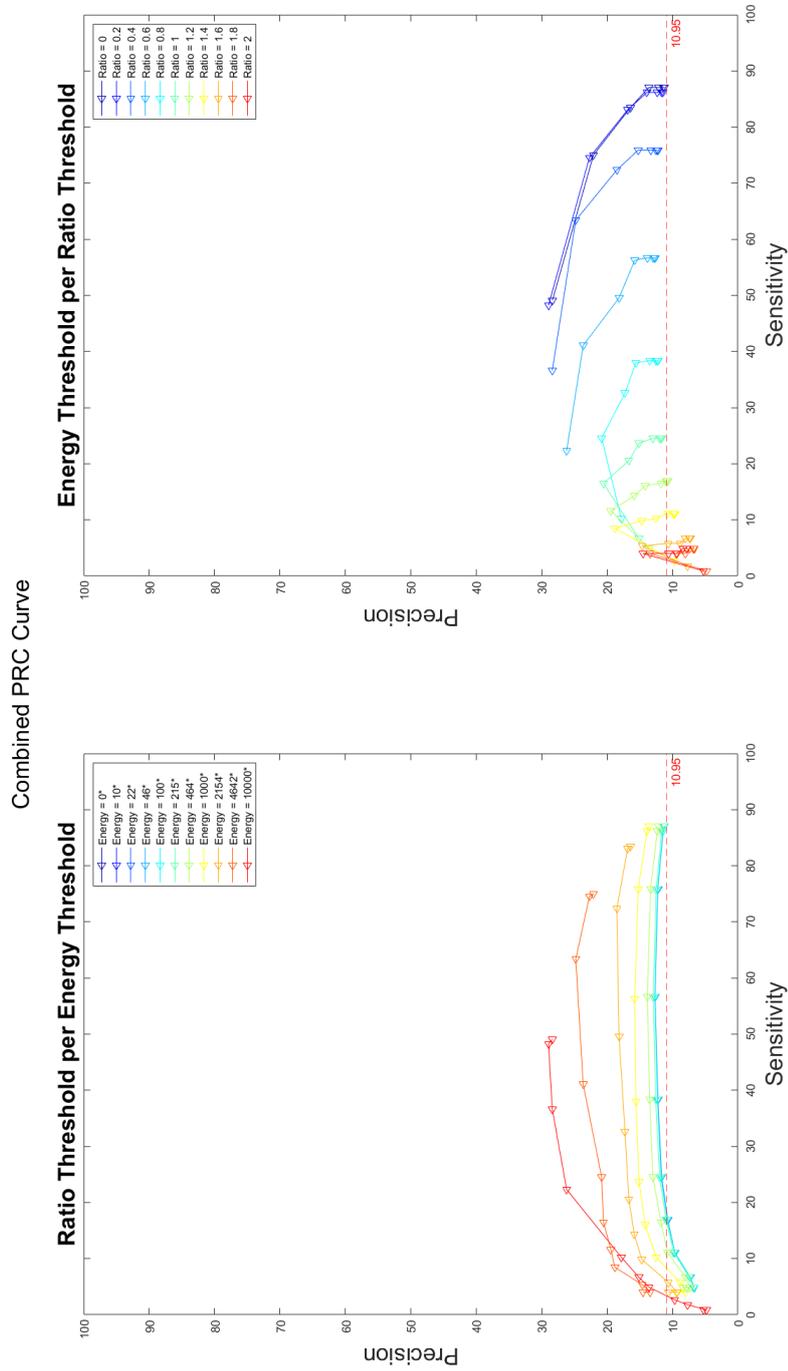


Figure D.4: Solution Space for Referenced Energy Feature of Validation Set

Bibliography

- [1] H. J. Niemarkt, P. Andriessen, J. W. Pasman, H. J. Vles, L. J. Zimmermann, and S. B. Oetomo, "Analyzing EEG maturation in preterm infants : The value of a quantitative approach," *Journal of Neonatal-Perinatal Medicine 1*, vol. 1, pp. 131–144, 2008.
- [2] H. C. Glass, A. T. Costarino, S. A. Stayer, C. M. Brett, F. Cladis, and P. J. Davis, "Outcomes for Extremely Premature Infants," *Anesthesia & Analgesia*, vol. 120, no. 6, pp. 1337–1351, jun 2015.
- [3] World Health Organisation, "Preterm birth," 2017. [Online]. Available: www.who.int/mediacentre/factsheets/fs363/en/ [Accessed: 2017-12-07]
- [4] A. Waelput and H. M., "Sterfte naar zwangerschapsduur in 2016," 2018. [Online]. Available: <https://www.volksgezondheidszorg.info/onderwerp/vroeggeboorte-en-laag-geboortegewicht/cijfers-context/sterfte#!node-sterfte-naar-zwangerschapsduur> [Accessed: 2018-08-16]
- [5] W. Engle, "Age Terminology During the Perinatal Period," *PEDIATRICS*, vol. 114, no. 5, pp. 1362–1364, nov 2004. [Online]. Available: <http://pediatrics.aappublications.org/cgi/doi/10.1542/peds.2004-1915>
- [6] J. Stiles and T. L. Jernigan, "The basics of brain development," *Neuropsychology Review*, vol. 20, no. 4, pp. 327–348, 2010.
- [7] M. Teplan, "Fundamentals of EEG measurement," *Measurement Science Review*, vol. 2, no. 2, pp. 1–11, 2002.
- [8] M. J. Aminoff, *Aminoff's Electrodiagnosis in Clinical Neurology*, 6th ed. Saunders, 2012.
- [9] S. J. Counsell, "Magnetic resonance imaging of preterm brain injury," *Archives of Disease in Childhood - Fetal and Neonatal Edition*, vol. 88, no. 4, pp. 269F–274, jul 2003. [Online]. Available: <http://fn.bmj.com/cgi/doi/10.1136/fn.88.4.F269>
- [10] M. Woo, "Planets Abound: Caltech-led astronomers estimate that at least 100 billion planets populate the galaxy," 2013. [Online]. Available: <http://www.caltech.edu/content/planets-abound> [Accessed: 2018-09-13]
- [11] J. M. O'Toole, G. B. Boylan, S. Vanhatalo, and N. J. Stevenson, "Estimating functional brain maturity in very and extremely preterm neonates using automated analysis of the electroencephalogram," *Clinical Neurophysiology*, vol. 127, no. 8, pp. 2910–2918, 2016. [Online]. Available: <http://dx.doi.org/10.1016/j.clinph.2016.02.024>
- [12] MindEdge, "Robomageddon: The Skills for the Future Study," 2018. [Online]. Available: <https://mindedge.com/page/future-skills-2018> [Accessed: 25-06-2019]
- [13] F. Cabitza, I. Ortopedico, R. Rasoini, C. S. Medicina, and G. F. Gensini, "Unintended Consequences of Machine Learning in Medicine," *JAMA*, vol. 318, no. 6, pp. 517–518, 2017.
- [14] K. Siau and W. Wang, "Building Trust in Artificial Intelligence , Machine Learning , and Robotics," *Cutter Business Technology Journal*, vol. 31, no. 2, pp. 47–53, 2018.
- [15] E. St. Louis, L. Frey, J. Britton, J. Hopp, P. Korb, M. Koubeissi, W. Lievens, and E. Pestana-Knight, *Electroencephalography (EEG): An Introductory Text and Atlas of Normal and Abnormal Findings in Adults, Children, and Infants*. American Epilepsy Society, 2016. [Online]. Available: <https://www.ncbi.nlm.nih.gov/books/NBK390354/>
- [16] N. V. Thakor, "Biopotentials and electro-physiology measurement," *The measurement, Instrumentation, and Sensors Handbook*, pp. 71–74, 1999.

- [17] G. H. Klem, H. Otto Lüders, H. H. Jasper, and C. Elger, "The ten twenty electrode system of the international federation," *Electroencephalography and clinical neurophysiology*, vol. 10, no. 2, pp. 370–375, 1958.
- [18] W. Klonowski, "Everything you wanted to ask about EEG but were afraid to get the right answer," *Non-linear Biomedical Physics*, vol. 3, pp. 1–5, 2009.
- [19] S. Vanhatalo, J. Voipio, and K. Kaila, "Full-band EEG (FbEEG): An emerging standard in electroencephalography," *Clinical Neurophysiology*, vol. 116, no. 1, pp. 1–8, 2005.
- [20] M. S. Scher, M. Sun, D. A. Steppe, R. D. Guthrie, and R. J. Scwabassi, "Comparisons of EEG spectral and correlation measures between healthy term and preterm infants." *Pediatric neurology*, vol. 10, no. 2, pp. 104–8, mar 1994. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/8024658>
- [21] E. Pavlidis, R. O. Lloyd, S. Mathieson, and G. B. Boylan, "A review of important electroencephalogram features for the assessment of brain maturation in premature infants," *Acta Paediatrica, International Journal of Paediatrics*, vol. 106, no. 9, pp. 1394–1408, 2017.
- [22] M. S. Scher, D. A. Steppe, and D. L. Banks, "Prediction of Lower Developmental Performances of Healthy Neonates by Neonatal EEG-Sleep Measures," *Pediatric Neurology*, vol. 14, no. 2, pp. 137–144, 1996.
- [23] M. Tolonen, J. M. Palva, S. Andersson, and S. Vanhatalo, "Development of the spontaneous activity transients and ongoing cortical activity in human preterm babies," *Neuroscience*, vol. 145, no. 3, pp. 997–1006, 2007.
- [24] L. C. Katz and C. J. Shatz, "Synaptic activity and the construction of cortical circuits." *Science (New York, N.Y.)*, vol. 274, no. 5290, pp. 1133–8, nov 1996. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/8895456>
- [25] K. Whitehead, R. Pressler, and L. Fabrizi, "Characteristics and clinical significance of delta brushes in the EEG of premature infants," *Clinical Neurophysiology Practice*, vol. 2, pp. 12–18, 2016. [Online]. Available: <https://doi.org/10.1016/j.cnp.2016.11.002>
- [26] J. M. O. Toole and G. B. Boylan, "NEURAL: quantitative features for newborn EEG using Matlab," 2017. [Online]. Available: <http://arxiv.org/abs/1704.05694>
- [27] B. Girod, R. Rabenstein, and A. Stenger, *Signals and Systems*. John Wiley & Sons, LTD, 2001.
- [28] J. Blanton and L. Euler, *Introduction to Analysis of the Infinite*. Springer New York, 1988.
- [29] A. E. Hramov, A. A. Koronovskii, V. A. Makarov, A. N. Pavlov, and E. Sitnikova, *Wavelets in Neuroscience*, ser. Springer Series in Synergetics. Berlin, Heidelberg: Springer Berlin Heidelberg, 2015. [Online]. Available: <http://link.springer.com/10.1007/978-3-662-43850-3>
- [30] F. Kaiser, "On a simple algorithm to calculate the 'energy' of a signal," *International Conference on Acoustics, Speech, and Signal Processing*, vol. 1, pp. 381–384, 1990.
- [31] F. Hlawatsch, T. G. Manickam, R. L. Urbank, and W. Jones, "Smoothed pseudo-Wigner distribution, Choi-Williams distribution, and cone-kernel representation: Ambiguity-domain analysis and experimental comparison," *Signal Processing*, vol. 43, no. 2, pp. 149–168, 1995.
- [32] F. Hlawatsch and G. F. Boudreaux-Bartels, "Linear and quadratic time-frequency signal representations," *IEEE Signal Processing Magazine*, vol. 9, no. 2, pp. 21–67, 1992.
- [33] S. Mallat, *A Wavelet Tour of Signal Processing*, 3rd ed. Oxford: Elsevier Inc., 2009.
- [34] G. Kaiser, *A Friendly Guide to Wavelets*. Boston: Birkhäuser Boston, 2011. [Online]. Available: <http://link.springer.com/10.1007/978-0-8176-8111-1>
- [35] I. Daubechies, *Ten Lectures on Wavelets*. Society for Industrial and Applied Mathematics, jan 1992. [Online]. Available: <http://epubs.siam.org/doi/book/10.1137/1.9781611970104>

- [36] A. Aldroubi and M. Unser, *Wavelets in Medicine and Biology*, 1st ed. Boca Raton FL, USA: CRC Press, 1996.
- [37] M. Misiti, Y. Misiti, G. Oppenheim, and J.-M. Poggi, *Wavelet Toolbox™ Getting Started Guide*. Natick: The Mathworks, Inc, 2018. [Online]. Available: https://nl.mathworks.com/help/pdf/doc/wavelet/wavelet_gs.pdf
- [38] P. S. Addison, "Introduction to redundancy rules: the continuous wavelet transform comes of age," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 376, no. 2126, p. 20170258, aug 2018. [Online]. Available: <http://rsta.royalsocietypublishing.org/lookup/doi/10.1098/rsta.2017.0258>
- [39] D. Alvarez-Estevéz, "European Data Format." [Online]. Available: <https://www.edfplus.info/> [Accessed: 17-12-18]
- [40] B. Zonjy, "Online EDF Viewer," 2017. [Online]. Available: <https://bilalzonjy.github.io/EDFViewer/EDFViewer.html> [Accessed: 17-12-18]
- [41] A. Delorme and S. Makeig, "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *Journal of Neuroscience Methods*, vol. 134, pp. 9–21, 2004.
- [42] J. Nitschke, G. A. Miller, and E. W. C. Iii, "Digital filtering in EEG/ERP analysis : Some technical and empirical comparisons," vol. 30, no. 1, pp. 54–67, 1998.
- [43] MathWorks, "FIR Filter Design." [Online]. Available: <https://nl.mathworks.com/help/signal/ug/fir-filter-design.html> [Accessed: 2019-08-13]
- [44] E. W. Cook and G. A. Miller, "Digital filtering: background and tutorial for psychophysicologists." *Psychophysiology*, vol. 29, no. 3, pp. 350–67, may 1992. [Online]. Available: <http://www.ncbi.nlm.nih.gov/pubmed/1626044>
- [45] MathWorks, "Design Digital Filters." [Online]. Available: <https://nl.mathworks.com/help/signal/ref/designfilt.html> [Accessed: 2019-08-13]
- [46] M. P. Wachowiak, R. Wachowiak-Smolíková, M. J. Johnson, D. C. Hay, K. E. Power, and F. M. Williams-Bell, "Quantitative feature analysis of continuous analytic wavelet transforms of electrocardiography and electromyography," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 376, no. 2126, p. 20170250, aug 2018. [Online]. Available: <http://rsta.royalsocietypublishing.org/lookup/doi/10.1098/rsta.2017.0250>
- [47] Mathworks, "MATLAB Runtime." [Online]. Available: <https://nl.mathworks.com/products/compiler/matlab-runtime.html> [Accessed: 2019-09-24]
- [48] J. J. Newson, "The Remarkable Inconsistency of EEG Frequency Band Definitions," 2018. [Online]. Available: <https://sapienlabs.co/the-remarkable-inconsistency-of-frequency-band-definitions/> [Accessed: 2019-04-23]
- [49] R. M. Rangayyan, *Biomedical Signal Analysis*, 2nd ed. Hoboken, New Jersey: John Wiley & Sons, Inc., 2015.
- [50] D. Tax, "One-class classification," Ph.D. dissertation, Delft University of Technology, 2001.
- [51] D. Dimitriadis, A. Potamianos, and P. Maragos, "A Comparison of the Squared Energy and Teager-Kaiser Operators for Short-Term Energy Estimation in Additive Noise," vol. 57, no. 7, pp. 2569–2581, 2009.
- [52] J. M. O. Toole, A. Temko, and N. Stevenson, "Assessing instantaneous energy in the EEG : a non-negative , frequency-weighted energy operator," *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 2, no. 3, pp. 3288–3291, 2014.
- [53] T. Saito and M. Rehmsmeier, "The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets," *PLoS ONE*, vol. 10, no. 3, 2015. [Online]. Available: <http://dx.doi.org/10.1371/journal.pone.0118432>

-
- [54] R. Kumar, "Receiver Operating Characteristic (ROC) Curve for Medical Researchers," *Indian Pediatrics*, vol. 48, pp. 277–287, 2011.
- [55] S. Narkhede, "Understanding AUC - ROC Curve," 2018. [Online]. Available: <https://towardsdatascience.com/understanding-auc-roc-curve-68b2303cc9c5> [Accessed: 2-10-19]
- [56] S. McCann, "Average precision," 2011. [Online]. Available: <https://sanchom.wordpress.com/tag/average-precision/> [Accessed: 2-10-19]
- [57] J. M. O. Toole, G. B. Boylan, R. O. Lloyd, R. M. Goulding, S. Vanhatalo, and N. J. Stevenson, "Detecting bursts in the EEG of very and extremely premature infants using a multi-feature approach," *Medical Engineering and Physics*, vol. 45, pp. 42–50, 2017. [Online]. Available: <http://dx.doi.org/10.1016/j.medengphy.2017.04.003>