UCL - University College London

Department of Statistical Sciences

Modelling neonatal electroencephalogram time series

A thesis submitted for the degree of Doctor of Philosophy

Author:
Simon Wallace

Supervisors:
Prof. Sofia C. Olhede
Prof. Maria Fitzgerald
Dr. Lorenzo Fabrizi

I, Simon Wallace, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signed: _____

Abstract

Creating a model for brain activity is a highly complex task; this is especially true in modelling neonatal electroencephalogram (EEG) signals. Whereas previous work is motivated by improving seizure detection, this research focuses on describing the development of these complicated multivariate signals. Using data collected from inpatients at University College London Hospital at different degrees of prematurity, we propose a model for background and somatosensory response neonatal EEG signals and subsequently make inferences about the observed EEG signals using this model.

We construct a univariate model for neonatal EEG by analysing the second order properties of these signals, taking into account time segments which have time-heterogeneous second order properties. To do so we utilise time, frequency and time-frequency domain methods. The presented univariate model is combined with a time domain correlation structure to generate a multivariate representation which is possible, in part, due to the resolution of the data. Furthermore, the parameters and signal components are best described by taking into account not only the age at which testing occurred, but also the age at which an infant was born. This research has attempted to create a model that is not only descriptive of somatosensory responses, but also applicable in other avenues of similar research.

We propose to use generalised linear models to describe the age dependence of the observed time series, and use these models to simulate EEG observations. When modelling characteristics of the estimated parameters, all models require the age pairing - age at birth and age at test - as variables. Combined with an appropriate time domain correlation structure, this allows us to achieve suitable estimates of observed signal structure.

The model class presented is a flexible and accurate representation of neonatal background and somatosensory response electroencephalogram signals, and can be used to describe similar multivariate observations.

Table of Contents

Acknowledgements

First I would like to thank my supervisors for their valuable guidance throughout this research; Lorenzo for helping me understand EEG, Maria for asking me the difficult questions, and Sofia for keeping me on track as well as pointing me in the right directions. Even when it looked like all was lost, Sofia kept me calm and guided me through fixing what seemed unfixable. I am eternally grateful for giving me the opportunity to undertake this amazing journey.

To my friends who listened to me talk about Statistics, proofread this thesis and helped me relax when needed. I can't promise that I still won't talk Statistics, but I won't ask you to look over any formulae again.

To Alex, who found this Ph.D. position and encouraged me to apply for it, you are amazing. You helped keep me sane and reminded me to eat, clean up the masses of Pepsi Max bottles that surrounded me when coding, and stop every now and then to relax.

Finally, to my family for their unwavering love and support. My parents have always made sure that I have had the ability to take advantage of every available opportunity. I would not be the person I am if they didn't push and support me in everything I do. I am exceptionally lucky to have such wonderful parents who see the value in education.

I couldn't have done this if it wasn't for you all.

"Aim for brevity while avoiding jargon."

E. W. Dijkstra

List of Abbreviations, Notations and Terms

Abbreviations

ADC	*Analogue to digital conversion*
ADP	*Automated detection procedure*
dB	*Decibel*
EEG	*Electroencephalography*
FDR	*False discovery rate*
FFT	*Fast Fourier transform*
GLM	*Generalised linear model*
Hz	*Hertz*
PMA	*Post Menstrual Age in weeks at birth*
PMAT	*Post Menstrual Age in weeks at test*
SDF	*Spectral density function*

Notation

α	*Significance level*
Δt	*Sampling rate of a continuous stochastic process*
δ	*Smoothness parameter determining the self similarity of a process*

γ	*Euler-Mascheroni constant*
$\Gamma(\cdot)$	*Gamma function*
$\hat{p}_{\mathrm{B}}$	*Probability of the binary variable B being equal to 1,* $P(B_1)$
$\hat{p}_{\mathrm{M}}$	*Probability of the binary variable M being equal to 1,* $P(M_1)$
$\hat{p}_{\mathrm{S}}$	*Probability of the binary variable S being equal to 1,* $P(S_1)$
$\hat{S}_X^{(d)}(f)$	*Direct spectral estimate of a stochastic process* X
$\hat{S}_X^{(p)}(f)$	*Periodogram of a stochastic process* X
$\{\underline{\varphi}, \underline{\vartheta}\}$	*Autogressive moving average process process parameters*
$\mathrm{B}_{\{0,1\}}$	*Binary variable indicating whether* $\hat{\sigma} \geq 0.0007$
$\mathrm{M}_{\{0,1\}}$	*Binary variable indicating whether* $\hat{\delta} > 1.5$ *or* $\hat{d} \neq 0$
$\mathrm{MCA_i}$	*Principal component i of PMA and PMAT to avoid multicollinearity*
R_{DB}	*Recordings contain delta brush response to noxious or tactile stimuli*
R_{PC}	*Recordings contain specific response to noxious or tactile stimuli*
$\mathrm{S}_{\{0,1\}}$	*Binary variable indicating the shape of estimated spectral density*
μ	*Mean of a stochastic process*
σ^2	*Variance of a process*
ε	*Random noise component*
d	*Range parameter in Matérn process*
$G(f)$	*Gain function of a linear time invariant filter*
$H(f)$	*Transfer function of a linear time invariant filter*
h_t	*Taper*
$K_\nu(\cdot)$	*Modified Bessel function of the second kind of order* ν
$l_W(\theta)$	*Whittle likelihood*

n_{DB}^{N} *Number of electrodes expressing delta brush occurrence in response to noxious stimuli*

n_{DB}^{T} *Number of electrodes expressing delta brush occurrence in response to tactile stimuli*

n_{DB}^{SP} *Number of electrodes expressing spontaneous delta brush occurrence*

n_G *Number of distinct electrode groupings in a set of recordings*

$S_X(f)$ *Spectral density function of a stochastic process X*

$S_X(f;\theta)$ *Parametric spectral density function*

X_t *Stochastic process X at time t*

p_{DB}^{N} *Probability of delta brush occurrence in response to noxious stimuli*

p_{DB}^{T} *Probability of delta brush occurrence in response to tactile stimuli*

p_{DB}^{SP} *Probability of spontaneous delta brush occurrence*

ARMA(p,q) *Autogressive moving average process of order p,q*

fBm(δ) *Fractional Brownian motion process with smoothness parameter δ*

fBmARMA(δ,p,q) *Convolution of a Fractional Brownian motion process with smoothness parameter δ and ARMA process with order p, q*

Matérn(δ,d) *Matérn process with smoothness parameter δ*

MatérnARMA(δ,d,p,q) *Convolution of a Matérn process with smoothness parameter δ and ARMA process with order p, q*

PMA *Variable used to refer to estimated post menstrual age at birth*

PMAT *Variable used to refer to post menstrual age at test*

Statistical Terminology

Aliasing - *The effect of sampling a continuous stochastic process on the frequency domain representation*

Multicollinearity - *Where predictor variables in a regression are highly correlated*

Stationarity - *A finite stochastic process whose distribution is dependent solely upon time differences*

Time-heterogeneous process - *A stochastic process whose properties change with time*

Time-homogeneous process - *A stochastic process whose properties do not change with time*

Weak stationarity - *A finite stochastic process where the mean and covariance of the distribution is dependent solely upon time differences*

EEG and Medical Terminology

Background EEG - *Baseline level of activity seen in an awake patient*

Gestational age - *Time elapsed since the first day of the last menstrual period and birth*

Neonatal - *Relating to newborn infants*

NMDA - *N-methyl-D-aspartate*

Nociceptive - *Pain as a result of stimulation of nerve cells*

Post menstrual age - *Gestational age plus time since birth*

Premature - *Infant born before 37 weeks gestational age*

Somatosensory - *Relating to a sensory stimulation*

List of Tables

List of Figures

Chapter 1

Introduction

The work in this thesis is a contribution to applied time series analysis. In particular methods are developed to describe electrical activity in infant electroencephalography (EEG). EEG observations are modelled in terms of frequency domain representations so that brain activity can be associated with different temporal scales. The challenge of applying time series methods to the data sets that will be studied in this thesis is that, whilst adult EEG observations are very well understood, we shall model and estimate representations of infant EEG signals. One of the main reasons this is difficult is that infant EEG is extremely heterogeneous. In adult EEG the signals analysed are the result of averaging several electrodes around the site of interest. Whereas in infants this is not possible due to the fact that the infant brain is significantly different from the adult brain [4,5]. This thesis will focus on the analysis of background EEG signals. The main purpose of such models is to be able to discover abnormal, or stimulation associated activity. In essence, it helps us build a null model understanding of what an EEG signal would show if nothing interesting was happening.

Activity in the brain is the result of the polarisation and depolarisation of billions of neurons instigated by an influx of ions across the cell membrane [1–3]. This results in voltage fluctuations which are measureable using electroencephalography. First implemented in the early 20^{th} century, electroencephalography (EEG) records the changing electrical activity by recording the output of multiple electrodes; the resultant signals form a multivariate time series. These are inherently time-heterogeneous over long epochs, however they contain time-homogeneous segments [1–3].

To build a good model of background EEG signals we need to develop methods of describing time series observations that are not changing significantly in their generation across time. Such behaviour will be called "time-homogenous". In contrast, time observations that are the consequence of stimulation or temporally localised activity will be time-heterogeneous by their very nature of generation. The challenging statistics problems is to propose models that can reproduce both time-homogenous and time-heterogeneous structure so that we can assess how unlikely or abnormal certain temporal characteristics are. To focus our efforts, we shall answer the following questions:

1. Can we describe six seconds of background EEG using a time-homogeneous model?
2. Can we adapt such a model to incorporate time-heterogeneous signal components?
3. Are the properties of neonatal EEG dependent upon the age of the infant?
4. Given the time-heterogeneous nature of EEG signals, can we identify when the baseline activity changes?
5. Do previous modelling techniques concur with the presented model?

All the presented concepts in this thesis are given to enable answers to these questions to be obtained.

This and the subsequent Chapter are designed to provide an understanding of the motivation and methods of this research. More priority is given to describing the statistical concepts employed, however the electroencephalogram background presented is enough detail to understand the outlined material.

1.1 Previous Research

The aim of this research is to describe, model and simulate neonatal EEG signals focusing upon how the signals change across development. We do so by analysing signals recorded from inpatients at University College London Hospital, by the Department of Neuroscience, Physiology and Pharmacology at University College London [6].

The developments and breakthroughs in measuring activity allow us to study and discover more about the brain, especially in the weeks after birth. Premature births are increasing globally [7, 8], and as a result of medical advancement in the field of paediatrics, the survival rate of premature (neonatal) infants has also increased [7, 8]. This has enabled research into the early stages of brain activity to be undertaken, and has been vital in improving the diagnosis and treatment of neonatal infants [4, 9, 10]. The effect that premature birth has on the development of the infant, and the consequences of invasive intensive care procedures on the development of the central nervous system, are critical problems warranting further investigation.

Premature birth has been shown to correlate with long-term neurological problems [11–13], as well as more immediate issues such as increased intercranial pressure [14]. Additionally, infants born prior to 33 weeks gestation have been shown to have decreased brain volume later in life [15] as well as learning and developmental issues [16, 17]. The survival rates for neonatal infants (neonates) on the brink of viability has increased, however they are at risk of complications. These complications are not only present whilst under observation, but persist throughout life [13, 18]. It has been shown that the procedures undertaken whilst in neonatal intensive care can affect the function and structure of the brain [19].

The functional circuits of the brain form in the first stages of pregnancy [1, 2]; however, the final trimester is a crucial stage in the development of the neonatal brain [20–22]. Whilst the brain and spinal cord are developed in the first six weeks of gestation [23], it is the final trimester where the crucial stages of development occur [20, 21]. In this period, the weight of the brain triples and the cerebellum increases by a factor of 30 [20, 21]. Any premature brain damage is more than just loss of tissue - neurological issues are increasingly recognised to be trans-synaptic events - in which damage in one part of the brain affects another region that is synaptically connected [21]. It has been illustrated within developing neural circuits that if disruption occurs in one element, another will generate similar activity [22]. Brain activity in the weeks following birth is crucial and affects subsequent growth of the brain [15], and has long-term effects on future activity [11, 12, 23, 24].

The main focus of neonatal EEG modelling has been on seizure detection, with multiple approaches outlined [25–46] and assessed [47,48]. Whilst an important area of research, it has been made clear in recent years that a vital area has been overlooked: pain response development. Whereas oral sucrose is seen as sufficient management for neonatal pain [49–51], it has only recently been discovered that this is an unsuitable analgesic; although the behavioural effects of painful stimuli are reduced, the response within the brain is not [5].

Neonates can undergo many procedures in the weeks after birth, which can illicit somatosensory responses such as the nociceptive (pain) response [24]. The behavioural response to these procedures was shown to change if an infant was born prematurely [52]. Infants born prematurely expressed less mature behavioural responses than those tested within four days of birth at full term [52]. This illustrates the possible effect of prematurity and invasive procedures in the weeks after birth. Additionally, the time frame immediately pre or post birth has been shown to be a critical period, during which the underlying neuronal circuitry is vulnerable to long-term neurological development problems [11, 53–56]

If appropriate pain management techniques are not administered, the consequences could affect an infants development and cause issues with pain response in later life [57–59]. Based upon parental perception, it has been reported that at 18 months ex-preterm neonates were less sensitive to pain than ex-full term infants [11, 60]. Specifically it has been proposed that exposure to repetitive pain causes excessive NMDA/excitatory amino acid activation causing excitotoxic damage to the developing neurons [11]. Epidemiological studies have correlated neonatal complications with abnormal adult behaviour and in mammalian subjects, alterations in the adult brain have been correlated with such complications [11]. Combined with the behavioural experiments previously undertaken, this portrays a concerning pattern with regards to the effect of painful stimuli evoked in neonatal infants.

It has been shown that the response to somatosensory stimuli, such as touch or pain, changes as the infant matures [4]. During brain development, specifically during the formation of functional circuits, it has been shown that there is a transient sponta-

neous neuronal bursting activity, known as a delta brush [61, 62]. Present in infant responses prior to 35-37 weeks post menstrual ages, as well as in background EEG, this activity is defined as a time-varying or time-heterogeneous process with simultaneous activity in the frequency ranges [0.5,1.5]Hz and [8,25]Hz [4, 46, 62, 63]. These delta brushes are an important component of neonatal EEG that have been majoritively overlooked, but are necessary to include in signal analysis [46]. EEG signals are inherently time-heterogeneous multivariate processes, due to the changing brain waves present at different times [1–3]. The aim of this research is to analyse the development of neonatal EEG signals and we shall do this by utilising time, frequency and time-frequency domain techniques with primary focus on the frequency domain.

1.2 EEG characteristics and analysis methods

We are motivated to use the frequency and time-frequency domains by the composition of EEG signals and the corresponding brain waves [1–3]. Brain waves are classified by the frequency bands that the wave is in, and correspond to different brain states. Figure 1.1 demonstrates such waves, and provides the corresponding frequency bands [1–3]

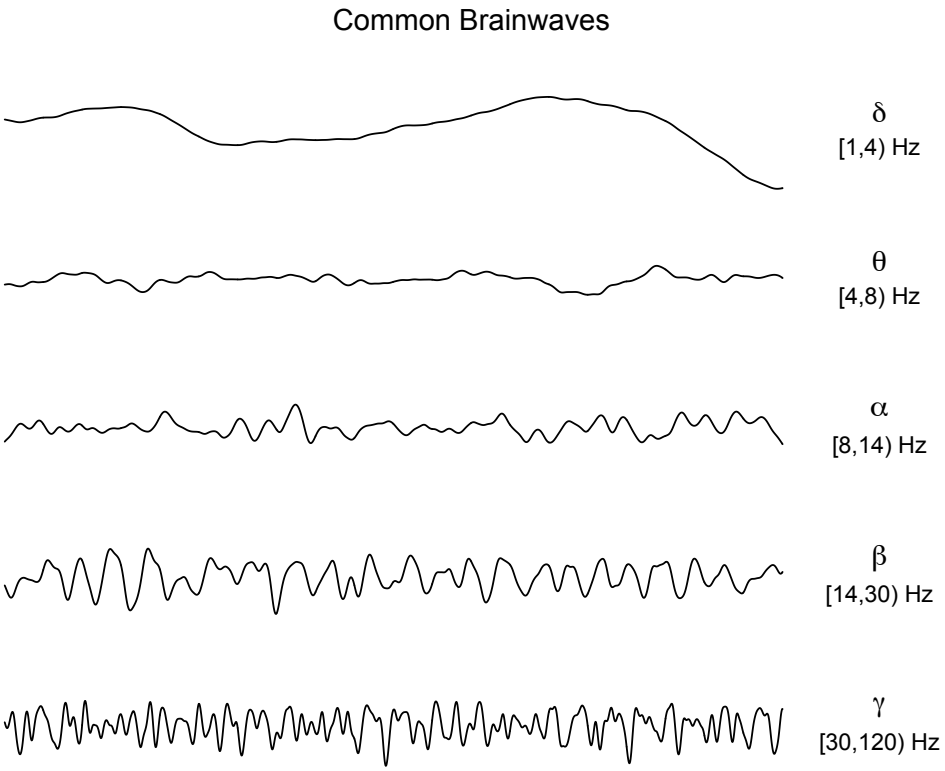
Utilising the frequency domain enables us to describe the signal's covariance structure, taking into account the waves present. The use of the frequency domain to analyse EEG is a standard method of analysis [46, 64, 65]. The frequency domain describes a signal's covariance structure by describing its spectrum [66, 67]. However, EEG signals are time-heterogeneous, and thus require that we describe how the covariance changes over time [1–3]. This can be done using the time-frequency domain [68]. EEG signals can be viewed and analysed as multiple time-homogeneous segments (quasi-stationary) [1–3, 61, 62]. We shall identify these segments by utilising the time-frequency domain as the frequency domain is best utilised if analysing time-homogeneous segments. Utilising frequency domain analysis on time-heterogeneous signals describes the frequency composition, however we lose information with regards to how the structure changes over time.

When recording EEG, electrodes are usually placed according to the International 10-20 system and the analogue signal recorded; these signals are stored digitally and therefore

must undergo conversion from analogue to digital format [1, 2]. This conversion procedure defines the resolution of the signals, according to how many times per second the signal is sampled, i.e. a 2000Hz signal is sampled every 0.0005s or 2000 times a second. This sampling has an aliasing effect on the signal and as such anti-aliasing filters are often applied [1,2,69]. Additionally, other filters are applied to "remove" unwanted components - or artefacts - from the signal during this conversion [1–3]. A high-pass filter would remove artefacts such as those from movement, whereas a low-pass filter would remove artefacts such as breathing [1–3, 66, 67]. These signals are stored digitally once referenced against a designated electrode, such as FCz in the data set presented [4, 6].

Electrodes are named according to the position of the electrode on the scalp. The letters refer to the frontal, temporal, occipital and parietal lobes; the C stands for central, and

Figure 1.1: Common brainwaves and their corresponding frequency bands



This figure provides illustrative example of the five common brainwaves and gives the corresponding frequency bands that define these brainwaves.

is solely for reference. Odd numbers are on the left hemisphere, even numbers on the right hemisphere and Z refers to electrodes placed on the midline [1–3].

It has been shown that in response to noxious and tactile stimuli, infants are more likely to express a delta brush response prior to 35-37 weeks post menstrual age at test [4]. The effect of such responses to stimuli is concerning, considering the long-term neurological and behavioural effects of pain in neonates [12, 52, 58, 60].

The maturational stages of neonatal EEG are well classified [61, 62]. These classifications illustrate that the delta brush activity is not only a response to somatosensory stimuli, but can also occur spontaneously in background EEG recording. Background EEG is a classification of EEG signal segments that refers to the baseline electrical activity seen in an awake patient [1–3]. With neonates however, background EEG changes across development with spontaneous delta brushes decreasing in frequency at around 37 weeks post menstrual age [61]. We can identify possible time-homogeneous segments within background signals by investigating the segments between delta brushes, known as inter-burst intervals [61, 62].

The time heterogeneity of background EEG signals has been modelled and recreated utilising fractional Brownian motion, creating time-heterogeneous signals by appending simulated segments with different parameters whilst ignoring the contribution of delta brushes to the time heterogeneity of the signals [70]. The time-heterogeneous nature of neonatal EEG has not been accurately identified, with no understanding of how to recreate the time-heterogeneous nature of the signal's parameters. Given that we know EEG signals are time-heterogeneous, it would be interesting to find the maximum/minimum lengths of the component time-homogeneous segments.

As mentioned, the majority of neonatal EEG modelling has focused upon seizure detection and has proved crucial in developing early detection procedures. The multiple possible models proposed have not been used clinically, due in part to the large variability in the infant [48]. Many analysis techniques have been proposed such as: Duffing oscillators [70], linear time invariant filters [33], Kalman filters [71], adapted probabilistic modelling [41, 44], autoregressive modelling [72] and causal modelling [73]. There is

little agreement in modelling neonatal EEG signals with a diverse and eclectic range of models employed. Many proposed models have placed importance on the signal to be detected and the baseline activity is a secondary concern.

With such a heavy focus upon creating a neonatal seizure detection procedure, it is clear that the modelling of neonatal background EEG is an important task and clinical problem. An accurate neonatal background EEG model has wide ranging applications to neonatal brain research, such as the development of cortical activity in response to somatosensory stimulation or as a baseline to be used in abnormal component detection.

The multivariate nature of EEG has been described previously using the multivariate autoregressive process [74]. However, little has been done when simulating the observed correlation between electrodes, with simulation methods presented not addressing this structure [71, 75]. Recordings obtained from surface electrodes are affected by the attenuation of the different layers in the head, resulting in the amalgamation of signals, and as such an artificial correlation structure [1–3]. Although it has been shown that neonatal EEG is very focal by nature due to the high conductivity of skull tissue [76], the spatial relationship of neonatal EEG is of analytical interest with recommendations to use a high number of scalp electrodes to adequately capture this relationship [77].

When utilising frequency domain analysis methods, the aim is to describe a signal's covariance structure by its spectrum, which can contain long range and/or short range dependence. The spectrum is a description of how the variance of a signal is distributed over different frequencies. Therefore by modelling the spectrum of a signal, we describe the second order properties.

Short range dependence (short memory) in a signal is where the immediate time points have an effect on the covariance structure, and long range dependence (long memory) is where more than the immediate time points have an effect on the signal. By analysing the auto covariance sequence, which is the covariance of the signal with itself at different pairs of time points, we can identify the type of dependence present. Long range dependence illustrates slow decay to zero in this sequence, whereas short range dependence illustrates fast decay. Therefore when utilising frequency domain methods, we aim to

describe the structure of the covariance sequence utilising models which describe the type of dependence present.

1.3 Structure and contribution of thesis

From six-second recording of neonatal EEG we explore these questions; and determine partial answers for the unique data set provided by UCLH. [6]. We shall build, estimate, analyse and simulate from a model constructed to explain the underlying observed signal characteristics.

In Chapter 2, we describe the methods that we shall employ stating relevant theorems and providing illustrative examples to enable an understanding of the statistical methods employed throughout this thesis. We then define our model in Chapter 3, refining it from observations made from the data and obtaining parameter estimates. Chapter 4 investigates the characteristics of these estimates, in an attempt to produce relevant parameter distributions; from which accurate realisations could be simulated. Having performed the analyses univariately up to this point, in Chapter 5 we analyse the multivariate characteristics of the signals in the time domain. Utilising this multivariate representation we simulate signals from our model, and compare the time domain characteristics against the observed data and competing models. Following this we summarise our conclusions before outlining possible avenues of future research, ways in which this research could be extended and the limitations of this research.

The modelling and description of the activity within the brain is an ongoing task for researchers, and will be for the foreseeable future. As such this thesis whilst applying statistical methods to best describe the data set, and building upon previous work, will hopefully serve as starting point for the development of the observed features.

This research has made several contributions to the understanding of this field. First we have illustrated short and long range dependence in the covariance structure of EEG recorded at high resolutions. Additionally, for EEG signals of up to six seconds in length, we can utilise a time-homogeneous parameter model, i.e. the same parameters across the signal segment. This parameterisation can be used even if time-heterogeneous

components are present. Furthermore we have shown that we can extend our model to incorporate time-heterogeneous components such as the delta brush.

Extending upon previous research, we have found evidence to suggest that the pairing of age at birth and age at test is a better description of neonatal somatosensory stimuli response development than just age at test. Similarly, this research concludes that the signals recorded are less correlated in the time domain if obtained from a premature infant or tested close to birth. Whilst information is lost modelling univariately, we can recreate a set of EEG recordings by taking into account the time domain correlation structure of similar recordings thus obtaining a multivariate representation from univariate estimates.

Chapter 2

Methodology

In the introduction we provided the motivation for utilising the frequency and time-frequency domains to analyse our signals. Now we shall introduce and describe the statistical tools utilised in this thesis.

We shall model the EEG observations by describing the first two moments of the signal via the frequency domain. This modelling will take different forms if the underlying process is assumed to be stationary. We start by describing methods that can applied to signals with a similar second order structure to those observed. EEG signals can be analysed utilising a number of methods [1–3, 78], with advancements in the recording process and analytical software, further understanding of these signals can be obtained. One such method that has benefitted from this advancement is frequency analysis, which we shall outline in this chapter.

By analysing a signal in the frequency domain we can determine which frequencies are present within the signal. Considering the composition of EEG as an amalgamation of different brain waves, this provides an intuitive description that can describe the type of brain waves present within a signal [1–3, 78]. In order to utilise frequency domain methods to analyse a signal, we must first transform the signal from the time domain into the frequency domain. There are several transforms that can be utilised to achieve this task and we shall utilise the Fourier transform to obtain the frequency domain structure of the signals. The frequency domain allows us to describe the second order properties of a signal by describing the signal's spectrum, and it is of use to have a way

in which we can assess the suitability of a fitted spectrum. By describing a signal in terms of its spectrum, we are describing its covariance structure in the time domain; since our processes are Gaussian this is sufficient to describe the structure of the process.

The way in which a spectrum can be fitted falls into the two traditional estimation methods, parametric and non-parametric. Parametric estimates assume a finite number of parameters and a specified distribution, as opposed to non-parametric which does not assume any distribution and allows an infinite number of parameters.

EEG signals are highly complex signals whose properties inherently change over time. To best describe the spectrum of signals whose properties change over time we must utilise time-frequency domain methods. The time-frequency domain describes how a signal's spectrum changes across time intervals and can be used to identify periods where the characteristics of a signal remain constant. Periods where the activity changes are referred to as time-heterogeneous, and require different approaches to their time-homogeneous counterparts. Such segments are typically identified prior to analysis to allow appropriate analysis to be conducted.

EEG signals are usually filtered to remove unwanted components that can mask features within the signal [1–3]. The signals presented are filtered to remove unwanted artefacts such as breathing or limb movement. Furthermore, as we are analysing a segment of EEG signals they will contain a phenomenon known as aliasing, and as such a filter has been applied to limit the aliasing effect on the signal; Aliasing relates to the sampling of continuous signals and how this affects the frequency domain representation of the signal. The filtering that is applied in the preprocessing affects the spectrum of a signal as it acts as a convolution. As such we need to apply the filtering function to the function that we wish to estimate. Therefore taking into account the contribution of the filter on the signals structure.

The outlined approaches are utilised to analyse the presented data set in this thesis; from which we aim to obtain a description of the second order properties of six second neonatal EEG signals across a range of developmental ages. In order to assess the suitability of an estimated structure for these signals, we construct a procedure based

upon the properties of the signal in the frequency domain. Following this we investigate frequency domain methods by which we can estimate the structure of the signal, and the effect that the preprocessing of the signals has upon this structure. Finally we outline ways in which we can identify time-heterogeneous segments within our signals. Throughout this chapter we utilise simulated data to illustrate these concepts.

2.1 Fourier transform

The Fourier transform, is one way to decompose a signal from the time domain into the frequency domain [66, 67]. Similarly to the Fourier series, wave forms in the time domain are decomposed and represented by a series of sine and cosine waves. The Fourier transform results in a complex-valued signal, representing the amplitude and phase of the component frequencies [66, 67]. Definitions of the continuous and discrete Fourier transforms are given in Equations 2.1.1 and 2.1.2 respectively. The discrete Fourier transform is important as it is used in the fast Fourier transform (FFT). The FFT is one way in which the frequency domain representation of a signal can be obtained using analytical packages such as R or MATLAB. [66, 67, 79]:

$$\mathcal{X}(f) = \int_{-\infty}^{\infty} (X(t) - \mu) e^{-i2\pi ft} dt, \tag{2.1.1}$$

$$\mathcal{X}_k = \Delta t \sum_{n=0}^{N-1} (X_n - \mu) e^{-i2\pi kn/N} = \Delta t \sum_{n=0}^{N-1} (X_n - \mu) e^{-i2\pi f_k n \Delta t}, f_k \equiv \frac{k}{N\Delta t}, k = 0, \ldots, N-1. \tag{2.1.2}$$

We can see from these definitions that theoretically, the transform is performed over an infinite range. However, in practice this is impossible; as shown the FFT decomposes the n point signal, not an infinite length signal. This truncation has the same effect as convolving the infinite signal with a finite length rectangular window. As such, we need to know the effect that this has upon the frequency domain representation.

Consider two signals $X^{\circ}(t)$ and X_n, where X_n is an n point sample from the infinite length $X^{\circ}(t)$, sampled every Δt seconds/points. These signals have the frequency domain representations $\mathcal{X}^{\circ}(f)$ and $\mathcal{X}(f)$ respectively and the relationship between them is [66, 67]:

$$\mathcal{X}(f) = \sum_{k=-\infty}^{\infty} \mathcal{X}^{\circ}\left(f + \frac{k}{\Delta t}\right), \quad -\frac{1}{2\Delta t} \leq f \leq \frac{1}{2\Delta t}. \tag{2.1.3}$$

From this Equation 2.1.3, we see that the Fourier transform of X_n, at a frequency f, is the sum of the Fourier transform of $X(t)$ across the frequencies $f + \frac{k}{\Delta t}$. As a result, we cannot obtain $\mathcal{X}^{\circ}(f)$ from $\mathcal{X}(f)$ because frequencies outside the range $\left[-\frac{1}{2\Delta t}, \frac{1}{2\Delta t}\right]$ are folded into $\mathcal{X}(f)$. This is called aliasing and is an often overlooked aspect of frequency domain analysis [66] - Figure 2.3 shows how aliasing can affect the signal's structure.

In Equation 2.1.3 the bounding interval for f is known as the Nyquist interval. This interval, defined by the sampling rate of the signal, determines the frequencies over which a discrete signal can be analysed; in our dataset $\Delta t = 0.0005$ indicating a Nyquist interval of $[-1000, 1000]$Hz. The highest frequency with no aliases is defined as the Nyquist frequency, and is the point about which the folding of frequencies occurs, this is the bound of the Nyquist interval [66, 67].

We can utilise the Fourier transform only for a specific subset of signals; specifically stationary signals [66, 67]. We can however, use the Fourier transform for some non-stationary signals, provided certain conditions are satisfied. The reason behind this constriction of signals is as follows and we start with the spectral representation theorem [66, 67].

Theorem 2.1.1 (**Spectral Representation Theorem**). *Let X_n be a real valued discrete parameter stationary process with zero mean, with increments Δt. There exists an orthogonal process $Z(f)$ defined on the interval $[-\frac{1}{2\Delta t}, \frac{1}{2\Delta t}]$, such that:*

$$X_n = \int_{-\frac{1}{2\Delta t}}^{\frac{1}{2\Delta t}} dZ(f) e^{i2\pi f n \Delta t}, \forall n \in \mathbb{Z}. \tag{2.1.4}$$

The process Z(f) has the properties:

1. $E\{dZ(f)\} = 0 \ \ \forall f \in [-\frac{1}{2\Delta t}, \frac{1}{2\Delta t}]$.
2. $E\{|dZ(f)|^2\} = dS^{(I)} \ \ \forall f \in [-\frac{1}{2\Delta t}, \frac{1}{2\Delta t}]$ *Where $dS^{(I)}(f)$ is the integrated spectrum of X_n.*
3. *For any two frequencies $f, f' \in [-\frac{1}{2\Delta t}, \frac{1}{2\Delta t}]$, since Z(f) is an orthogonal process.*

$$Cov\{dZ(f'), dZ(f)\} = E\{dZ^*(f')dZ(f)\} = 0.$$

This theorem shows any discrete stationary process can be written as an infinite sum of complex exponentials. Consider the auto covariance sequence (ACVS), s_τ, using

the spectral representation theorem, specifically property 3, the ACVS can be written as [66, 67]:

$$s_\tau = \iint_{-\frac{1}{2\Delta t}}^{\frac{1}{2\Delta t}} E\{dZ^*(f')dZ(f)\}e^{i2\pi(f-f')n\Delta t}e^{i2\pi f\tau\Delta t}df.$$

Through Theorem 2.1.1 and property 1, the only contribution to the integral occurs when $f = f'$, therefore:

$$s_\tau = \int_{-\frac{1}{2\Delta t}}^{\frac{1}{2\Delta t}} E\{|dZ(f)|^2\}e^{i2\pi f\tau\Delta t}df = \int_{-\frac{1}{2\Delta t}}^{\frac{1}{2\Delta t}} dS^{(I)}(f)e^{i2\pi f\tau\Delta t}. \tag{2.1.5}$$

Where the differential of the integrated spectrum, $S^{(I)}(f)$, is the spectral density function, $S(f)$. Therefore the ACVS and the SDF form an FFT pairing, i.e. the FFT of the ACVS is the SDF and vice versa. These equations illustrate what we want to achieve by modelling in the frequency domain; we want to describe the ACVS of a signal through its spectrum [66, 67].

Stationarity is important due to property 3 of the spectral representation theorem. If we were dealing with a non-stationary series we would have correlated increments, and we could not define the spectrum as cleanly [66,67]. We would have to look at the correlated increments between frequencies, nor is it guaranteed that the spectrum would be a finite process. If we have a non-stationary process, under the definition of weak stationarity, we might be able to utilise this representation. Such a case is fractional Brownian motion, which is a non-stationary process due to the covariance being dependent upon time, however has a defined spectral density function [67]. Fractional Brownian motion (fBm) is a generalisation of Brownian motion whose increments are not required to be independent; the covariance structure of fBm are described in Equation 3.2.2 [80].

If the parameters of any process change over time, so are therefore time-heterogeneous; we cannot utilise this time-homogeneous description and the frequency domain. Instead we must utilise the time-frequency domain to suitably describe the behaviour of a signal's second order properties.

2.2 Estimating the second order properties of a stochastic process

For now let us focus upon estimating a signal's spectrum, starting with data visualisation. Given a signal's frequency domain representation, $\mathcal{X}(f)$, we can obtain a naive estimator of the spectrum by taking the absolute value squared: the periodogram [66,67].

Definition 2.2.1 (**Periodogram**). *The periodogram for a signal X_n is defined by:*

$$\hat{S}_X^{(p)}(f) \equiv \frac{\Delta t}{N}\left|\sum_{n=1}^{N}(X_n-\mu)e^{-i2\pi f n\Delta t}\right|^2 = \frac{1}{N\Delta t}\left|\mathcal{X}(f)\right|^2.$$

Which has properties:

1. $Var\{\hat{S}_X^{(p)}(f)\} = \begin{cases} S_X^2(f) & 0<|f|<\frac{1}{2\Delta t} \\ 2S_X^2(f) & |f|=0, |f|=\frac{1}{2\Delta t} \end{cases}$
2. $Cov\{\hat{S}_X^{(p)}(f), \hat{S}_X^{(p)}(f')\} = 0, \quad 0 \leq |f'| < |f| \leq \frac{1}{2\Delta t}$

The periodogram is a sensible starting point for frequency analysis as it can be viewed as a diagnostic plot for the signal's covariance structure. The periodogram is the sample variance of the FFT, like the SDF is the variance of the orthogonal increment process $dZ(f)$ in the Spectral Representation Theorem. Thus, the periodogram is a method of moments estimator of the spectrum. [66, 67]. Whilst it does not tell us the exact relationship it does provide an insight into what is needed to obtain such a description. From inspection of the periodogram the type of spectrum can be become obvious; for example, red processes - such as fractional Brownian motion - are difficult to visualise unless on a decibel (dB) scale; furthermore on this scale they show evidence of decay from $f = 0$.

Fractional Brownian motion has different properties depending upon the Hurst parameter, which describes the amount of self similarity within the process. When the Hurst parameter is greater than one we obtain an integrated processes, the periodogram of such processes illustrates high bias due to leakage [66]. As well as susceptible to issues due to leakage, the periodogram is an inconsistent estimator as the variance does not asymptotically tend to 0 [66,67]. This can be rectified simply through application of the periodogram's variance stabilising function, the logarithm [67]. A by product of taking

logarithms is the Euler-Mascheroni constant, γ, which is something we need to be aware of when plotting on a decibel scale - $\log_{10}$.

$$E\{\log(\hat{S}_X^{(p)}(f))\} = \log\{S_X(f)\} - \gamma, \quad Var\{\log(\hat{S}_X^{(p)}(f))\} = \frac{\pi^2}{6}.$$

Then let the non normal random variable $\epsilon^\ell(f)$ be defined for $|f| < \frac{1}{2\Delta t}$ as [67]:

$$\epsilon^\ell(f) = \log\left(\frac{\hat{S}_X^{(p)}(f)}{S_X(f)}\right) + \gamma,$$

which is equal in distribution to:

$$\epsilon^\ell(f) = \log(\chi_2^2) + \gamma - \log(2).$$

Therefore if we let:

$$Y^{(p)}(f) \equiv \log(\hat{S}_X^{(p)}(f)) + \gamma,$$

Then for $|f| < \frac{1}{2\Delta t}$:

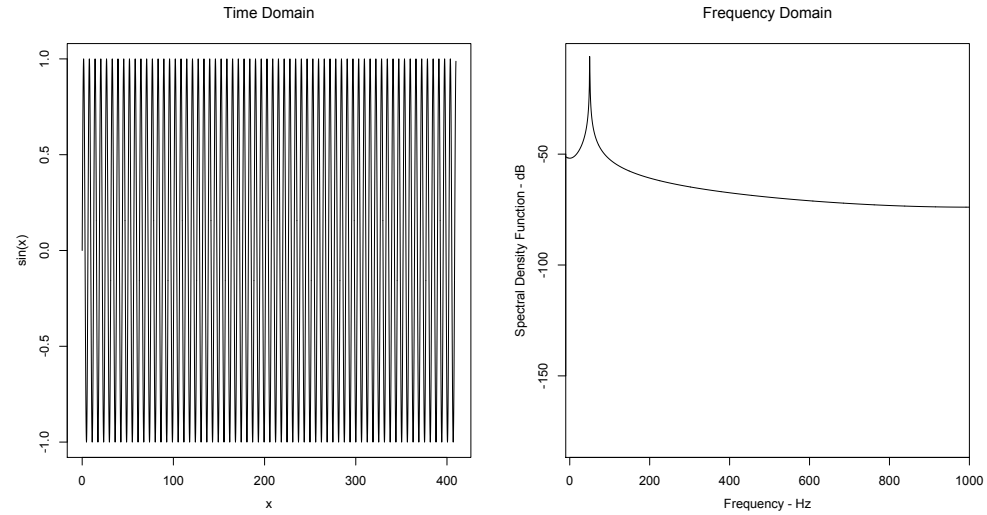
$$Y^{(p)}(f) = \log(S_X(f)) + \epsilon^\ell(f).$$

Therefore the log periodogram plus a constant (Euler-Mascheroni Constant γ) can be written as the true log SDF plus the non normal noise variable $\epsilon^\ell(f)$. This information has been presented for two purposes, first we can identify the asymptotic distribution of the periodogram as χ_2^2 and as previously mentioned we need to take this into account when plotting on a log scale, such as the decibel scale. [67]

The periodogram is a biased estimator of the spectrum, and is severely affected by leakage which is the result of the finite length sequence in the fast Fourier transform [66, 67, 79]. The expectation of the periodogram is equal to a convolution of the Fourier transform of a rectangular window with the true spectrum [66, 67, 79]; an issue to be aware of is that the FFT assumes the signal to be periodic and repeats the signal accordingly. The resultant sharp discontinuity between the start and end of the signal, as well as the abrupt way in which the rectangular taper goes to zero outside the interval, results in leakage [66, 67]. We have already presented this type of leakage discussed above: aliasing. Leakage is a result of the signal's energy being blurred across frequencies, and is visible in the sidelobes of the Fejér kernel. The energy in the sidelobes of Fejér's kernel is the cause of leakage within an estimator, as sidelobes transfer power from one region of the spectrum to another [66, 67].

To illustrate this we shall simulate a 50Hz sine wave sampled at 2000Hz (i.e. $\Delta t = 0.0005s$), and plot the periodogram of the signal - Figure 2.1. We should see a solitary spike at 50Hz however we can see energy around 50Hz and across the spectrum, which is leakage. We can reduce the sidelobes of Fejér's kernel by applying a taper/windowing function to our signal [66, 67, 81, 82].

Figure 2.1: 50Hz Simulated Signal



The plot on the left is the time domain representation of a 50Hz sine wave sampled every 0.0005s, the plot on the right is the raw periodogram - a naive estimator the spectrum. We can see in the frequency domain plot evidence of leakage due to the energy around the spike at 50Hz. As the signal is only a 50Hz sine wave we should not see any energy at these frequencies, however due to the effect of Fejér's kernel we see blurring/leakage across them.

2.2.1 Bias reduction techniques

A taper is a suitable sequence of real valued constants, which smoothes the transition of the rectangular window to zero; therefore reducing the sidelobes of Fejér's kernel and reducing leakage [66]. Application of a taper, h_n, results in the direct spectral estimate [66]:

Definition 2.2.2 (**Direct Spectral Estimator**). *The direct spectral estimator of a signal X_n is defined by:*

$$\hat{S}_X^{(d)}(f) = \Delta t \left| \sum_{n=1}^{N} h_n X_n e^{-i2\pi f n \Delta t} \right|^2 .$$

We obtain the smoothed spectral window of the taper applied from:

$$\mathcal{H}(f) = \Delta t \left| \sum_{n=1}^{N} h_n e^{-i2\pi f n \Delta t} \right|^2 .$$

Finally as a result of the orthogonal increments property of the periodogram the expected value of this estimator is:

$$E\{\hat{S}_X^{(d)}(f)\} = \int_{-\frac{1}{2\Delta t}}^{\frac{1}{2\Delta t}} \mathcal{H}(f - f')S(f')df'.$$

This illustrates that the spectral window of a taper is a convolution of Fejér's kernel, with hopefully reduced sidelobes. This is illustrated by setting $h_n = \frac{1}{\sqrt{N}}, 1 \leq n \leq N$; as the direct spectral estimate is then equal to the periodogram, and the spectral window of the taper is equal to the Fejér kernel [66, 67, 82].

One possible taper which we can use is the cosine taper [66, 82]:

Definition 2.2.3 (**Cosine Taper**). *The cosine taper is defined as:*

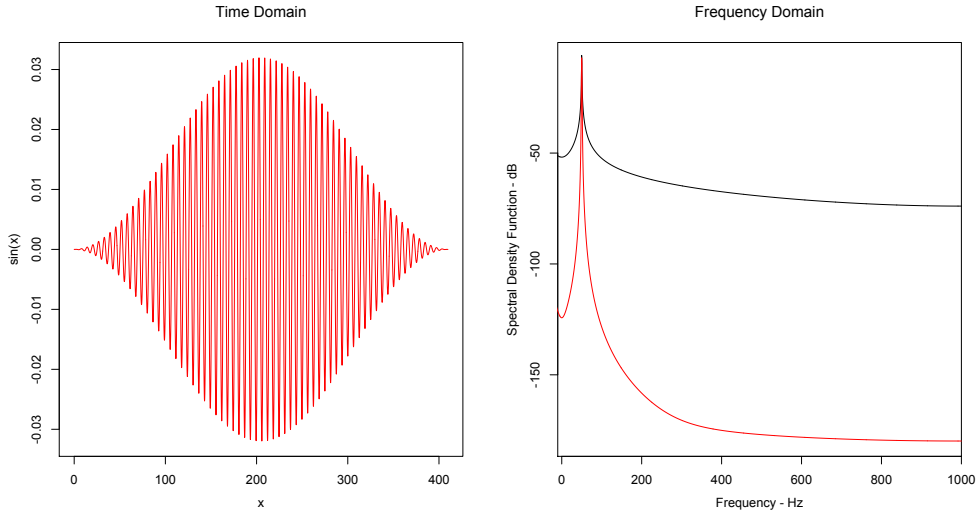
$$h_n = \begin{cases} \frac{C}{2}[1 - \cos(\frac{2\pi n}{\lfloor pN \rfloor + 1})], & 1 \leq n \leq \frac{\lfloor pN \rfloor}{2} \\ C, & \frac{\lfloor pN \rfloor}{2} < n < N + 1 - \frac{\lfloor pN \rfloor}{2} \\ \frac{C}{2}[1 - \cos(\frac{2\pi(N+1-n)}{\lfloor pN \rfloor + 1})], & N + 1 - \frac{\lfloor pN \rfloor}{2} \leq n \leq N \end{cases}$$

Where p is the tapering ratio - $p \in [0, 1]$ - the higher the value of p the more tapered the signal is.

If we set $p = 1$ we obtain the Hanning window. The cosine taper reduces leakage within the estimator by reducing the sidelobes of the Fejér kernel, which causes leakage within an estimator. Furthermore, when dealing with segments of a signal the FFT has issues with regards to the sharp discontinuity between the start and end of the signal. The FFT assumes that the signal is periodic with period N and treats it as such, which is not accurate in practice. The cosine taper smoothes the transition at the ends of the signal to zero, removing this source of bias within the periodogram [66, 67]. Figure 2.2

shows the effect that this window has had on the simulated 50Hz signal, the spectral leakage has been reduced.

Figure 2.2: 50Hz Cosine Tapered Simulated Signal

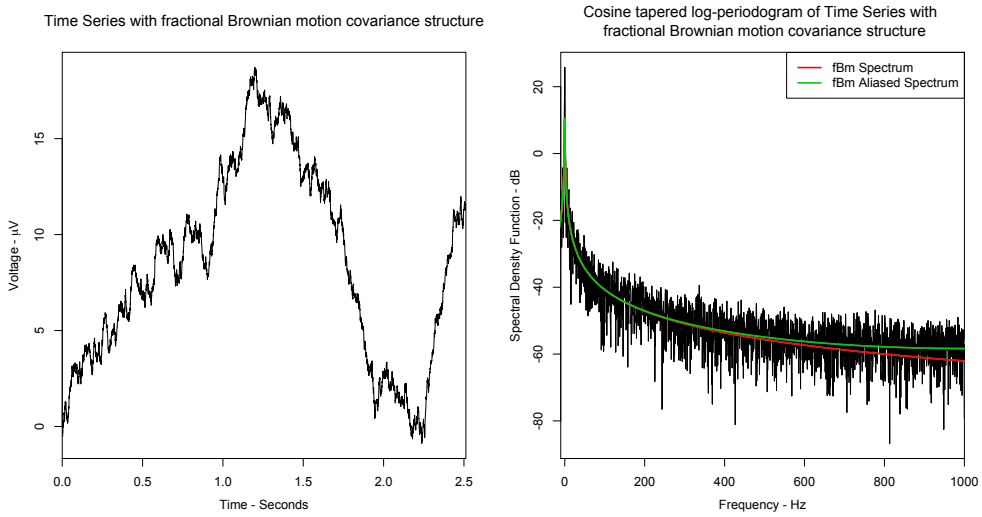


The plot on the left shows the full - $p = 1$ - cosine tapered signal presented in Figure 2.1. The plot on the right shows the corresponding direct spectral estimator - red line - against the raw periodogram from Figure 2.1 - black line. We see the same spike at 50Hz however we have also reduced the leakage as can be seen by the lower energy at the other frequencies when comparing the direct spectral estimator and the raw periodogram.

Let us illustrate the effect that leakage can have on a simulated fractional Brownian motion process. Figure 2.3 shows the time domain representation and the cosine tapered periodogram of a fractional Brownian motion process. Notice on the decibel scale we can see clear evidence of decay in the spectrum; furthermore, the plot illustrates the effect of aliasing and shows that unless the signal undergoes a process to minimise the effect of aliasing, we need the spectrum to reflect this [67, 83].

In the periodogram definition, several properties were given of this naive estimator of the spectrum. The second property, asymptotic independence, allows estimates of the spectral density function to be created by smoothing across frequencies. As mentioned previously, tapering obtains a less bias estimate by smoothing across frequencies but at the cost of increased variance [66, 67, 84]. Multitaper estimation is asymptotically

Figure 2.3: Simulated fractional Brownian motion process illustrating leakage



The plot on the left is a time domain representation of a fractional Brownian motion process. The plot of the right is the cosine tapered direct spectral estimator of the simulated fractional Brownian motion process. In the direct spectral estimator we can see the effect of aliasing in the spectral density functions plotted. The green line is the spectral density function taking into account the effect of aliasing, and fits the process better especially at high frequencies when compared to the non-aliased spectrum.

consistent if the spectrum is continuous, however if the signal is finite bias is increased due to blurring.

Multitapering, first outlined by Thomson in 1982 [85], obtains a spectral estimate by averaging several orthogonal tapers. The multitapered spectral estimate is the average of K direct spectral estimators and is defined by [66, 67, 85]:

Definition 2.2.4 (**Multitaper Estimate**). *The multitapered estimate of a signal X_n using k tapers is defined by:*

$$\hat{S}_X^{(mt)k}(f) = \Delta t \left| \sum_{n=1}^{N} h_{n,k}(X_n - \mu) e^{-i2\pi f n \Delta t} \right|^2, \qquad \hat{S}_X^{(mt)}(f) = \frac{1}{K} \sum_{k=1}^{K} \hat{S}_X^{(mt)k}(f).$$

Where $\{h_{t,k}\}$ is the data taper for the k^{th} direct spectral estimator $\hat{S}_X^{(mt)k}(\cdot)$ and has the properties

$$E\{\hat{S}_X^{(mt)k}(\cdot)\} = \int_{-\frac{1}{2\Delta t}}^{\frac{1}{2\Delta t}} \mathcal{H}_k(f - f')S(f')df',$$

$$E\{\hat{S}_X^{(mt)}(\cdot)\} = \int_{-\frac{1}{2\Delta t}}^{\frac{1}{2\Delta t}} \bar{\mathcal{H}}(f-f')S(f')df', \quad \bar{\mathcal{H}}(f) = \frac{1}{K}\sum_{k=0}^{K-1}\mathcal{H}_k(f). \tag{2.2.1}$$

The behaviour of the sidelobes dictates the bias of the estimate; therefore, the K spectral windows must provide adequate protection against leakage, if we are to obtain a suitable estimate with respect to minimising leakage. Utilising more tapers results in a reduced variance, however at the cost of bias [66]. The more tapers utilised the smaller the window over which the spectra is calculated; if a wide window is used then the estimator will not be as smooth, it will have high variance. However if too narrow a window is used then the estimator will lose information contained within the signal, it will have high bias. When using a multitaper estimate it is important to strike a balance and choose a value for the width of the window, W, and K which has a sufficient trade off between the two. The rate at which the variance of $\hat{S}^{(mt)}(\cdot)$ decreases as the number of eigenspectra K increase. The change in variance of $\hat{S}^{(mt)}(\cdot)$ as K increases can be expressed as [66, 84]:

$$\begin{aligned} Var\{\hat{S}^{(mt)}(f)\} &= Var\left\{\frac{1}{K}\sum_{k=0}^{K-1}\hat{S}^{(mt)}(f)\right\} \\ &= \frac{1}{K^2}\sum_{j=0}^{K-1}\sum_{k=0}^{K-1} Cov\{\hat{S}^{(mt)j}(f), \hat{S}^{(mt)k}(f)\}. \end{aligned}$$

We shall illustrate the multi tapering approach utilising sinusoidal multi tapers [86]; the bandwidth of the k$^{\text{th}}$ taper is $\frac{1}{N}$ centred around the frequency $\frac{k}{2N}$, as such the bandwidth can be altered by adding or removing tapers [86]. Therefore sinusoidal tapers do not have the parameter W, unlike other approaches such as discrete prolate spheroidal sequence tapers [66, 67, 87–91]. The k$^{\text{th}}$ sinusoidal taper is defined as:

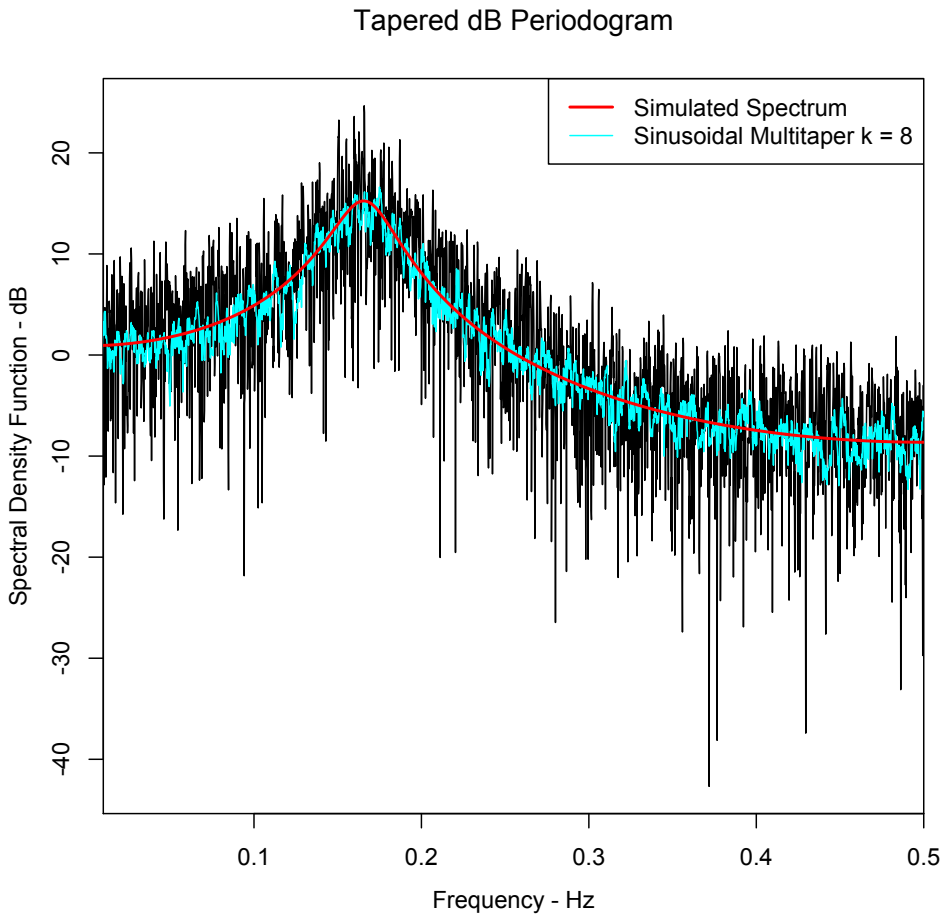
Definition 2.2.5 (**Sinusoidal Multitaper**). *The k^{th} sinusoidal multitaper is defined by:*

$$h_{n,k} = \sqrt{\frac{2}{N+1}}\sin\left(\frac{\pi k n}{N+1}\right).$$

We shall estimate the spectral density of a simulated AR(2) process with parameters: $\sigma = 1, \phi_1 = 0.9, \phi_2 = -0.8$ and K=8. The estimated spectral density function from sinusoidal multi tapering and the simulated spectrum can be seen in Figure 2.4.

As demonstrated, we obtain a more refined non-parametric estimate of the spectrum from multi-tapering. Whilst averaging over several spectral density estimates has produced a closer approximation to the spectrum - reduced bias - we have done so at the

Figure 2.4: Multitaper estimate of simulated AR(2) process



This plot illustrates the effectiveness of averaging over multiple spectral density estimates in reducing the bias of an estimator. Previously we have utilised only one taper, which results in an estimate that looks similar to the periodogram - averaging over multiple estimates results in an estimator that is more refined and closer to the simulated spectrum. This reduction in bias comes at the cost of increased variance, and is a tradeoff accounted for when multitapering.

cost of increased variance. The multivariate taper is one way in which we can get a more interpretable estimate of the spectral density function. To generate estimates of the neonatal EEG signals within our dataset we are going to use the univariate tapered signal in an estimation procedure that produces a parametric estimate of the spectrum.

2.3 Automated detection procedure for the suitability of estimated spectra

Before we start fitting spectral density function estimates, it would be useful to have an objective way in which we can assess the fit of an estimated spectrum. Especially given then size of data set we are to analyse; as such we shall define a method. In order to do so we need to first define the residuals of a signal given an estimated SDF.

Definition 2.3.1 (**Spectral residuals**). *For a signal $X(t)$ with direct spectral estimate $\hat{S}_X^{(d)}(f)$ we shall define the residuals of the signal given an estimated set of parameters $\hat{\underline{\theta}}$ and objective function $S_X(f;\underline{\theta})$ by:*

$$\hat{\epsilon}(f) = \frac{\hat{S}_X^{(d)}(f)}{S_X(f;\hat{\underline{\theta}})}.$$

In a signal where the estimated spectral density function is a fair approximation to the variance of the Fourier transformed signal we would expect to see no trend in the residuals. If trend is remaining in the residuals then the estimated spectrum is not a suitable description of the signal. In the estimation of the neonatal EEG signals contained in the presented data set, we cannot visually inspect each signal for lack of trend. As such we need to create a detection procedure which will be an automated way to determine whether the signal's spectrum has been described. In order to do so we need to utilise the asymptotic distribution of the periodogram.

We have already identified the asymptotic distribution of the periodogram in the definition of the Euler-Mascheroni constant and as a result of the periodogram's properties. From the definition of the periodogram, the variance stabilisation proof and standard likelihood theory we obtain the asymptotic distribution as χ_2^2 [66, 67, 109]. Using this distributional result we can construct a hypothesis test for the fit of an estimated spectrum.

Definition 2.3.2 (**Automated detection procedure for a fitted spectral density function**). *For a signal $X(t)$ with spectral residuals $\hat{\epsilon}(f)$ we would expect for an adequately fitting spectrum that:*

$$T = 2\sum_{j=1}^{J} \hat{\epsilon}(f) \sim \chi^2_{2J}.$$

Therefore, we can construct a hypothesis test, with a null hypothesis of an adequately described spectrum. i.e.

$$H_0 : T \sim X^2_{2J}$$

If J is sufficiently large, i.e. $J > 50$, then we can test this using the normal approximation to the chi-squared distribution. Since the number of frequencies we are analysing shall always be large enough we can obtain a p-value easily by:

$$P_{GoF} = 2\left[1 - \Phi\left(\frac{|T - 2J|}{\sqrt{4J}}\right)\right]. \tag{2.3.1}$$

Since we are assessing multiple infants, we want to reduce the number of Type II errors obtained; these errors reflect null hypotheses that are failed to be rejected, when they should be rejected. As such we shall implement false discovery rate analysis (FDR) by the Benjamini-Hochberg procedure [92].

Definition 2.3.3 (**Benjamini-Hochberg False Discovery Rate: Independent Tests**). *Consider m independent tests with null hypotheses $H_1, ...H_m$ and corresponding p values $p_1, ..., p_m$. Then order these p values in ascending order denoted $p_{(1)}, ..., p_{(m)}$, for a given α find the largest k such that:*

$$p_{(m)} \leq \frac{k}{m}\alpha.$$

Then reject all $H_{(i)}$ $\forall$ $i = 1, .., k$ i.e. reject the null hypothesis with significance level $\frac{k}{m}\alpha$ If no such k exists then fail to reject all m hypotheses

This ensures that the expected false discovery rate is less than a given α. The p value shall be calculated through the following equation. Although, implementing false discovery rate analysis can increase Type II errors; not rejecting the null hypothesis when it should be rejected.

We have implemented FDR analysis because of the number of signals that we are analysing and we want to ensure that we are not rejecting signals that should not

be rejected. This application of FDR is to help ensure that during future analyses that the estimates are adequate descriptions of the procedure. Furthermore we have utilised the Benjamini-Hochberg procedure over other methods, such as Bonferroni correction, because of sample size and we do not want to reduce statistical power as a result of an increase in Type II errors.

2.4 Whittle Estimation

Now we have a way in which the suitability of an estimated spectrum can be assessed, we shall introduce the way in which a non-parametric estimate can be used to fit a parametric estimate: Whittle estimation.

The Whittle likelihood, first introduced by Peter Whittle in 1953 [93], constructs and approximation to the time domain log likelihood in a manner similar to weighted least squares. For a non-parametric estimator of the SDF, such as the periodogram $\hat{S}_X(f)$, and parametric SDF, $S_X(f;\underline{\theta})$, the Whittle likelihood is defined as [93]:

$$l_W(\underline{\theta}) = -\int_{-\frac{1}{2\Delta t}}^{\frac{1}{2\Delta t}} \left(\frac{\hat{S}_X^{(d)}(f)}{S_X(f;\underline{\theta})} + \log\{S_X(f;\underline{\theta})\} \right) df. \tag{2.4.1}$$

From this, the parameters can be estimated by minimising this likelihood [93]:

$$\hat{\underline{\theta}} = \arg\min_{\theta\in\Theta} l_W(\theta).$$

The Whittle likelihood approximates the time domain likelihood of a Gaussian time series. By minimising the negative of this function efficient estimates are found. The Whittle likelihood constructs an estimate in a similar manner to weighted least squares, by minimising the sum (or theoretically the integral) of the spectral residuals [93]. Care use be taken however with traits in the signal, such as in our case electrical interference. As with other concepts explained, the above formulation is infeasible as we are integrating across all frequencies. We can approximate this equation however, through a Riemann approximation; which is an approximation of the area described by an integral.

$$l_W^{(R)}(\underline{\theta}) = -\frac{2}{N\Delta t}\sum_{k=0}^{M}\left(\frac{\hat{S}_X(f_k)}{S_X(f_k;\underline{\theta})} + \log\{S_X(f_k;\underline{\theta})\} \right), \quad M = \begin{cases} \frac{N}{2} & Neven \\ \frac{N-1}{2} & Nodd \end{cases} \tag{2.4.2}$$

Riemann approximations become more accurate the finer you make the approximation, in the above equation the larger than M is, to the point that the approximation is effectively equal to the actual function. However, as suitable non-parametric estimates such as the periodogram, are not smooth we cannot utilise this argument. However, we can say that if N is large enough, we get a decent enough approximation. Since the distance between frequencies becomes smaller, and the approximation becomes finer.

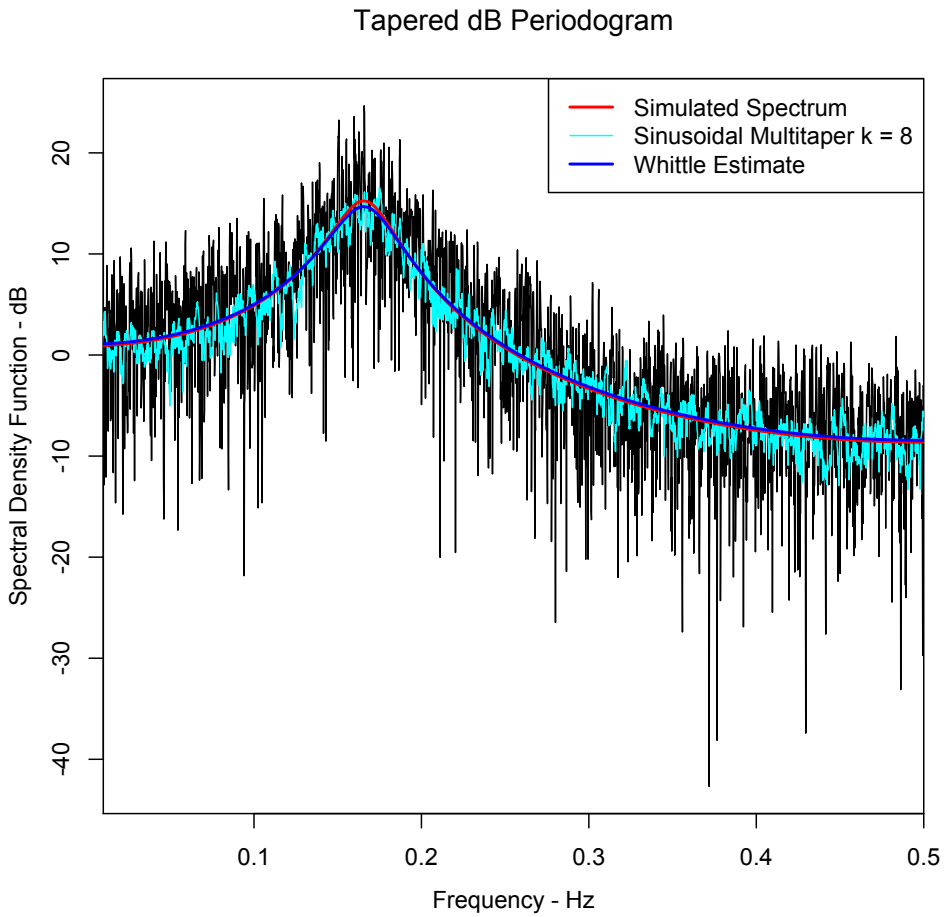
We can implement Whittle estimation with any well defined SDF and either the periodogram or a univariate tapered signal; multitapered signals should not be used since they increase the bias of the estimator with no reduction in the variance of the estimates obtained from the Whittle estimation. Due to the formulation of the likelihood, Whittle estimation is affected by high bias, and can produce unsatisfactory estimates. With spectral density functions such as fractional Brownian motion, we need to be careful about their contribution to the likelihood at certain frequencies. Consider the frequency representation of a fractional Brownian motion process; at $f = 0$ there is an infinite peak in the periodogram due to the process' underlying generating mechanism. This frequency cannot be included in Whittle estimation due its formulation being similar to that of weighted least squares.

We can illustrate the effectiveness with a simple demonstration using an autoregressive process of order two. In Figure 2.5 we can see the cosine tapered periodogram of a simulated AR(2) process with variance 1 and parameters 0.9, -0.8. In this plot, the red line is the simulated spectrum, the light blue line is the multi-taper estimate and the blue line is the Whittle estimate; furthermore the table of estimates shows the accuracy of the estimated parameters.

Table 2.4.1: AR(2) Whittle Estimates

	Simulated Value	Whittle	Whittle 95% C.I.
σ	1	1.01231	(0.99252, 1.03210)
ϕ_1	0.9	0.88969	(0.87252, 0.90686)
ϕ_2	-0.8	-0.78392	(-0.80110, -0.76676)

Figure 2.5: Whittle estimate of simulated AR(2) process



This plot illustrates the sinusoidal multitaper estimate outlined in Figure 2.4 and the Whittle estimate obtain by utilising an AR(2) objective function. We find that when the objective function matches the underlying generating mechanism of the signal that the estimate obtained is accurate - Table 2.4.1

Whittle estimation is an elegant and effective way of fitting a parametric spectrum from a non-parametric estimator of the spectrum. Furthermore, confidence intervals for the estimates are easily constructed from the Hessian matrix. However, there are limitations that can cause this estimation procedure to produce unsatisfactory estimates; mainly spikes and bias in the non-parametric estimator. Both of these show the issues in the formulation of the Whittle likelihood. Perhaps the most important issue is the bias and leakage in the periodogram, whilst easily rectified it is an important further illustration

of how bias affects estimation of the spectrum.

One of the major strengths of Whittle estimation is the flexibility with regards to the estimates that can be obtained. For any process whose SDF is well defined, we can obtain a parametric estimate; however, should we omit frequencies - due to worrying components such as bias or electrical interference - we obtain a semi-parametric estimate [83]. Should there be any alteration to the signal, such as filtering, we can take this into account in the estimation by applying the same transformation to the covariance sequence, and as such the SDF.

2.5 Linear time invariant filters

During the signal recording process, unwanted components can be present. These unwanted components could be electrical interference or recording of unwanted movements, such as eyes/limb movements, known as artefacts [1–3]. We can minimise, and almost entirely remove, the effect of these components through filtering [1–3, 66].

Filters can be described by the frequencies which they omit or include. For example a band-pass filter minimises the contribution of all frequencies, except the band specified. Conversely a band-stop filter minimises the contribution of these frequencies within the specified band. Similarly a low-pass filter minimises the effect of frequencies after a specified cutoff, leaving lower frequencies along; and a high-pass leaves higher frequencies along minimising the contribution of lower ones [66]. There are more filters than those mentioned, for example if we were to convolve a stochastic process with an ARMA(p,q) SDF we would be filtering the process according the characteristics of the ARMA(p,q) process. By convolving an ARMA spectrum onto another process we impose a short term covariance structure onto the time series in question. We will utilise this method when simulating a time series with both a long term and short term covariance structure.

The way in which a filter is applied is by its transfer function, $H(\cdot)$; this is applied to a signal X, resulting in the filtered signal Y [66]. We can also describe a filter by its gain function, which we apply to a signal in the frequency domain; as such a convolution [66].

The ideal stop band or pass band filter has values of only zero or one; with unfiltered frequencies equal to one, and filtered frequencies equal to zero. However, we cannot apply a filter of this type, and as such multiple ways to approximate this have been defined. A filter is assessed by whether it has ripples in the pass or stop band and how quickly it excludes frequencies, the roll off, with tradeoffs between these factors being assessed when deciding which to use. The strength of a filter and how fast the roll off occurs can be changed by increasing the order of a filter [66].

We present the gain functions for the Butterworth filter [94]. These are outlined because a first order Butterworth, $n = 1$, is the best approximation of the type of filter implemented in the analysed dataset. The documentation of the preprocessing did not reference the filter implemented. Requesting further information about the device used to filter the signals resulted in a technical document. Upon consultation the Butterworth first order filter was identified as the best approximation of the applied filter. This is an approximation due to the fact that we could not obtain confirmation from the manufacturer of the device that this is the filter implemented during preprocessing.

Definition 2.5.1 (**Low-pass Butterworth Filter**). *A low-pass Butterworth filter of order n with cut off frequency f_{c_L} and DC gain G_0 has a gain function defined as:*

$$\left|\mathrm{H}(f)^2\right| = \frac{G_0^2}{1 + \left(\frac{f}{f_{c_L}}\right)^{2n}}. \tag{2.5.1}$$

Definition 2.5.2 (**High-pass Butterworth Filter**). *A high-pass Butterworth filter of order n with cut off frequency f_{c_H} and DC gain G_0 has a gain function defined as:*

$$\left|\mathrm{H}(f)^2\right| = \frac{G_0^2}{1 + \left(\frac{f_{c_H}}{f}\right)^{2n}}. \tag{2.5.2}$$

Definition 2.5.3 (**Band-pass Butterworth Filter**). *A band-pass Butterworth filter of order n with cut off frequencies f_{c_L} and f_{c_H} and DC gain G_0 has a gain function defined as:*

$$\left|\mathrm{H}(f)^2\right| = \frac{G_0^2}{\left[1 + \left(\frac{f_{c_H}}{f}\right)^{2n}\right]\left[1 + \left(\frac{f}{f_{c_L}}\right)^{2n}\right]}. \tag{2.5.3}$$

Definition 2.5.4 (**Band-stop Butterworth Filter**). *A band-stop Butterworth filter of order n with cut off frequencies f_{c_L} and f_{c_H} and DC gain G_0 has a gain function defined as:*

$$\left|\mathrm{H}(f)^2\right| = \frac{G_0^2}{\left[1+\left(\frac{f_{c_L}}{f}\right)^{2n}\right]\left[1+\left(\frac{f}{f_{c_H}}\right)^{2n}\right]}. \tag{2.5.4}$$

In the above equations DC gain refers to the gain at zero frequency, and the frequencies f_{c_L} and f_{c_H} are the cut off frequencies for the low and high pass filters respectively. As $n \to \infty$ the filter approaches the ideal rectangular filter, as such the higher the value of n the stronger the effect of the filter shall be. The filters discussed are all linear time invariant filters as their characteristics are time-homogeneous.

We can estimate a filtered signal easily by Whittle estimation, so long as we know the filter applied. For a filtered signal X which has a direct spectral estimator $\hat{S}_X^{(d)}(f)$, if we know the gain function $G(f)$ we can estimate the spectral density function by minimising the following likelihood.

$$l_W(\underline{\theta}) = -\int_{-\frac{1}{2\Delta t}}^{\frac{1}{2\Delta t}} \left(\frac{\hat{S}_X^{(d)}(f)}{G(f)S_X(f;\underline{\theta})} + \log\{G(f)S_X(f;\underline{\theta})\} \right) df. \tag{2.5.5}$$

2.6 Continuous Wavelet Transform

The methods presented so far have all been time-homogeneous; i.e. can be used for signals whose properties do not change with time. However, EEG signals are known to be time-heterogeneous and contain many different time-heterogeneous components [1–3]. As such, we need to outline ways in which to assess and visualise time-heterogeneous signals.

The way in which we can assess how the second order properties of a signal change across time, is to decompose a signal into the time-frequency domain. Two ways in which we can do this are short-time Fourier transform [66, 68] and wavelet decomposition [67, 95]. The first performs the Fourier transform over a series of windows in time, and the latter decomposes the signal into wavelets. We shall implement wavelet decomposition due to its preferable time and frequency localisation properties, as well as its robustness to noisy signals [95].

The wavelet transform is a way in which we can decompose a time series into the time frequency domain, and can obtain finer detail and resolution than the STFT. The way in which we consider the time frequency decomposition of a signal is through the scalogram, in a similar manner to studying the periodogram in the frequency domain. Similarly to the periodogram, the behaviour of the signal in the time-frequency domain can be masked by high variance which can be visible in the scalogram unless further smoothing is implemented.

The scalogram is constructed through the use of wavelets, which can be considered local to a time point and frequency. Through the use of a family of wavelets, which are specific functions, and translating and scaling the wavelet function we can isolate signal behaviour associated with the localised time point and frequency using the wavelet transform.

The continuous wavelet transform takes a signal, $X(t)$, and a possibly complex-valued wavelet function, $\psi(t)$, then for a scale and time point the continuous wavelet transform is defined as [67, 95]:

$$\mathcal{X}(t_i, a_j) = \frac{1}{\sqrt{|a_j|}} \int_{-\infty}^{\infty} X(t)\psi^* \left(\frac{t - t_i}{a_j} \right) dt. \tag{2.6.1}$$

The continuous wavelet transform is convolution of the signal with a set of functions generated by the so called mother wavelet, which can be computed using the FFT.

Earlier in this thesis we introduced the concept of obtaining a more refined spectral density estimate through the creation of a set of orthogonal tapers - multitapering. The result of multi tapering an estimate that is more interpretable due to reduced variance. This approach can be utilised to obtain a more interpretable scalogram utilising wavelets, for example the Morse wavelet as shall be utilised. If we create orthogonal wavelets and average the resulting scalograms we can obtain a lower variance scalogram estimate - in a similar manner to multitapering. The time-frequency decomposition obtained from the Morse wavelet decomposition will allow such a scalogram to be obtained. [67, 95]

Since we want to obtain a reduced variance scalogram, therefore allowing a more accurate detection of any time heterogeneous components within our signal, we will utilise

such a set of orthogonal continuous wavelets - generalised morse wavelets. [95]

Generalised Morse wavelets are eigenfunction wavelets that are suitable for use in time-heterogeneous SDF estimation as a result of average time-scale eigenscalograms. Eigenscalograms are scalograms where the wavelet is an eigenfunction derived from a time-frequency concentration problem. We obtain a set of orthogonal eigenfunctions, and the first K can be ordered by the corresponding eigenvalue have orders $k = 0, ..., K-1$. From these K eigenscalograms can be computed and the resulting time-varying spectrum estimate is obtained from their average. These wavelets are defined by their order k - the corresponding eigenvalue order- and are dependent upon a pair of parameters (β, γ); in the analysis that follows the wavelets will be β=8, γ=3. When $\gamma = 1$ the zeroth order wavelet is known as a Cauchy wavelet. These wavelets have superior energy concentration over Hermites wavelets when $\gamma > 1$ and are easily computed using the FFT. [95].

From this time-frequency decomposition, we can obtain the scalogram, which is a measure of the energy at a scale a_j and time point t_i.

$$S_X(t_i, a_j) = |\mathcal{X}(t_i, a_j)|^2. \tag{2.6.2}$$

The scales can be related to a set of pseudo-frequencies, and can be localised to a specific interval of interest. We can obtain the pseudo-frequencies from the scales by the following definition:

$$f_a = \frac{f_c}{a\Delta t}, \tag{2.6.3}$$

where f_c is the central frequency of the wavelet.

The scales that we will analyse the scalogram over have been chosen due to the analytical range of interest - in this case the areas where delta brush activity occurs $[0.5, 1.5]$Hz and $[8, 25]$Hz. We are defining the scales using Equation 2.6.3 from pseudo-frequencies going from 0.04Hz to 31.17Hz in increments of 0.21Hz. Using $f_c = 0.23$Hz and $\Delta t = 0.0005$ we obtain scales that can be used to obtain the scalogram. The increments utilised by the pseudo-frequencies define the resolution of the scalogram accordingly. These values have been chosen to identify delta brush activity as mentioned, which are time heterogeneous

components present within neonatal EEG signals. These waveforms can be present in background EEG or somatosensory responses dependent upon the infants age.

2.7 Identification of time-heterogeneous signal segments

Delta brush, or neuronal bursting, activity is an identified response to stimuli [4, 96] as well as a spontaneous occurrence in background EEG in neonatal infants [61, 62]. As such we need a way in which to detect this time-heterogeneous signal component, so it does not affect the fitting of any future model. Ways in which this has previously been done involves either automated detection [4, 97] or visual inspection by an expert [98].

Time heterogeneity in a signal will result in a poor estimate of a signal's SDF, resulting in an inadequately fitted spectral density function. Furthermore, it is an unsuitable estimate because it does not take into account the time-heterogeneous nature of the signal. We can still use the outlined methods to analyse such a signal, but first we must segment the signal into time-heterogeneous and homogeneous segments; however we need a way in which to do this. One way in which we could do so is through visual inspection [96], although in our case this is impractical due to the sample size and would require a trained expert to identify such varying waveforms. Another, more accurate way, is to test the proportion of energy in the scalogram against a baseline [4].

Both spontaneous and event related neuronal bursts, also known as delta brushes, have been identified according to their time-frequency characteristics. Specifically a high frequency ripple between 8 and 25Hz occurring simultaneously as low frequency activity between 0.5 and 1.5Hz [4, 61, 96]. The scales of the scalogram have been chosen to highlight these areas of activity as they must occur simultaneously in order the activity to be classified as a delta brush.

In order to detect delta brush activity we shall test each frequency in the scalogram against a signal that is time homogeneous. We have constructed such a baseline from the signals identified as time-homogeneous and have been adequately described by our model. Utilising a simple proportion test we shall test to see whether the proportion of energy at a frequency and time point in the signal of interest is the same as the time

homogeneous baseline; rejection of the null hypothesis that $p_0 = p_s$ - where p_0 is the homogeneous baseline and p_s is the signal we want to test for the presence of delta brushes. This shall identify a region with possible delta brush activity before refining the region further utilising criteria defined from the numerical derivative of the signal. The procedure to refine the interval is the same that was utilised by Fabrizi et al. [4] and shall be outlined.

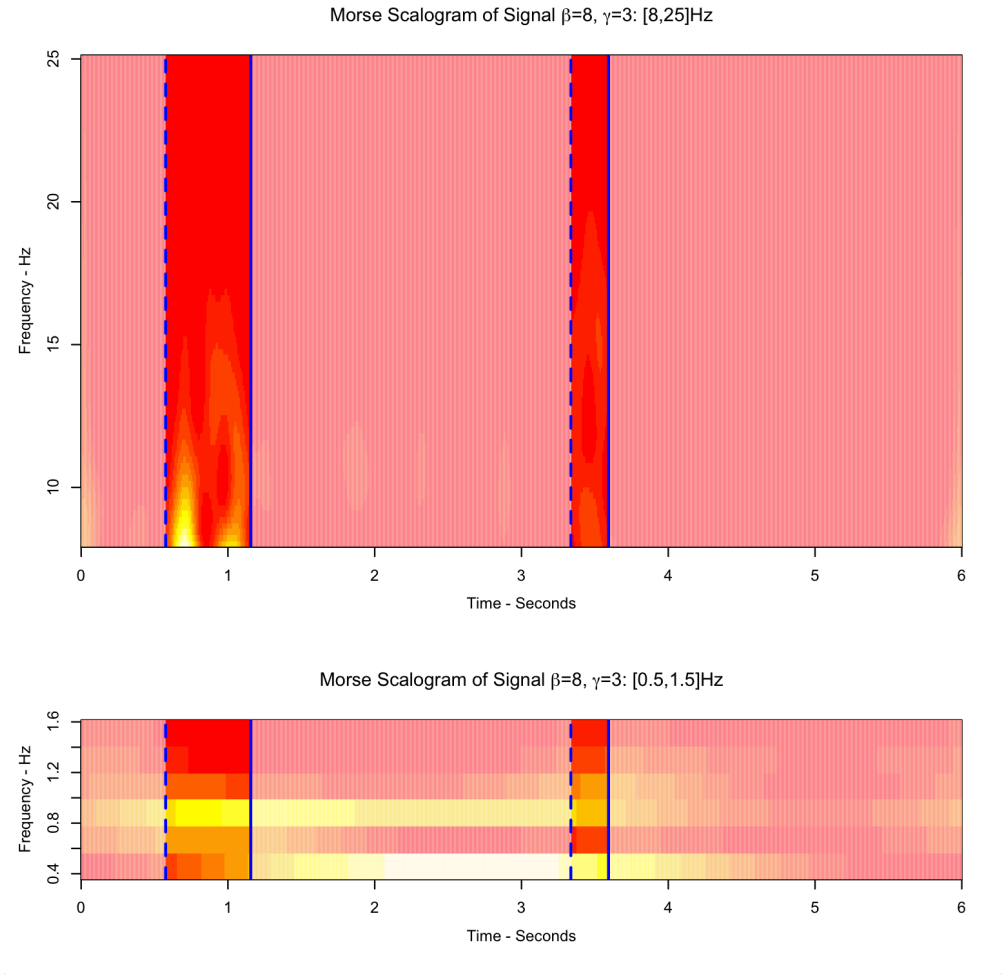
To refine the delta brush region - $[X(t_1), X(t_2)]$ - within a signal $X(t)$ it was first low-pass filtered at 2Hz using a 4th order Butterworth filter, before the numerical first and second order derivatives - $X'(t)$ and $X''(t)$ respectively - were obtained. Delta brushes are identified as negative deflections and therefore we determined the start and end point of the delta brush as follows. The start - onset - was defined as the time point, t_{on}, preceding t_1 where: $X'(t_{on} - \Delta t) > 0$ and $X'(t_{on} + \Delta t) < 0$, or $X'(t_{on}) < 0$ and $X''(t_{on} - \Delta t) > 0$ and $X''(t_{on} + \Delta t) < 0$. The trough - t_{tr} - of the delta brush is a time point that is needed in defining the end of the delta brush and is defined by a point following t_{on} where $X'(t_{tr} - \Delta t) < 0$ and $X'(t_{tr} + \Delta t) > 0$ The end - offset - was defined as the time point, t_{off}, following t_{tr} where $X'(t_{on} - \Delta t) > 0$ and $X'(t_{on} + \Delta t) < 0$, or $X'(t_{on}) < 0$ and $X''(t_{on} - \Delta t) < 0$ and $X''(t_{on} + \Delta t) > 0$ [4].

It is important that, unlike previously outlined detection procedures, our baseline remains constant across infants and time points. As such we shall utilise the same time homogeneous baseline between sets of signals therefore helping to ensure that we are detecting time-heterogeneous components fairly across all infants and recordings.

Using a signal from our data set we shall illustrate the results of this method. The image presented in Figure 2.6 illustrates the ability of the detection procedure to identify regions with significant time heterogeneity from the baseline. The significant level that was used in Figure 2.6, and shall be used throughout the detection analysis, is $\alpha = 0.1$.

The signals analysed also contain specific responses to nociceptive and tactile stimuli, as such to obtain estimates we need to be able to detect the presence of these components. Fabrizi et al. 2011 illustrated how the nociceptive and tactile somatosensory stimuli can be represented by a weighted principal component [4].

Figure 2.6: Detection procedure for time-heterogeneous delta brush within signal



The outlined delta brush detection procedure returns intervals where there is simultaneous activity in the regions [0.5,1.5]Hz and [8,25]Hz. This plot illustrates two regions where at least one frequency in each range has a higher proportion of energy than a time-homogenous baseline constructed from the presented dataset.

The evoked specific responses with regards to noxious and tactile stimuli have been outlined and illustrated by Fabrizi et al. [4]. As such we shall use these identified waveforms to classify whether a specific response is present. These waveforms were identified through principle component analysis, and the way in which they were identified will have bearing upon how we detect their presence. Fabrizi et al considered time points as observations and recordings (epochs) as variables. The recordings were aligned to

correct for latency jitter, by maximising the normalised inner product of the individual epoch against an iteratively constructed grand average of all the epochs analysed. Tactile stimuli was recognised as occurring 50-300ms post stimulation (with a maximum -50 to +50ms jitter allowed) and nociceptive stimuli occurring 300-700ms post stimulation (maximum allowed jitter -50 to 100ms). From this procedure the waveforms were identified that are associated with noxious and tactile stimulation, and these shall be used in our analysis. With these waveforms identified we can ascertain the presence of these waveforms in a given signal utilising the singular value decomposition of the waveforms and the signals.

For a signal X, the singular value decomposition, SVD_X, is a linear decomposition such that:

$$SVD_X = U_X D_X V_X^T. \tag{2.7.1}$$

Where the diagonal entries of D_X are non-negative real numbers and are the singular values of X. The parameter U_X is a matrix whose columns contain the left singular vectors of X, and the parameter V_X is a matrix whose columns contain the right singular vectors of X.

In order to get the weight of a principal component in a signal, we also need the singular value decomposition of the data that generates the principal component. We shall utilise the defined principal components for Fabrizi et al. 2011 [4], which show the tactile and noxious responses are the first and second principal components of the relevant data respectively. For the aligned epochs, obtained by maximising the normalised inner product of individual epochs with a grand reference electrode constructed of an iterative average of all epochs, we can obtain the relevant principal components, W. Using the corresponding SVD of the defined specific responses, S_W, we can obtain the weight of the i^{th} principal component in X, $a_X^{(i)}$, by:

$$a_X^{(i)} = \frac{U_W^T[,i]X}{D_W[i]}. \tag{2.7.2}$$

Therefore we can obtain the fitted value of the principal component. $\hat{W}_X^{(i)}$, within the signals by:

$$\hat{W}_X^{(i)} = a_X^{(i)} SVD_W[,i]. \tag{2.7.3}$$

In order to obtain the weight of the principal component within the signal we utilised the following procedure.

1. For nociceptive stimuli response the epoch was filtered between 0.5 and 8Hz as this is the frequency band containing most of the principal components energy
2. The signal was aligned by maximising the inner product of the identified waveform and the signal in the ranges outlined taking into account latency jitter.
3. The weight of the principal component was calculated utilising the SVD of the signal and waveform as outlined in Equation 2.7.3.

A visualisation of the waveforms identified with specific tactile and nociceptive responses can be seen in Figures 3.5 and 3.6 respectively.

To ensure a specific response was identified in the region where a somatosensory response is typically evoked - central and central parietal region - we confirmed the presence of the waveform visually; as an automated system was prone to false positives with regards to specific response presence. After detecting a response in this region we analysed for a presence at other electrode sites. Whilst this method requires visual inspection of the data it ensures, for the purposes of testing our model, that we appropriately identify the specific response to stimuli.

Chapter 3

Estimation of neonatal EEG signal parameters

As outlined in the introduction, modelling neonatal EEG is a rapidly expanding field, whose primary focus seems to be on seizure detection. However, little has been done on modelling the development of these signals or the presence of delta brushes within them. In this Chapter we present and fit the presented model for background and somatosensory response EEG signals.

The data set at our disposal is recorded at a higher resolution than previously analysed [4, 6]. Whereas previous research suggests fractional Brownian motion [75], or other processes [33, 41, 44, 70–73], none have made mention of the integrated nature of the signals and that we might require both a long and short range covariance sequence to describe these signals. Inspection of the periodograms illustrates high bias and leakage associated with such processes. This observation could be due to the fact that we are analysing a higher signal resolution than previous research [4, 75]. As illustrated in the methodology section, we will favour cosine tapering to address this integration and leakage [81]. As well as the evident long-term covariance sequence, we present a model with the short term covariance described by an autoregressive moving average process.

We define our fitting procedure for time-homogeneous signal segments using a subset of our data identified as time-homogeneous by an EEG practitioner, such as the inter-burst interval. A time-heterogeneous process is one whose generating mechanism for

the process does not change across time but the covariance structure does, whereas a time-homogeneous process' covariance structure remains constant across time. We test three fitting procedures on the cosine tapered data: fBm only, simultaneous or stepwise fBmARMA(δ,p,q) estimation. The stepwise procedure is performed by estimating the long range covariance structure first, then estimating the short term covariance structure from the resulting spectral residuals.

During the estimation procedure a signal is classified as "adequately" described if it fails to the reject the null hypothesis of the automated detection procedure, outlined in Chapter 2, which states that the residuals come from a χ^2_2 distribution. If this holds then the trend contained within the signal has been removed and therefore adequately described. Upon inspection of the results from the automated detection procedure we find the stepwise model to be the best fitting procedure. From this we improve our model by investigating rejected signals and adapting the long-term covariance sequence to the Matérn process [99]. The inclusion of this provides greater flexibility to our model, whilst still agreeing with previous findings, as fBm is allowable as a limiting case of the Matérn process.

Modelling these signals utilising a time-homogeneous covariance structure can aid in the diagnosis of the type of signal. This is because a time-homogeneous covariance structure will be an inadequate description of a time-heterogeneous process, due to the changing covariance structure. In this analysis such an occurrence would be when a delta brush is present in the signal. As such before progressing to somatosensory response data we investigate the possibility that the time-heterogeneous delta brush is present in some of the rejected signals. Whilst classified by an EEG practitioner as time-homogeneous background/IBI signal segments, the developmental stages of neonatal EEG means that delta brush components could be present and remain unobserved [61, 62]. Utilising the detection procedure outlined in the methodology, and adapted from previous techniques [4], we can identify these segments. Then utilising band-stop filters, we subtract the identified time-heterogeneous delta brush frequencies, and fit our suggested model to the remaining covariance structure. Finally we apply the outlined fitting procedure to somatosensory response data from tactile and noxious stimuli.

Neonatal EEG signals are highly complex, and time-heterogeneous signal components can be present which are not described by delta brush or somatosensory specific activity. This could be an unidentified waveform or an issue from the recording procedure given the highly sensitive and delicate nature of the subjects within our sample. The results of our estimation show that 69% of the overall data set are adequately described by our model; the proportion of adequately described signals increases to 76%, once signals with unidentified time-heterogeneous components were removed.

3.1 Proposed model for neonatal electroencephalogram signals

To start this Chapter we present models that we are going to fit to the data, in an attempt to describe the typical variation in neonatal EEG signals. We propose that we can model time-homogeneous EEG signals, recorded at 2000Hz, using a special case of the unobserved components model [100].

Definition 3.1.1 (**Time-homogeneous Model**). *For an infant, i, and electrode, e, the variation typical in time-homogeneous EEG, $X_{i,e}(t_k)$, can be described by the unobserved components model:*

$$X_{i,e}(t_k) = Z_{i,e}(t_k) + \varepsilon_{i,e}(t_k), \quad \varepsilon_{i,e}(t_k) \sim N(0,1). \tag{3.1.1}$$

Where $Z_{i,e}(\cdot)$ is a time-homogeneous process that adequately describes the covariance structure and $\varepsilon_{i,e}(\cdot)$ is random noise.

The resolution of a signal is defined by the number of times a continuous signal is sampled per second, a 2000Hz signal is sampled every 0.0005 seconds or 2000 times a second. Neonatal EEG contains time-heterogeneous components, and is inherently time-heterogeneous in nature; as such, we need to propose a model that takes time heterogeneity into account.

Definition 3.1.2 (**Additive time-heterogeneous Model**). *For an infant, i, and electrode, e, the variation typical in time-heterogeneous EEG, $X_{i,e}(t_k)$, can be described by the unobserved components model:*

$$X_{i,e}(t_k) = Z_{i,e}(t_k) + Y_{i,e}(t_k) + \varepsilon_{i,e}(t_k), \quad \varepsilon_{i,e}(t_k) \sim N(0,1). \tag{3.1.2}$$

Where $Z_{i,e}(\cdot)$ is a time-homogeneous process that adequately describes the covariance structure, $Y_{i,e}(\cdot)$ is a time-heterogeneous signal component and $\varepsilon_{i,e}(\cdot)$ is random noise.

With this additive time-heterogeneous model, it suggests a simple way in which we can remove the time-heterogeneous component, resulting in a time-homogeneous signal:

$$X_{i,e}(t_k) - Y_{i,e}(t_k) = Z_{i,e}(t_k) + \varepsilon_{i,e}(t_k). \tag{3.1.3}$$

Furthermore, this formulation can be used for any additional signal components such as somatosensory specific responses. A somatosensory specific response as described and estimated in this thesis is one that exhibits the waveforms outlined in Fabrizi et al. [4] rather than a delta brush in response to somatosensory stimuli. There are other somatosensory responses than those estimated however, the focus will lie on noxious and tactile as identified by Fabrizi et al [4]. These models describe time-homogeneous and time-heterogeneous signal components. However, there is more that needs to be added to provide a full model for the data. Whilst background activity is present for all electrodes, the time-heterogeneous components may not be.

The probability of a delta brush, or somatosensory response, occurring in response to stimuli has been well outlined [4], and we can obtain simultaneous delta brush expression locations straightforwardly. Using this we can create an indicator variable, based upon age, that determines the presence of time heterogeneity at an electrode. As such we present the following model we can use to describe the neonatal EEG, which takes into account the behaviour of time-heterogeneous components.

Definition 3.1.3 (**Additive neonatal electroencephalogram model**). *For an infant, i, and a set of electrodes of size n_e, the variation typical in time-heterogeneous EEG, $\underline{X}_i(t_k)$, can be described by the unobserved components model, where:*

$$\underline{X}_i(t_k) = [X_{i,1}(t_k), \ldots, X_{i,n_e}(t_k)]^T, \ \underline{Y}_i(t_k) = [Y_{i,1}(t_k), \ldots, Y_{i,n_e}(t_k)]^T, \tag{3.1.4}$$

$$\underline{Z}_i(t_k) = [Z_{i,1}(t_k), \ldots, Z_{i,n_e}(t_k)]^T, \ \underline{\varepsilon}_i(t_k) = [\varepsilon_{i,1}(t_k), \ldots, \varepsilon_{i,n_e}(t_k)]^T. \tag{3.1.5}$$

Let us define an indicator variable such that:

$$\xi_{i,e}(t_k) = \begin{cases} 1 & \textit{if time heterogeneity is present at electrode } e \textit{ and time point } t_k \\ 0 & \textit{else} \end{cases}, \tag{3.1.6}$$

where

$$\underline{\xi}_i^T(t_k) = [\xi_{i,1}(t_k), \ldots, \xi_{i,n_e}(t_k)]\,. \tag{3.1.7}$$

We define a model for electroencephalogram recordings with n additive time-heterogeneous components by:

$$\underline{X}_i(t_k) = \underline{Z}_i(t_k) + \sum_{j=1}^{n} \underline{\xi}_i^j(t_k)\underline{Y}_i^j(t_k) + \underline{\varepsilon}_i(t_k), \quad \underline{\varepsilon}_i(t_k) \sim N(0, \underline{\mathrm{I}}_{n_e}). \tag{3.1.8}$$

If at all time points $\xi_{i,1}(t_k) = \ldots = \xi_{i,n_e}(t_k) = 0$ we obtain a time-homogeneous model.

We have presented additive models throughout this section however this is not the only way in which this model can be constructed. It is possible that a time heterogeneous component could be multiplicative to the baseline level of activity rather than additive. Such a model is presented in Model 3.1.4.

Definition 3.1.4 (**Multiplicative neonatal electroencephalogram model**). *For an infant, i, and a set of electrodes of size n_e, the variation typical in time-heterogeneous EEG, $\underline{X}_i(t_k)$, can be described by the unobserved components model, where:*

$$\underline{X}_i(t_k) = [X_{i,1}(t_k), \ldots, X_{i,n_e}(t_k)]^T\,,\ \underline{Y}_i(t_k) = [Y_{i,1}(t_k), \ldots, Y_{i,n_e}(t_k)]^T\,, \tag{3.1.9}$$

$$\underline{Z}_i(t_k) = [Z_{i,1}(t_k), \ldots, Z_{i,n_e}(t_k)]^T\,,\ \underline{\varepsilon}_i(t_k) = [\varepsilon_{i,1}(t_k), \ldots, \varepsilon_{i,n_e}(t_k)]^T\,. \tag{3.1.10}$$

Let us define an indicator variable such that:

$$\xi_{i,e}(t_k) = \begin{cases} 1 & \textit{if time heterogeneity is present at electrode } e \textit{ and time point } t_k \\ 0 & \textit{else} \end{cases}, \tag{3.1.11}$$

where

$$\underline{\xi}_i^T(t_k) = [\xi_{i,1}(t_k), \ldots, \xi_{i,n_e}(t_k)]\,. \tag{3.1.12}$$

We define a model for electroencephalogram recordings with n multiplicative time-heterogeneous components by:

$$\underline{X}_i(t_k) = \underline{Z}_i(t_k) \times \sum_{j=1}^{n} \underline{\xi}_i^j(t_k)\underline{Y}_i^j(t_k) + \underline{\varepsilon}_i(t_k), \quad \underline{\varepsilon}_i(t_k) \sim N(0, \underline{\mathrm{I}}_{n_e}). \tag{3.1.13}$$

The presented model class is flexible to other such approaches to the way in which the time-heterogeneous nature of the signal can be applied to a time-homogeneous baseline. Similarly the error - or random noise - component could have another operator

than the additive operator. Furthermore we could change the operator that applies the time-heterogeneous component to the signal dependent upon the time-heterogeneous component in question; i.e. different time-heterogeneous components will require different models such as the additive or multiplicative models presented.

3.2 Procedure to obtain fitted values for the proposed model

Now that we have outlined the model we are going to use, we can proceed to fitting the model to the signals in the data set. To start, we shall provide the descriptive statistics of the data set. The signals have all been recorded at 2000Hz, $\Delta t = 0.0005$, and undergone band-pass filtering with a frequency band of $[0.05, 70]$Hz and a first order Butterworth filter. The filter at 70Hz has been implemented as an anti-aliasing filter, as such we shall implement non aliased spectral density functions. Filtering minimises the contribution of unwanted components of EEG such as breathing or limb movement, known as artefacts. Done as part of the preprocessing procedure it is important that we take this into account when estimating the long term covariance structure so that we gain an accurate estimate.

The 10769 six second recordings were recorded from inpatients at University College London Hospital and have an age range of $[24, 41.57]$ gestational weeks at birth and $[28.43, 45.29]$ gestational weeks at test [6]. Gestational weeks is a measurement based upon the time since the woman's last menstruation. Background segments were obtained from 44 infants, tactile stimuli responses were obtained from 45 infants and noxious stimuli responses were obtained from 47 infants.

Ethical approval was obtained from the University College Hospital ethics committee, and informed written parental consent was obtained prior to each study. The study conformed to the standards set by the Declaration of Helskini guidelines [101]. The infants underwent the clinical test only once, however the length of the recording was longer than the six second interval analysed. When obtaining noxious stimuli response, this was done as part of a clinically required heel lance to obtain a blood sample. Noxious stimuli were not administered solely for the purpose of obtaining data. Tactile

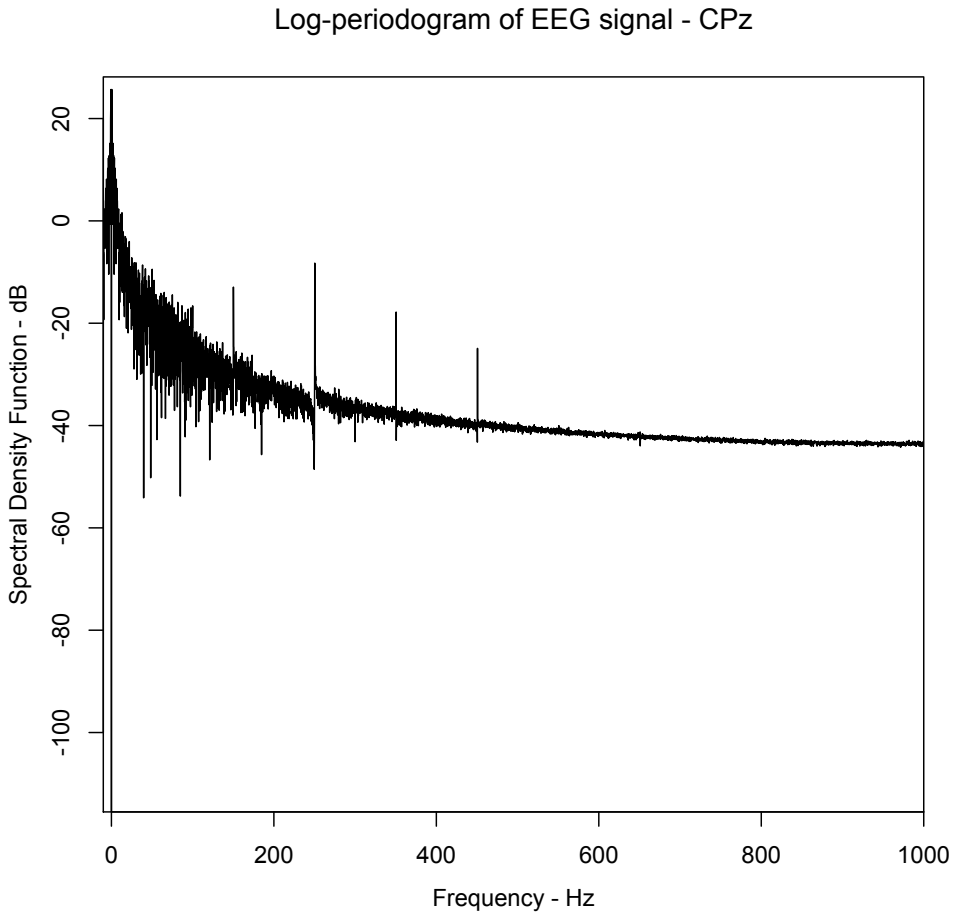
stimulation was applied through use of a tendon hammer against the heel of the infants. Whereas the noxious stimuli could be applied only once, the tactile stimuli was able to be repeated and was in some cases. The multiple six second segments were taken from a singular recording far enough apart to not be affected by the previous tests, as decided by an EEG practitioner.

3.2.1 Time-homogeneous signals

First, we shall visually inspect the periodogram of a signal to gain an initial understanding about the covariance structure to fit by Whittle estimation. The periodogram in Figure 3.1 is indicative of signals within the presented dataset, and shows evidence of leakage, indicating a possible integrated process [66, 67]. Specifically the decay in the variance of the periodogram at high frequencies is evidence of leakage. As observed for nonstationary infinite random functions that are strongly nonstationary, the periodogram is strongly biased and the periodogram illustrates a large amount of bias and the variance of the periodogram is decreasing as the frequency increases [81, 102]. This large amount of leakage is indicative of a disjoint signal and a possible integrated process, this shall be considered and investigated further. tWe cannot obtain a reliable estimate of the covariance structure from the periodogram when we have large bias, as evidenced in Figure 3.1, since the fit at higher frequencies will be affected.

The fit of a spectrum will be assessed by the automated detection procedure outlined in the methodology. Whilst the description at all frequencies is important, the suitability of the fit will be assessed based upon the fitting in the passband which in the presented dataset is the interval $[0.05, 70]$Hz; this raises a question about how we proceed in the analysis. Since the leakage within the presented signals only affects the higher frequencies, we could limit the estimation only to the passband in the raw periodogram, and fit only a long-term covariance structure. This is an entirely unsuitable and unacceptable approach to analysing the data, as we are ignoring the effect of 93% of the frequencies on the spectrum; therefore, we shall fit the spectrum across all frequencies. Whist assessing the fit only in the passband is similar to ignoring the other frequencies, we are giving analytical weight to the interval of interest in EEG applications, and incorporating the evident need for describing these high frequencies that effect the signal's structure. In

Figure 3.1: Periodogram representative of neonatal electroencephalogram signal properties



This plot illustrates the raw periodogram of a signal from within the presented data set. This behaviour is evident throughout the data set and is possibly indicative of an integrated process. This evidence is seen in the variance of the periodogram and how it reduces at higher frequencies. Attempting to utilise Whittle estimation would result in substandard estimates as the construction of the Whittle likelihood would result in a poor fit at the higher frequencies, as such this needs to be addressed before estimation.

order for us to do this, we need to reduce the leakage in the estimator; this shall be done utilising the Hanning window [81], as outlined in the methodology.

The Hanning window - or full cosine taper - reduces leakage by reducing the sidelobes of the Fejér kernel, which causes a transfer in energy from one region to another. The

Hanning window also smoothes the way in which the signal goes to 0 at the extremities reducing signal discontinuity problems with the FFT.

We shall begin building the time-homogeneous model structure by analysing 1358 signals identified as time homogenous background/inter-burst intervals by an EEG practitioner, dependent upon the age of the infant. Previous research has utilised fractional Brownian motion (fBm) as a description of the long-term covariance, so this shall be the starting point for our model building. The fractional Brownian motion spectrum is defined by [67]:

Definition 3.2.1. *The spectral density function of a fractional Brownian motion covariance structure is defined as:*

$$S_X(f) = \frac{\sigma^2}{|f\Delta t|^{2\delta}}, \quad \delta = H + 0.5,\ \delta \in [0.5, 1.5]\,. \tag{3.2.1}$$

Where H is measure of the self similarity of the process and Δt is the sampling rate.

Which has associated covariance structure [67, 80]:

Definition 3.2.2. *The covariance of a fractional Brownian motion process is defined as:*

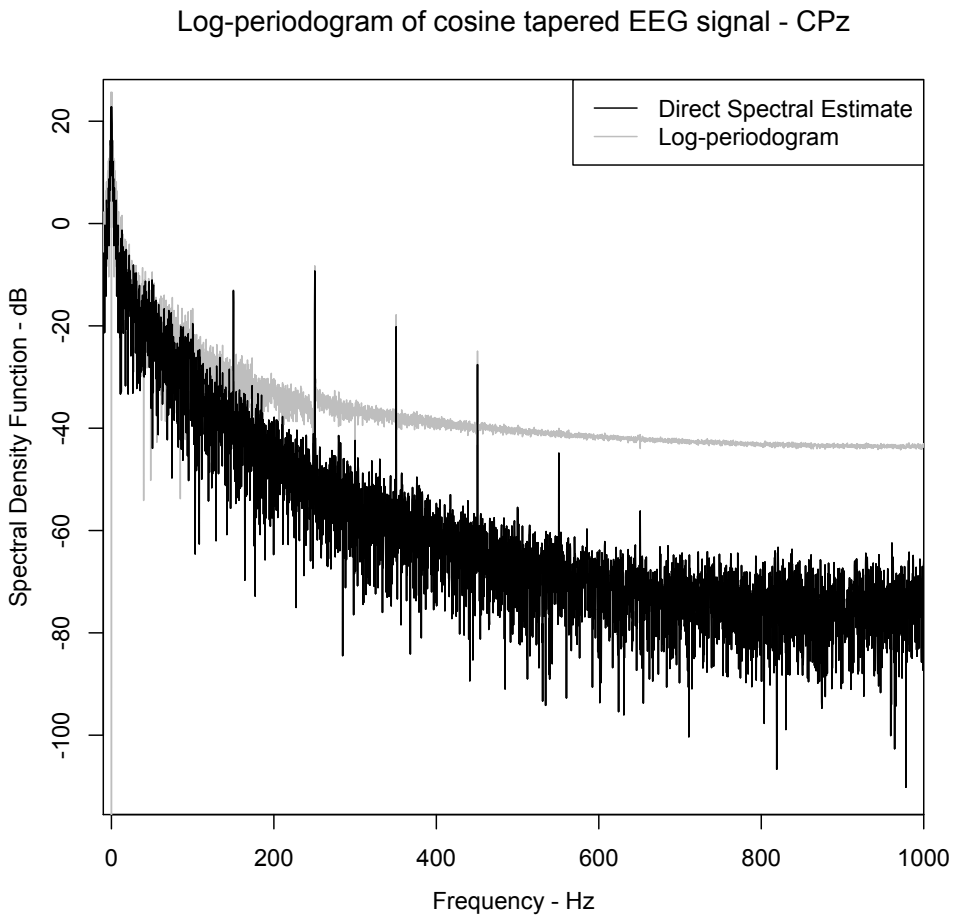
$$\text{Cov}\,\{t, s\} = \frac{C\sigma^2}{2}\left(|t|^{2H} + |s|^{2H} - |t-s|^{2H}\right),\ H = \delta - 0.5,\ \sigma \in \mathbb{R}. \tag{3.2.2}$$

Where H is measure of the self similarity of the process and C is defined by,

$$C = \frac{\Gamma(2-2H)\cos(\pi H)}{\pi H(2H-1)}.$$

This illustrates the non-stationarity of fractional Brownian motion as the covariance sequence is dependent upon time. Whilst fBm is a non-stationary process we can fit estimates using Whittle estimation because it has a time homogeneous structure and a defined spectral density function. It is a time-homogeneous process because the structure does not change with respect to time. If we difference fBm a sufficient number of times we obtain a stationary process. This cannot be done for the time heterogeneous stochastic process $X_t = \sigma_t \epsilon_t$, where σ_t is a deterministic time-varying function and ϵ_t is stationary.

Figure 3.2: Cosine tapered periodogram representative of neonatal electroencephalogram signal properties



This plot shows the direct spectral estimate of the signal presented in Figure 3.1 utilising the full cosine taper. As can be seen the variance in the periodogram remains constant due to reduced bias and leakage. Furthermore the taper used is addressing problems with the FFT assuming the signal is infinitely repeating and as such the sharp discontinuity between the start and end of the signal. We can utilise Whittle estimation effectively using this direct spectral estimate; however we must take into account spikes seen at 50Hz and resonances of 50Hz. This is due to electrical interference and if left unaddressed would affect the estimate obtained. Due to the construction of the Whittle likelihood we would see the estimated MatérnARMA class spectrum "dragged" to account for the contribution of these frequencies to the likelihood.

At this point it is important to note another component of the periodogram that could affect estimates; electrical interference [1–3]. As mentioned the signals have undergone band-pass filtering, which shall be reflected in the estimation of the long-term covariance sequence, however they have not undergone notch filtering to remove electrical interference. In the UK, this is seen by spikes in the periodogram at 50Hz, and also at resonant frequencies. As such, we shall omit 2.5Hz either side of 50Hz and its resonances to avoid our estimate being affected. This results in the estimate we obtain no longer being parametric, instead we obtain a semi-parametric estimate of the spectral density function [102].

The resulting estimates are semi-parametric due to the removal of frequencies in the estimation of the parameters. As a result of this removal of frequencies we are assuming that the underlying generating mechanism of the signal behaves the same in the removed frequencies; we are not fitting the parametric model to all of the data. We are aware that they do not behave the same due to electrical interference, however the assumption is with respect to the process' generating mechanism.

A purely fBm estimation is unsurprisingly a completely unsatisfactory description with the automated detection procedure - Section 2.3 - finding only 1.04% of the signals are adequately described. The untapered periodogram is illustrative of the possible need for a short term covariance sequence, in addition to the long-term covariance. This is because the spectral estimate is not showing a simple decay but more a complex high frequency structure. In order to address this we shall implement the autoregressive moving average process as a description of the short range covariance, which has the spectral density function:

Definition 3.2.3 (**Autoregressive Moving Average Spectral Density Function**)**.** *The spectral density function of an autoregressive moving average process of order (p,q) is defined as:*

$$S_X(f) = \sigma^2 \frac{\left|1 - \sum_{k=1}^{q} \theta_k e^{-i2\pi i f \Delta t}\right|^2}{\left|1 - \sum_{j=1}^{p} \phi_j e^{-i2\pi j f \Delta t}\right|^2}.$$

Where $(\underline{\theta}, \underline{\phi})$ are the parameters of the moving average and autoregressive components

respectively.

We determine the order of the ARMA(p, q) process by implementing a grid search procedure based upon Akaike's information criterion corrected for finite size (AICc). The AICc of a model is defined by [103]:

Definition 3.2.4 (AICc). *Akaike's information criterion corrected for finite sample size is defined as:*

$$AICC = 2n - 2\ell^{(w)} + \frac{2n(n+1)}{M-n-1}, \quad \ell_W = -\sum_{k=0}^{M}\left(\frac{\hat{S}_X^{(d)}(f_k)}{S_X(f_k;\underline{\theta})} + \log\{S_X(f_k;\underline{\theta})\}\right).$$

Where $\ell^{(w)}$ is the Whittle likelihood defined in Equation 2.4.2, $n = ||\underline{\theta}||$ and $\underline{\theta}$ is the parameters used in the objective function $S_X(f_k;\underline{\theta})$.

There are two ways in which we investigated fitting the fBmARMA model to the data presented, simultaneous and two step estimation. These titles refer to how we fit the long term and short term covariance sequence - before both estimations we have implemented the Hanning window to obtain a direct spectral estimator of the spectrum.

There are two ways in which we investigated fitting the fBmARMA model to the data presented, simultaneous and stepwise. These titles refer to how we fit the long term and short term covariance sequence - before both estimations we have implemented the Hanning window to obtain a direct spectral estimator of the spectrum.

In the simultaneous procedure we estimate both the long and short range covariance sequence at the same time utilising the spectral density function, Definition 3.2.8, the Butterworth passband filter with frequency cutoffs at 0.05Hz and 70Hz and Whittle estimation. We obtain the best description of the process utilising an AICc reduction, limiting the ARMA(p,q) component to a maximum order of (3,3), we chose the model with the lowest AICc as calculated from Equation 3.2.4.

In the stepwise procedure we first estimate the long term covariance sequence utilising the fBm spectral density function and the Butterworth passband filter to take into account the filtering that the signal has undergone during preprocessing. Once we have obtained the long term covariance estimate we obtain the spectral residuals. Following

this we estimate the short term ARMA(p,q) utilising the same AICc approach as the simultaneous model.

Having tested both, we proceed with the stepwise procedure over the simultaneous procedure; as the automated detection procedure outlined in Section 2.3 finding 75.09% and 23.32% of the signals are described by the fitted spectra respectively. The image in Figure 3.3 illustrates the effectiveness of the two step fitted spectrum, and the frequencies over which we have performed the estimation.

The stepwise estimation procedure induces bias because we are not allowing all parameters to change simultaneously. The ARMA(p,q) estimation is performed upon a set of residuals where the estimated long term covariance sequence has been described. Whilst this method of estimating the spectra induces bias in our estimates, it is evidently a better way to obtain descriptions of the signals based upon the proportion of signals described. It appears also that the simultaneous procedure had convergence problems, indicated by the values of $\hat{\delta}$ being equal to 0.5 or 1.5. since these are the boundary values of δ in fractional Brownian motion.

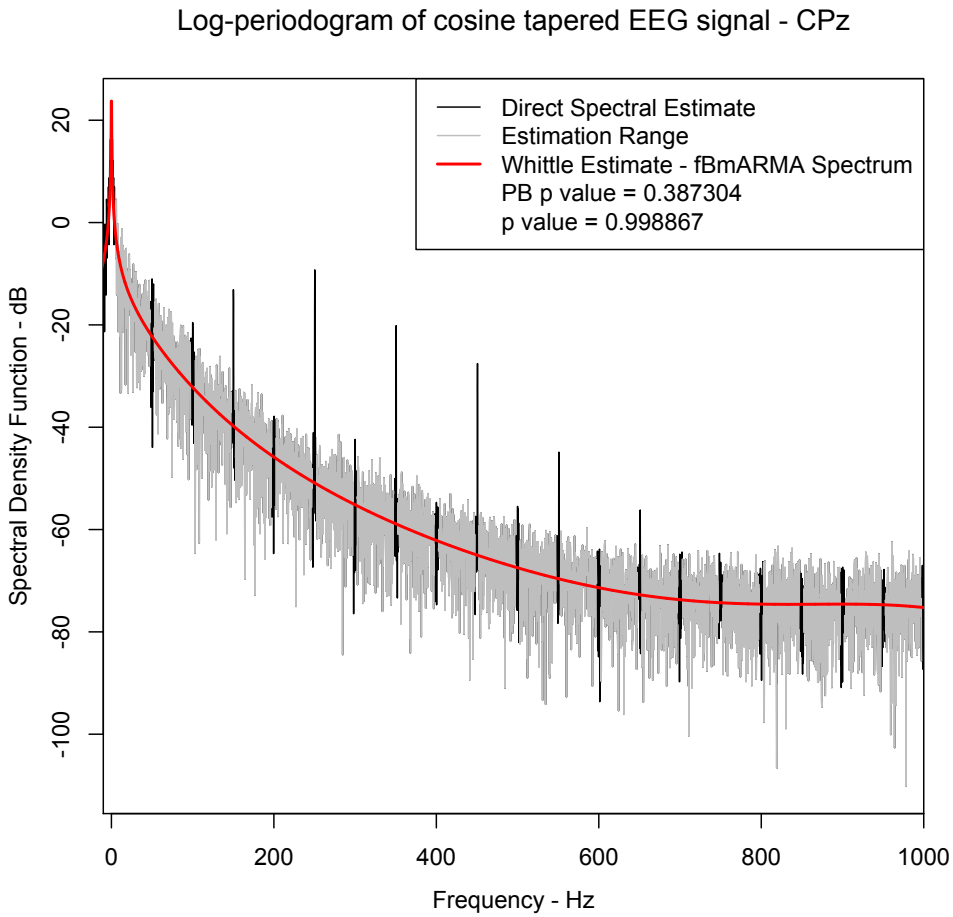
Although we have described 75% of the background signals identified as time homogeneous, it would be prudent to assess why we have not been able to describe the other 25%; since after fitting spectra to the background EEG signals, we progress to the rest of our data set: somatosensory response data.

First we shall look at whether fractional Brownian motion is a suitable description of the long range covariance, or does the limitation of the Hurst parameter value result in inadequate descriptions of the signal's covariance. We shall look at utilising the Matérn covariance structure [99] to explain signals not describable by fractional Brownian motion.

Definition 3.2.5. *The covariance sequence of a Matérn process is defined as:*

$$\operatorname{Cov}\{t,s\} = \frac{\vartheta^2 \pi^{\frac{1}{2}}}{2^{\nu-1}\Gamma\left(\nu+\frac{1}{2}\right)\beta^{2\nu}} \left(\beta|\tau|\right)^{\nu} K_{\nu}\left(\beta|\tau|\right) \tag{3.2.3}$$
$$\tau = t-s,\ \nu \in [0,\infty],\ \beta \in (0,\infty],\ \vartheta \in \mathbb{R}$$

Figure 3.3: Illustration of fitting fBm(δ)ARMA(p,q) process



This plot illustrates the result of our estimation procedure and the frequencies over which we have performed the estimation. As can be seen in the plot the estimated spectrum is a good description of the sample variance of the Fourier transform, and this is supported by the associated p-values. The p-values relate to the ADP implemented to assess fit where the null hypothesis is an adequately described signal. We have removed frequencies either side of 50Hz and its resonances to avoid the effect of the electrical interference spikes on our signal. This results in a semi-parametric signal due to the assumption that the frequencies behave the same in the signal's generating mechanism as in our model.

Where $K_H(\cdot)$ is the modified Bessel function of the second kind.

The Matérn process is a stationary process, whose limiting behaviours are exceptionally useful in time series analysis. These behaviours become clearer if we reparameterise to

utilise the parameters outlined in Definition 3.2.2 and by looking at the corresponding spectral density function.

Definition 3.2.6. *The covariance sequence of a Matérn process reparameterised to utilise parameters similar to fractional Brownian motion is defined as:*

$$Cov\{t,s\} = \frac{\sigma^2}{2^{H-1}\Gamma\left(H+\frac{1}{2}\right)d^{2H}}\left(d|\tau|\right)^H K_H\left(d|\tau|\right), \tag{3.2.4}$$
$$\tau = t-s,\ H \in [0,\infty],\ d \in (0,\infty],\ \sigma \in \mathbb{R}.$$

Definition 3.2.7. *The spectral density function of a Matérn process reparameterised to utilise parameters similar to fractional Brownian motion is defined as:*

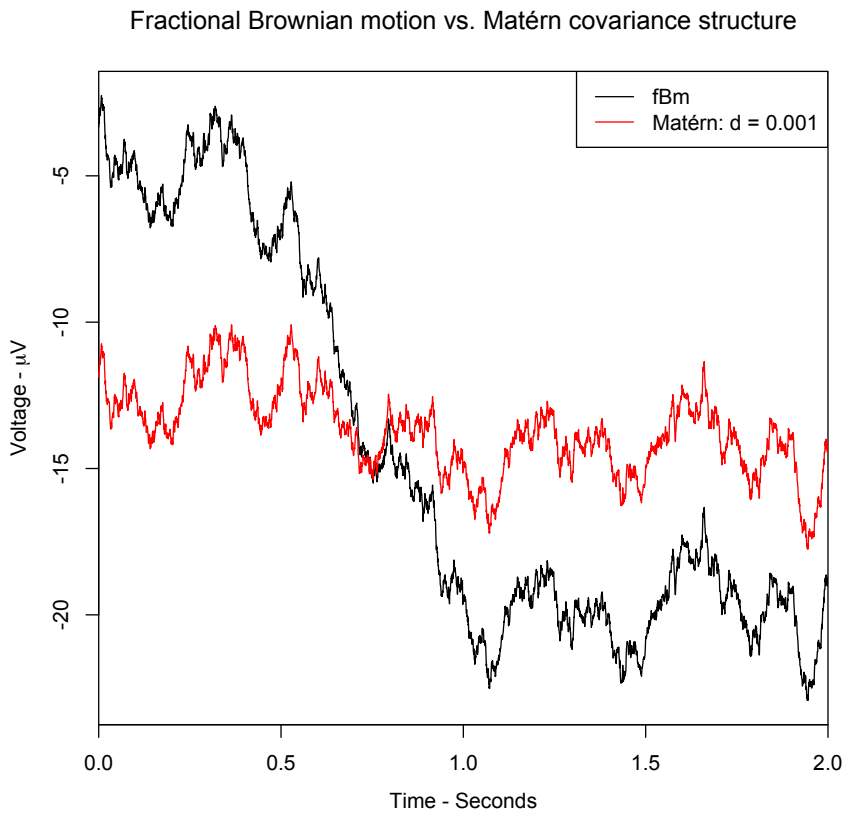
$$S_X(f) = \frac{\sigma^2}{\left(|f\Delta t|^2 + d^2\right)^{\delta}}, \tag{3.2.5}$$
$$\delta = H + 0.5,\ \delta \in [0.5,\infty],\ d \in (0,\infty].$$

Consider the range, or distance, parameter d, as $d \to \infty$ the limiting behaviour is white noise, $S_X(f) \to \sigma^2$. However as $d \to 0$ the limiting behaviour is fractional Brownian motion, $S_X(f) \to \frac{\sigma^2}{|f\Delta t|^{2\delta}}$. Figure 3.4 illustrates the effect the distance parameter has on the time domain representation of a signal with a distance parameter $d = 0.001$.

We can see from Figure 3.4 that the distance parameter makes the fractional Brownian motion process behave more like white noise, i.e. reduces the correlation of the increments. The formulation of the Matérn covariance in Definition 3.2.6 is not the standard parameterisation, however we have utilised the parameters from Definition 3.2.2 to illustrate the similarity between these two processes. One of the benefits of the Matérn process is that the smoothness parameter is not restricted at the upper bound, as under fractional Brownian motion the δ parameter is restricted to the range $[0.5, 1.5]$.

Applying this Matérn extension to the long-term covariance structure in our stepwise estimation model describes another 7% of the overall time homogeneous signals, and 28% of the previously rejected signals. As such the Matérn flexibility is beneficial for signals inadequately described by fBm. Therefore if the fBmARMA(δ, p, q) parameterisation is rejected by the automated detection procedure, we shall re estimate the signal using MatérnARMA(δ, d, p, q).

Figure 3.4: Fractional Brownian motion vs. Matérn covariance structure



The Matérn covariance sequence allows fBm as a limiting case, by letting $d \to 0$*, and as such is useful as a description of a process, as it allows for a wider range of signals to be described in comparison to just using an fBm covariance structure. The spectral density functions are also similar between fBm and Matérn, with the Matérn SDF containing an extra parameter - the range parameter. This plot shows the effect of the additional range parameter in the Matérn covariance function on the time domain structure of a signal. In comparison to the fBm signal - black line - the signal with distance parameter* $d = 0.001$ *- red line - behaves less like a random walk and stays closer to the mean than the fBm. As the distance parameter increases the signal will behave more like a white noise process, as the distance parameter decreases the signal will behave more like a fractional Brownian motion process.*

Definition 3.2.8. *The spectral density function for a process with a Matérn long range covariance and autoregressive moving average short range covariance is defined as:*

$$S_X(f) = \frac{\sigma^2}{(|f\Delta t|^2 + d^2)^{\delta}} \frac{\left|1 - \sum_{k=1}^{q} \theta_k e^{-i2\pi i f \Delta t}\right|^2}{\left|1 - \sum_{j=1}^{p} \phi_j e^{-i2\pi j f \Delta t}\right|^2}, \tag{3.2.6}$$

Where $(\underline{\theta}, \underline{\phi})$ are the parameters of the moving average and autoregressive components respectively and $\delta = H+0.5, \quad \delta \in [0.5, \infty]$, $d \in (0, \infty]$.

This is one of the first pieces of research to use the Matérn covariance structure to describe neonatal electroencephalogram signals. If we can limit the use of Matérn to describe a signal's long range covariance, we might be able to gain an insight as to why these signals require this covariance structure.

Before proceeding to apply this procedure to somatosensory response data, we shall try one more analysis technique to the remaining signals yet to be described: time-heterogeneous signals.

3.2.2 Time-heterogeneous signals

Although the signals analysed up to now have been identified as time-homogeneous by an EEG practitioner, it is possible that misclassification occurred. Furthermore, a common component of neonatal background EEG is the time-heterogeneous delta brush. Therefore using the detection procedure outlined in Chapter 2 we identified signals that contained delta brush activity and then estimated them appropriately. As a result we described another 3% of the presented signals without stimuli.

Below we illustrate the procedure, by which we fit the time-heterogeneous model to signals identified as containing time-heterogeneous signal segments by an EEG technician.

We identify time-heterogeneous regions in two manners dependent upon the classification of the signal being analysed. If the signal segment is attempting to detect a response to a somatosensory stimuli then it has been focused around the time of experiment

- i.e. the time when the stimulation was administered. If we are looking for time heterogeneous regions in any signal we are looking at the scalogram for one type of time heterogeneity in particular - delta brush activity. We have explained the delta brush in detail in the introduction and the estimation and identification of regions where they are present is done by testing the proportion of energy - proportion test - within the defined delta brush regions - [0.5,1.5]Hz and [8,25]Hz - against a time homogeneous baseline constructed from our adequately estimated time homogeneous signals.

Non-specific delta brush somatosensory response signal estimation

Once we have identified the regions containing delta brush activity, we can obtain the delta brush component by implementing band-stop filters on the delta brush frequencies. We then "remove" the delta brush activity from the signal by implementing a stop-band filter on the delta brush region of activity then subtracting this from our initial signal as prescribed in Model 3.1.3. Once subtracted from the relevant segments, this should allow the time-homogeneous model to fitted. The only difference between the model for the time-heterogeneous segments is that we allow the MatérnARMA covariance from the beginning, due to the removal of frequencies. We present the estimates of a signal with a delta brush identified as present in Figure 3.5. This figure illustrates the effectiveness of the estimation procedure in the segments, and raises questions as to how we determine the suitability of a signal's description, when the signal has multiple segments.

Specific somatosensory response identification and signal estimation

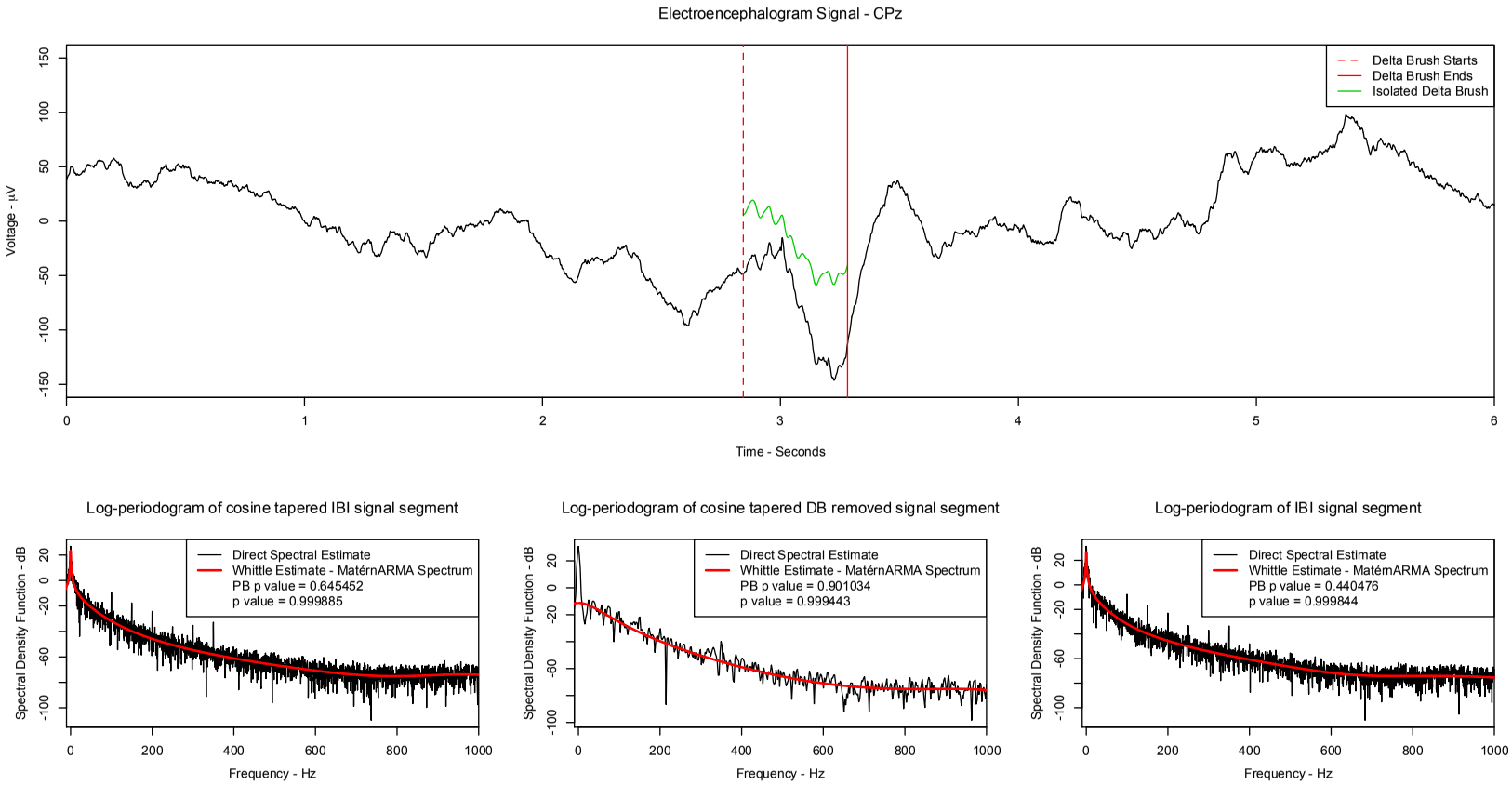
Once signals which express specific somatosensory responses to nociceptive and tactile stimuli are identified we can fit the model to these signals. We illustrate the time-heterogeneous estimation procedure with a nociceptive somatosensory response in Figure 3.6.

If a signal is identified as containing time-heterogeneous components we split the signal into n time-homogeneous segments and m time heterogeneous segments. We estimate the time-homogeneous segments as outlined in the previous section and test the suitability of the estimated spectrum utilising our automated detection procedure described in Definition 2.3.2. With respect to the time heterogeneous segments, we first attempt to

"remove" the time-heterogeneous components from the segments resulting theoretically in a time-homogeneous segment. By assuming the components are orthogonal these can be extracted non parametrically using PCA [105]. From there we implement the time-homogeneous estimation procedure as outlined.

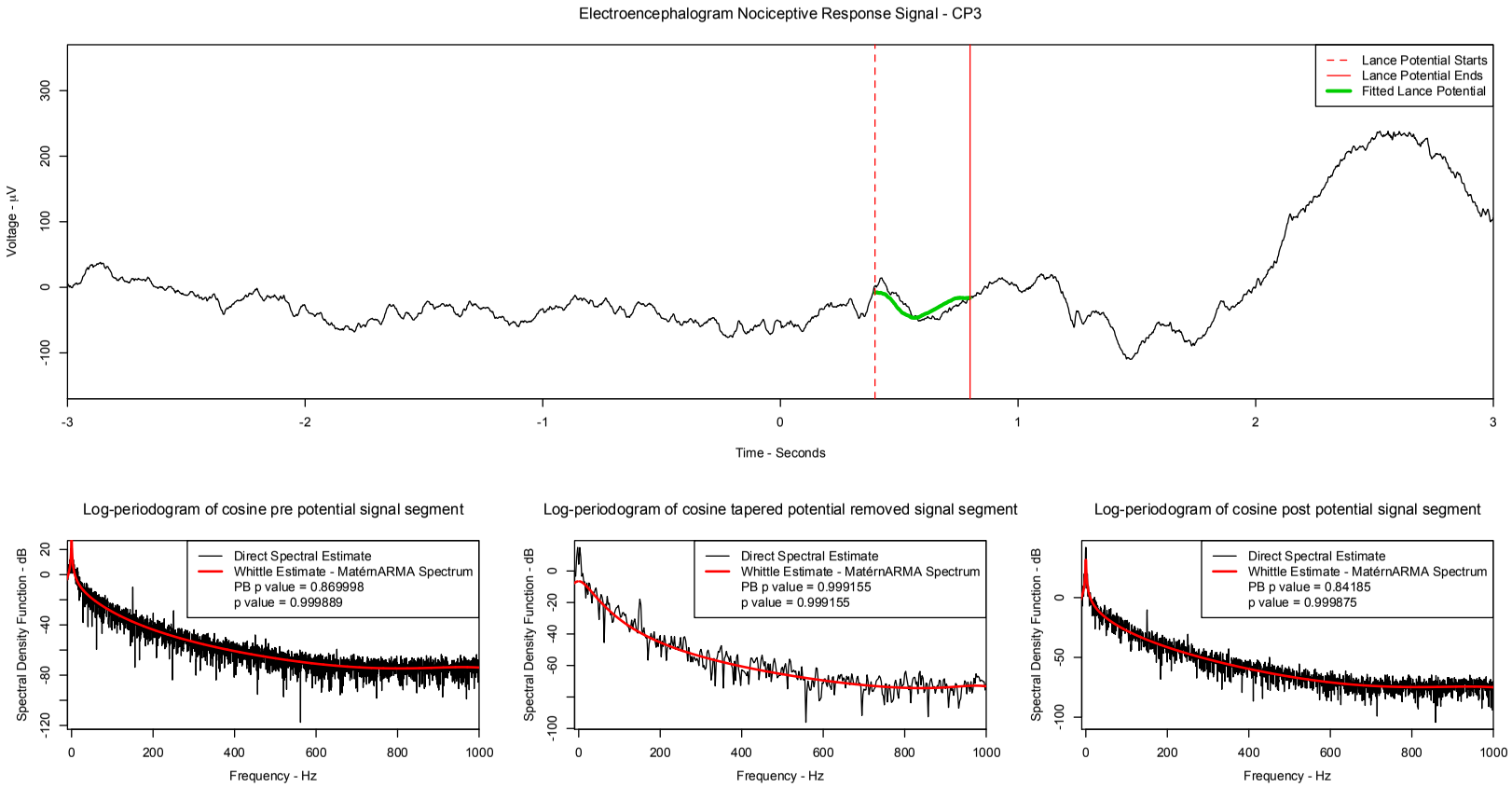
We intend to perform analysis on these estimates in the next chapter and so we want all of the segments to be adequately described by our fitting. Therefore if one segment rejects the null hypothesis outlined in the ADP procedure - Definition 2.3.2 - then we will classify the entire signal as having failed to be described. This will be punitive with respect to the total number of signals within our data set that we have described using our model, however it ensures that we can utilise these estimates freely in resulting analyses.

Figure 3.5: Illustration of time-heterogeneous estimation procedure - delta brush activity



The top plot shows the time domain representation with the interval outlined - between red lines - and the isolated delta brush - green line. We have done this through the use of a bandstop filter constructed to isolate the delta brush activity which is then removed to estimate the underlying structure. The bottom plots show the estimated SDFs for each region where all segments are adequately described.

Figure 3.6: Illustration of time-heterogeneous estimation procedure - somatosensory specific activity

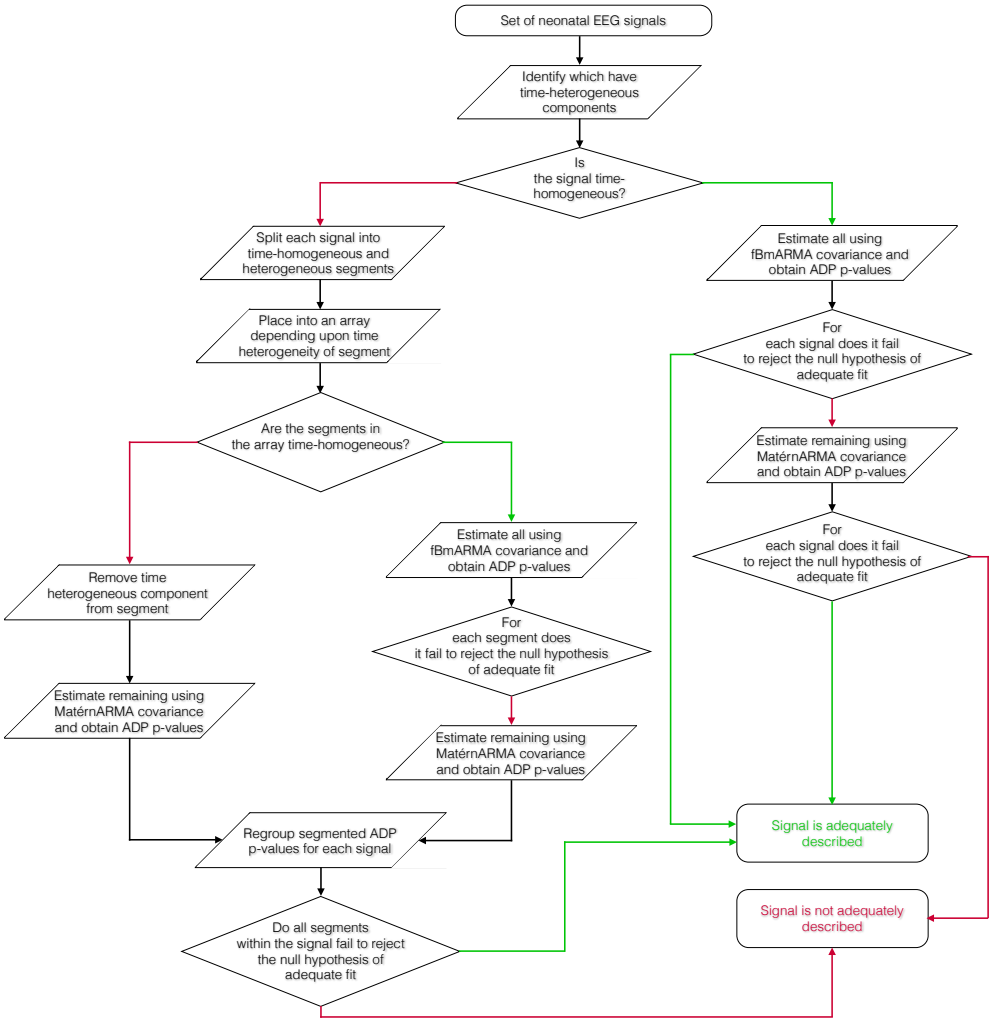


This plot shows a signal with a specific reaction to a touch stimuli - green line in the outlined red region. The bottom plots show the estimated SDFs for each region - pre, during and post. The pre and post activity regions are time homogeneous estimations as outlined previously. With regards to estimating time-heterogeneous segments we "remove" the time-heterogeneous activity to theoretically leave a time homogeneous model which we can estimate using the outlined method. All segments have been adequately described by our process

3.2.3 Estimation procedure

Throughout this chapter we have illustrated the methods by which we obtained parameter estimates for our model, utilising signals recorded with no stimuli administered. The estimation procedure was refined in order to obtain a methodology to apply to future signals, such as signals recorded with a somatosensory stimulus response. Taking into account the fact that we wish to limit the use of the Matérn covariance, such that we can attempt to identify why this structure is necessary. The flowchart in Figure 3.7 is an algorithmic overview of the procedure taken to estimate the parameters of the model presented in Model 3.1.

Figure 3.7: Flowchart describing the estimation process for neonatal EEG covariance structure



3.3 Results and Conclusions

Using the outlined model we attempted to estimate the parameters from the signals. Noxious and tactile response data underwent delta brush detection, to define non-specific responses, before analysing for the presence of waveforms associated with specific response in the somatosensory region of the brain. We assume the observed signal is an aggregation of unobserved components. By assuming the components are orthogonal these can be extracted non parametrically using PCA. [105]. Once the characteristics of the signal were obtained, the parameters of the model were fitted; the suitability of estimates was ascertained by the automated detection procedure, $\alpha = 0.05$, with false discovery rate analysis to reduce type II errors - Definition 2.3.3.

In the entire data set 69% of the signals were adequately described by the presented model, Definition 3.1.3:

$$\underline{X}_i(t_k) = \underline{Z}_i(t_k) + \underline{\xi}_i^T(t_k)\underline{Y}_i(t_k) + \underline{\varepsilon}_i(t_k), \quad \underline{\varepsilon}_i(t_k) \sim N(0, \underline{\mathrm{I}}_{n_e}). \tag{3.3.1}$$

Upon inspection of the rejected signals, to determine why the model rejects 31%, problems were noticed with the time-heterogeneous components. We have detected and fitted two departures from time homogeneity: delta brushes and somatosensory specific responses. It appears that in some of the rejected signals, 9%, there are unidentified time-heterogeneous components. Therefore, there is an argument that we cannot use these signals to assess the model validity.

If we remove these unsuitable signals from the analysis, the presented model describes 76% of the suitable signals. However, given the flexibility of the presented model, it is possible that if we were to identify the time-homogeneous structure present, as in the case of delta brushes, these signals could be described.

Previous models have focused solely upon the pass-band when estimating and analysing neonatal EEG. Whilst understandable due to its analytical importance, ignoring or sampling to remove the higher frequencies ignores their contribution to the structure of the signal. The presented model was fitted across all frequencies - avoiding electrical interference - and assessed for suitability of fit within the passband. This approach takes into

account the full structure of the signal whilst placing importance upon the frequencies of interest to EEG practitioners.

Whilst we do not know the motivations behind the previous estimations performed, analysing in low resolution allows for faster estimation to be performed within the frequency domain. Furthermore it has been common in previous research to disregard frequencies outside the analytical interval of interest. This could be because the components being detected are easily identifiable at low resolution, however if higher resolutions are required we have to take into account the short range covariance.

However, the presented model and estimation procedure is not without flaws. We have introduced bias into our estimates by performing the estimation in two steps, and most importantly we have ignored the multivariate nature of the data. The presented signals were estimated univariately due to computational intensities, to estimate a set of 17 electrodes multivariately, we would have to estimate each cross-spectrum for a total of 153 fitted spectra for each set of recordings. Taking into account the number of recordings that we have this corresponds to around 90000 fitted spectral density functions. Whilst this is the preferred way in which to estimate these signals, taking into account the length and resolution of the signals, it is wholly impractical to perform and would be a task for an analytic team.

The outlined approach is a flexible parameterisation of neonatal EEG; furthermore we could adapt this model to take into account other components not analysed in this research. The inability to describe the remaining 24% of the dataset could be due to an unidentified time-heterogeneous component as we only took delta brush activity, nociceptive and tactile responses into account when fitting. However, given the time-heterogeneous characteristics we could extend our model to take these into account. Overall 76% of the signals adequately described in the passband is evidence that the presented model can be used to describe neonatal EEG signals.

Chapter 4

Analysis of estimated model parameters

Having obtained satisfactory estimates for 76% of suitable signals within our data set, we now proceed to analyse the parameter estimates. Our eventual goal is to construct a model that can reproduce the variability of the parameters seen in the data.

Several authors have found indications that the EEG characteristics of prematurely born infants evolve over time, such as Stevenson et al. [75]. We, to be specific about the nature of change, will attempt to capture the change-point indicating the time point of evolution within the presented dataset. We do so by comparing the parameters of signals with somatosensory specific responses and up to two delta brush components. We are utilising these time heterogeneous segments to look at whether the time heterogeneity "triggers" a change in the underlying parameters of the time-homogeneous segments surrounding these regions. The values outside of these regions, and therefore baseline activity, were tested to see whether there was a change in the parameters. From visual inspection of graphs, Wilcoxon signed ranks tests and Kolmogorov-Smirnov tests, we were able to determine that there was no change in the parameters. As a consequence of this conclusion, we reduced the number of parameters per signal, such that only one set of parameters is present per electrode per signal. From this we investigate the properties of, and relationships between, the reduced estimated parameters. Finding that the estimated values of the variance has a bimodal distribution, and the value of this has an apparent relationship to the values of the long and short term covariance structure.

Taking these observations we construct three binary parameterisations based upon the bimodality of the variance, parameterisation of the long-term covariance structure and parameterisation of the short term covariance structure. These binary parameters are described using generalised linear models. These reflect the relationships seen in the parameters, with an apparant link between the parameterisations of long and short term covariance structures, depending upon the value of the variance. Additionally we find that the parameterisation of the long-term covariance has an effect on the parameterisation of the short term covariance, which is intuitive given the stepwise estimation procedure from which these estimates have been obtained. Utilising the binary parameter models we construct several multivariate normal distributions, dependent upon the predicted binary parameters.

Finally, we model probability of delta brushes occurring and the number of electrodes expressing delta brush activity, both spontaneously and as a response to stimuli. With regards to the probability of occurrence we find that age at birth affects the probability of occurrence, not just age at test. We find that not only do infants before 35-37 weeks age at test express delta brush responses at more electrodes, in response to noxious or tactile stimuli, but also infants tested close to the age of birth. We notice also, that in response to tactile stimuli, the expected number of electrodes expressing delta brushes is in a smaller range of values. Finally, we have identified that spontaneous delta brush activity is expected at more electrodes when infants are tested close to birth. These findings are concurrent with current theory, but utilise the age pairing of birth and test instead of just age at test.

4.1 Time-heterogeneous parameters

Electroencephalogram signals are by definition time-heterogeneous [1–3]. Previously, short time intervals have been simulated with different parameters, in order to recreate this time heterogeneity [75]. However, we have been able to adequately describe up to six seconds with a time-homogeneous model or a segmented approach to take into account time heterogeneity within the signal; with the descriptions working for multiple segments within one time course. A point at which the baseline parameters might change, is after a delta brush or a specific response. Utilising the estimated parameters,

we shall now analyse the parameter values from time-heterogeneous estimates to test whether the parameters change after the observation of a time-heterogeneous component. Signals with up to two delta brushes or a specific response were analysed, as the sample size becomes too small for signals with more than two delta brushes present.

We investigate whether the parameters are the same by visual inspection and Wilcoxon signed ranks test [104]. Although estimates from the Whittle likelihood are asymptotically normal, this method will allow for departures from this due to the sample to be accounted for. We would expect the estimates from Whittle estimation to be asymptotically normal, but only if the underlying parameter distribution is normal. Across the EEG signals the distribution of the parameter values is not a delta function, thus the observed estimates have a distribution which is a convolution of the parameter distribution in the infant population. Observed parameters have a distribution that is a convolution between the estimation error distribution and the original distribution of the sample parameters. So whilst the error distribution under Whittle estimation is normal, the estimated parameters distribution might not be due to the original distribution of the parameters. Because of the number of possible autoregressive moving average combinations, we reduce the dimensionality by analysing the first and second principal components of these estimates [105].

We will now text whether the parameters of the time-homogeneous segments before and after a time-heterogeneous component are the same. This shall be done in two ways; first we shall look to see whether they are the same using the paired non-parametric Wilcoxon signed ranks test. Second, we shall look at whether the distribution of the parameters is the same. This is to ascertain, if they are not equal, could they be random realisations from the same distribution.

Table 4.1.1: Comparison of estimated spectrum parameters pre and post delta brush utilising Wilcoxon signed ranks and Kolmogorov-Smirnov tests

Test	σ	δ	ARMA PC1	ARMA PC2
Wilcoxon	0.8476	0.1847	0.7716	0.3401
Kolmogorov-Smirnov	0.4806	0.2202	0.5316	0.5316

Table 4.1.2: Comparison of estimated spectrum parameters pre and inter two delta brush utilising Wilcoxon signed ranks and Kolmogorov-Smirnov tests

Test	σ	δ	ARMA PC1	ARMA PC2
Wilcoxon	0.6053	0.5538	0.8849	0.0694
Kolmogorov-Smirnov	0.7417	0.5862	0.7911	0.1661

Table 4.1.3: Comparison of estimated spectrum parameters inter and post two delta brush utilising Wilcoxon signed ranks and Kolmogorov-Smirnov tests

Test	σ	δ	ARMA PC1	ARMA PC2
Wilcoxon	0.7495	0.2506	0.7933	0.1587
Kolmogorov-Smirnov	0.9918	0.2794	0.8790	0.2469

Table 4.1.4: Comparison of estimated spectrum parameters pre and post specific reaction utilising Wilcoxon signed ranks and Kolmogorov-Smirnov tests

Test	σ	δ	ARMA PC1	ARMA PC2
Wilcoxon	0.7624	0.7004	0.5043	0.8069
Kolmogorov-Smirnov	0.4189	0.8336	0.6679	0.9010

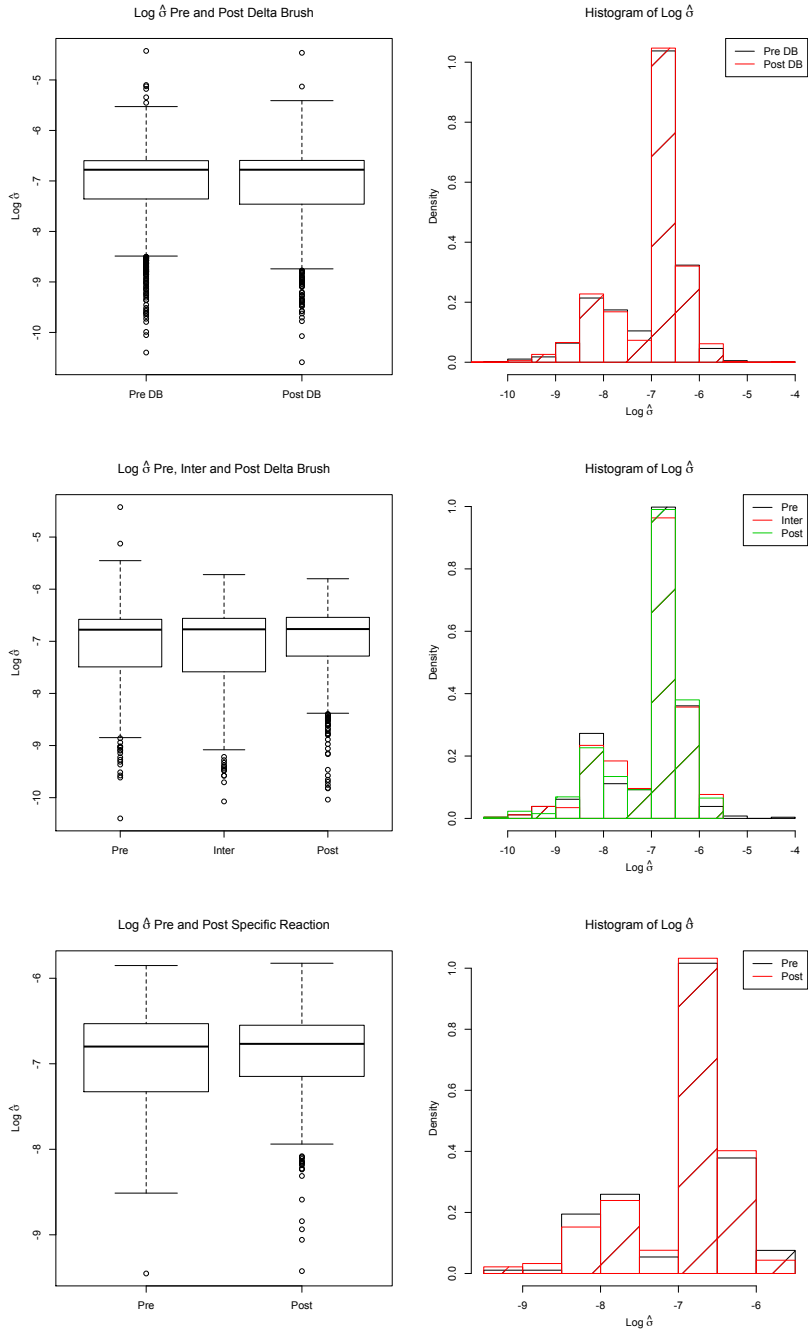
We have plotted the histograms of the pre-post, pre-inter-post parameters where relevant, and by visual inspection they appear to be identically distributed - Figures 4.1, 4.2, 4.3 and 4.4. A conclusion supported by the p values of the Kolmogorov-Smirnov test for distributional similarity [106, 107], and the Wilcoxon signed ranks tests - Tables 4.1.1, 4.1.2, 4.1.3 and 4.1.4. We also checked the parameterisation with respect to Matérn or fBm, and we found that: 82% of specific responses, 96% of signals with one delta brush and 93% of signals with two delta brushes remained the same parameterisation across all segments. These analyses indicate that we cannot detect a change-point in the estimated parameters of these six second recordings. Furthermore, we have illustrated that we can use the same long range covariance parameters across all six seconds. This also suggests that the time heterogeneity identified is an additional component that can

be added to the time-homogeneous baseline.

From these analyses we obtain an answer to the question of interest: At what point does the baseline activity change for neonatal EEG signals? We do not know, and we cannot ascertain this from our estimates. What we can infer from this analysis, is that in six seconds the parameters remain the same and time-heterogeneity can be included at relevant points, as prescribed by our model.

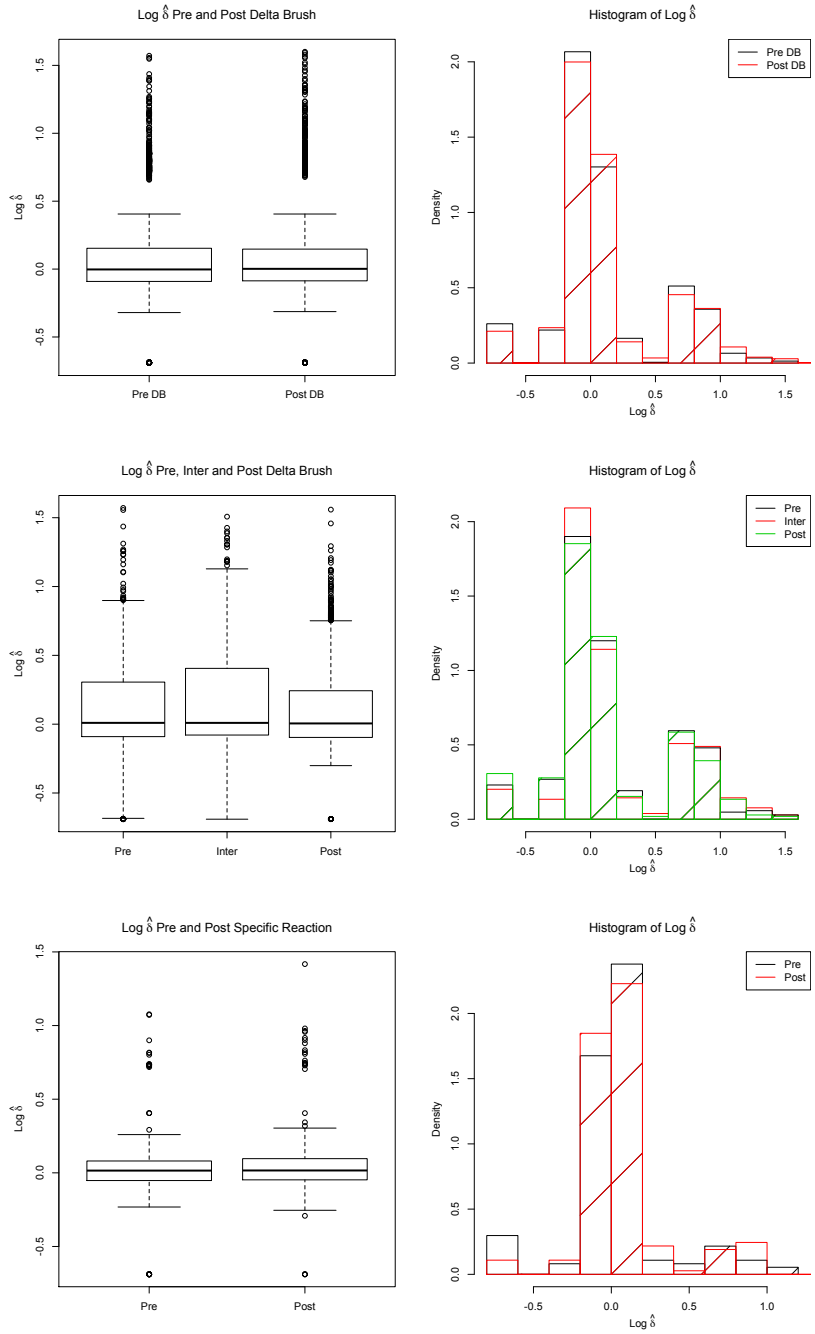
As a result of this analysis, we are going to reduce the number of parameters per signal, so only one set of parameters is present per electrode and signal; by taking the mean of the parameters when necessary.

Figure 4.1: Histograms of $\hat{\sigma}$ between time heterogeneous segments



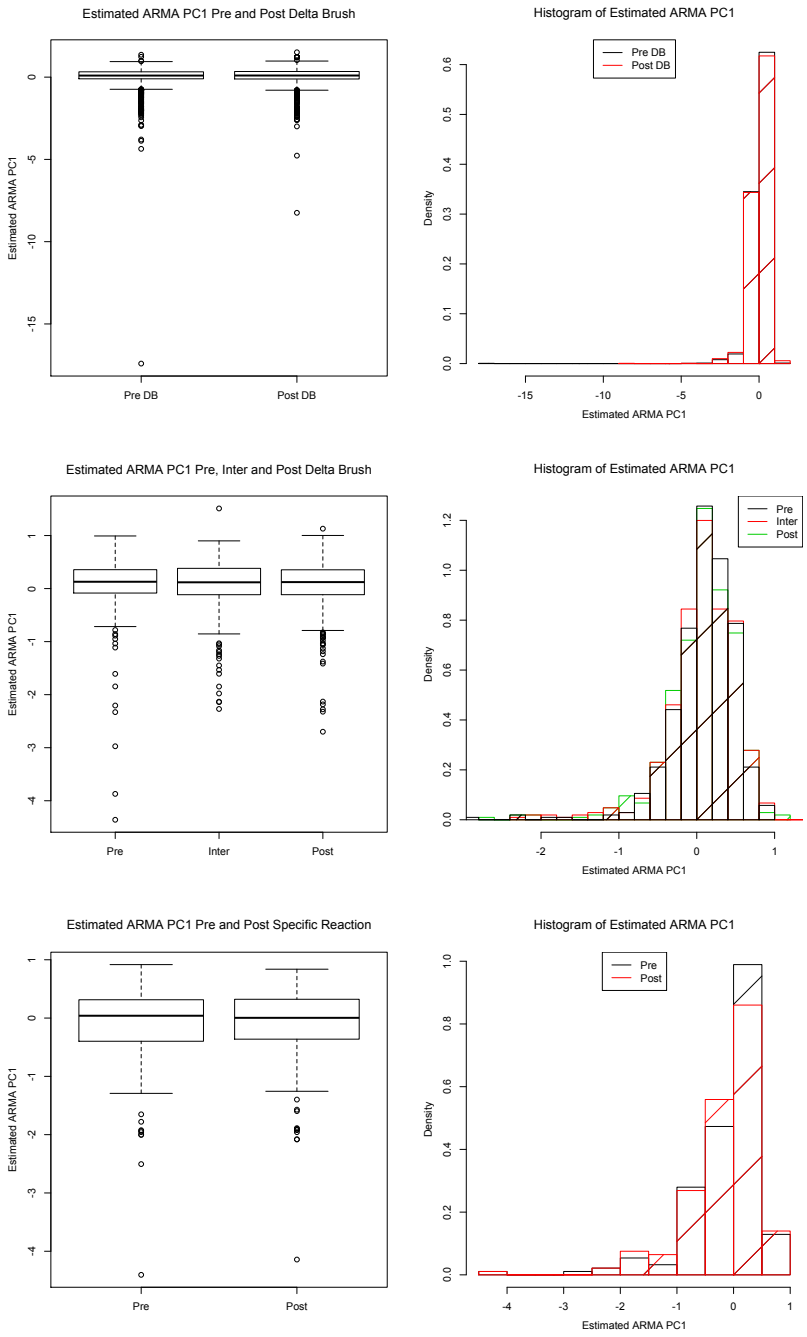
This figure shows histograms and box plots for $\hat{\sigma}$ between time heterogeneous segments and shows the distributional similarities between them - an analysis that is confirmed by the Kolmogorov-Smirnov test results. This when analysed with the results of the Wilcoxon signed ranks tests illustrates that we cannot reject the hypothesis that the parameters are the same between time heterogeneous segments

Figure 4.2: Histograms of $\hat{\delta}$ between time heterogeneous segments



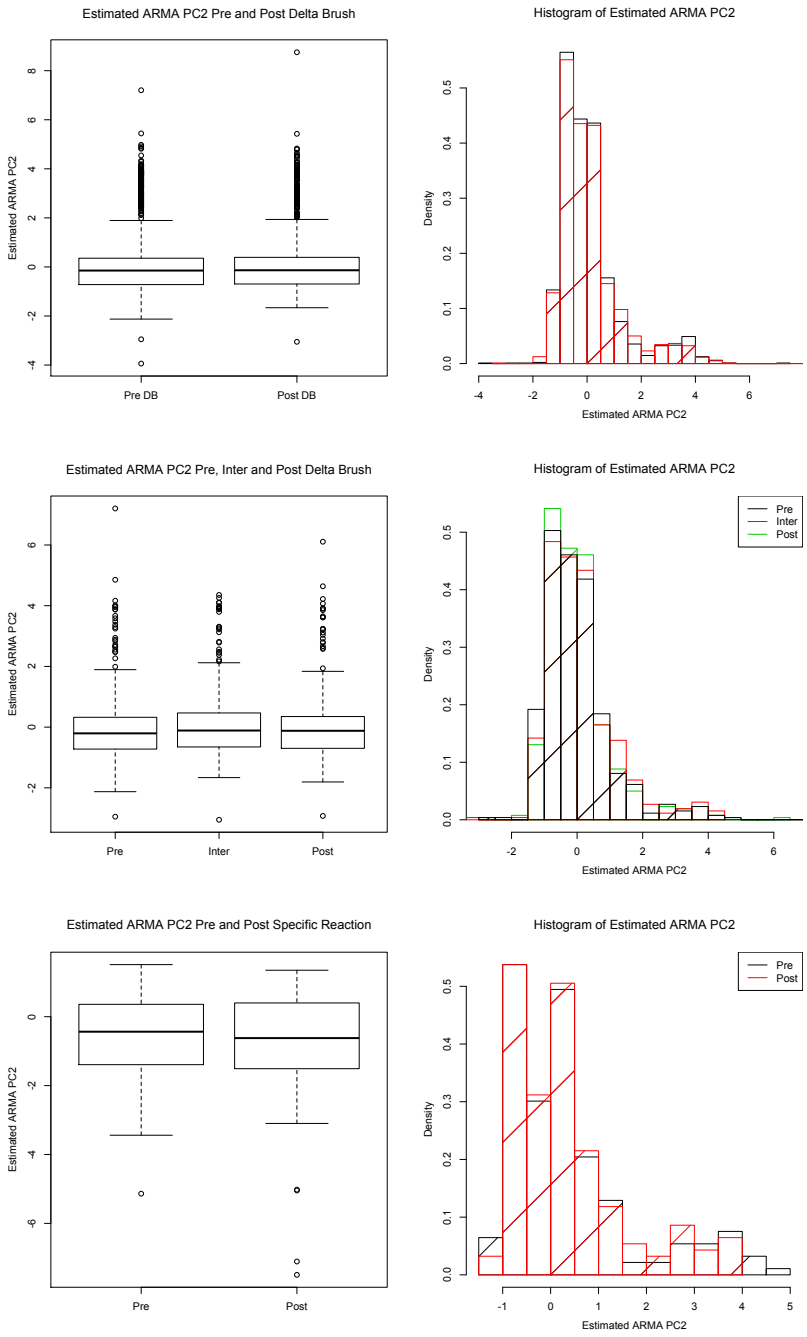
This figure shows histograms and box plots for $\hat{\delta}$ between time heterogeneous segments and shows the distributional similarities between them - an analysis that is confirmed by the Kolmogorov-Smirnov test results. This when analysed with the results of the Wilcoxon signed ranks tests illustrates that we cannot reject the hypothesis that the parameters are the same between time heterogeneous segments

Figure 4.3: Histograms of ARMA PC1 between time heterogeneous segments



This figure shows histograms and box plots for the first principal component description of the ARMA parameters between time heterogeneous segments and shows the distributional similarities between them - an analysis that is confirmed by the Kolmogorov-Smirnov test results. This when analysed with the results of the Wilcoxon signed ranks tests illustrates that we cannot reject the hypothesis that the parameters are the same between time heterogeneous segments.

Figure 4.4: Histograms of ARMA PC2 between time heterogeneous segments



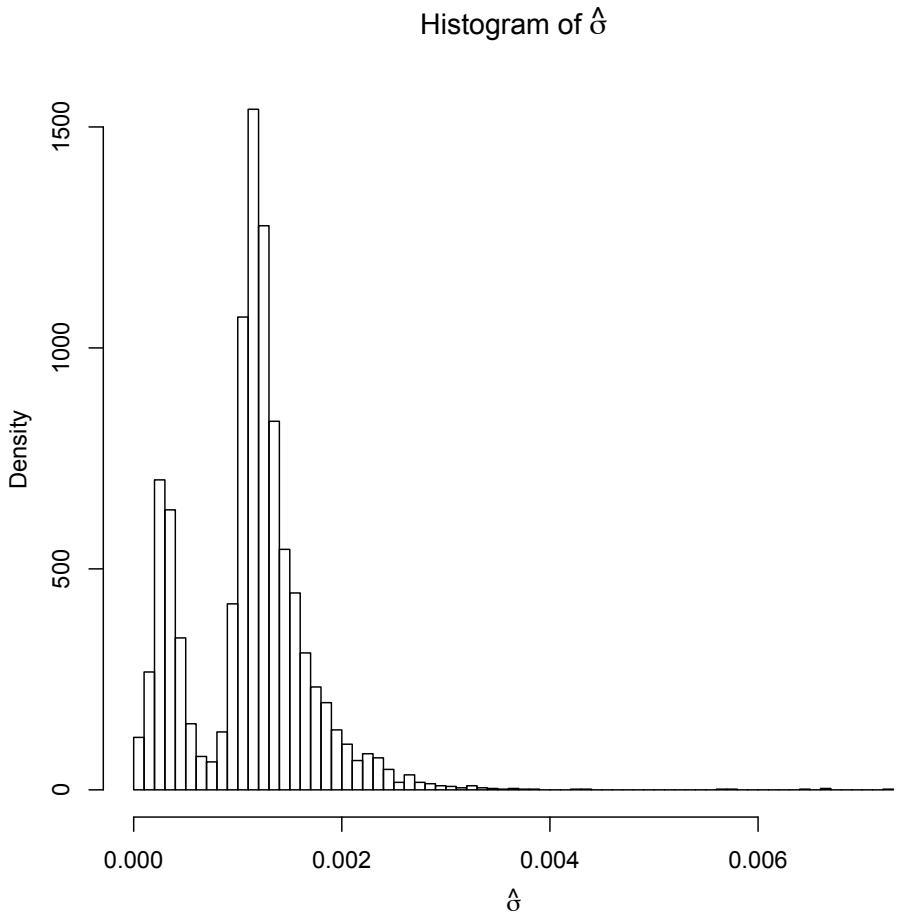
This figure shows histograms and box plots for the second principal component description of the ARMA parameters between time heterogeneous segments and shows the distributional similarities between them - an analysis that is confirmed by the Kolmogorov-Smirnov test results. This when analysed with the results of the Wilcoxon signed ranks tests illustrates that we cannot reject the hypothesis that the parameters are the same between time heterogeneous segments.

4.2 Inter-parameter relationships

Previous research has looked at the distribution of δ [75], however since we have a different covariance structure, containing short term covariance, we need to incorporate this also. First we shall assess the inter-parameters relationships and the behaviour of the parameters, starting with the standard deviation, σ.

Figure 4.5 shows the histogram of $\hat{\sigma}$ as it is an intuitive starting point for assessing the pattern in the data.

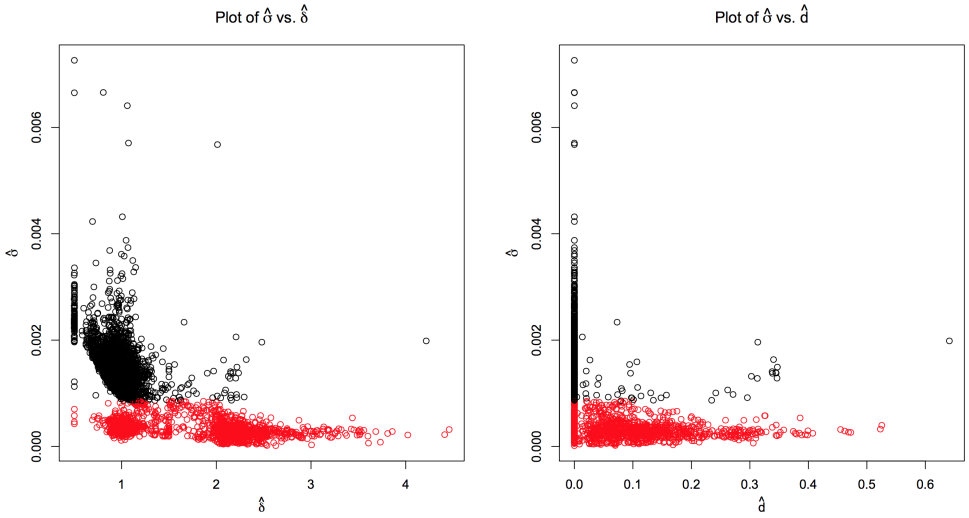
Figure 4.5: Histogram of $\hat{\sigma}$ illustrating bimodality



This figure illustrates that we have a possibly bimodal distribution of $\hat{\sigma}$ and as such it is a sensible starting point for analysis into the variable relationships

The histogram in Figure 4.5 appears to illustrate evidence of bimodality, with the two distributions separated at $\hat{\sigma} = 0.0007$. The amount of bimodality appears slight, however it could become more evident looking at the relationship with other parameters. Figure 4.6 shows a plot of $\hat{\sigma}$ vs. $\hat{\delta}$ with the possible separation between the two distributions illustrated.

Figure 4.6: Plots of $\hat{\sigma}$ vs. $\hat{\delta}$ and $\hat{d}$ coloured according to bimodality



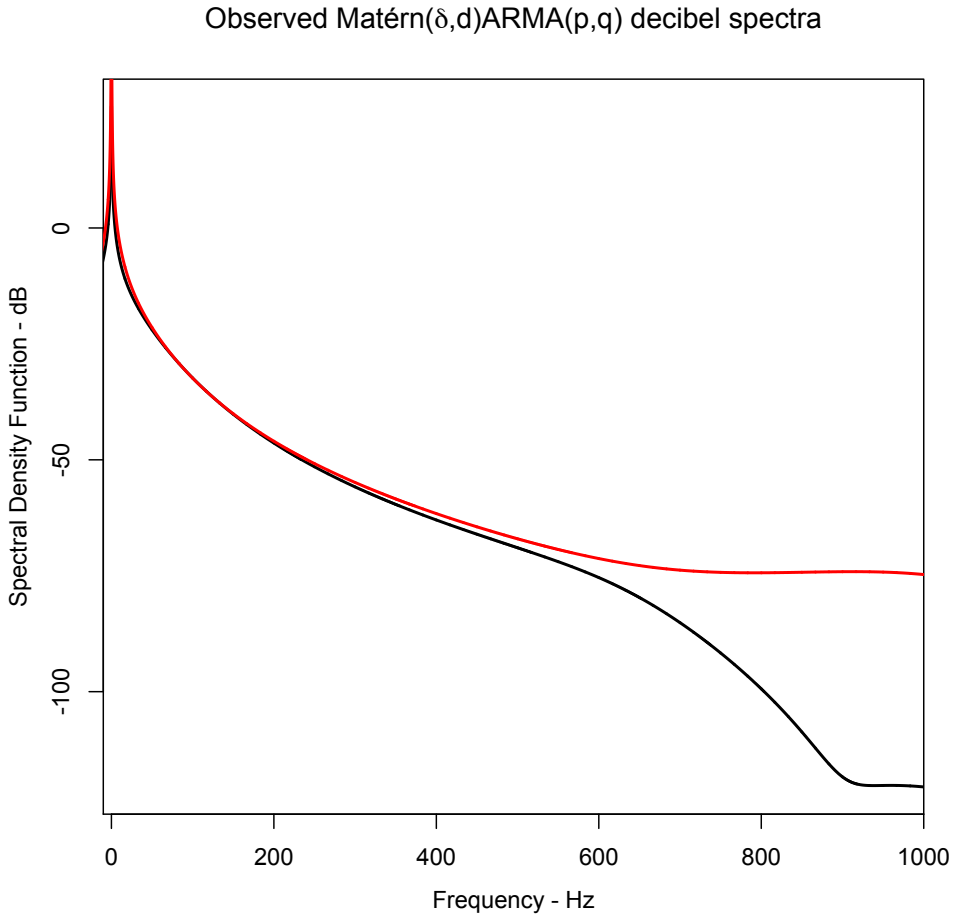
This plot illustrates that the bimodality of $\hat{\sigma}$ could correlate to some behaviours with respect to whether the process is best described using a Matérn parameterisation over fBm. With values in the minor peak possibly being better described by the Matérn process

The bimodality is more evident here, especially when looking at its effect of the values of $\hat{\delta}$. We see from Figure 4.6, that the Matérn parameterisation , $\hat{\delta} > 1.5$ or $\hat{d} > 0$, has some relationship with respect to the bimodality: the Matérn parameterisation appears to become more likely when $\hat{\sigma} < 0.0007$.

Now let us investigate the fitted ARMA parameters, $\underline{\hat{\varphi}}$ and $\underline{\hat{\vartheta}}$; these are more difficult to visualise due to the changing values of $\hat{p}$ and $\hat{q}$. Therefore, we shall assess these parameters based upon the shape of the spectral density function. On inspection of the estimates, it became clear that two distinct shapes of spectra appeared, Figure 4.7, and we noticed the main difference appears to be the behaviour in the passband and at high frequencies. This difference was first noticed when fitting the time-homogeneous

model and realised it might affect the delta brush detection; as such we created two time-homogeneous detection baselines based upon the shape of the raw periodogram.

Figure 4.7: Illustrative plots of observed MatérnARMA(δ,d,p,q) spectral density shapes



This plot illustrates the two shapes of spectral density function that classify the parameter Shape. For each set of parameters we have assessed which they are more similar to in shape by which is more similar at $f > 600$ as this is where the most noticeable difference occurs. The parameter Shape = 0 if more similar to the black line and 1 if more similar to the red

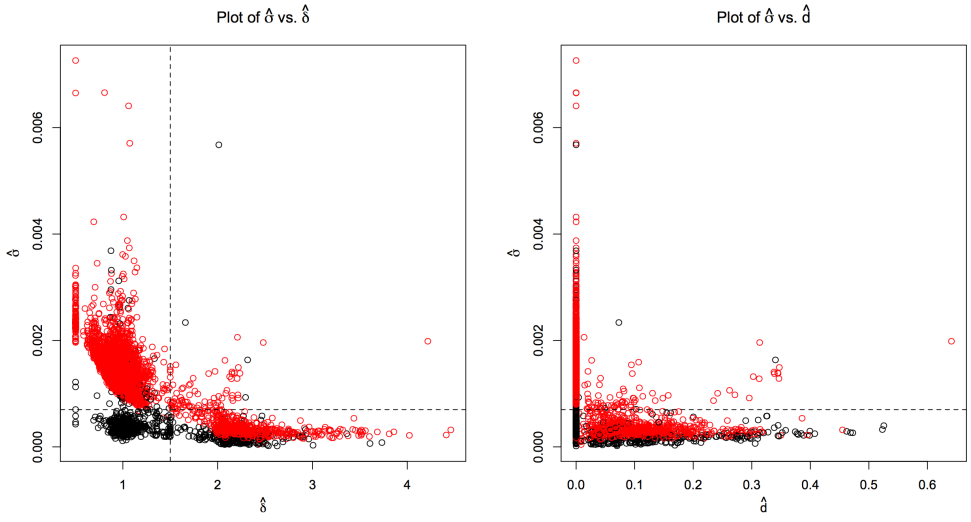
Having classified the spectral density functions accordingly, we present $\hat{\sigma}$ vs. $\hat{\delta}$ with the points coloured to reflect the spectral shape. Figure 4.8 illustrates some very interesting behaviour amongst all parameters. It appears the shape of the spectral density function

is strongly related to Matérn and Bimodality. The behaviour when the covariance is Matérn and $\hat{\sigma} < 0.0007$ is slightly concerning, but appears to be distinct.

So, from the estimated parameters we have extracted three binary parameters: Matérn, Shape and Bimodality, and have illustrated the effect that they have on the data and each other. Next, the relationship of these three variables with respect to age shall be assessed. The plot below shows the box plots of age at birth and test, segregated into the eight possible groupings based upon the values of the three binary variables. The value of the binary parameter Matérn is equal to 1 if the long-term covariance parameterisation is Matérn and equal to 0 otherwise. Bimodality is equal to 1 if $\hat{\sigma} < 0.0007$, and Shape is equal to 1 if the raw periodogram is similar to the first spectral density function illustrated in Figure 4.7 - as indicated by the red line.

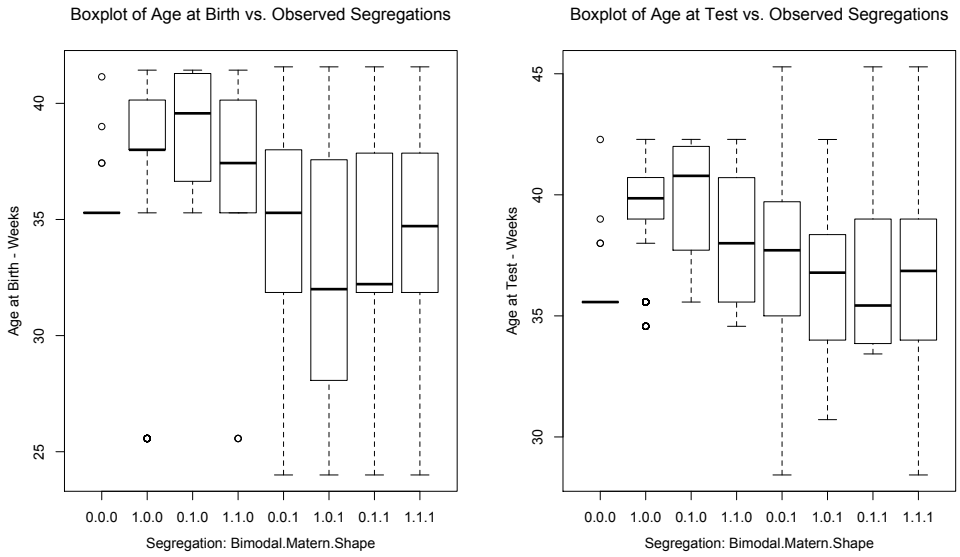
These box plots appear to illustrate a relationship with age, and raise a worrying feature in the segregated data. It would appear for the grouping Shape = 0 and Bimodality = 0 we do not have enough data to segregate according to Matérn when modelling.

Figure 4.8: Plots of $\hat{\sigma}$ vs. $\hat{\delta}$ and $\hat{d}$ coloured according to spectral shape



This plot illustrates the behaviour of parameters when classified by the variable Shape, we have also indicated the values of the Matérn parameterisation using dotted lines. We see for values of $\hat{\delta} < 1.5$ & $\hat{d} = 0$ the classification seems to be distinct. However outside these values this relationship is less clear and as such the classification is not as simple.

Figure 4.9: Boxplots of ages by observed segregations



This figure shows some interesting behaviour with regards to the parameter Shape and the age at birth at which it is occurs. We notice that the distribution of Shape = 0 - black line Figure 4.7 - is skewed towards higher ages at birth. But for Shape = 1 - red line Figure 4.7 - the distribution is less skewed. We also notice the issue with lack of points for certain classifications due to their occurrence within the sample.

4.3 Modelling the estimated parameters

Having assessed the inter-parameter relationships, we now attempt to fit a series of models to the parameters, with the aim of producing similar data for simulation.

The previous section illustrated the possible relationship between age, and the three binary variables we defined based upon the characteristics of the estimated model parameters. Now, we want to describe the variables' relationship with age and other possible explanatory variables; this shall be done using a generalised linear model with a logit link function. i.e. binomial regression. Taking into account the effect of the segregated variables on each other we obtain the following models by stepwise regression.

Model 4.3.1. *The generalised linear models to describe the probability of the binary variables being equal to 1 have the general form:*

$$\begin{aligned}\text{Bimodal} \sim \text{MCA}_1^{\text{C}} \times \text{MCA}_2^{\text{C}} + \text{Type} + \text{Electrode} + \text{MCA}_1^{\text{C}} : \text{Type} + \\ \text{MCA}_1^{\text{C}} : \text{Electrode} + \text{MCA}_2^{\text{C}} : \text{Electrode}, \quad (4.3.1)\end{aligned}$$

$$\begin{aligned}\text{Matern} \sim \text{MCA}_1^{\text{C}} \times \text{MCA}_2^{\text{C}} + \text{Type} + \text{Bimodal} + \text{Electrode} + \\ \text{MCA}_1^{\text{C}} : \text{Type} + \text{Bimodal} : \text{Electrode} + \text{Bimodal} : \text{Type}, \quad (4.3.2)\end{aligned}$$

$$\begin{aligned}\text{Shape} \sim \text{MCA}_2^{\text{C}} \times (\text{Bimodal} : \text{Matern} + \text{Bimodal} : \text{Type} + \text{Matern} : \text{Type}) + \\ \text{MCA}_1^{\text{C}} \times (\text{Bimodal} : \text{Matern} + \text{Bimodal} : \text{Type}) + \\ \text{Bimodal} \times \text{Matern} \times \text{Type}. \quad (4.3.3)\end{aligned}$$

The models with corresponding estimated coefficients are described in Appendix A.

Where MCA_1^{C} and MCA_2^{C} are the principal components defined in Model 4.3.2; since these variables are strongly correlated we use the principal components to avoid issues with multicollinearity [105]. The variable Type is whether the signal is time-homogeneous background, delta brush or somatosensory specific stimuli response activity.

We require a multicollinearity adjustment because we have strong positive correlation between age at birth and age at test ($\hat{\tau} = 0.678$). This correlation is partially induced by the fact that we cannot test before an infant is born and by the fact that clinicians will not allow testing for medical reasons. We have many tests performed within the immediate weeks after birth, again this could be induced because of medical pressures, testing procedure or discharge time. Utilising solely age at test ignores the length of time infants have been ex utero, which could have an effect on development.

Model 4.3.2 (**Linear transformation of ages to avoid multicollinearity**). *In order to avoid issues with multicollinearity the linear transformation of the parameters PMA and PMAT shall be used. For the subset of the data describing the characteristics of the estimated parameters these are defined as:*

$$\text{MCA}_1^{\text{C}} = -0.904741 \times \text{PMA} - 0.425962 \times \text{PMAT},$$

$$\mathrm{MCA}_2^{\mathrm{C}} = -0.425962 \times \mathrm{PMA} + 0.904741 \times \mathrm{PMAT}.$$

Where PMA is post menstrual age at birth and PMAT is post menstrual age at test.

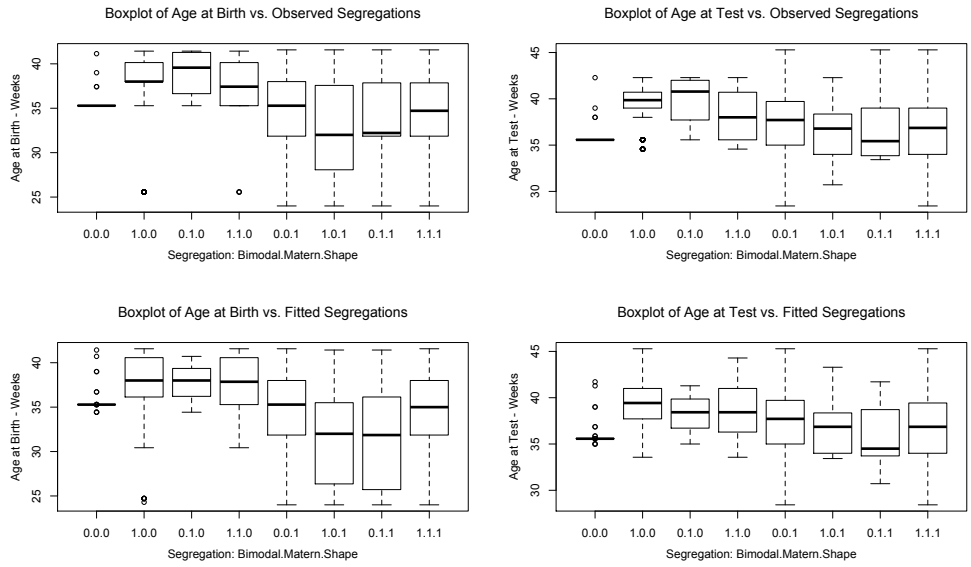
Since the presented models are too complicated to give all parameters, we shall highlight some interesting features from the variable coefficients. Generalised linear models with the logic link function fit the probability of the response variable being equal to 1, as such we shall describe the parameters in terms of increase or decrease in this probability.

From the estimated coefficients, presented in Appendix A, we notice the following effects on the estimated parameter characteristics described by these binary variables. First we find across all GLMs that as the age at birth increases, the probability of the binary variables being equal to 1 decreases. We also find that only the Shape variable being equal to 1 decreases as age at test increases. This suggests an interesting feature with respect to development, as it would appear that infants with a low age at birth and high age at test are more likely, under our model, to have a small variance and a covariance structure not describable by the fBmARMA covariance model. A conclusion further supported since when the bimodal variable is equal to 1 in the Matérn GLM, the probability increases. When we look at the plots in the previous section, compared with the results from the above models, we also see the behaviour described by the generalised linear models.

We find that across all GLMs, that delta brush responses increase the probability of the binary variables being equal to 1, in comparison to time-homogeneous signals. For somatosensory specific responses, the probability decreases in comparison to time-homogeneous signals, in all GLMs except the Shape GLM.

To check that these models can be used to recreate the structure seen in the estimates, we obtained fitted estimates and shall check these against the observed statistics. We do so by comparing the estimated and actual box plots of age at birth and test, segregated into the eight possible groupings based upon the values of the three binary variables. Figure 4.10 illustrates that we can obtain a suitably similar distribution to the observed values. With the extracted binary variables described, we can now look to describe the estimated parameters.

Figure 4.10: Boxplots of ages by observed and fitted segregations



This plot shows the results of our modelling by comparing the box plots of ages vs the observed segregations. We see a similar distribution from the fitted values within our model indicating a capturing of the behaviour. This does not suggest a perfect description but rather we can realise a decent approximation to the observed behaviour

The way in which we shall produce estimates for simulation is via several multivariate normal distributions, since under the Whittle likelihood, estimates obtained are asymptotically normal due to the estimated error distribution under Whittle estimation following a normal distribution.

However, we must take the observed behaviour of our estimates into account when constructing multivariate distributions. We shall do so by creating multiple distributions based upon the values of the observed binary parameters outlined. The estimated mean vectors and covariance matrices from the fitted model parameters have been obtained and these are given in Appendix B.

An issue that is noticeable at this juncture is the number of parameters from which the distribution is estimated, with some segregations having very few observations from which a distribution might be fitted. Furthermore, whilst the estimates might be asymp-

totically normal, the sample from which the distribution is estimated might not be. It should be noted that as a result of some of the groupings having few observations, some of the simulated parameters may be inappropriate; however, in other groupings adequate simulated parameters are obtained.

Whilst this approach may produce adequate realisations of the estimated parameters, we have not produced a structure that can adequately describe neonatal electroencephalogram signals. In Chapter 1 we noted that these signals are multivariate, and due to computational intensities we estimated them as univariate. In order to produce similar signals we must recreate a suitable multivariate structure.

4.4 Modelling the presence and occurrence of delta brushes

A key component of neonatal EEG signals is the delta brush, as such any simulation must take into account its presence. Before using the models to realise similar signals, and assessing their suitability, we shall analyse the behaviour of the delta brush components, starting with spontaneous occurrence. As with the parameter models we shall utilise the principal components of the age variables to avoid multicollinearity issues.

Model 4.4.1 (**Linear transformation of ages to avoid multicollinearity**). *In order to avoid issues with multicollinearity the linear transformation of the parameters PMA and PMAT shall be used. For the subset of the data describing the signal's structure these are defined as:*

$$\mathrm{MCA}_1^{\mathrm{S}} = -0.88207 \times \mathrm{PMA} - 0.47113 \times \mathrm{PMAT},$$

$$\mathrm{MCA}_2^{\mathrm{S}} = -0.47113 \times \mathrm{PMA} + 0.88207 \times \mathrm{PMAT}.$$

Where PMA is post menstrual age at birth and PMAT is post menstrual age at test.

We require a multicollinearity adjustment because we have strong positive correlation between age at birth and age at test ($\hat{\tau} = 0.653$). The same reasoning applied for the use of a multicollinearity adjustment in Section 4.3 still applies. The linear transformation is different to Model 4.3.2 due to the use of different samples of our data. The inclusion of age at birth differs from previous research into this developmental stage [4]. However, this parameter cannot be overlooked; previous research has shown behaviourally that the

difference between age at birth and test has an effect on the response to stimuli [52, 60] and as such we shall investigate this utilising non behavioural measurements of stimuli response. Furthermore it has been shown that the weeks preceding birth have an effect upon the response to stimuli [12, 14, 108].

The models fitted are illustrated using heat maps and we have indicated the areas that we cannot have data for as, under the testing procedure, an infant cannot be tested before they are born. Furthermore we have indicated areas which have not been represented in the sample.

This raises an issue that can affect the results of our analysis, and as such affects the conclusions that we can draw. We have evidence that both age at birth and age at test are important explanatory variables, however we do not have data points for many of the pairings. As such, we can construct hypotheses from the trends that we discover, however a more representative sample is required to confirm or reject these hypotheses.

4.4.1 Spontaneous delta brush activity

Spontaneously occurring delta brushes are a feature of neonatal EEG across development, becoming less common by 39-41 weeks post menstrual ages [61]. First, we want to determine if the probability of occurrence changes with the age at test and age at birth. We did so by fitting a binomial generalised linear model, and found that there is a relationship between spontaneous delta brush occurrences and ages.

Model 4.4.2 (**Binomial GLM of spontaneous delta brush activity**). *The probability of an infant expressing delta brush activity spontaneously, $\hat{p}^{SP}$, is modelled by the logit function,*

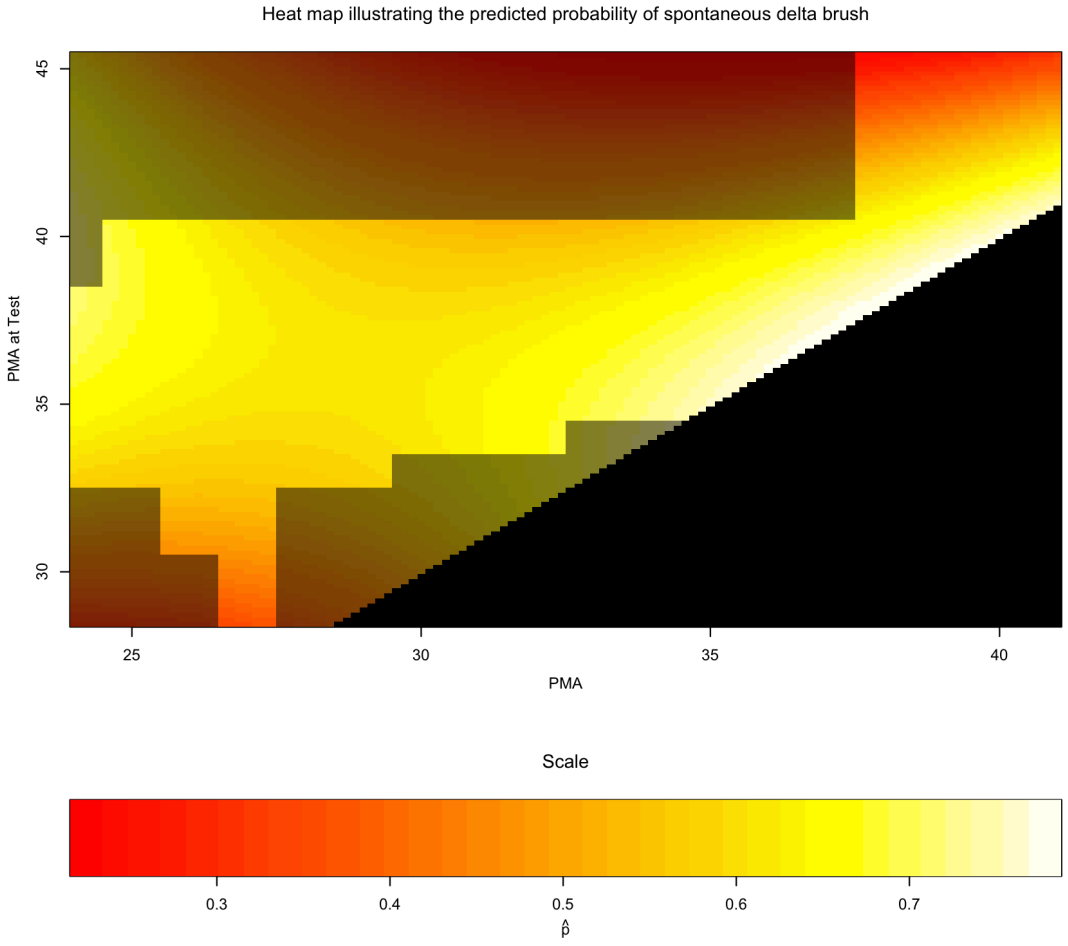
$$\log\left(\frac{\hat{p}^{SP}}{1-\hat{p}^{SP}}\right) = 0.72508 - 0.05093\times \mathrm{MCA}_1^{\mathrm{S}} - 0.16907\times \mathrm{MCA}_2^{\mathrm{S}} + 0.03369\times \mathrm{MCA}_1^{\mathrm{S}} \times \mathrm{MCA}_2^{\mathrm{S}}.$$

Interaction significant at $\alpha = 0.1$

Figure 4.13 shows a heat map illustrating the probability of spontaneous delta brushes occurring and is congruent with the underlying theory with regards to spontaneous delta brush occurrence [61]. We can see that infants born prematurely have a higher probability of expressing spontaneous delta brushes. We also see that infants tested close

to birth have a similarly high probability of expressing delta brushes spontaneously. The probability of spontaneous delta brushes being present in an infant's recording reduces at around 40 weeks at test depending upon the age of the infant at birth; thus showing the possible effect of the noxious procedures that premature infants undergo [21].

Figure 4.11: Heatmap of the fitted probability of an infant expressing spontaneous delta brush



This figure shows the predicted probability of a spontaneous delta brush occurring. The greyed out areas indicate areas where data points are not present within the sample.

With a model to describe the probability of spontaneous delta brushes occurrences, we shall proceed to modelling the number of electrodes that express such activity. We shall model this observed behaviour utilising the binomial distribution, by describing the proportion of electrodes that simultaneously express delta brush activity.

By viewing electrodes with simultaneous delta brush activity as successes and the number of electrodes as the number of trials, we can describe the proportion of electrodes expressing spontaneous delta brush activity, p_n^{SP}, using a binomial generalised linear model, as outlined in Model 4.4.3.

Model 4.4.3 (**Binomial GLM for simultaneous spontaneous delta brushes**). *The proportion of electrodes simultaneously expressing spontaneous delta brush activity, $\hat{p}_n^{SP}$, is modelled by the logit function,*

$$\log\left(\frac{\hat{p}_n^{SP}}{1-\hat{p}_n^{SP}}\right) = -0.95378 - 0.02847 \times \mathrm{MCA}_1^{\mathrm{S}} - 0.09 \times \mathrm{MCA}_2^{\mathrm{S}} + 0.01112 \times \mathrm{MCA}_1^{\mathrm{S}} \times \mathrm{MCA}_2^{\mathrm{S}}$$

Interaction significant at $\alpha = 0.07$

From Model 4.4.3 and Figure 4.12 we can see that infants tested close to birth are more likely to have a higher number of electrodes expressing simultaneous delta brush activity, and fewer electrodes expressing simultaneous delta brush activity after 40 weeks gestational age at test.

Whilst the model describing the probability of spontaneous delta brush occurrence suggests that infants born prior to 28 gestational weeks at test are more likely to express spontaneous delta brushes, we see from Figure 4.12 that these infants could express them at fewer electrode sites. This is an important realisation to take into account when simulating a set of neonatal EEG signals and could be due to the development of the brain, specifically the growth and increase in neural connections that occur in the brain over development.

4.4.2 Delta brush response to somatosensory stimuli

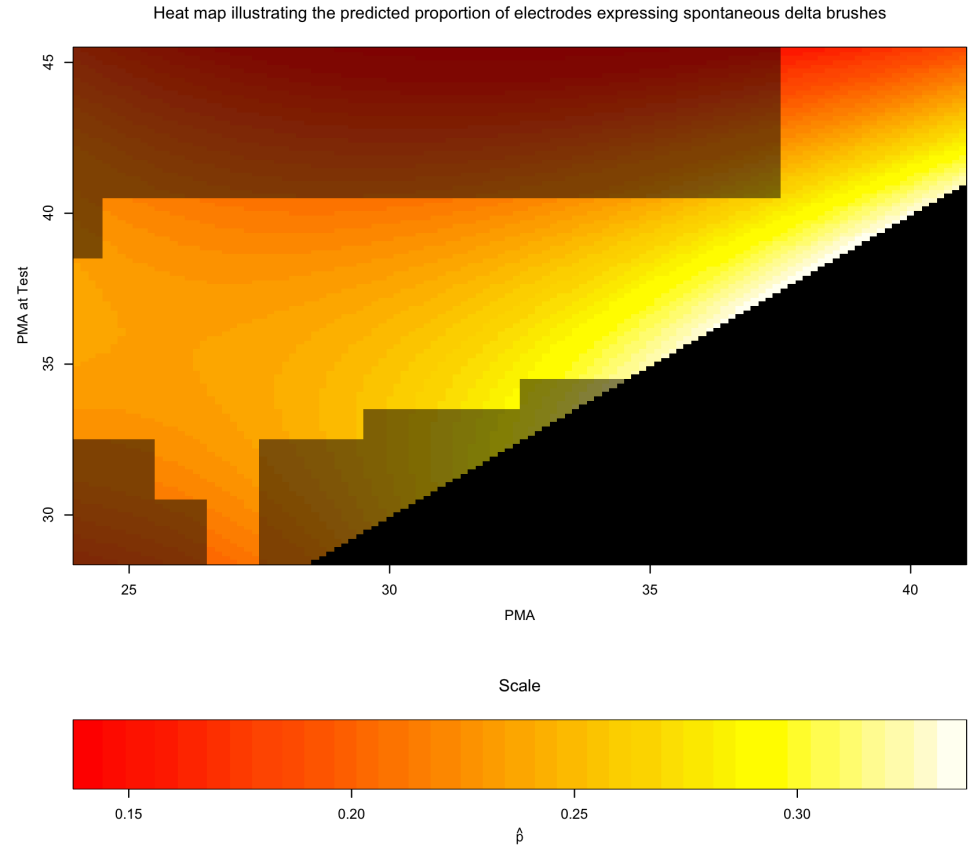
Now we shall analyse the presence of delta brushes in response to noxious and tactile stimuli; this shall be done with the methods outlined and demonstrated with spontaneous delta brushes. Before discussion of the results from our models we shall present the models and fitted heat maps illustrating the modelled behaviour.

First we present the probability of delta brushes occurring in response to noxious and tactile stimuli, then the proportion of electrodes that express delta brushes shall be described. As with the previous model fitting, we shall utilise stepwise regression to fit

our models. Whilst stepwise regression does not guarantee terms significant at a given significance level, we obtain a model that includes the relevant terms that best describes the pattern within the data. Finally, whilst the presented sample is representative with respect to age at test or age at birth, when we look at the age pairings we do not have a suitable sample; future research is required to test the conclusions obtained.

Model 4.4.4 (Binomial GLM of nociceptive delta brush response). *The probability of an infant expressing delta brush activity in response to nociceptive stimuli, $\hat{p}^N$,*

Figure 4.12: Heatmap of predicted expected number of electrodes, simultaneously expressing spontaneous delta brush activity



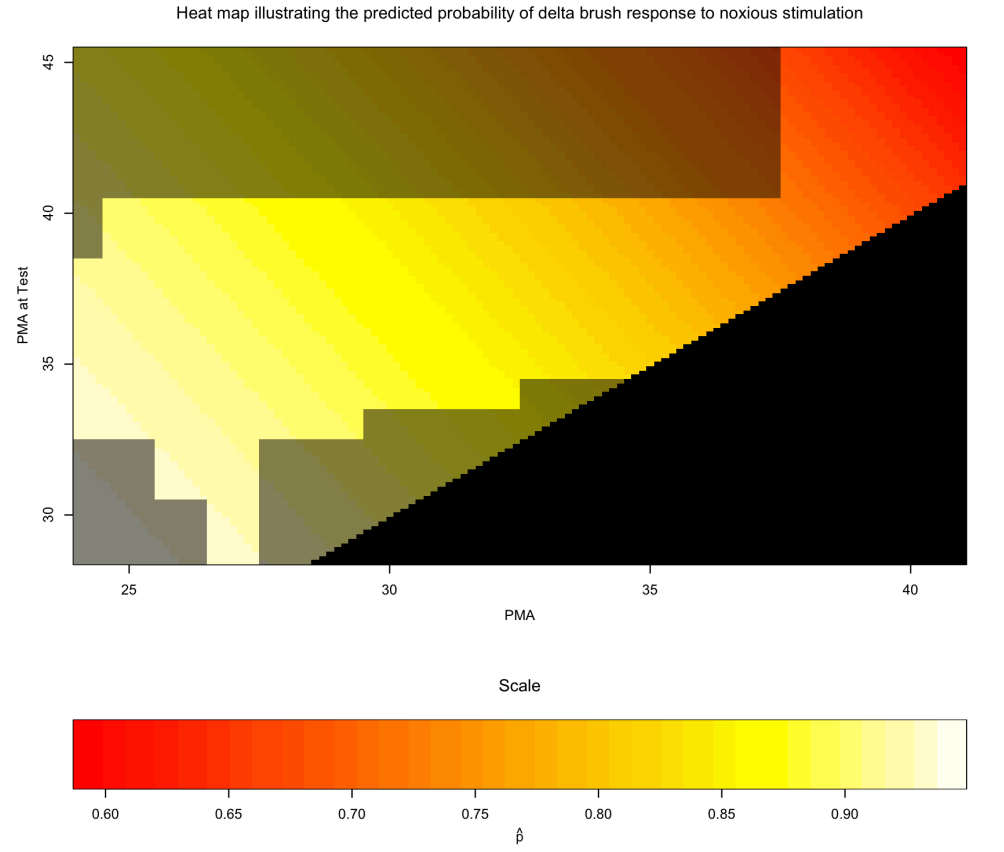
This figure shows the predicted expected number of electrodes, simultaneously expressing spontaneous delta brush activity. The greyed out areas indicate areas where data points are not present within the sample.

is modelled by the logit function,

$$\log\left(\frac{\hat{p}^N}{1-\hat{p}^N}\right) = 1.481 + 0.11046 \times \text{MCA}_1^{\text{S}}$$

Significant at $\alpha = 0.12$

Figure 4.13: Heatmap of the probability of noxious stimulation expressing non-specific response

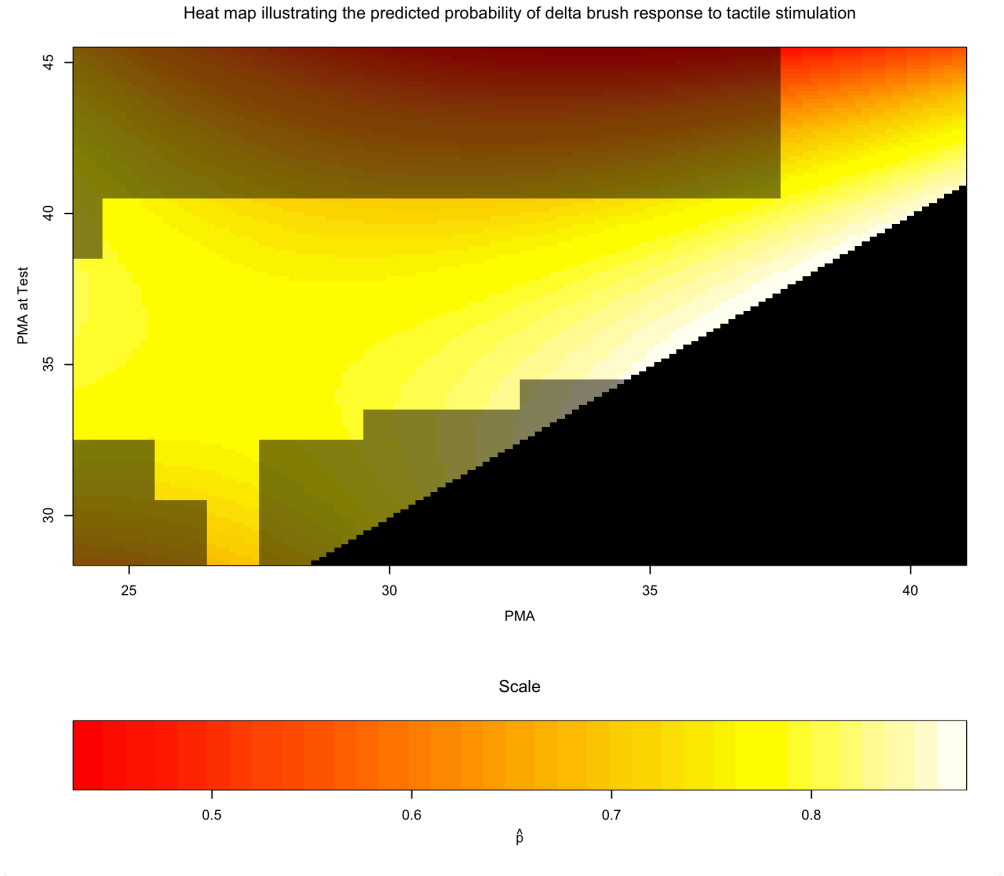


This figure shows the predicted probability of noxious stimulation expressing non-specific response. The greyed out areas indicate areas where data points are not present within the sample.

Model 4.4.5 (Binomial GLM of tactile delta brush response). *The probability of an infant expressing delta brush activity in response to tactile stimuli,* $\hat{p}^T$*, is modelled by the logit function,*

$$\log\left(\frac{\hat{p}^T}{1-\hat{p}^T}\right) = 1.49012 - 0.03358 \times \text{MCA}_1^{\text{S}} - 0.1658 \times \text{MCA}_2^{\text{S}} + 0.02414 \times \text{MCA}_1^{\text{S}} \times \text{MCA}_2^{\text{S}}$$

Figure 4.14: Heatmap of the probability of tactile stimulation expressing non-specific response



This figure shows the predicted probability of tactile stimulation expressing non-specific response. The greyed out areas indicate areas where data points are not present within the sample.

Interaction significant at $\alpha < 0.05$

Models 4.4.4 and 4.4.5 are in agreement with previous findings about the maturation of somatosensory stimuli response [4]. The probability of a non-specific response to stimuli starts to decrease past 37 weeks age at test, however we do not find as strong a result as previous research.

The probability for a delta brush being expressed in response to tactile stimuli, appears to be similar for almost all infants until 37-39 weeks, when we see a decrease in this probability. We notice however, that it is not only the age at test that affects the

probability of a delta brush response; for both noxious and tactile stimuli, infants tested close to birth, or born prematurely, are more likely to express a non-specific reaction. This indication of a less mature response to such stimuli is congruent with behavioural studies into noxious stimuli [52, 60]; however we find a similar response to babies tested close to birth, in contradiction to some behavioural observations [52]. Next, we present the models that describe the proportion of electrodes that express delta brush responses to somatosensory stimuli.

Model 4.4.6 (**Binomial GLM simultaneous nociceptive delta brush response**)**.** *The proportion of electrodes simultaneously expressing delta brush activity in response to nociceptive stimuli, $\hat{p}_n^N$, is modelled by the logit function,*

$$\log\left(\frac{\hat{p}_n^N}{1-\hat{p}_n^N}\right) = -1.18499 + 0.00784 \times \mathrm{MCA}_1^{\mathrm{S}} - 0.19393 \times \mathrm{MCA}_2^{\mathrm{S}} + 0.02835 \times \mathrm{MCA}_1^{\mathrm{S}} \times \mathrm{MCA}_2^{\mathrm{S}}$$

Interaction significant at $\alpha < 0.05$

Model 4.4.7 (**Binomial GLM simultaneous tactile delta brush response**)**.** *The proportion of electrodes simultaneously expressing delta brush activity in response to tactile stimuli, $\hat{p}_n^T$, is modelled by the logit function,*

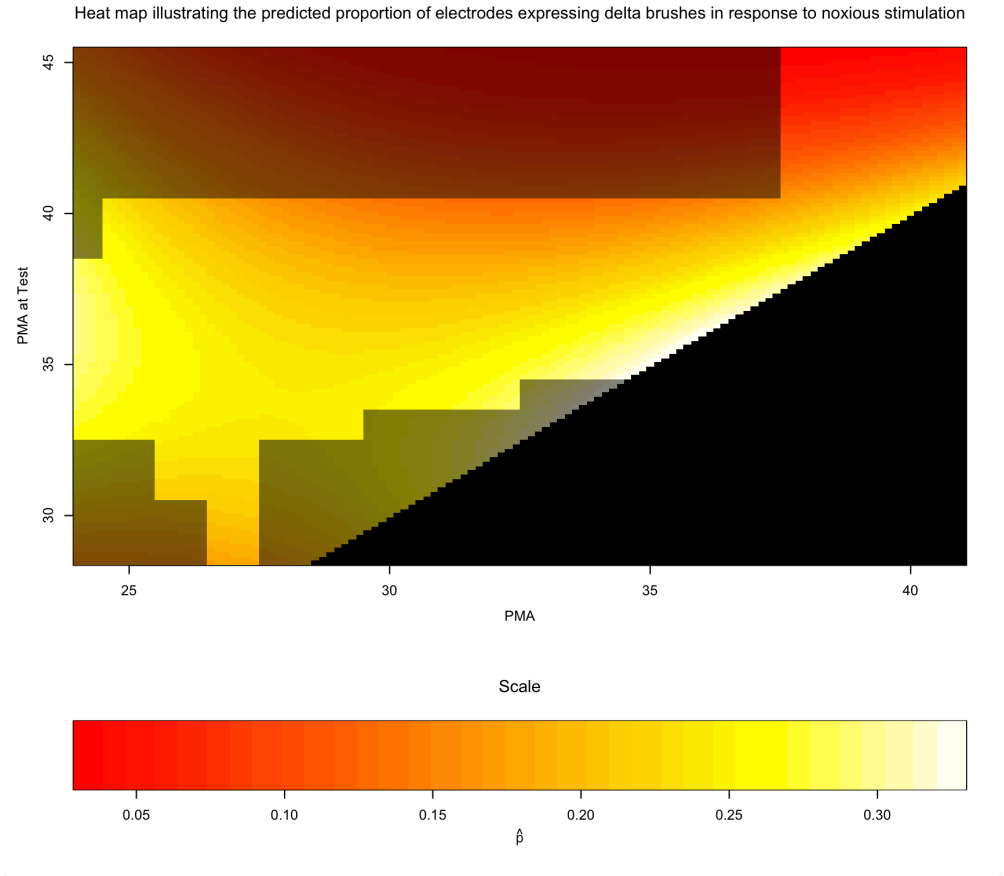
$$\log\left(\frac{\hat{p}_n^T}{1-\hat{p}_n^T}\right) = -0.92191 - 0.00004 \times \mathrm{MCA}_1^{\mathrm{S}} - 0.02739 \times \mathrm{MCA}_2^{\mathrm{S}} + 0.00486 \times \mathrm{MCA}_1^{\mathrm{S}} \times \mathrm{MCA}_2^{\mathrm{S}}$$

Interaction significant at $\alpha < 0.05$

From Models 4.4.6 and 4.4.7, we find again that both age at test and age at birth are important explanatory variables describing the proportion of electrodes that express delta brush responses. For noxious stimuli, we find that infants born prematurely or tested close to birth, have a larger proportion of electrodes expressing simultaneous delta brush activity. Furthermore we find that post 37 weeks age at test this proportion decreases across all ages.

With regards to tactile stimuli response, we find a less distinct result than noxious stimuli response; the range of fitted proportions is smaller and we see that any infant tested at 33-37 weeks has a higher proportion of electrodes simultaneously expressing delta brushes. This was seen in the probability of delta brush occurrence in response to tactile stimuli, and seems to indicate a difference in the way which neonates respond to

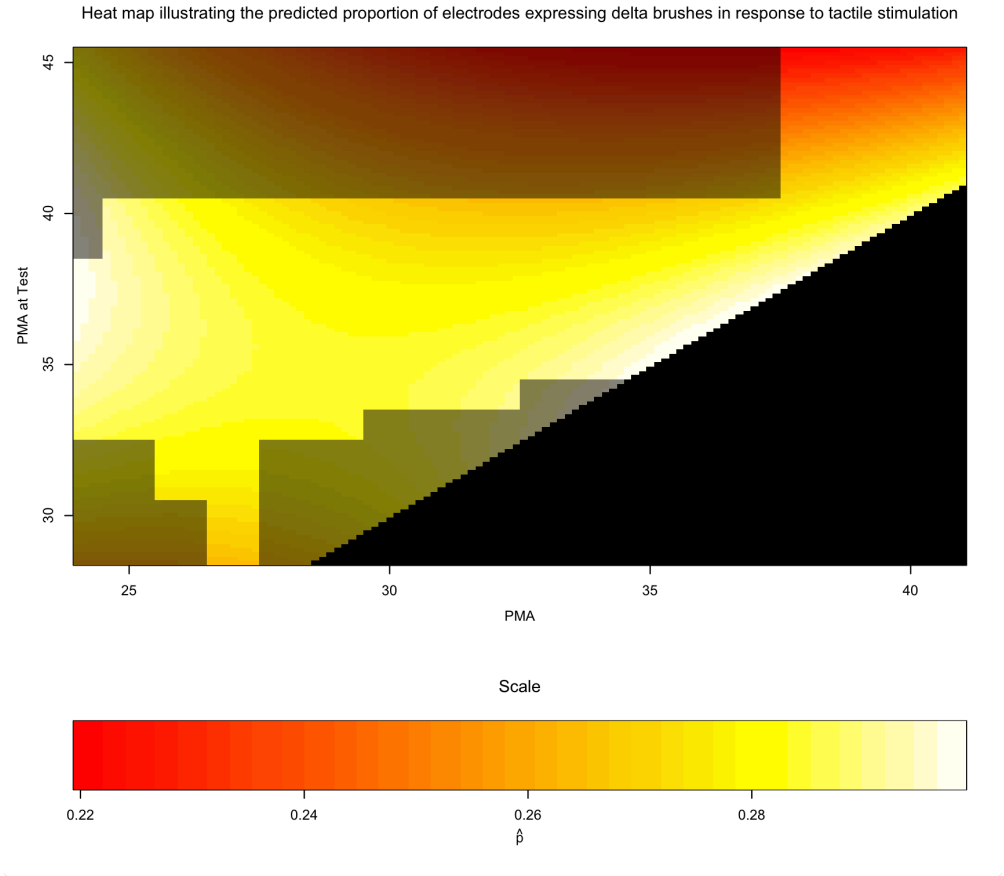
Figure 4.15: Heatmap of predicted proportion of electrodes, simultaneously expressing non-specific nociceptive stimuli response



This figure shows the predicted proportion of electrodes, simultaneously expressing non-specific nociceptive stimuli response. The greyed out areas indicate areas where data points are not present within the sample.

noxious and tactile stimuli. These results seem to indicate that infants tested close to birth or born prematurely are more likely to express a non-specific response to stimuli, and this response is present at at a higher proportion of electrodes than infants tested outside this range.

Figure 4.16: Heatmap of predicted proportion of electrodes, simultaneously expressing non-specific tactile stimuli response



This figure shows the predicted proportion of electrodes, simultaneously expressing non-specific tactile stimuli response. The greyed out areas indicate areas where data points are not present within the sample.

4.5 Results and Conclusions

From the estimates obtained for the model presented in Definition 3.1.3 we have identified some key features that must be taken into account when simulating neonatal EEG signals.

First we analysed whether we could detect the point at which parameters change within a signal. The inherent time-heterogeneous nature of neonatal EEG is poorly understood, with previous attempts to model such activity arbitrarily selecting parameter change-points. By analysing the parameters after identified time heterogeneous components, we

found no significant difference in the estimated parameters from the time-homogeneous segments. Our inability to identify the parameters change-point could be due to the fact that we are only analysing a short segment; the change-point might become evident over a longer interval.

Using this conclusion about the parameterisation of the signals, we reduced the set of time homogeneous parameters per recording; such that we had only one set of parameters, per infant and electrode. Given the previous result we assume that the time-heterogeneous segments are additions to a time-homogeneous baseline.

From this we were able to discern three behaviours in the signal that are seemingly dependent upon age; the value of the estimated variance within a bimodal distribution, the shape of the ARMA(p,q) spectral density function and whether the long range covariance was best described by a Matérn or fBm process. Analysing these parameters using a GLM with a logit link function showed a dependence upon the age pairing, gestational age at birth and gestational age at test. Additionally we found that the value of the variance has an effect on the other parameters, whereas this has previously been overlooked.

The way in which models have been estimated and analysed previously has focused upon the long range covariance sequence - specifically the Hurst parameter - and have ignored the variance parameter as just affecting the scale of the signal. We have found some evidence to suggest that the value of the variance affects other values in the spectral density function and should not be ignored.

From these analyses several multivariate Gaussian distributions were constructed to describe the parameters - which are grouped according to the three outlined parameter characteristics.

Finally, in order to realise similar signals from our estimated parameters we must include time-heterogeneous components appropriately; such a component is neuronal bursting activity, or delta brush. Analysing the occurrence of spontaneous delta brush activity we found a dependence upon age at birth and age at test. We noticed similar behaviour with regards to probability of spontaneous delta brush occurrence in infants born pre-

maturely and tested close to birth. Looking at the proportion of electrodes that express such activity showed a higher proportion only at infants tested close to birth.

The occurrence of delta brush activity in response to nociceptive and tactile stimuli was also investigated, showing dependence upon the age pairing - age at birth and age at test - with respect to both probability of occurrence and proportion of electrodes expressing delta brush activity. These show a similar behaviour to spontaneous delta brush responses, with a higher probability of expressing delta brush responses seen in infants born prematurely and tested close to birth. In contrast to spontaneous delta brushes, we see a similar proportion of electrodes expressing delta brush activity with infants born prematurely and tested close to birth.

This Chapter has highlighted some interesting behaviours and characteristics of neonatal EEG - especially with regards to the occurrence of delta brush activity - however we are limited in the conclusions that we can draw. This is as a result of the sample which we are analysing, which is unrepresentative as highlighted in the heat maps throughout. Throughout this Chapter we have looked at the age pairing - age at birth and age at test - in contrast to previous work which has focused upon age at test. Whilst the sample analysed is representative with respect to age at birth or age at test, it is not representative when looking at the aforementioned age pairing. As a result of this, further research is required to test these hypotheses suitably.

Chapter 5

Simulation of similar neonatal EEG signals

Up to this point in the research presented, we have analysed one signal at a time independently of any other signals recorded at the same time - i.e. univariately. A set of electroencephalogram signals contains multiple electrodes, each having an effect and relationship with the other signals within the set - i.e. multivariate. Up until now we have overlooked the multivariate nature of these highly complex signals. Although our univariate analyses has been successful in describing the second order structure, we have lost a large amount of information with regards to the interaction of electrodes; which would increase the accuracy of any simulated set of signals. Since this Chapter discusses simulation methods, it is crucial we attempt to regain this multivariate nature present within the signals.

The most obvious multivariate nature of these signals, evident from visual inspection, is the time domain correlation structure between electrodes; which has been documented in previous research [76, 77]. This structure was estimated by analysing the principal component weights of the signals within a recording [105], and classifying signals according to these weights. From this we were able to obtain a representation of the time domain correlation structure, represented by a set of adjacency matrices. Taking infants with 17 recorded electrodes, we were able to model the number of distinct groups of electrodes using a binomial generalised linear model. We can use this GLM to obtain the number groups expected for simulated infant, given the age at birth and the age at test,

and obtain a correlation structure by sampling from the set of observed adjacency matrices. This analysis describes an interesting feature of these signals: prematurely born infants and infants tested close to birth have a weaker correlation structure, i.e. more distinct electrodes. This is congruent with underlying theory, considering the growth and increase in neural connections that occur in the brain over development.

The last aspect we shall discuss is the way in which we can simulate a set of signals from an infants age at birth, and age at test. The focus of our simulations shall be background EEG activity as this illustrates the time-heterogeneous nature of our model and allows direct comparisons to be drawn. In order to assess our approach, we shall compare the resulting simulated signals against another approach outlined in previous literature.

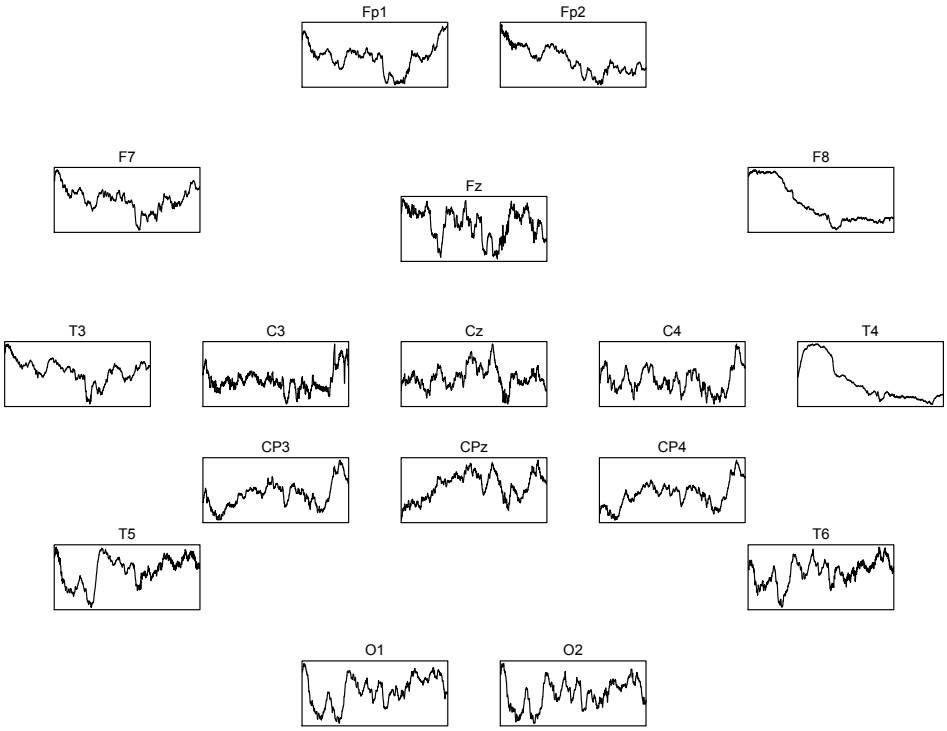
5.1 Modelling time domain correlation

The model presented in Chapter 3 provides an adequate description for 76% of suitable signals, and has been modelled in Chapter 4 using a group of multivariate distributions. We need to simulate full signals in order to assess the suitability of the predicted parameters.

Since we have estimated these multivariate signals univariately; we have lost information about the relationship between signals [109, 110]. If we had have estimated the signals' spectrum multivariately, we would have obtained a measure of association between the signals as described by the cross spectral density function. In order to provide adequate realisations from our model we must attempt to regain the inter-electrode recording structure. Visual inspection of the recorded signals shows that there is a large amount of similarity among signals. Whilst the parameters of the covariance sequence for the signals within a recording set are different, the underlying Gaussian stochastic process of the signals is similar; this is demonstrated in the plotted signals in Figure 5.1.

Figure 5.1 illustrates the multivariate structure we are wanting to capture with this procedure, the correlation between signals across time. If we wanted to capture the correlation between the covariance sequence parameters during estimation we would

Figure 5.1: Full set of electroencephalogram recordings from an infant PMA = 35.57, PMAT = 37.85



This figure illustrates a full set of neonatal electroencephalogram signals. There is visual evidence of a similar structure between electrodes as is to be expected given the multivariate nature of these signals.

require the use of multivariate Whittle estimation. Utilising multivariate Whittle estimation would allow us to ascertain not only a measure of the correlation between time domain signals, but additionally the components within the covariance sequences.

When simulating, a Gaussian stochastic process with orthogonal increments is used [75, 80, 111], which corresponds to ϵ_t - $\epsilon_t \sim \mathrm{N}(0, 1)$ - in Model 3.1; we shall refer to this as the random normal basis, as it is the basis to which we convolve the covariance structure to obtain simulated signals. Figure 5.1 illustrates that we should utilise the same random normal basis for some signals, in order to recreate the signals appropriately. Therefore, we need a means by which we can assess the similarity of the signals in the time domain. We require that for a set of EEG signals, we can obtain groupings of simi-

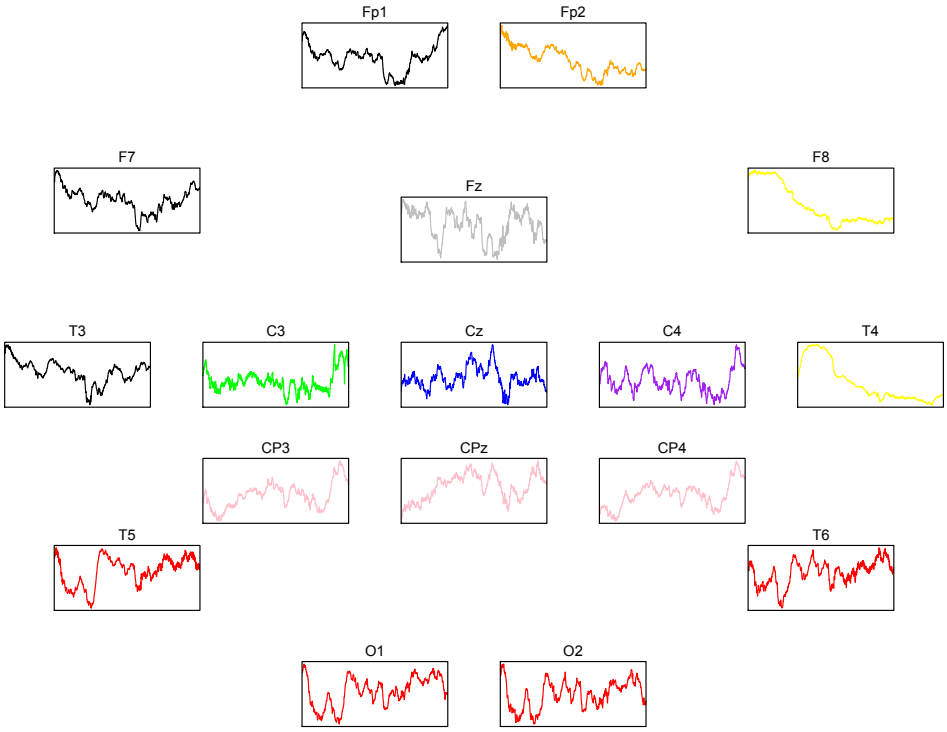
lar signals in the time domain. Whilst the correlation coefficient could be used [112], we are going to use principal component analysis to obtain an estimate of these groupings.

Definition 5.1.1 (**User tuned algorithm to determine the number of electrode groupings**). *The algorithm by which we shall estimate the number of distinct groups present within a set of signals is:*

1. *Scale the time series to avoid issues with regards to PCA and scale.*
2. *Obtain the linear decomposition of the time series by PCA - princomp in R.*
3. *Obtain the number of significant principal components as prescribed by Kaiser's criterion -* pc_n
4. *For each loading* $l_j^i; i = 1, ..., 17, j = 1, ..., pc_n$
   If $l_j^i > 0.1, l_{j_G}^i = 1$
   Else if $l_j^i < -0.1, l_{j_G}^i = -1$
   Else $l_{j_G}^i = 0$
5. *For each electrode* $e_i; i = 1, ..., 17$
   If $\underline{l_{j_G}^i} \notin G_{INF}$
   add $\underline{l_{j_G}^i}$ *to* G_{INF} *Where* G_{INF} *is the proposed groupings of electrodes*
6. $n_G = ||G_{INF}||$

Definition 5.1.1 outlines the algorithm used to obtain the estimated groupings present within the analysed signals. The value 0.1 was chosen as a minimum to have either negative or positive assignment so as to stop low loadings affecting the grouping algorithm. Taking the observed signals, once scaled, we obtain the loadings of up to the first three principal components. Using Kaiser's criterion to determine how many of the three sets of loadings to use [105], we then obtain groups depending upon the similarity of the loadings. Kaiser's criterion suggests the number of components to use based upon the eigenvalues of the loadings; the number of suggested loadings is such that the eigenvalues are greater than one. The results of the groupings on the data presented above can be seen in Figure 5.2.

Figure 5.2: Full set of electroencephalogram recordings from an infant PMA = 35.57, PMAT = 37.85 - Coloured according to identified groupings



This plot illustrates the result of the algorithmic procedure to identify groupings as outlined in Definition 5.1.1. We can see evidence of a similarity in the signal structure among the groupings. We can also notice slight similarities between the groupings as well, although this is not as similar as within the colour coded groups. The autonomous and algorithmic nature of this process removes any bias in the analysis with regards to different analysts identifying different structural patterns.

Figure 5.2 shows that the grouping procedure identified nine different groupings present in the observed signals, corresponding to an adjacency matrix comprised of nine disjoint graphs. This shows that utilising a different white noise parameters for each simulated electrode is, most likely, an unrealistic recreation of the data. As previously described EEG signals contain a highly complex multivariate structure [76,77]; utilising a different basis for each electrode is further shown to be an inadequate simulation method.
By utilising n_E independent Gaussian stochastic processes to serve as a basis for the covariance structures, we do not recreate the similarity between signals recorded at dif-

ferent electrode sites. In the presented example, rather that simulating 17 independent white noise signals, to serve as a random normal basis to which we convolve our estimated covariance sequence, we would simulate nine independent Gaussian stochastic processes. This is an improvement upon previous simulation techniques [75, 112], but has room for improvement itself.

The benefit of this analysis is that we now have a measure of the association between signals recorded at different electrode sites within an infant. First, this provides a multivariate structure that we can apply to our univariate estimates by attempting to recreate the observed similarity between signals recorded at different electrode sites. Second, we have a statistic by which we can assess the similarity of signals recorded at electrode sites across development.

Taking the proportion of groups, $\hat{p}_n^G$ in a set of signals as data points, we obtain a set of proportions we can model using a binomial generalised linear model. We are looking at the proportion of groups instead of the number of electrodes so that we do not get a suggested number of groupings that is outside the maximum number of electrodes allowable - in our data set that value is 17. We then fit this model using stepwise procedures and the age pairings as explanatory variables, again utilising principal component analysis to avoid multicollinearity issues in the regression. The linear transformation of the variables age at birth and age at test are given in Model 4.4.1. The resultant linear model is below with the interaction term significant at $\alpha = 0.06$, and is supported by stepwise regression.

Model 5.1.1 (**Binomial GLM for expected number of electrode groupings**)**.** *The number of distinct underlying Gaussian stochastic processes in an infant - expressed as a proportion of the number of electrodes within a recording - is modelled by the logit function,*

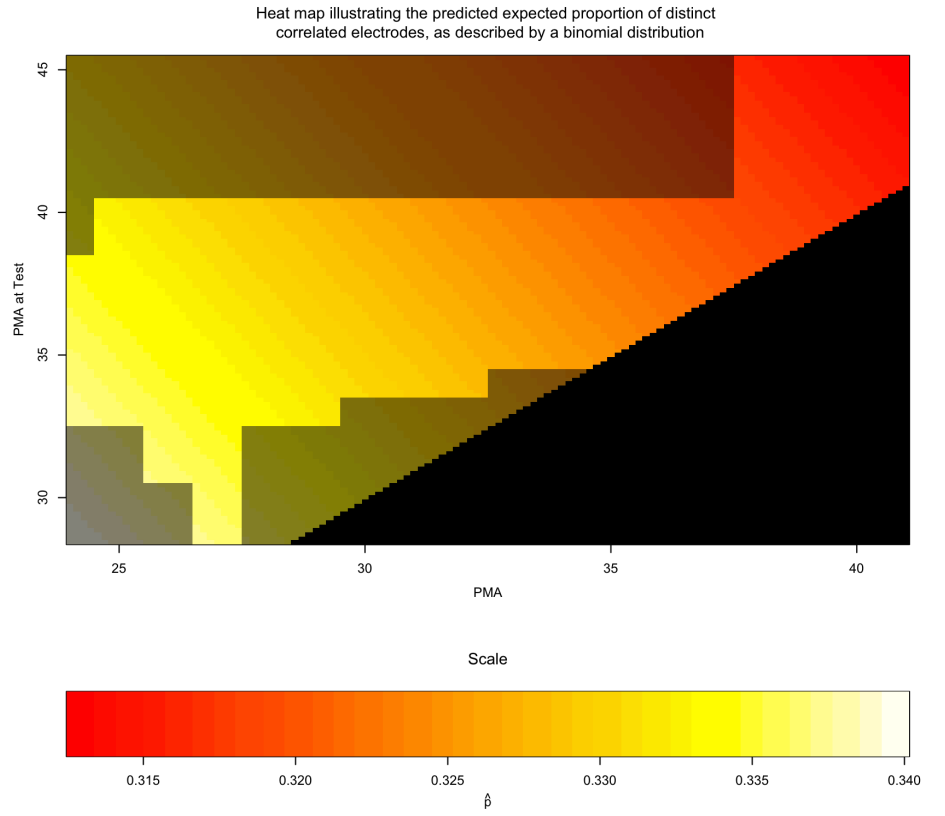
$$\log\left(\frac{\hat{p}_n^G}{1-\hat{p}_n^G}\right) = -0.73279 - 0.00543 \times \mathrm{MCA}_1^{\mathrm{S}}$$

Interaction significant at $\alpha = 0.06$

As a result of this we can obtain the expected proportion of correlated electrodes, given an age pairing and simulate from the binomial GLM to obtain the number of distinct

white noise signals to simulate; Figure 5.3 shows predicted values from Model 5.1.1

Figure 5.3: Heatmap of predicted expected proportion of electrode groupings



This figure shows the predicted p expected proportion of electrode groupings. The number of electrode groupings identified by the outlined algorithmic procedure is divided by the total number of electrodes present. This is done to allow the use of logistic regression and to ensure that any proposed grouping is not larger than the total number of electrodes within a recording. The greyed out areas indicate areas where data points are not present within the sample.

Similarly to the delta brush heat maps, we have indicated the areas that we cannot have data and also areas for which have not been represented in the sample; these sections of the graph are indicated by a dark overlay in Figure 5.3. From the heat map in Figure 5.3 we can see a trend in the number of distinct groupings across development. The colouration of Figure 5.3 indicates that if an infant is born prematurely, or if they are tested close to birth, then the predicted proportion of distinct groupings is higher; and

becomes lower as the infant matures. This could be due to the formation of regions in the brain, however as mentioned, more investigation is required to confirm or reject the hypothesis of greater similarity between electrode sites as an infant matures. Figure 5.3 shows that in infants tested between 33 and 37 weeks post menstrual age at test, we have reason to investigate this further. This is because infants born extremely prematurely and at the brink of viability, 25-27 weeks [61], have a higher expected value for groupings. This lessens the later an infant is born, and increases again the shorter the time frame between birth and test - according to our model.

Having identified trends and constructed models for neonatal development in response to somatosensory stimuli, as well as having constructed a multivariate framework which we can apply to our univariate estimates. We shall now simulate similar signals, given a set of ages, and compare them against observed signals and selected current techniques for simulating neonatal EEG.

5.2 Simulating a similar set of neonatal electroencephalogram signals

The simulation of structurally similar neonatal EEG signals is possible, however a direct comparison to observed signals is grossly impractical and extremely difficult to be accurate due to the unobserved random noise component described by ϵ_t in our model outlined in Definition 3.1. We can however, compare the properties of simulated signals against observed signals, and compare against simulated signals from the estimated parameters.

Having constructed models for the time domain correlation structure and the estimated parameters, we shall assess their ability to simulate similar neonatal electroencephalogram signals. We shall compare our models against the time-varying parameter model proposed in previous research [75], and determine the suitability of our simulation model. Due to the resolution at which previous modelling techniques were performed, we shall sample our signals to be comparable.

The estimation and analysis of neonatal EEG signals has primarily been to focus upon

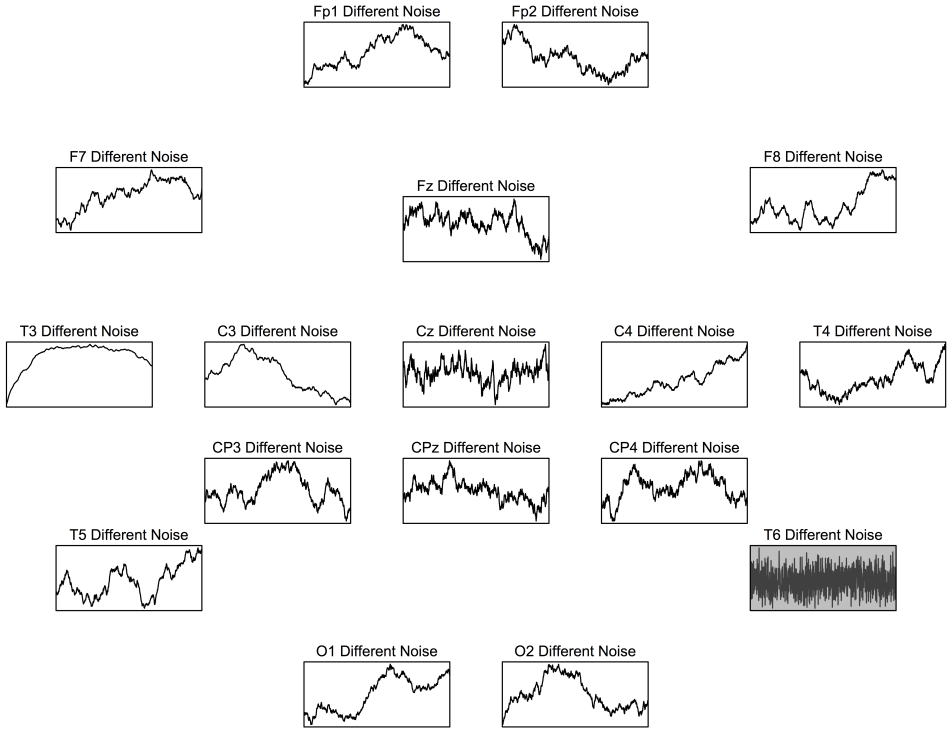
seizure detection. Simulated signals have been utilised as a baseline against which comparisons can be performed to ascertain whether a seizure is present within a signal. The aim of simulating signals is similar in our case, we wish to simulate a set of signals that can be utilised as a baseline. The advantage of our model is that it is obtained from high resolution data and as such can be sampled for lower resolution signals, the inverse is not true for current models.

First the time domain correlation structure shall be tested, this shall be done by simulating the following levels of association between the underlying Gaussian process: All the same random noise, different random noise for each electrode and the proposed random noise according to the estimated structure - as obtained in Section 5.1; simulations of these can be seen in Figures 5.4, 5.5 and 5.6. To ensure we are just testing the appropriateness of the random noise structure, we shall simulate signals with the estimated parameters. The estimates for the covariance structure were obtained according the procedure outlined in Chapter 3; all the electrodes are adequately described by the proposed model apart from T6, which is reflected in the plots.

Figures 5.4 and 5.5 show that utilising the same white noise structure does not produce signals with a similar structure to the observed signals, neither does utilising a different noise structure for all electrodes. The grouped structure provides a decent approximation, however as mentioned earlier, we still have disjoint signals. Although, there is room for improvement, it is clear that the best white noise structure takes into account the correlation between electrodes rather than assuming they are all the same, or all different. The spatial nature of neonatal EEG has been studied and found that a high number, and spatial density, of electrodes is required to capture this spatial correlation [77].

Now that we have demonstrated the need for a correlated white noise structure when simulating we shall proceed to assess the suitability of the simulation method. Focusing upon background EEG, we shall now present the method by which we can realise signals with a similar structure to those observed.

Figure 5.4: Signals simulated from estimated parameters with different random bases PMA = 35.57, PMAT = 37.85



This figure illustrates that utilising a different $\epsilon(t)$ when simulating does not recreate the structure of the signals as presented in Figure 5.1. The parameters used were estimated using the outlined methods in this thesis and all apart from T6 were deemed adequately described by the outlined ADP.

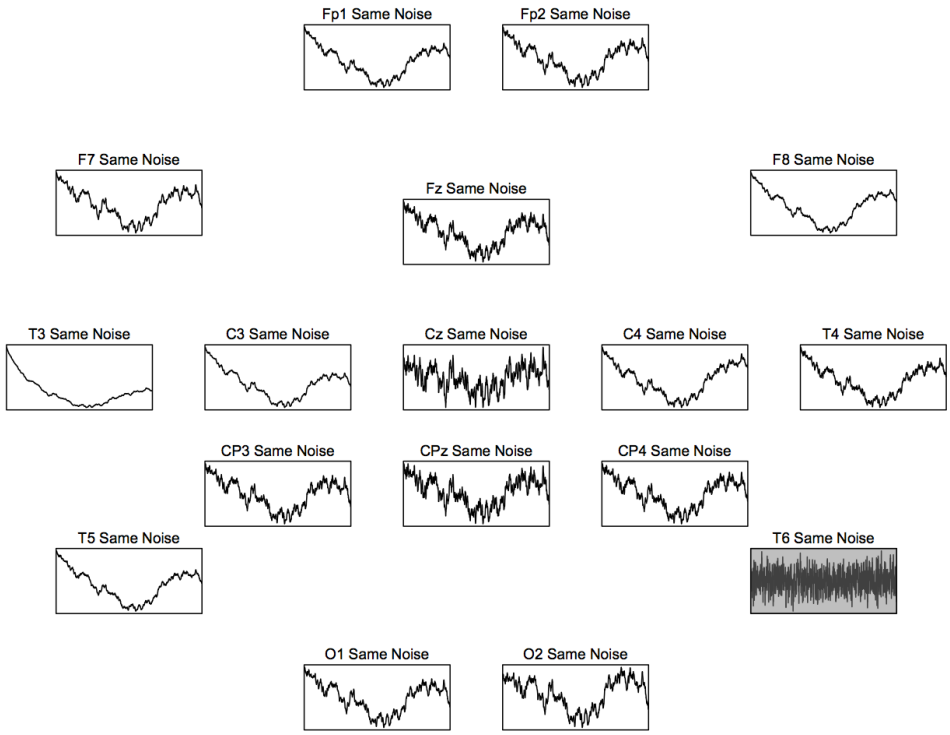
Theorem 5.2.1 (**Background EEG Simulation Procedure**)**.**

1. *For an age at test and age at birth pairing obtain the predicted values from the principal component reductions to avoid multicollinearity presented in Model 4.3.2 and Model 4.4.1. Denoted MCA^C and MCA^P respectively, these are used in the structure simulations and parameter simulations respectively*

2. *Obtain a time domain correlation structure for the random noise, according to Model 5.1.1. After obtaining n_G from Model 5.1.1 we then sample from observed structures.*

3. *Obtain the probability of spontaneous delta brush occurrence and the number of*

Figure 5.5: Signals simulated from estimated parameters with the same random base PMA = 35.57, PMAT = 37.85



This figure illustrates that utilising the same random normal base when simulating does not recreate the structure of the signals as presented in Figure 5.1. Therefore this approach should not be used when simulating a set of neonatal EEG signals. The parameters used were estimated using the outlined methods in this thesis and all apart from T6 were deemed adequately described by the outlined ADP.

electrodes expressing delta brushes, according to Models 4.4.2 and 4.4.3 respectively.

4. *For the structures obtained include delta brush activity appropriately in the random noise components.*

5. *Obtain the probabilities of the electrodes being within the segregated groups: Bimodal, Matern and Shape, according to Sections A.1, A.2 and A.3. From these the binary parameters can be obtained by drawing from a Bernoulli distribution.*

6. *Use the obtained binary parameters, electrodes, and* MCA^P *to obtain predicted*

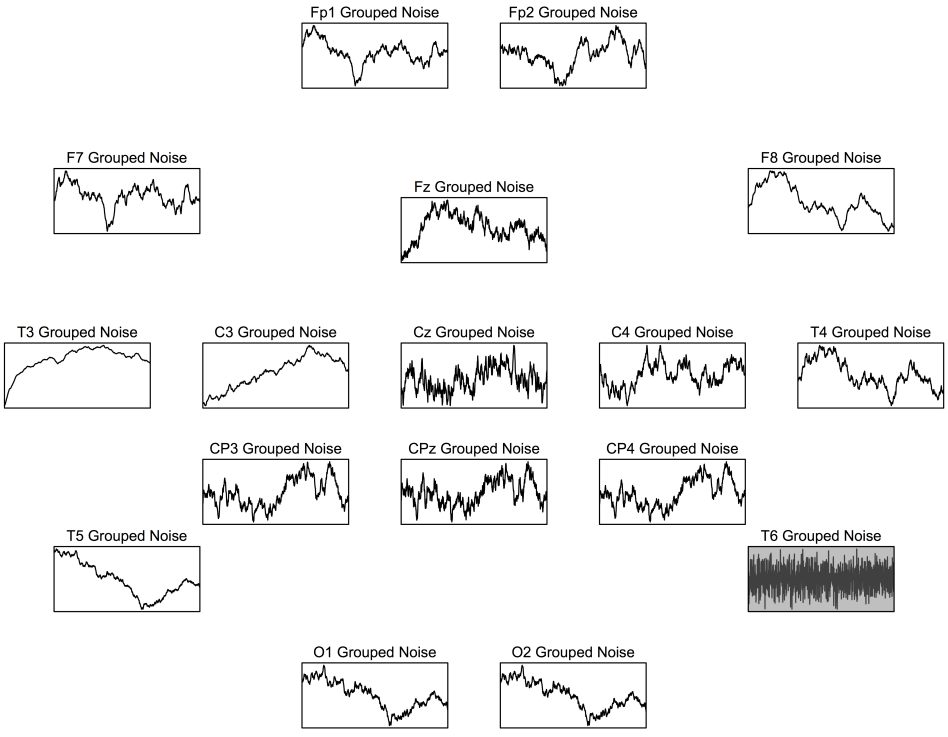
parameter values from the segregated parameter distributions.

7. *Utilising the Cholesky decomposition method, simulate the long range covariance structure as described by the non aliased Matérn parameterisation.*

8. *Apply the short term covariance structure, as described by the autoregressive moving average process, and the band-pass filter to the signals.*

In order to compare the proposed simulation against another outlined method for simulating neonatal electroencephalogram signals, we have presented a method to simulate background EEG signals. Although we have focused upon background EEG simulation, this procedure can be adapted easily to simulate other time-heterogeneous components;

Figure 5.6: Signals simulated from estimated parameters with grouped random base PMA = 35.57, PMAT = 37.85

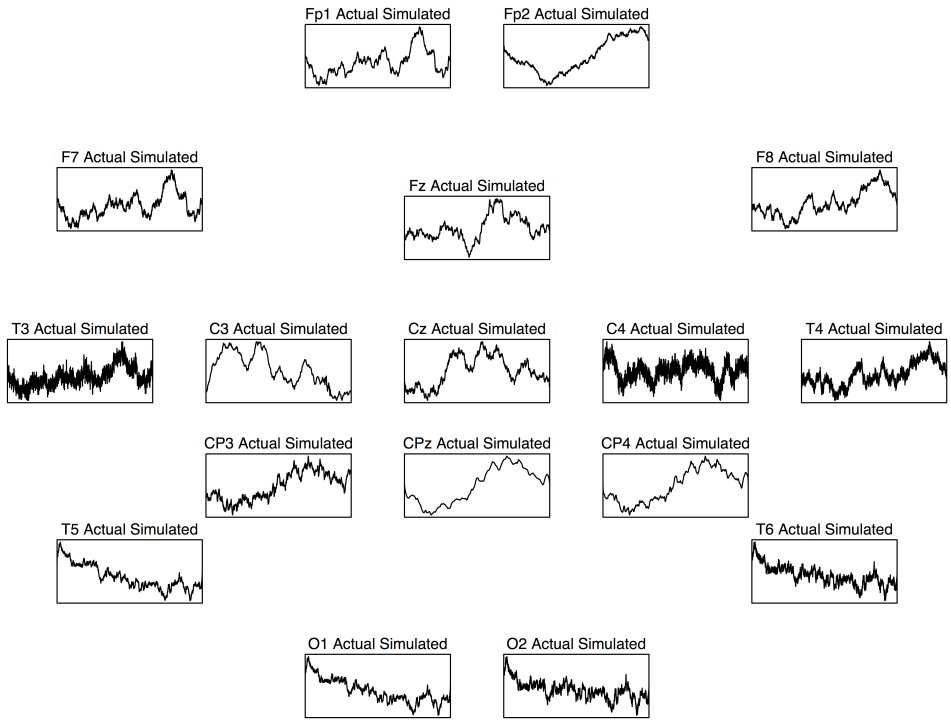


This figure illustrates that utilising a grouped $\epsilon(t)$ when simulating is a better approximation to the structure of the signals as presented in Figure 5.1 than Figures 5.4 and 5.5. The parameters used were estimated using the outlined methods in this thesis and all apart from T6 were deemed adequately described by the outlined ADP.

such as somatosensory stimuli response or possibly neonatal seizure activity.

The signals simulated in Figures 5.4, 5.5 and 5.6 were created utilising this method. We shall now assess the simulation of parameters from our proposed distributions by utilising the simulated parameters to generate signals. Figure 5.7 illustrates signals simulated with the estimated signal structure outlined in Figure 5.2.

Figure 5.7: Signals simulated from simulated parameters with estimated signal structure PMA = 35.57, PMAT = 37.85

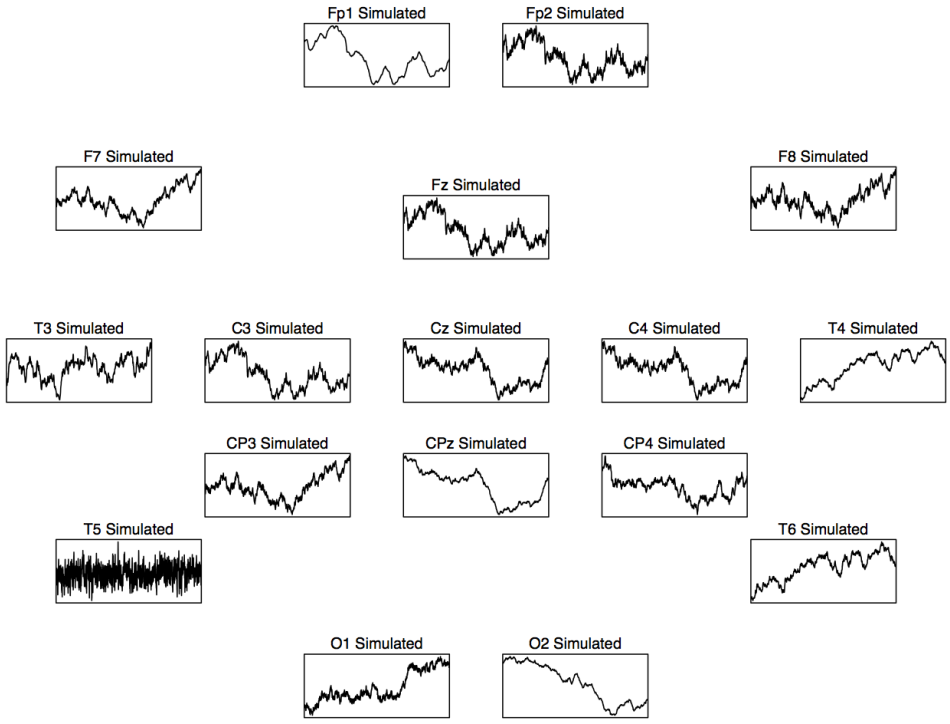


This figure shows parameters simulated from our suggested parameter distributions - Appendix B - and the actual structure estimated in Figure 5.2. We see issues regarding the proposed distributions realising parameters due to the number of estimates from which they have been constructed.

Figure 5.7 highlights an issue with the proposed distributions utilised in electrodes T3 and C4, which is the number of estimates from which the parameters are constructed is not large enough to produce reliable estimates. This issue is highlighted by the structure of the simulated signals which is not similar to that observed in Figure 5.1; further analysis is required to refine these distributions.

Figure 5.8 illustrates a fully simulated set of signals where the structure and parameters are estimated. Utilising the age pairing 38 weeks at birth and 45.29 weeks at test, we utilised the proposed models to simulate a set of signals recorded at 17 electrodes. Apart from the simulated parameter issue outlined, the structure simulated seems appropriate and is similar to the observed signals in Figure 5.1.

Figure 5.8: Fully simulated signals PMA = 38, PMAT = 45.29



This figure illustrates a set of signals fully simulated utilising an age pairing of 38 weeks at birth and 45.29 weeks at test. Upon visual inspection these display traits and characteristics evidenced throughout our data set, and as illustrated in Figure 5.1. At electrode site T5 we see issues due to the estimated parameter distribution, however as a set of data it appears to capture the nature of the signals analysed throughout.

The problem with comparing the structure of EEG signals is the underlying inter-patient variation. Having estimated the parameters in the frequency domain, we have described the underlying structure, however EEG analysis is also performed within the time domain. Apart from T5 in Figure 5.8 the signals produced are visually similar to the EEG signals presented in Figure 5.1. Comparing these in the frequency domain could result

in the signals being dissimilar due to a different underlying generating mechanism, albeit from a relevant distribution. If we were to look at correlation between a simulated and observed electrode that would be inappropriate also due to the differences in the random noise of the structure; as such we are relying upon visual similarity to assess appropriateness.

Figure 5.8 illustrates that similar signals can be realised, however it shows issues with regards to some parameterisations being unsuitable. This is further seen in Appendix C which contains further simulations across a range of age pairings. However we have managed to recreate the time-heterogeneous structure by including the delta brush component, often overlooked in other neonatal EEG models. Whilst we have this time heterogeneity accounted for, it is clear that we still need to find when the parameters change in order to adequately simulate neonatal EEG.

5.3 Comparison against time-varying simulation method

One approach to simulate signals with similar characteristics is to utilise time-varying fractional Brownian motion [75, 112] since it is understood that EEG signals are quasi-stationary - i.e. comprised of time-homogeneous segments [1–3]. Whilst this addresses the quasi stationarity, it ignores the time-heterogeneous delta brush component. Most, if not all, of the previously presented models for neonatal background EEG omit the delta brush when producing simulated signals [37, 39, 70, 75, 112–114].

The time-varying fractional Brownian motion process draws values for the Hurst parameter from a Gamma(α, β) distribution, where $\alpha = 7.82$ and $\beta = 7.44$ [75]. We shall simulate six second intervals at 2000Hz, changing parameters every 0.5 seconds to directly compare against a signal from our model. After this we shall sample both signals to 64.5Hz, to provide a more appropriate estimate for the time-varying simulation process. So we can assess the effect of the time-varying parameters, we shall use the same underlying Gaussian process for both simulated processes. Figure 5.9 shows the 2000Hz signals and the signals sampled to 64.5Hz. The observed signal presented was failed to be described by our model, and the simulated parameters were generated according to the outlined procedure.

At 2000Hz the presented model more closely resembles the signal due to its similar characteristics with regards to the self similarity of the process, as indicated by the roughness of the signal. However, at 64.5Hz both the time-varying parameter and the presented model are both close approximations to the signal. This is because at this resolution the short term covariance structure does not have an effect on the process. In our estimations of the underlying covariance sequence, the short term covariance was described using an autoregressive moving average process up to order (3,3). Therefore if we subsample at 500Hz or lower we might not see the effect of any short term covariance on the signal.

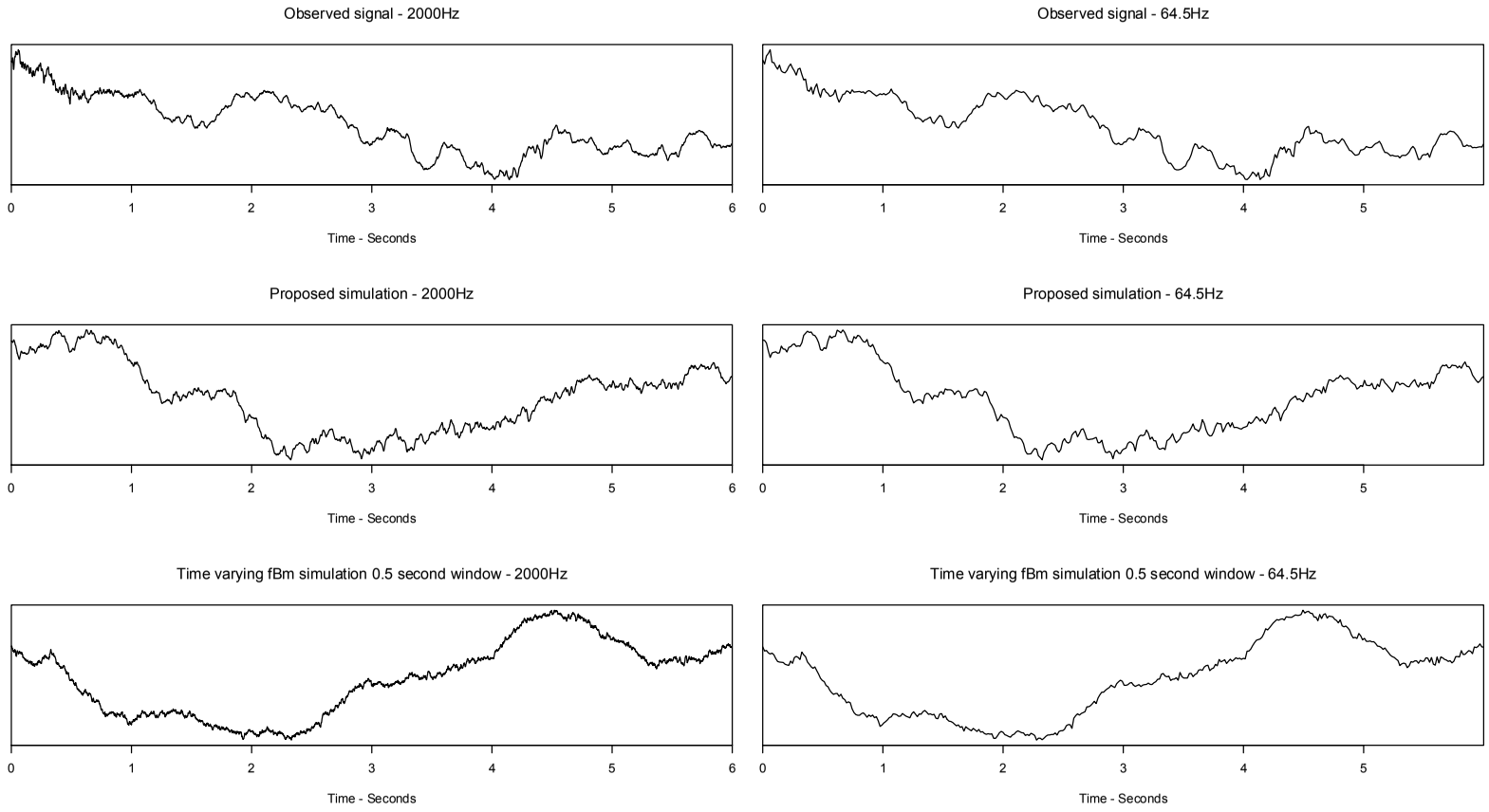
Despite this the presented model is technically a better representation, as we have better represented the time heterogeneity of neonatal background EEG, by taking into account the presence of delta brushes. The time-varying fBm model however, takes into account the changing parameters albeit randomly, and is a better description of the first second of the observed signal. We can see some possible evidence of a changing parameterisation in this first second of the observed signal at 2000Hz, again observable by change in the signals' self similarity, however this is less pronounced at 64.5Hz.

A combination of these models, taking into account the changing parameterisation of the signal, the short term covariance as well as the delta brush component would be better. However we have not been able to identify the point at which the parameters change, and it is better to present the time-homogeneous parameterisation with delta brush added then a random change in parameters. Our model could be used for such a parameterisation we just need to find where we change the parameters and to take care in doing so.

This comparison highlights an important advantage of our model and process over other methods; we can sample our 2000Hz simulated signals to obtain other recording resolutions, however other processes cannot accurately describe a higher resolution than they have been fitted on. This comes down to a fundamental property observed from the data, the short term dependence; this has not been fitted previously and, according to our fitting, possibly not present in resolutions lower than 500Hz due to the order of the fitted ARMA(p, q) process.

The time-varying nature of the parameters could be recreated utilising our parameter distributions, as we could draw realisations in a similar manner to the time-varying simulation. However, extreme care would have to be taken at the time points where the parameters are appended due to the short term covariance structure. Furthermore, an understanding of when to change the parameters, rather than an arbitrary change-point, would be preferred

Figure 5.9: Comparison of presented model and time-varying fractional Brownian motion against an observed signal



This figure illustrates an EEG signal from within the analysed dataset (top), and simulated signals according to the outlined procedure (middle) and a previously suggested method (bottom) at 2000Hz and 64.5Hz. Whilst both simulated signals appear to recreate the observed signal at 64.5Hz, the previously suggested method fails to recreate the high frequency detail of the signal at 2000Hz.

5.4 Results and Conclusions

Throughout this thesis univariate methods have been employed successfully to describe these highly complex multivariate signals. However, we have overlooked the multivariate nature by not taking this into account whilst performing the estimations. In order to realise similar signals to those recorded, we must address this lack of information.

In Chapter 4 we analysed the proportion of electrodes that express delta brushes simultaneously or in response to nociceptive or tactile stimuli. In this Chapter we utilise the same methodology to describe the amount of time domain correlation seen in a set of EEG signals. By obtaining a measure of similarity from principal component analysis we analysed how the similarity between activity at electrode sites changed across development.

We found that as an infant matures the similarity between recorded activity at different electrode sites increased; we also noticed that infants born prematurely have more distinct electrodes than those born closer to term. As mentioned in Chapter 4 we lack a representative sample with regards to the age at birth and age at test pairing utilised, as such we require more data points in order to test the results obtained. The distinction of the recorded activity at electrode sites is crucial to incorporate into any simulated data. Given a description of the association between activity at electrode sites - as well as information about the behaviour of delta brush activity contained within Chapter 4 - we were able to define a simulation method that produces signals exhibiting similar behaviours to the signals observed throughout this thesis.

A key aspect of the research undertaken is the high resolution at which the signals are recorded. This enabled the detection of the short range covariance structure present within these signals. Previous simulation methods have been described at a lower resolution and as such did not detect this component of the covariance structure. This has an effect on the strength of the simulation procedure as we can sub-sample to reduce the resolution, however a higher resolution cannot be described. Combining the results obtained throughout this thesis we have been able to realise a set of signals with a similar structure to the observed neonatal EEG signals.

Chapter 6

Conclusions and Future Work

This thesis outlines: a model that describes 76% of analysable signals, a description of the time domain correlation structure and focuses upon the development of neonatal electroencephalogram signals. In this Chapter, we summarise the results obtained, present the limitations of the research and answer the questions of interest outlined at the very start of the thesis. From these conclusions we discuss possible extensions and future avenues of work.

6.1 Conclusions

The major contributions this research has made to the understanding of neonatal electroencephalogram signals are twofold; first that there is a long and short term covariance structure evident in these recordings. Where previously recordings were obtained at 64-200Hz, as this was the area of interest, we focused upon modelling the signals at as high a resolution as possible. In most situations 2000Hz is the highest resolution utilised for electroencephalography, as it provides detailed information, as well as flexibility with regards to subsampling. This resolution allowed us to identify the previously unnoticed short term covariance structure present within neonatal EEG. Secondly, our presented model takes into account the time heterogenous properties that are observed in neonatal background electroencephalography signals. Whilst our focus has been upon spontaneous neuronal bursting activity - "delta brush" - this model could also be adapted to multiple time-heterogeneous components.

This flexible and detailed model - constructed using the time, frequency and time-frequency domains - was able to describe 69% of all the signals in the presented data set; using only three time-heterogeneous components: delta brush as well as the noxious and tactile somatosensory specific responses. Of the unexplained signals 9% of the signals had an unidentified time-heterogeneous component, which was unsuitable to be modelled by the outlined and implemented approach. This could have been an error in the recording process, however we removed these signals from analysis as they were unsuitable for future analyses. With these signals omitted from the data set our simple, flexible and detailed model described 76% of the data. Furthermore this model can be viewed as a first step, and a robust baseline that can be improved upon with future developments.

Our model admitted the Matérn process as the description of the long range covariance sequence, previously described using fractional Brownian motion. The Matérn process is a stationary process that admits fBm as a special case, furthermore the inclusion of the distance parameter affects the mean reversion of the process. This development allows for flexibility with regards to the range of possible parameterisations, and has not previously been used. With the identification of the Matérn process being able to describe neonatal EEG, as well as the evidence of an additional short term covariance structure, we wanted to gain an understanding of why these values occurred. Unfortunately, due to a poor sample with regards to the age at birth and age at test pairing, we are limited with the strength of conclusions that we can draw. A representative sample means that we can make inferences about the population from which it has been drawn. However, due to the sample being analysed there would be a large amount of extrapolation and as such further testing is required to confirm or reject the results obtained.

The inherent time-heterogeneous nature of neonatal electroencephalography is poorly understood, with seemingly arbitrary changes in parameters reflecting this change. We attempted to determine the point at which the parameters change, by assessing the behaviour after the identified time-heterogeneous components in our signals. However we found that the parameters were not significantly different, following these components. Whilst disappointing that we were not able to identify the point at which the parameters change, the outlined approach and presented model would be able to be utilised.

From the parameters we were able to discern three behaviours in signals that are seemingly dependent upon age; the value of the estimated variance within a bimodal distribution, the shape of the autoregressive moving average spectral density function and whether or not the long range covariance was best described by the Matérn process. Although in previous research the value of the variance is largely overlooked, we found that the value of the variance had an effect on the other parameters. Previous models have ignored the contribution of the variance to the model, choosing to focus solely upon the long term covariance structure of the signals. Whilst the signals may be scaled by an EEG practitioner, the effect of the variance on the other parameters shows that this is important to take into account and should not be viewed as an arbitrary value that affects only the scale of the time series.

The age relationship seen in these parameters is interesting, given that the possible relationship is not just dependent upon the age at test, but also the age at birth. We found similar results in infants who were tested close to birth and those born prematurely. This is one of the first pieces of research to study non-behaviourally the effect of time since birth on responses.

Utilising these parameter characteristics we constructed several multivariate Gaussian distributions. Estimates obtained by minimising the Whittle likelihood are asymptotically normal. However, the observed parameters have a distribution that is a convolution between the estimation error distribution and the original distribution of the sample parameters. So whilst the error distribution under Whittle estimation is normal, the estimated parameters distribution might not be due to the original distribution of the parameters. As the original distribution is unknown, unsuitability in the simulated parameters could be due to the MVN distributions being too severe a departure in some cases from the original unknown parameter distribution. Due to the complexity and variability of the data set the distribution could require non parametric modelling. Whilst the estimated Gaussian distributions might produce unsatisfactory estimates for some segregations, we obtain suitable parameterisations for segregations with a suitable number of observations.

This research has estimated the properties of neonatal electroencephalogram signals uni-

variately, however in order to realise similar signals we had to recreate the similarity of time series between electrode sites. To obtain an estimate of the time domain correlation present within the signals, we analysed the loadings of a principal component analysis performed on the recordings of an infant. This method looked for similarity between the loadings of the significant principal components at electrode sites - significance of the principal components was determined by Kaiser's criterion. If electrodes displayed similarity in the loadings then we grouped them together, finishing the grouping procedure once all electrodes were placed into a group. This approach is a user tuned clustering of signals based upon principal component analysis, whilst it may only be an approximation of the time domain correlation structure of the time series', however it illustrates an important point when simulating a set of neonatal EEG signals which cannot continue to be ignored.

In obtaining this framework, we noticed a pattern in the number of distinct electrode groupings; mainly that recordings at electrode sites become less distinct, i.e. more similar, as development occurs. Again, we noticed similar results in infants born prematurely as well as infants tested close to birth. We found that we can recreate the time domain correlation structure by utilising the same underlying Gaussian stochastic process across certain electrodes when simulating, an aspect overlooked in previous research. In regards to the spontaneous occurrence of delta brushes within the data and delta brush responses to somatosensory stimuli we observe a similar dependence upon the ages at birth and test, whereas previous research focused solely upon age at test. It would be of interest to see whether the pattern obtained with regards to the multivariate nature of the data being dependent upon both age at birth and test are obtained given a more representative sample.

Simulating from the predicted parameters, generated from the presented multivariate distributions seen in Appendix B, we see issues with regards to the scale of the signals. Given that other simulation methods ignore the variance [75, 112] we compare our method against others according to the way in which the structure has been described. Simulating similar signals at 2000Hz we find that a time-varying fractional Brownian motion, with parameters changing every 0.5 seconds, does not describe the signal's structure as well as our limited distributional parameterisations. Furthermore, our approach

better describes the signal's behaviour as we include the time-heterogeneous delta brush component. When sampled to 64.5Hz, both models seem to provide adequate realisations; this is because the lack of a short term covariance does not affect the simulated signal at this resolution, however the delta brush component is still missing. Whilst the time-varying fractional Brownian motion omits features from the simulated signal, the changing parameters is a recognised weakness in our approach. However, randomly changing the parameters is not a sensible approach to describing these complicated signals; we would find out why the parameters change, or at least how long the parameters stay the same for. By including the effect of the delta brushes in the simulated signals we have produced an inherently better model, as we have not ignored an inherent property of neonatal electroencephalography.

With any research it is important to identify the limitations of the conclusions drawn, the main limitation in our research is that the sample is not representative; specifically with regards to the age at birth and age at test pairings. When looking at the age pairing we do not have a representative sample, as we are missing data points for key age pairings. We are missing these due to experimental restrictions, such as consent issues or doctors allowing the requisite heel prick procedure to performed.

Although we have a large data set the lack of a representative sample in this regard limits the strength of the conclusions that we can draw, and future research is required to confirm our findings. Additionally we have estimated these multivariate signals utilising a univariate approach so we could obtain results in a reasonable time frame. As a result, we have lost information with regards to the interaction between signals at different electrode sites. Finally, in our estimation we utilised a stepwise approach, estimating the long-term covariance and then the short term covariance, inducing unwanted bias into our estimates. Future research should limit these issues to improve upon the accuracy of any obtained results.

6.2 Future Research

No research is perfect, and there are always ways in which we can improve upon the results obtained. This thesis has presented the progress made in modelling only a small

part and a short time window of neonatal brain activity. The inherent beauty of modelling electroencephalogram signals, such as the ones presented, is that there is always more that we can uncover. Much like the previous research outlined in the introduction, this research can provide a stepping stone for future research avenues, some possible directions of which are outlined now.

First, with regards to our model and estimation procedure, we could model bipolar montages, as analysing such montages is a common technique, possibly utilising bivariate or complex Whittle estimation. In such an approach we would let one signal be the real component, and the other being the complex component; this approach has been successfully employed in modelling oceanographic currents, and could be of interest here [83]. The obvious extension of this would be to model the entire signal multivariately, however given the number of electrodes that can be recorded, this might be infeasible as it would be incredibly computationally intensive.

Additionally in any future estimation of parameters, the Matérn covariance would be used from the start; instead of only implementing this covariance structure should fBm fail to be an adequate description. Care would have to be taken for signals where the range parameter is equal to zero, $d = 0$, however as we are fitting in the frequency domain this could be set as a parameter value. Although in the covariance sequence a range parameter being equal to zero is infeasible, we could obtain estimates of this when fitting spectra. As $d \to \infty$ the process becomes less similar to a random walk process and more similar to white noise. Therefore as the distance parameter increases the process will stay around the mean more. As such it would be interesting to see the way in which the distance parameter changes across windows within a longer signal.

Finally in terms of modelling approaches, improving the way in which we describe delta brushes would be of interest. Delta brushes are a commonly overlooked component of neonatal EEG when modelling signals, especially when creating a baseline for seizure detection. In the outlined research we implemented a band-pass filter over the range of possible frequencies; it would be of interest if we modelled the delta brushes in the time-frequency domain, either by inverting the wavelet transform, or by providing a model that could assess the change in the activity across development. The parameterisation

approach for delta brushes would fit in our model, as it is flexible with regards to time heterogenous components. However, we might require an approach other than addition to adequately include these components.

Analysing the estimated parameters of our model, we found that we have no reason to change parameters during these six second signal segments; which is counter intuitive with regards to the expected properties of electroencephalogram signals [1–3, 61]. Whereas previous simulation methods have utilised time-varying parameter processes, at arbitrary lengths, it would be of interest to see if the time at which the parameters change could be identified and what could cause the change in parameters. Whilst our model describes the overlooked delta brush time heterogeneity, the constant parameters across a longer length of signal are completely counterintuitive to the properties of electroencephalogram signals. Furthermore, our modelling of the estimated parameters could also be improved by taking into account the possible departures from a multivariate Gaussian distribution.

Given the univariate estimates, we constructed a multivariate framework in order to produce similar signals. When simulating, we sampled from adjacency matrices created from the signals within out sample. Instead of sampling from those obtained from observed signals, it would be of interest to see if we could simulate similar adjacency matrices. We require that the adjacency matrices form n_G disjoint graphs from a matrix of probabilities. One way in this might be achieved is utilising Monte Carlo Markov chain in adjacency matrices, similar to the approach of image noise reduction [115]. We can represent the grouping of electrodes as a symmetrical adjacency matrix, where a value of 1 in an off diagonal element illustrates that the electrodes belong to the same grouping. Furthermore, we can ascertain transition probabilities according to the groupings identified within the data set. It is possible therefore that we could simulate a random adjacency matrix utilising the transition probabilities and the Bernoulli distribution. From this we stage we could perform MCMC in a similar manner to image noise reduction to obtain n_G disjoint graphs. [115]. This approach might be applied similarly with regards to the simultaneous expression of delta brushes. These problems could also be addressed through modelling the signals using an n dimensional multivariate estimation approach.

Our model provided a univariate description of the signal, and we then applied a multivariate framework to simulate similar signals; this is another commonly overlooked aspect of the data when modelling. However, due to the sample being poor with regards to the age at birth and age at test pairing analysed we cannot draw solid conclusions about the multivariate framework obtained. In the majority of the analyses performed, both age at birth and age at test had an effect on the signal's properties; with premature infants showing similar behaviour to infants tested close to birth. One interesting aspect identified was the time domain correlation within the signals. However this is described, testing whether the formation of regions within the brain is dependent upon the ages at birth and test would be an interesting avenue to pursue. This would be easily tested, as we could observe an infant across the weeks after birth, without having to perform the painful or tactile stimuli.

To summarise, our research has possibly raised more questions than it has answered; due primarily to the observation that the ages at birth and test affect the signals. We are limited in the conclusions that we can draw due to our data set forming an unrepresentative sample, with respect to the age pairings. Given more time and a larger, more representative data set, these are the questions of interest that we would pursue.

1. How does the inherent time-heterogeneous nature of neonatal EEG signal parameters occur?

   The signal characteristics will not remain the same for long periods, different waves will be present changing the signal. We are analysing six seconds, which has been identified as quasi-stationary in electroencephalography.

2. Does the way in which parameters change across time remain the same across all developmental ages?

   It would be of interest to analyse the interval lengths of time-homogeneous segments, specifically how the inherent time-heterogeneous nature of neonatal EEG signals changes across development.

3. Does prematurity affect the development of regions within the brain?

   It is understood that as development occurs the growth and increase in neural

connections that occur in the brain over this time. However, we have illustrated that there is less correlation between activity recorded at electrode sites from premature infants. Therefore it would be of interest to analyse whether prematurity affects this formation of regions within the brain.

4. What further insights into the development of neonatal EEG will be obtained from modelling bivariately or mutlivariately?
   A recognised weakness of our analysis is the loss of information from estimating these highly complex multivariate signals univariately. It would be of interest to see what further information could be obtained by utilising a bivariate or multivariate estimation procedure. This could improve the detection of time domain correlated signals utilising information contained within the cross spectral density functions.

5. Could we apply the other time-heterogeneous signal components to this modelling technique?
   9% of our signals contain non delta brush time-heterogeneous activity. Can we adapt our model to describe this? Is this seizure activity, or time heterogeneity describable by a time-heterogeneous MatérnARMA process.

6. Can we apply this modelling procedure to other areas of electroencephalographic research?
   Our defined model is flexible with regards to time-heterogeneity, as such it would be of interest to apply it to other clinical problems.

6.3 Summary

We shall finish this research the way in which we started, by outlining the questions of interest alongside the answers discovered over its course.

1. Can we describe six seconds of background EEG using a time-homogeneous model?
   Yes with a but; if background EEG contains spontaneous delta brush activity, as is a possibility given the ages of the infants analysed, we require additional components. However, we have found that we can describe up to six seconds with a time-homogeneous model, where appropriate.

2. Can we adapt such a model to incorporate time-heterogeneous signal components?
   Yes. When we include time-heterogeneous signal components into our model we end up being able to describe 76% of the analysable signals and 69% of all the signals. It is possible that we would be able to describe more of the presented signals if we could identify the time heterogeneity present

3. Are the properties of neonatal EEG dependent upon the age of the infant?
   Maybe. Unlike previous research into neonatal somatosensory stimuli development, we focus upon the pairing age at birth and age at test and we do not have enough pairings to draw solid conclusions about the estimated parameters. It is common place to look solely at age at test or the difference between age at test and age at birth - known as post natal age. This results in a loss of resolution with respect to development as it is implying that any infant tested at a given age will have a similar response regardless of how long the gestational period was. Neonatal infants undergo procedures and are exposed to many stimuli, as such the thought that they would exhibit the same responses as an infant tested at the same age, but born a day previously, is counter intuitive. We have shown that certain signal characteristics and parameter characteristics could be dependent upon the age pairings. We have also shown that the probability of a delta brush response to noxious and tactile stimuli could require the outlined age pairing.

4. Given the time-heterogeneous nature of EEG signals, can we identify when the baseline activity changes?
   We don't know. But we have shown that in six seconds, even with time-varying components interspersed, the baseline activity remains the same in those signals adequately described by our model.

5. Do previous modelling techniques concur with the presented model?
   Yes and No. Previous models do not look at as high a resolution as we have, our signals show strong evidence that at this resolution both long and short term covariance structures are required, which is only observable at high resolution. We agree that fractional Brownian motion is a good description of the long range covariance structure, however Matérn should be implemented for greater flexibility especially since fractional Brownian motion is a limiting case. Furthermore, the

time-varying parameters should only be applied in combination with delta brush activity and a knowledge of when the parameters change.

Appendix A

Parameter characteristic generalised linear models

In the following sections we present the generalised linear models with a logit link function that describe the probability of a parameter characteristic occurring. The fitting procedure and general form was outlined in Chapter 4, this Chapter gives specific parameter values.

The values of this GLM, can then be used to obtain a binary variable indicating: whether the signal is Matérn, the variance is bimodal, or the overall spectral density shape. These estimated parameter characteristics are then utilised in constructing parameter estimates by segregating the data according to these values. Which is done in Appendix B.

A.1 Bimodal GLM

$$\log\left(\frac{\hat{p}_B}{1-\hat{p}_B}\right) = \begin{bmatrix} -0.64073 \\ -0.01773 \\ -0.17819 \\ 0.14273 \\ -0.25223 \\ -0.18543 \\ -0.29395 \\ -0.75767 \\ -0.95710 \\ -0.28700 \\ -0.34810 \\ -0.10224 \\ -0.66445 \\ -0.80608 \\ -0.69558 \\ -1.10127 \\ -0.98185 \\ -1.09377 \\ -0.95988 \\ -0.80777 \\ -0.89930 \\ 0.01414 \\ 0.00372 \\ -0.13070 \\ -0.03191 \\ -0.03383 \\ -0.04185 \\ -0.01394 \end{bmatrix}^T \begin{bmatrix} \mathrm{c}_0 \\ \mathrm{MCA}_1^{\mathrm{C}} \\ \mathrm{MCA}_2^{\mathrm{C}} \\ \mathrm{R}_{\mathrm{DB}} \\ \mathrm{R}_{\mathrm{PC}} \\ \mathrm{F7} \\ \mathrm{T3} \\ \mathrm{T5} \\ \mathrm{O1} \\ \mathrm{Fp2} \\ \mathrm{F8} \\ \mathrm{T4} \\ \mathrm{T6} \\ \mathrm{O2} \\ \mathrm{C3} \\ \mathrm{Cz} \\ \mathrm{C4} \\ \mathrm{CPz} \\ \mathrm{Fz} \\ \mathrm{CP3} \\ \mathrm{CP4} \\ \mathrm{MCA}_1^{\mathrm{C}} \times \mathrm{MCA}_2^{\mathrm{C}} \\ \mathrm{MCA}_1^{\mathrm{C}} \times \mathrm{R}_{\mathrm{DB}} \\ \mathrm{MCA}_1^{\mathrm{C}} \times \mathrm{R}_{\mathrm{PC}} \\ \mathrm{MCA}_1^{\mathrm{C}} \times \mathrm{F7} \\ \mathrm{MCA}_1^{\mathrm{C}} \times \mathrm{T3} \\ \mathrm{MCA}_1^{\mathrm{C}} \times \mathrm{T5} \\ \mathrm{MCA}_1^{\mathrm{C}} \times \mathrm{O1} \end{bmatrix} + \begin{bmatrix} -0.02322 \\ 0.02745 \\ -0.06518 \\ -0.06544 \\ -0.09636 \\ -0.03619 \\ -0.09461 \\ -0.10796 \\ -0.04950 \\ -0.07190 \\ -0.04786 \\ -0.06403 \\ 0.00136 \\ 0.23387 \\ 0.10300 \\ 0.00793 \\ -0.04738 \\ 0.01352 \\ 0.12944 \\ 0.07034 \\ -0.06987 \\ 0.09286 \\ 0.07576 \\ 0.15364 \\ 0.08926 \\ -0.08033 \\ 0.07507 \\ 0.21136 \end{bmatrix}^T \begin{bmatrix} \mathrm{MCA}_1^{\mathrm{C}} \times \mathrm{Fp2} \\ \mathrm{MCA}_1^{\mathrm{C}} \times \mathrm{F8} \\ \mathrm{MCA}_1^{\mathrm{C}} \times \mathrm{T4} \\ \mathrm{MCA}_1^{\mathrm{C}} \times \mathrm{T6} \\ \mathrm{MCA}_1^{\mathrm{C}} \times \mathrm{O2} \\ \mathrm{MCA}_1^{\mathrm{C}} \times \mathrm{C3} \\ \mathrm{MCA}_1^{\mathrm{C}} \times \mathrm{Cz} \\ \mathrm{MCA}_1^{\mathrm{C}} \times \mathrm{C4} \\ \mathrm{MCA}_1^{\mathrm{C}} \times \mathrm{CPz} \\ \mathrm{MCA}_1^{\mathrm{C}} \times \mathrm{Fz} \\ \mathrm{MCA}_1^{\mathrm{C}} \times \mathrm{CP3} \\ \mathrm{MCA}_1^{\mathrm{C}} \times \mathrm{CP4} \\ \mathrm{MCA}_2^{\mathrm{C}} \times \mathrm{F7} \\ \mathrm{MCA}_2^{\mathrm{C}} \times \mathrm{T3} \\ \mathrm{MCA}_2^{\mathrm{C}} \times \mathrm{T5} \\ \mathrm{MCA}_2^{\mathrm{C}} \times \mathrm{O1} \\ \mathrm{MCA}_2^{\mathrm{C}} \times \mathrm{Fp2} \\ \mathrm{MCA}_2^{\mathrm{C}} \times \mathrm{F8} \\ \mathrm{MCA}_2^{\mathrm{C}} \times \mathrm{T4} \\ \mathrm{MCA}_2^{\mathrm{C}} \times \mathrm{T6} \\ \mathrm{MCA}_2^{\mathrm{C}} \times \mathrm{O2} \\ \mathrm{MCA}_2^{\mathrm{C}} \times \mathrm{C3} \\ \mathrm{MCA}_2^{\mathrm{C}} \times \mathrm{Cz} \\ \mathrm{MCA}_2^{\mathrm{C}} \times \mathrm{C4} \\ \mathrm{MCA}_2^{\mathrm{C}} \times \mathrm{CPz} \\ \mathrm{MCA}_2^{\mathrm{C}} \times \mathrm{Fz} \\ \mathrm{MCA}_2^{\mathrm{C}} \times \mathrm{CP3} \\ \mathrm{MCA}_2^{\mathrm{C}} \times \mathrm{CP4} \end{bmatrix}$$

A.2 Matérn GLM

$$\log\left(\frac{\hat{p}_M}{1-\hat{p}_M}\right) = \begin{bmatrix} -5.37140 \\ 0.08900 \\ -0.02249 \\ 0.47643 \\ 1.84684 \\ 6.30424 \\ -1.39201 \\ -0.03813 \\ -0.73410 \\ -15.21064 \\ 0.67202 \\ 1.81969 \\ 0.18869 \\ -0.87831 \\ -0.88771 \\ -15.01520 \\ -1.67519 \\ -14.99819 \\ -1.53963 \\ 0.06778 \\ -1.58737 \\ -0.27992 \end{bmatrix}^T \begin{bmatrix} \mathrm{c}_0 \\ \mathrm{MCA}_1^{\mathrm{C}} \\ \mathrm{MCA}_2^{\mathrm{C}} \\ \mathrm{R_{DB}} \\ \mathrm{R_{PC}} \\ \mathrm{B}_1 \\ \mathrm{F7} \\ \mathrm{T3} \\ \mathrm{T5} \\ \mathrm{O1} \\ \mathrm{Fp2} \\ \mathrm{F8} \\ \mathrm{T4} \\ \mathrm{T6} \\ \mathrm{O2} \\ \mathrm{C3} \\ \mathrm{Cz} \\ \mathrm{C4} \\ \mathrm{CPz} \\ \mathrm{Fz} \\ \mathrm{CP3} \\ \mathrm{CP4} \end{bmatrix} + \begin{bmatrix} -0.02745 \\ 0.11695 \\ -0.03409 \\ 1.42813 \\ 0.16918 \\ 0.58810 \\ 14.81049 \\ -0.90365 \\ -1.73012 \\ 0.41036 \\ 0.57422 \\ 0.44560 \\ 14.60411 \\ 1.49700 \\ 14.34827 \\ 1.34516 \\ -0.16781 \\ 1.08496 \\ 0.10472 \\ -0.17729 \\ -2.46030 \end{bmatrix}^T \begin{bmatrix} \mathrm{MCA}_1^{\mathrm{C}} \times \mathrm{MCA}_2^{\mathrm{C}} \\ \mathrm{MCA}_1^{\mathrm{C}} \times \mathrm{R_{DB}} \\ \mathrm{MCA}_1^{\mathrm{C}} \times \mathrm{R_{PC}} \\ \mathrm{B}_1 \times \mathrm{F7} \\ \mathrm{B}_1 \times \mathrm{T3} \\ \mathrm{B}_1 \times \mathrm{T5} \\ \mathrm{B}_1 \times \mathrm{O1} \\ \mathrm{B}_1 \times \mathrm{Fp2} \\ \mathrm{B}_1 \times \mathrm{F8} \\ \mathrm{B}_1 \times \mathrm{T4} \\ \mathrm{B}_1 \times \mathrm{T6} \\ \mathrm{B}_1 \times \mathrm{O2} \\ \mathrm{B}_1 \times \mathrm{C3} \\ \mathrm{B}_1 \times \mathrm{Cz} \\ \mathrm{B}_1 \times \mathrm{C4} \\ \mathrm{B}_1 \times \mathrm{CPz} \\ \mathrm{B}_1 \times \mathrm{Fz} \\ \mathrm{B}_1 \times \mathrm{CP3} \\ \mathrm{B}_1 \times \mathrm{CP4} \\ \mathrm{R_{DB}} \times \mathrm{B}_1 \\ \mathrm{R_{PC}} \times \mathrm{B}_1 \end{bmatrix}$$

A.3 Shape GLM

$$\log\left(\frac{\hat{p}_S}{1-\hat{p}_S}\right) = \begin{bmatrix} 26.56483 \\ 3.29941 \\ -29.27632 \\ -402.69145 \\ 682.20751 \\ -14.49167 \\ 11.88175 \\ -3.21522 \\ -82.09257 \\ 407.20283 \\ 79.63035 \\ -1.87545 \\ -679.54564 \\ 16.44917 \\ -301.80697 \\ 1348.55447 \\ -11.81744 \\ -347.01516 \\ 334.56810 \\ -8.94544 \\ 82.38020 \\ -78.76856 \\ 2.40196 \\ 304.00171 \\ -1346.55626 \\ 347.52007 \\ -333.60461 \\ 6.97943 \\ 1.96662 \\ 2.17436 \end{bmatrix}^T \begin{bmatrix} c_0 \\ MCA_1^C \\ B_1 \\ M_1 \\ R_{DB} \\ R_{PC} \\ MCA_2^C \\ MCA_1^C \times B_1 \\ MCA_1^C \times M_1 \\ B_1 \times M_1 \\ MCA_1^C \times R_{DB} \\ MCA_1^C \times R_{PC} \\ B_1 \times R_{DB} \\ B_1 \times R_{PC} \\ M_1 \times R_{DB} \\ M_1 \times R_{PC} \\ B_1 \times MCA_2^C \\ M_1 \times MCA_2^C \\ R_{DB} \times MCA_2^C \\ R_{PC} \times MCA_2^C \\ MCA_1^C \times B_1 \times M_1 \\ MCA_1^C \times B_1 \times R_{DB} \\ MCA_1^C \times B_1 \times R_{PC} \\ M_1 \times B_1 \times R_{DB} \\ M_1 \times B_1 \times R_{PC} \\ MCA_2^C \times B_1 \times M_1 \\ MCA_2^C \times B_1 \times R_{DB} \\ MCA_2^C \times B_1 \times R_{PC} \\ MCA_2^C \times M_1 \times R_{DB} \\ MCA_2^C \times M_1 \times R_{PC} \end{bmatrix}$$

Appendix B

Fitted distributions for simulating neonatal EEG parameters

The estimated parameters of the covariance structure presented for neonatal electroencephalogram signals requires the use of a multivariate distribution. Utilising the asymptotic distribution of estimates obtained from minimising the Whittle likelihood, we fit several multivariate normal distributions based upon the characteristics of the estimated parameters.

Some of these distributions require further estimation in order to obtain estimates suitable for simulation, as outlined in the main body of the thesis. We present the mean vectors and covariance matrices for the distributions utilised, as well as the number of observations from which this distribution was obtained.

When Matérn = 0, the distance parameter is equal to 0; this is not allowable under the definition of the Matérn covariance, Definition 3.2.4. As such we must apply a fractional Brownian motion covariance structure when simulating from these distributions. The third element in the mean vectors as well as the third row and column in the covariance matrices have been left in these distributions, so as to illustrate the "dummy" parameter in these cases.

B.1 Matérn = 0, Bimodal = 0, Shape = 0

$$n_{000} = 44$$

$$\mu_{000} = \begin{bmatrix} 0.00146187 \\ 1.01979079 \\ 0.00000000 \\ 2.23376613 \\ 0.50592229 \\ 0.48573566 \\ 2.75261927 \\ 2.60560031 \\ 0.82791007 \end{bmatrix}$$

$$\Sigma_{000} = \begin{bmatrix} 0.00000066 & -0.00004147 & 0.00000000 & 0.00109007 & 0.00005794 & 0.00007445 & 0.00006559 & 0.00010984 & 0.00005087 \\ -0.00004147 & 0.04734783 & 0.00000000 & 0.02465484 & 0.02259090 & 0.02247657 & 0.00692640 & 0.00186632 & -0.00244829 \\ 0.00000000 & 0.00000000 & 0.00000000 & 0.00000000 & 0.00000000 & 0.00000000 & 0.00000000 & 0.00000000 & 0.00000000 \\ 0.00109007 & 0.02465484 & 0.00000000 & 2.44209901 & 0.29715384 & 0.28564585 & 0.20135261 & 0.32421688 & 0.14147097 \\ 0.00005794 & 0.02259090 & 0.00000000 & 0.29715384 & 0.12274140 & 0.06722111 & 0.01970833 & 0.03480096 & 0.01553578 \\ 0.00007445 & 0.02247657 & 0.00000000 & 0.28564585 & 0.06722111 & 0.06527069 & 0.01938924 & 0.03373344 & 0.01408507 \\ 0.00006559 & 0.00692640 & 0.00000000 & 0.20135261 & 0.01970833 & 0.01938924 & 0.06059857 & 0.07847362 & 0.03087291 \\ 0.00010984 & 0.00186632 & 0.00000000 & 0.32421688 & 0.03480096 & 0.03373344 & 0.07847362 & 0.12051623 & 0.05189345 \\ 0.00005087 & -0.00244829 & 0.00000000 & 0.14147097 & 0.01553578 & 0.01408507 & 0.03087291 & 0.05189345 & 0.02341861 \end{bmatrix}$$

B.2 Matérn = 0, Bimodal = 0, Shape = 1

$$n_{001} = 4875$$

$$\mu_{001} = \begin{bmatrix} 0.00136034 \\ 0.96187850 \\ 0.00000000 \\ 0.66987793 \\ 0.47129822 \\ -0.06935469 \\ -0.03651584 \\ -0.00620745 \\ -0.10028326 \end{bmatrix}$$

$$\Sigma_{001} = \begin{bmatrix} 0.00000016 & -0.00002602 & 0.00000000 & 0.00006846 & 0.00001552 & 0.00000380 & 0.00005067 & 0.00002343 & 0.00000618 \\ -0.00002602 & 0.01622797 & 0.00000000 & -0.00116996 & 0.00231996 & 0.00207361 & 0.00087401 & 0.00204067 & -0.00001301 \\ 0.00000000 & 0.00000000 & 0.00000000 & 0.00000000 & 0.00000000 & 0.00000000 & 0.00000000 & 0.00000000 & 0.00000000 \\ 0.00006846 & -0.00116996 & 0.00000000 & 0.19538312 & -0.05180342 & -0.08022406 & -0.11293734 & -0.02186548 & -0.00347929 \\ 0.00001552 & 0.00231996 & 0.00000000 & -0.05180342 & 0.11070166 & 0.01524362 & 0.08452775 & -0.00179045 & 0.01653803 \\ 0.00000380 & 0.00207361 & 0.00000000 & -0.08022406 & 0.01524362 & 0.10371497 & 0.08250548 & 0.04985155 & -0.00804048 \\ 0.00005067 & 0.00087401 & 0.00000000 & -0.11293734 & 0.08452775 & 0.08250548 & 0.17445869 & 0.05181953 & 0.01401043 \\ 0.00002343 & 0.00204067 & 0.00000000 & -0.02186548 & -0.00179045 & 0.04985155 & 0.05181953 & 0.04186361 & 0.00069598 \\ 0.00000618 & -0.00001301 & 0.00000000 & -0.00347929 & 0.01653803 & -0.00804048 & 0.01401043 & 0.00069598 & 0.00925240 \end{bmatrix}$$

B.3 Matérn = 0, Bimodal = 1, Shape = 0

$$n_{010} = 492$$

$$\mu_{010} = \begin{bmatrix} 0.00037416 \\ 1.05952515 \\ 0.00000000 \\ 0.95148583 \\ -0.01925940 \\ 0.47296451 \\ 2.35440485 \\ 1.81220603 \\ 0.42184077 \end{bmatrix}$$

$$\Sigma_{010} = \begin{bmatrix} 0.00000001 & -0.00000171 & 0.00000000 & 0.00000250 & 0.00001138 & 0.00000191 & -0.00000109 & 0.00000279 & 0.00000395 \\ -0.00000171 & 0.02845353 & 0.00000000 & 0.03746052 & -0.04041591 & -0.00278409 & -0.03694934 & -0.07366306 & -0.03896539 \\ 0.00000000 & 0.00000000 & 0.00000000 & 0.00000000 & 0.00000000 & 0.00000000 & 0.00000000 & 0.00000000 & 0.00000000 \\ 0.00000250 & 0.03746052 & 0.00000000 & 0.29142666 & -0.16010138 & 0.07195672 & -0.17174175 & -0.34419807 & -0.18365259 \\ 0.00001138 & -0.04041591 & 0.00000000 & -0.16010138 & 0.40985567 & -0.02776771 & 0.26876161 & 0.53422444 & 0.28364720 \\ 0.00000191 & -0.00278409 & 0.00000000 & 0.07195672 & -0.02776771 & 0.05554244 & -0.00942637 & -0.01859297 & -0.01045195 \\ -0.00000109 & -0.03694934 & 0.00000000 & -0.17174175 & 0.26876161 & -0.00942637 & 0.36005563 & 0.67993530 & 0.34299369 \\ 0.00000279 & -0.07366306 & 0.00000000 & -0.34419807 & 0.53422444 & -0.01859297 & 0.67993530 & 1.29951417 & 0.66365386 \\ 0.00000395 & -0.03896539 & 0.00000000 & -0.18365259 & 0.28364720 & -0.01045195 & 0.34299369 & 0.66365386 & 0.34362562 \end{bmatrix}$$

B.4 Matérn = 0, Bimodal = 1, Shape = 1

$$n_{011} = 4$$

$$\mu_{011} = \begin{bmatrix} 0.00059572 \\ 1.28340652 \\ 0.00000000 \\ 0.53020089 \\ -0.08409384 \\ 0.12329130 \\ 0.50117982 \\ -0.19085367 \\ -0.28442364 \end{bmatrix}$$

$$\Sigma_{011} = \begin{bmatrix} 0.00000000 & -0.00000937 & 0.00000000 & -0.00003023 & 0.00000854 & -0.00000729 & -0.00000402 & 0.00001737 & 0.00001693 \\ -0.00000937 & 0.04621533 & 0.00000000 & 0.05625430 & -0.00523271 & 0.03560545 & -0.01735795 & -0.10320673 & -0.07977830 \\ 0.00000000 & 0.00000000 & 0.00000000 & 0.00000000 & 0.00000000 & 0.00000000 & 0.00000000 & 0.00000000 & 0.00000000 \\ -0.00003023 & 0.05625430 & 0.00000000 & 0.58439225 & -0.39098863 & 0.15023849 & 0.37661196 & -0.23137239 & -0.30597604 \\ 0.00000854 & -0.00523271 & 0.00000000 & -0.39098863 & 0.32005963 & -0.10289051 & -0.32941572 & 0.12759992 & 0.19904645 \\ -0.00000729 & 0.03560545 & 0.00000000 & 0.15023849 & -0.10289051 & 0.06080297 & 0.08945318 & -0.12313527 & -0.12482556 \\ -0.00000402 & -0.01735795 & 0.00000000 & 0.37661196 & -0.32941572 & 0.08945318 & 0.35028700 & -0.08225575 & -0.16759610 \\ 0.00001737 & -0.10320673 & 0.00000000 & -0.23137239 & 0.12759992 & -0.12313527 & -0.08225575 & 0.29415781 & 0.25982761 \\ 0.00001693 & -0.07977830 & 0.00000000 & -0.30597604 & 0.19904645 & -0.12482556 & -0.16759610 & 0.25982761 & 0.25822179 \end{bmatrix}$$

B.5 Matérn = 1, Bimodal = 0, Shape = 0

$$n_{100} = 6$$

$$\mu_{100} = \begin{bmatrix} 0.00200164 \\ 1.83608506 \\ 0.09488228 \\ 26.05757682 \\ -11.18133348 \\ 13.75503342 \\ 2.14744434 \\ 1.28907169 \\ 0.07338345 \end{bmatrix}$$

$$\Sigma_{100} = \begin{bmatrix} 0.00000364 & 0.00022602 & -0.00006009 & 0.06754211 & -0.04900889 & 0.06137036 & -0.00044815 & -0.00106538 & -0.00067309 \\ 0.00022602 & 0.19427550 & 0.01499138 & 7.48640477 & -2.53061186 & 2.58673820 & -0.12598682 & -0.27079498 & -0.15872394 \\ -0.00006009 & 0.01499138 & 0.01603152 & -2.27178670 & 1.27953966 & -1.71994833 & 0.07354723 & 0.15766521 & 0.08948726 \\ 0.06754211 & 7.48640477 & -2.27178670 & 1633.7483 & -1045.6964 & 1296.9893 & -19.855695 & -44.699997 & -26.912500 \\ -0.04900889 & -2.53061186 & 1.27953966 & -1045.6964 & 744.83702 & -930.20676 & 8.75786461 & 20.615702 & 12.877744 \\ 0.06137036 & 2.58673820 & -1.71994833 & 1296.9893 & -930.20676 & 1164.3372 & -10.934300 & -25.713266 & -16.036131 \\ -0.00044815 & -0.12598682 & 0.07354723 & -19.855695 & 8.75786461 & -10.934300 & 0.56207085 & 1.20299197 & 0.68739698 \\ -0.00106538 & -0.27079498 & 0.15766521 & -44.699997 & 20.6157021 & -25.713266 & 1.20299197 & 2.58288042 & 1.48027894 \\ -0.00067309 & -0.15872394 & 0.08948726 & -26.912500 & 12.8777443 & -16.036131 & 0.68739698 & 1.48027894 & 0.85092725 \end{bmatrix}$$

B.6 Matérn = 1, Bimodal = 0, Shape = 1

$$n_{101} = 77$$

$$\mu_{101} = \begin{bmatrix} 0.00100779 \\ 1.85249284 \\ 0.11677929 \\ 1.06942855 \\ 0.70058849 \\ -0.14343897 \\ -0.08386558 \\ 0.10143097 \\ -0.03415402 \end{bmatrix}$$

$$\Sigma_{101} = \begin{bmatrix} 0.00000010 & 0.00006485 & 0.00001975 & 0.00025510 & 0.00019861 & 0.00004987 & 0.00002171 & 0.00000806 & -0.00000013 \\ 0.00006485 & 0.14099240 & 0.03419101 & 0.15916244 & 0.06846203 & -0.03536523 & 0.00775089 & -0.00619879 & 0.00208343 \\ 0.00001975 & 0.03419101 & 0.01374513 & 0.00303384 & -0.02418444 & -0.02601498 & 0.00321004 & -0.00420284 & 0.00029163 \\ 0.00025510 & 0.15916244 & 0.00303384 & 2.41235719 & 2.28999921 & 0.73085741 & -0.15864167 & 0.03111894 & -0.02041649 \\ 0.00019861 & 0.06846203 & -0.02418444 & 2.28999921 & 2.53188477 & 0.95652217 & -0.08206213 & 0.01261022 & -0.02160210 \\ 0.00004987 & -0.03536523 & -0.02601498 & 0.73085741 & 0.95652217 & 0.56403344 & -0.02294722 & 0.03908749 & -0.02670370 \\ 0.00002171 & 0.00775089 & 0.00321004 & -0.15864167 & -0.08206213 & -0.02294722 & 0.13628399 & 0.01075995 & 0.01082401 \\ 0.00000806 & -0.00619879 & -0.00420284 & 0.03111894 & 0.01261022 & 0.03908749 & 0.01075995 & 0.03385441 & -0.00156063 \\ -0.00000013 & 0.00208343 & 0.00029163 & -0.02041649 & -0.02160210 & -0.02670370 & 0.01082401 & -0.00156063 & 0.00832367 \end{bmatrix}$$

B.7 Matérn = 1, Bimodal = 1, Shape = 0

$$n_{110} = 277$$

$$\mu_{110} = \begin{bmatrix} 0.00018190 \\ 2.17232073 \\ 0.14272498 \\ 0.91945069 \\ -0.00219800 \\ 0.37738110 \\ 1.90334824 \\ 0.94236044 \\ -0.01277066 \end{bmatrix}$$

$$\Sigma_{110} = \begin{bmatrix} 0.00000001 & -0.00001083 & 0.00000459 & 0.00004736 & 0.00001529 & -0.00000013 & 0.00003003 & 0.00005980 & 0.00003093 \\ -0.00001083 & 0.09464633 & 0.01184901 & 0.02832193 & -0.00358862 & 0.00203113 & -0.01586032 & -0.03780928 & -0.02213350 \\ 0.00000459 & 0.01184901 & 0.01297103 & -0.04283038 & 0.01728337 & -0.01665367 & 0.05471075 & 0.10589614 & 0.05352562 \\ 0.00004736 & 0.02832193 & -0.04283038 & 2.37583795 & 0.15958544 & 0.49994324 & -0.32156212 & -0.64628565 & -0.34458467 \\ 0.00001529 & -0.00358862 & 0.01728337 & 0.15958544 & 0.19140517 & 0.07571504 & 0.14579068 & 0.28641135 & 0.14614626 \\ -0.00000013 & 0.00203113 & -0.01665367 & 0.49994324 & 0.07571504 & 0.18361327 & -0.08128027 & -0.16692200 & -0.08931324 \\ 0.00003003 & -0.01586032 & 0.05471075 & -0.32156212 & 0.14579068 & -0.08128027 & 0.48629408 & 0.92260223 & 0.45876614 \\ 0.00005980 & -0.03780928 & 0.10589614 & -0.64628565 & 0.28641135 & -0.16692200 & 0.92260223 & 1.79320171 & 0.90923491 \\ 0.00003093 & -0.02213350 & 0.05352562 & -0.34458467 & 0.14614626 & -0.08931324 & 0.45876614 & 0.90923491 & 0.46958287 \end{bmatrix}$$

B.8 Matérn = 1, Bimodal = 1, Shape = 1

$$n_{111} = 712$$

$$\mu_{111} = \begin{bmatrix} 0.00033612 \\ 2.29820459 \\ 0.10467164 \\ 0.74489614 \\ 0.57641231 \\ -0.26798045 \\ -0.50316722 \\ 0.07664778 \\ -0.01521496 \end{bmatrix}$$

$$\Sigma_{111} = \begin{bmatrix} 0.00000001 & -0.00001662 & -0.00000077 & 0.00000929 & 0.00000061 & 0.00000051 & 0.00001890 & -0.00000269 & -0.00000079 \\ -0.00001662 & 0.13486400 & 0.02014732 & 0.02433859 & -0.00343733 & -0.02859914 & -0.00419096 & 0.01215191 & 0.00285282 \\ -0.00000077 & 0.02014732 & 0.00447294 & 0.00313238 & -0.00008855 & -0.00506976 & 0.00313670 & 0.00190220 & 0.00134401 \\ 0.00000929 & 0.02433859 & 0.00313238 & 0.29060555 & -0.05760854 & -0.15276824 & -0.14825561 & 0.03941558 & 0.00014053 \\ 0.00000061 & -0.00343733 & -0.00008855 & -0.05760854 & 0.09365758 & 0.00730694 & 0.01475531 & -0.01750594 & 0.02547858 \\ 0.00000051 & -0.02859914 & -0.00506976 & -0.15276824 & 0.00730694 & 0.16479960 & 0.16190327 & -0.03795510 & -0.03180156 \\ 0.00001890 & -0.00419096 & 0.00313670 & -0.14825561 & 0.01475531 & 0.16190327 & 0.31542537 & -0.06424884 & -0.03693676 \\ -0.00000269 & 0.01215191 & 0.00190220 & 0.03941558 & -0.01750594 & -0.03795510 & -0.06424884 & 0.07801555 & 0.01320292 \\ -0.00000079 & 0.00285282 & 0.00134401 & 0.00014053 & 0.02547858 & -0.03180156 & -0.03693676 & 0.01320292 & 0.03396053 \end{bmatrix}$$

Appendix C

Simulated neonatal electroencephalogram signals

We present in this appendix a further eight realisations from our simulation model across a range of age at birth, age at test pairings.

As mentioned in the main body of the text there are some signals for which unrealistic parameterisations are generated. This is due in part to the number of observations from which the distribution is estimated, as well as the fact that the sample from which the distributions were estimated might not be muiltivariately normally distributed.

Figure C.1: Fully simulated signals PMA = 24 PMAT = 38

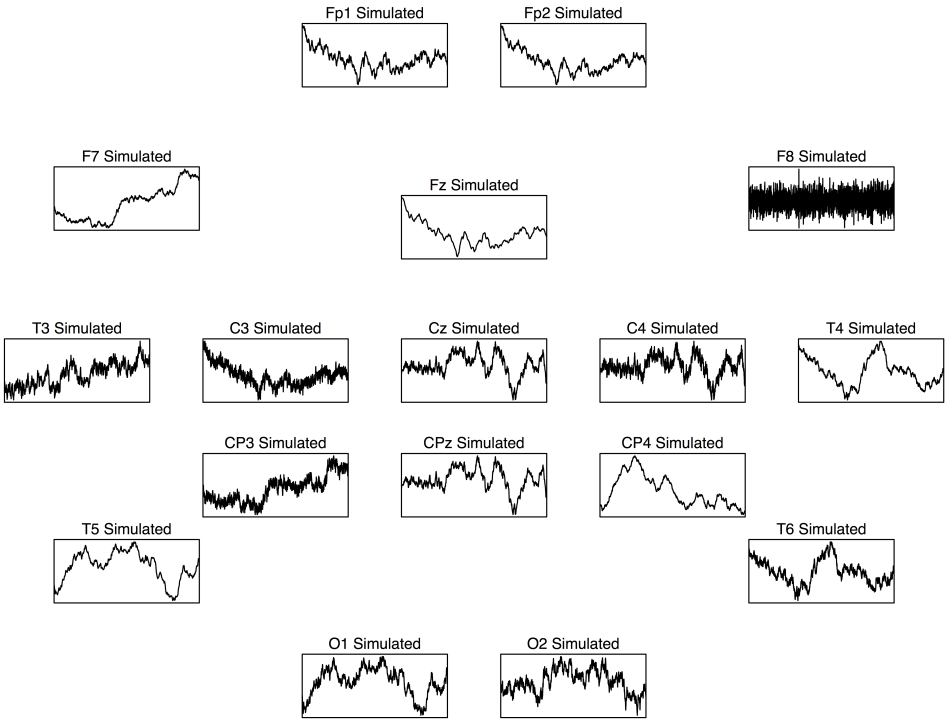


Figure C.2: Fully simulated signals PMA = 25.24 PMAT = 36.29

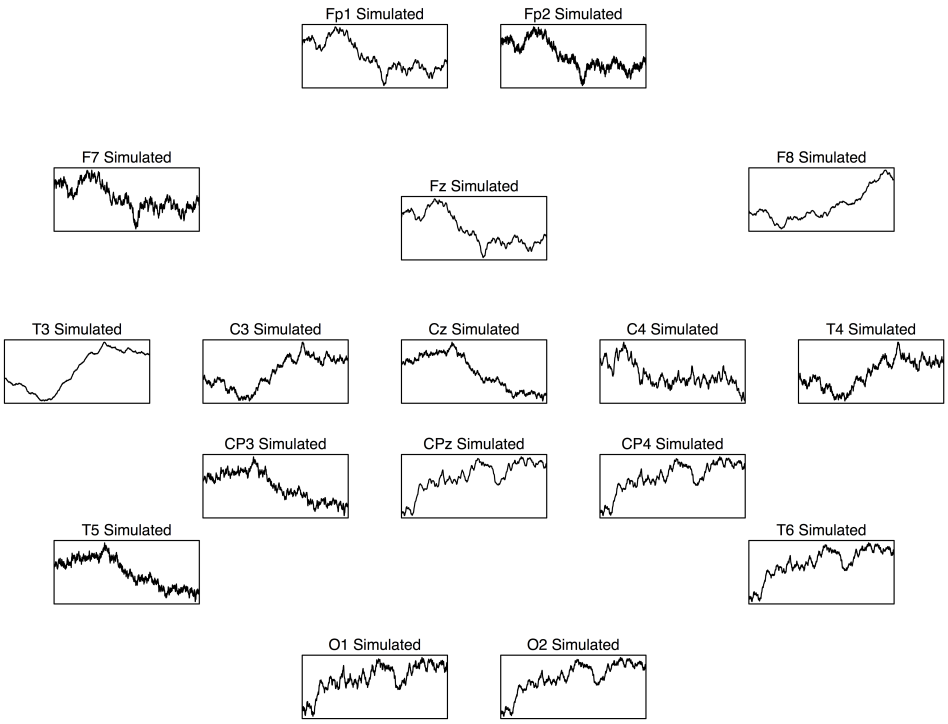


Figure C.3: Fully simulated signals PMA = 25.71, PMAT = 30.71

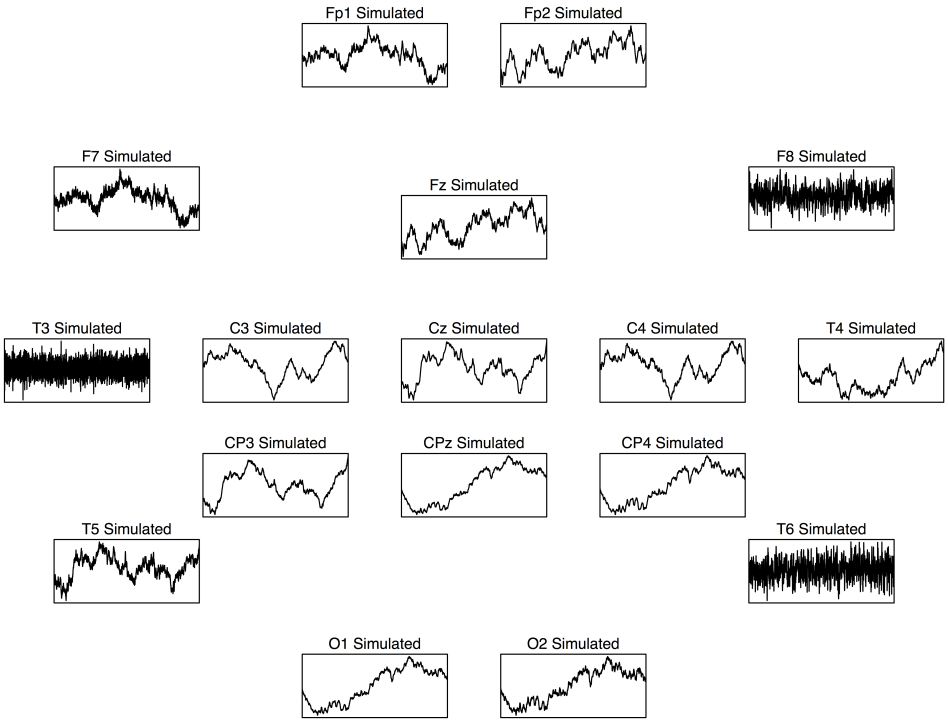


Figure C.4: Fully simulated signals PMA = 27, PMAT = 28.43

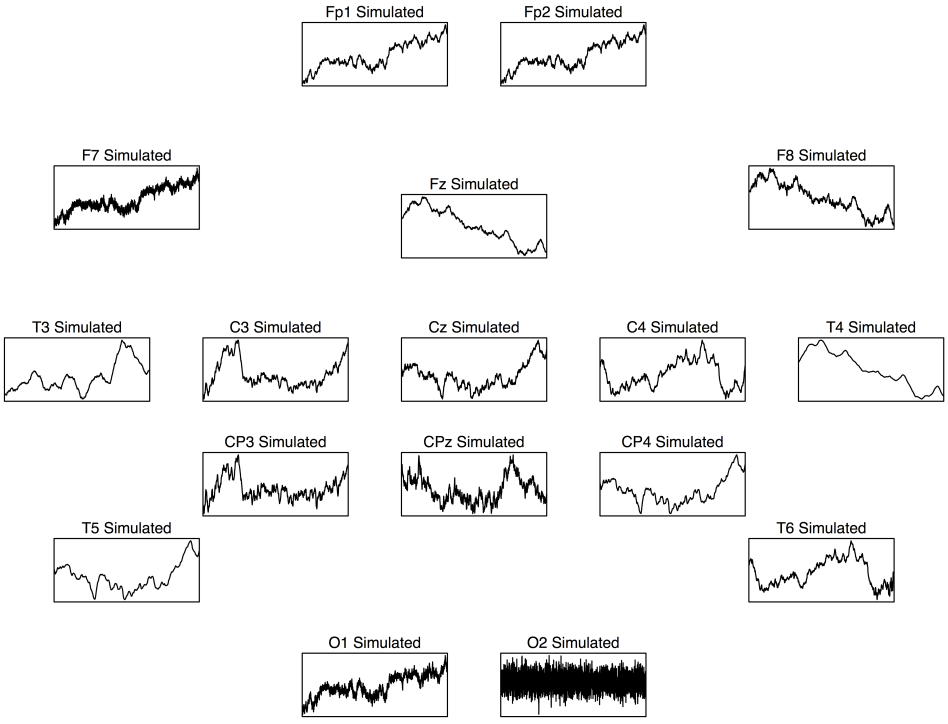


Figure C.5: Fully simulated signals PMA = 32.57, PMAT = 39.43

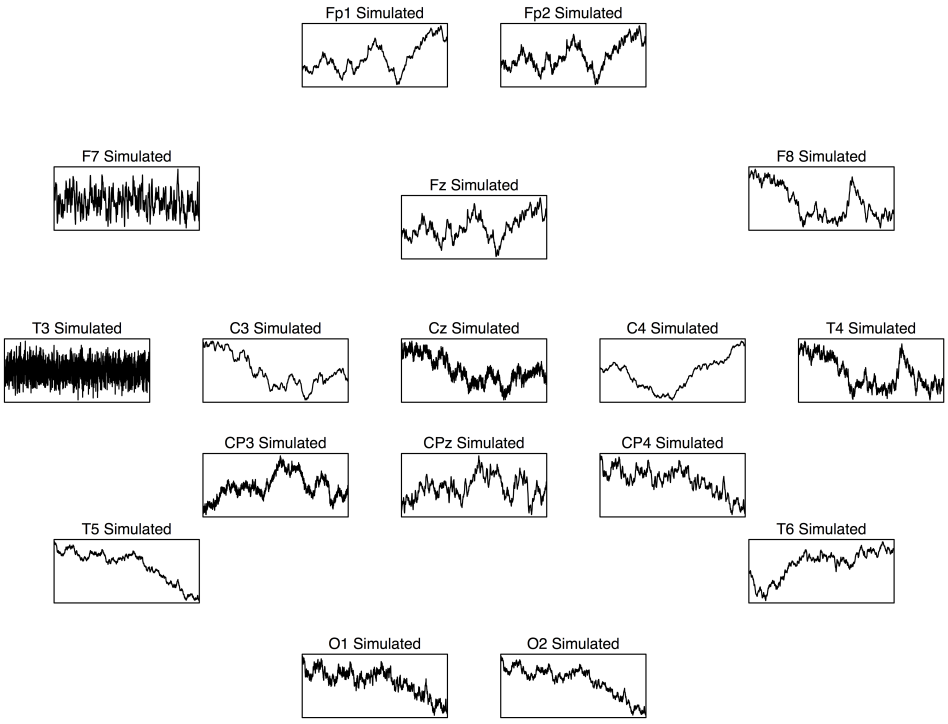


Figure C.6: Fully simulated signals PMA = 34.43, PMAT = 35

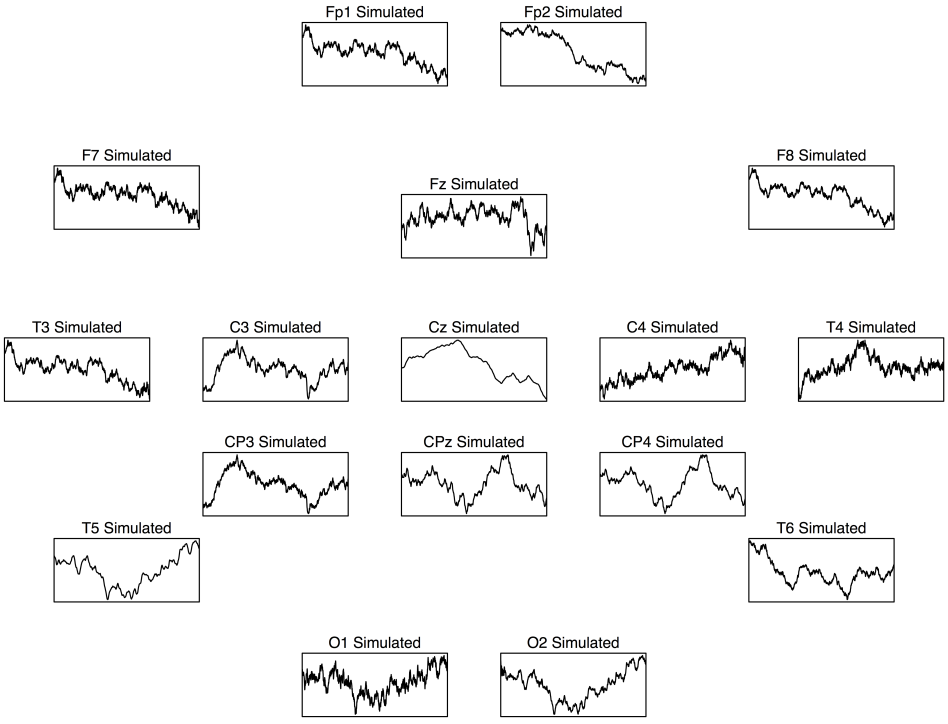


Figure C.7: Fully simulated signals PMA = 36.14, PMAT = 37

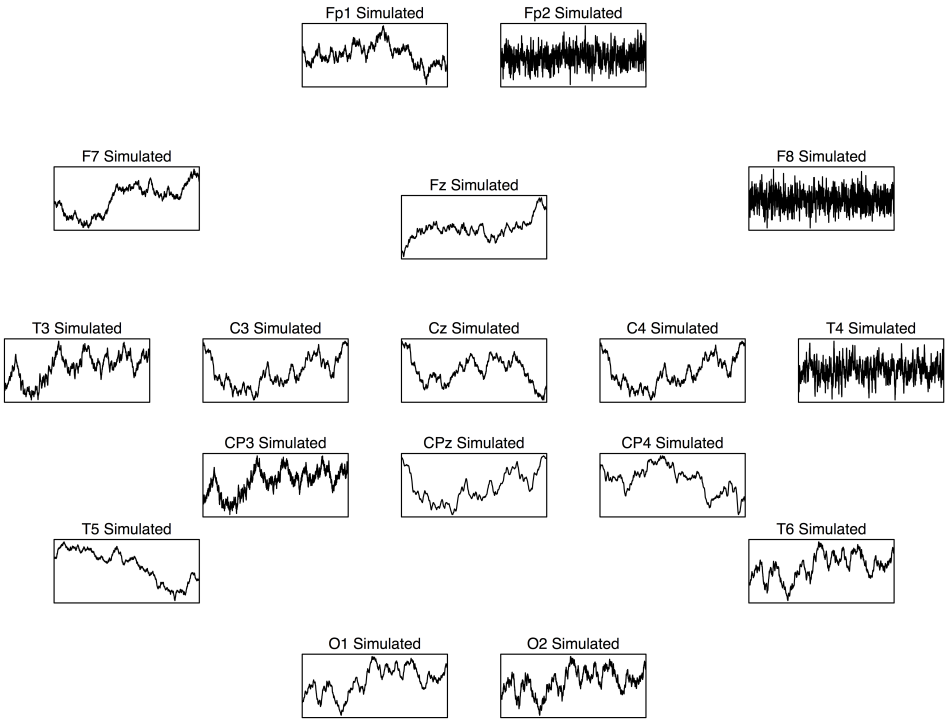
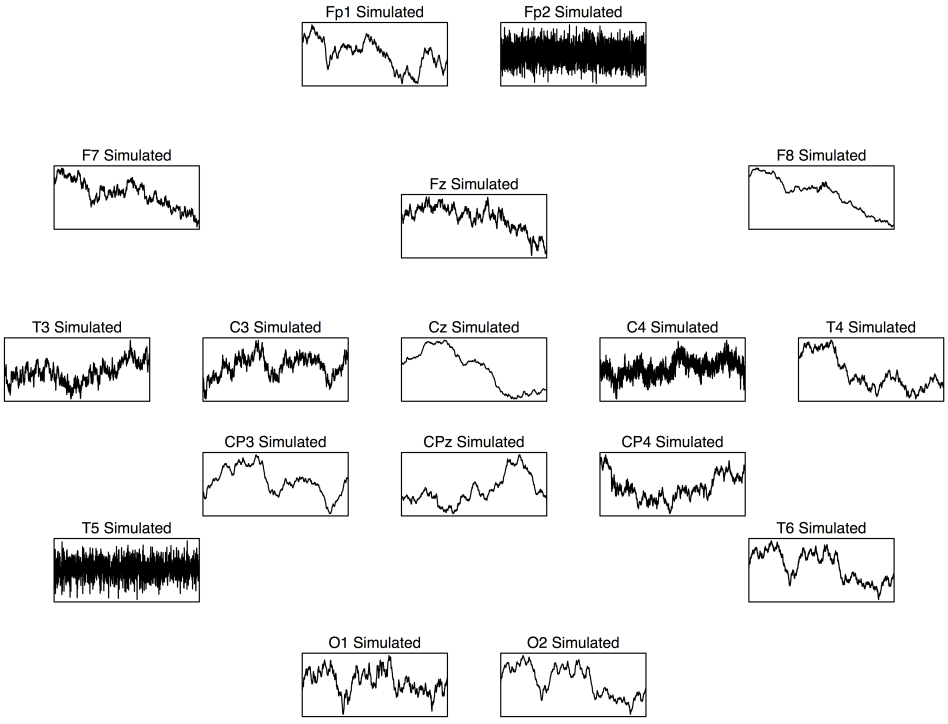


Figure C.8: Fully simulated signals PMA = 39, PMAT = 39



Bibliography

[1] S. Sanei and J. Chambers. *EEG signal processing*. John Wiley & Sons, 2007.

[2] P. L. Nunez and R. Srinivasan. *Electric fields of the brain: the neurophysics of EEG*. Oxford University Press, 2006.

[3] R. Cooper, C.D. Binnie, R. Billings, and R. Billings. *Techniques in clinical neurophysiology: a practical manual*. Elsevier Churchill Livingstone, 2005.

[4] L. Fabrizi, R. Slater, A. Worley, J. Meek, S. Boyd, S. Olhede, and M. Fitzgerald. A shift in sensory processing that enables the developing human brain to discriminate touch from pain. *Curr. Biol.*, 21(18):1552–1558, Sep 2011.

[5] R. Slater, L. Cornelissen, L. Fabrizi, D. Patten, J. Yoxen, A. Worley, S. Boyd, J. Meek, and M. Fitzgerald. Oral sucrose as an analgesic drug for procedural pain in newborn infants: a randomised controlled trial. *Lancet*, 376(9748):1225–1232, Oct 2010.

[6] A. Worley, L. Fabrizi, S. Boyd, and R. Slater. Multi-modal pain measurements in infants. *J. Neurosci. Methods*, 205(2):252–257, Apr 2012.

[7] World Health Organisation. Preterm birth, November 2014.

[8] Office for National Statistics. Gestation-specific infant mortality, 2012, October 2014.

[9] Z. A. Vesoulis and A. M. Mathur. Advances in management of neonatal seizures. *Indian J Pediatr*, 81(6):592–598, Jun 2014.

[10] P. Nevalainen, P. Rahkonen, E. Pihko, A. Lano, S. Vanhatalo, S. Andersson, T. Autti, L. Valanne, M. Metsaranta, and L. Lauronen. Evaluation of somatosen-

sory cortical processing in extremely preterm infants at term with MEG and EEG. *Clin Neurophysiol*, Jun 2014.

[11] K. J. Anand and F. M. Scalzo. Can adverse neonatal experiences alter brain development and subsequent behavior? *Biol. Neonate*, 77(2):69–82, Feb 2000.

[12] S. Chaudhari. Neonatal intensive care practices harmful to the developing brain. *Indian Pediatr*, 48(6):437–440, Jun 2011.

[13] T. Moore, E. M. Hennessy, J. Myles, S. J. Johnson, E. S. Draper, K. L. Costeloe, and N. Marlow. Neurological and developmental outcome in extremely preterm children born in England in 1995 and 2006: the EPICure studies. *BMJ*, 345:e7961, 2012.

[14] K. J. Anand. Clinical importance of pain and stress in preterm neonates. *Biol. Neonate*, 73(1):1–9, 1998.

[15] M. J. Benders, K. Palmu, C. Menache, C. Borradori-Tolsa, F. Lazeyras, S. Sizonenko, J. Dubois, S. Vanhatalo, and P. S. Huppi. Early Brain Activity Relates to Subsequent Brain Growth in Premature Infants. *Cereb. Cortex*, May 2014.

[16] R. W. Cooke. Are there critical periods for brain growth in children born preterm? *Arch. Dis. Child. Fetal Neonatal Ed.*, 91(1):17–20, Jan 2006.

[17] C. Nosarti, M. H. Al-Asady, S. Frangou, A. L. Stewart, L. Rifkin, and R. M. Murray. Adolescents who were born very preterm have decreased brain volumes. *Brain*, 125(Pt 7):1616–1623, Jul 2002.

[18] S. Saigal and L. W. Doyle. An overview of mortality and sequelae of preterm birth from infancy to adulthood. *Lancet*, 371(9608):261–269, Jan 2008.

[19] H. Als, F. H. Duffy, G. B. McAnulty, M. J. Rivkin, S. Vajapeyam, R. V. Mulkern, S. K. Warfield, P. S. Huppi, S. C. Butler, N. Conneman, C. Fischer, and E. C. Eichenwald. Early experience alters brain function and structure. *Pediatrics*, 113(4):846–857, Apr 2004.

[20] J. J. Volpe. Cerebellum of the premature infant: rapidly developing, vulnerable, clinically important. *J. Child Neurol.*, 24(9):1085–1104, Sep 2009.

[21] B. Patoine. The vulnerable premature brain, May 2010.

[22] A. G. Blankenship and M. B. Feller. Mechanisms underlying spontaneous patterned activity in developing neural circuits. *Nat. Rev. Neurosci.*, 11(1):18–29, Jan 2010.

[23] J. Stiles and T. L. Jernigan. The basics of brain development. *Neuropsychol Rev*, 20(4):327–348, Dec 2010.

[24] K. J. Anand, J. V. Aranda, C. B. Berde, S. Buckman, E. V. Capparelli, W. Carlo, P. Hummel, C. C. Johnston, J. Lantos, V. Tutag-Lehr, A. M. Lynn, L. G. Maxwell, T. F. Oberlander, T. N. Raju, S. G. Soriano, A. Taddio, and G. A. Walco. Summary proceedings from the neonatal pain-control group. *Pediatrics*, 117(3 Pt 2):S9–S22, Mar 2006.

[25] J. Gotman, D. Flanagan, J. Zhang, and B. Rosenblatt. Automatic seizure detection in the newborn: methods and initial evaluation. *Electroencephalogr Clin Neurophysiol*, 103(3):356–362, Sep 1997.

[26] A. Liu, J. S. Hahn, G. P. Heldt, and R. W. Coen. Detection of neonatal seizures through computerized EEG analysis. *Electroencephalogr Clin Neurophysiol*, 82(1):30–37, Jan 1992.

[27] A. Aarabi, F. Wallois, and R. Grebe. Automated neonatal seizure detection: a multistage classification system through feature selection based on relevance and redundancy analysis. *Clin Neurophysiol*, 117(2):328–340, Feb 2006.

[28] A. Aarabi, R. Grebe, and F. Wallois. A multistage knowledge-based system for EEG seizure detection in newborn infants. *Clin Neurophysiol*, 118(12):2781–2797, Dec 2007.

[29] R. Ahmed, A. Temko, W. Marnane, G. Boylan, and G. Lighbody. Dynamic time warping based neonatal seizure detection system. *Conf Proc IEEE Eng Med Biol Soc*, 2012:4919–4922, 2012.

[30] M. Alegre and E. Urrestarazu. Neonatal automated seizure detection: going ahead into clinical use. *Clin Neurophysiol*, 122(8):1480–1481, Aug 2011.

[31] G. B. Boylan, N. J. Stevenson, and S. Vanhatalo. Monitoring neonatal seizures. *Semin Fetal Neonatal Med*, 18(4):202–208, Aug 2013.

[32] P. J. Cherian, W. Deburchgraeve, R. M. Swarte, M. De Vos, P. Govaert, S. Van Huffel, and G. H. Visser. Validation of a new automated neonatal seizure detection system: a clinician's perspective. *Clin Neurophysiol*, 122(8):1490–1499, Aug 2011.

[33] M. S. Khlif, M. Mesbah, B. Boashash, and P. Colditz. Detection of neonatal seizure using multiple filters. In *Information Sciences Signal Processing and their Applications (ISSPA), 2010 10th International Conference on*, pages 284–287, 2010.

[34] H. Li and A. Jeremic. Neonatal seizure detection using blind distributed detection with correlated decisions. *Conf Proc IEEE Eng Med Biol Soc*, 2011:6580–6584, 2011.

[35] M. Mesbah, J. M. O' Toole, P. B. Colditz, and B. Boashash. Instantaneous frequency based newborn eeg seizure characterisation. *EURASIP Journal on Advances in Signal Processing*, 2012(1):1–11, 2012.

[36] F. Pisani, C. Spagnoli, E. Pavlidis, C. Facini, G. M. Kouamou Ntonfo, G. Ferrari, and R. Raheli. Real-time automated detection of clonic seizures in newborns. *Clin Neurophysiol*, 125(8):1533–1540, Aug 2014.

[37] K. K. Poh and P. Marziliano. Analysis of neonatal EEG signals using Stockwell transform. *Conf Proc IEEE Eng Med Biol Soc*, 2007:594–597, 2007.

[38] M. Quigg and D. Leiner. Limitations of single-channel EEG on the forehead for neonatal seizure detection. *J Perinatol*, 29(8):588, Aug 2009.

[39] M. Roessgen, A. M. Zoubir, and B. Boashash. Seizure detection of newborn EEG using a model-based approach. *IEEE Trans Biomed Eng*, 45(6):673–685, Jun 1998.

[40] L. S. Smit, R. J. Vermeulen, W. P. Fetter, R. L. Strijers, and C. J. Stam. Neonatal seizure monitoring using non-linear EEG analysis. *Neuropediatrics*, 35(6):329–335, Dec 2004.

[41] A. Temko, L. Marnane, G. Boylan, and G. Lightbody. Adaptive modelling of background eeg for robust detection of neonatal seizures. In *Biomedical Engineering and Sciences (IECBES), 2012 IEEE EMBS Conference on*, pages 46–51, 2012.

[42] A. Temko, N. Stevenson, W. Marnane, G. Boylan, and G. Lightbody. Inclusion of temporal priors for automated neonatal EEG classification. *J Neural Eng*, 9(4):046002, Aug 2012.

[43] A. Temko, G. Lightbody, G. Boylan, and W. Marnane. Online EEG channel weighting for detection of seizures in the neonate. *Conf Proc IEEE Eng Med Biol Soc*, 2011:1447–1450, 2011.

[44] A. Temko, G. Boylan, W. Marnane, and G. Lightbody. Robust neonatal EEG seizure detection through adaptive background modeling. *Int J Neural Syst*, 23(4):1350018, Aug 2013.

[45] M. J. van Putten. Neonatal seizure detection. *Clin Neurophysiol*, 119(11):2417–2418, Nov 2008.

[46] D. W. Chan, M. Yamazaki, T. Akiyama, B. Chu, E. J. Donner, and H. Otsubo. Rapid oscillatory activity in delta brushes of premature and term neonatal EEG. *Brain Dev.*, 32(6):482–486, Jun 2010.

[47] B. R. Greene, S. Faul, W. P. Marnane, G. Lightbody, I. Korotchikova, and G. B. Boylan. A comparison of quantitative EEG features for neonatal seizure detection. *Clin Neurophysiol*, 119(6):1248–1261, Jun 2008.

[48] M. A. Navakatikyan, P. B. Colditz, C. J. Burke, T. E. Inder, J. Richmond, and C. E. Williams. Seizure detection algorithm for neonates based on wave-sequence analysis. *Clin Neurophysiol*, 117(6):1190–1203, Jun 2006.

[49] B. Stevens, J. Yamada, G. Y. Lee, and A. Ohlsson. Sucrose for analgesia in newborn infants undergoing painful procedures. *Cochrane Database Syst Rev*, 1:CD001069, 2013.

[50] L. A. Hatfield, K. Chang, M. Bittle, J. Deluca, and R. C. Polomano. The analgesic properties of intraoral sucrose: an integrative review. *Adv Neonatal Care*, 11(2):83–92, Apr 2011.

[51] D. Harrison, L. Johnston, and P. Loughnan. Oral sucrose for procedural pain in sick hospitalized infants: a randomized-controlled trial. *J Paediatr Child Health*, 39(8):591–597, Nov 2003.

[52] C. C. Johnston and B. J. Stevens. Experience in a neonatal intensive care unit affects pain response. *Pediatrics*, 98(5):925–930, Nov 1996.

[53] M. A. Hofer. Early relationships as regulators of infant physiology and behavior. *Acta Paediatr Suppl*, 397:9–18, Jun 1994.

[54] L. A. Rosenblum and M. W. Andrews. Influences of environmental demand on maternal behavior and infant development. *Acta Paediatr Suppl*, 397:57–63, Jun 1994.

[55] M. Hack and D. W. Costello. Decrease in frequency of cerebral palsy in preterm infants. *Lancet*, 369(9555):7–8, Jan 2007.

[56] Z. D. Jiang, D. M. Brosi, and A. R. Wilkinson. Hearing impairment in preterm very low birthweight babies detected at term by brainstem auditory evoked responses. *Acta Paediatr.*, 90(12):1411–1415, Dec 2001.

[57] R. V. Grunau, M. F. Whitfield, J. H. Petrie, and E. L. Fryer. Early pain experience, child and family factors, as precursors of somatization: a prospective study of extremely premature and fullterm children. *Pain*, 56(3):353–359, Mar 1994.

[58] R. E. Grunau, M. F. Whitfield, and J. Petrie. Children's judgements about pain at age 8-10 years: do extremely low birthweight (¡ or = 1000 g) children differ from full birthweight peers? *J Child Psychol Psychiatry*, 39(4):587–594, May 1998.

[59] K. McCormack, P. Prather, and C. Chapleo. Some new insights into the effects of opioids in phasic and tonic nociceptive tests. *Pain*, 78(2):79–98, Nov 1998.

[60] R. V. Grunau, M. F. Whitfield, and J. H. Petrie. Pain sensitivity and temperament in extremely low-birth-weight premature toddlers and preterm and full-term controls. *Pain*, 58(3):341–346, Sep 1994.

[61] M. Andre, M. D. Lamblin, A. M. d'Allest, L. Curzi-Dascalova, F. Moussalli-Salefranque, T. S Nguyen The, M. F. Vecchierini-Blineau, F. Wallois, E. Walls-

Esquivel, and P. Plouin. Electroencephalography in premature and full-term infants. Developmental features and glossary. *Neurophysiol Clin*, 40(2):59–124, May 2010.

[62] E. Biagioni, M. F. Frisone, S. Laroche, B. A. Kapetanakis, D. Ricci, M. Adeyi-Obe, H. Lewis, N. Kennea, G. Cioni, F. Cowan, M. Rutherford, D. Azzopardi, and E. Mercuri. Maturation of cerebral electrical activity and development of cortical folding in young very preterm infants. *Clin Neurophysiol*, 118(1):53–59, Jan 2007.

[63] R. Khazipov and H. J. Luhmann. Early patterns of electrical activity in the developing cerebral cortex of humans and rodents. *Trends Neurosci.*, 29(7):414–418, Jul 2006.

[64] J. C. Woestenburg, M. N. Verbaten, and J. L. Slangen. The removal of the eye-movement artifact from the EEG by regression analysis in the frequency domain. *Biol Psychol*, 16(1-2):127–147, 1983.

[65] P. Valdes, J. Bosch, R. Grave, J. Hernandez, J. Riera, R. Pascual, and R. Biscay. Frequency domain models of the EEG. *Brain Topogr*, 4(4):309–319, 1992.

[66] D.B. Percival and A.T. Walden. *Spectral analysis for physical applications: multitaper and conventional univariate techniques*. Cambridge University Press, 1993.

[67] D.B. Percival and A.T. Walden. *Wavelet Methods for Time Series Analysis*. Cambridge Series In Statistical And Probabilistic Mathematics. Cambridge University Press, 2006.

[68] L. Cohen. *Time-frequency analysis*. Prentice Hall PTR, 1995.

[69] A. Worley, L. Fabrizi, S. Boyd, and R. Slater. Multi-modal pain measurements in infants. *J Neurosci Methods*, 205(2):252–7, 2012.

[70] N.J. Stevenson, M. Mesbah, G.B. Boylan, P.B. Colditz, and B. Boashash. A nonlinear model of newborn eeg with nonstationary inputs. *Annals of Biomedical Engineering*, 38(9):3010–3021, 2010.

[71] M.P. Tarvainen, J.K. Hiltunen, P.O. Ranta-aho, and Pasi A. Karjalainen. Estimation of nonstationary eeg with kalman smoother approach: an application to

event-related synchronization (ers). *Biomedical Engineering, IEEE Transactions on*, 51(3):516–524, 2004.

[72] N. Amir and I. Gath. Segmentation of eeg during sleep using time-varying autoregressive modeling. *Biological Cybernetics*, 61:447–455, 1989.

[73] C. C. Chen, S. J. Kiebel, and K. J. Friston. Dynamic causal modelling of induced responses. *Neuroimage*, 41(4):1293–1312, Jul 2008.

[74] B. L. Cheung, B. A. Riedner, G. Tononi, and B. D. Van Veen. Estimation of cortical connectivity from EEG using state-space models. *IEEE Trans Biomed Eng*, 57(9):2122–2134, Sep 2010.

[75] N Stevenson, M Mesbah, and B Boashash. Modelling newborn eeg background using a time-varying fractional brownian process. In *Proceedings of the 15th European Signal Processing Conference*, volume 1, pages 1246–1250, 2007.

[76] M. Odabaee, A. Tokariev, S. Layeghy, M. Mesbah, P. B. Colditz, C. Ramon, and S. Vanhatalo. Neonatal EEG at scalp is focal and implies high skull conductivity in realistic neonatal head models. *Neuroimage*, 96:73–80, Aug 2014.

[77] M. Odabaee, W. J. Freeman, P. B. Colditz, C. Ramon, and S. Vanhatalo. Spatial patterning of the neonatal EEG suggests a need for a high number of electrodes. *Neuroimage*, 68:229–235, Mar 2013.

[78] M. R. Nuwer. Quantitative EEG: I. Techniques and problems of frequency analysis and topographic mapping. *J Clin Neurophysiol*, 5(1):1–43, Jan 1988.

[79] J. Cooley, P. Lewis, and P. Welch. The finite fourier transform. *IEEE Transactions on audio and electroacoustics*, 17(2):77–85, 1969.

[80] R.J. Barton and H.V. Poor. Signal detection in fractional gaussian noise. *Information Theory, IEEE Transactions on*, 34(5):943–959, Sep 1988.

[81] C. Velasco and P. M. Robinson. Whittle pseudo-maximum likelihood estimation for nonstationary time series. *Journal of the American Statistical Association*, 95(452):1229–1243, 2000.

[82] F.J. Harris. On the use of windows for harmonic analysis with the discrete fourier transform. *Proceedings of the IEEE*, 66(1):51–83, Jan 1978.

[83] A. M. Sykulski, S. C. Olhede, J. M. Lilly, and J. J. Early. The whittle likelihood for complex-valued time series. *arXiv preprint arXiv:1306.5993*, 2013.

[84] A. T. Walden. A unified view of multitaper multivariate spectral estimation. *Biometrika*, 87(4):767–788, 2000.

[85] D.J. Thomson. Spectrum estimation and harmonic analysis. *Proceedings of the IEEE*, 70(9):1055–1096, Sept 1982.

[86] K. S. Riedel and A. Sidorenko. Minimum bias multiple taper spectral estimation. *Signal Processing, IEEE Transactions on*, 43(1):188–195, 1995.

[87] D. Slepian and H. O. Pollak. Prolate spheroidal wave functions, fourier analysis and uncertainty—i. *Bell System Technical Journal*, 40(1):43–63, 1961.

[88] H. J. Landau and H. O. Pollak. Prolate spheroidal wave functions, fourier analysis and uncertainty—ii. *Bell System Technical Journal*, 40(1):65–84, 1961.

[89] H. J. Landau and H. O. Pollak. Prolate spheroidal wave functions, fourier analysis and uncertainty—iii: The dimension of the space of essentially time-and band-limited signals. *Bell System Technical Journal*, 41(4):1295–1336, 1962.

[90] D. Slepian. Prolate spheroidal wave functions, fourier analysis and uncertainty—iv: extensions to many dimensions; generalized prolate spheroidal functions. *Bell System Technical Journal*, 43(6):3009–3057, 1964.

[91] D. Slepian. Prolate spheroidal wave functions, fourier analysis, and uncertainty—v: The discrete case. *Bell System Technical Journal*, 57(5):1371–1430, 1978.

[92] Y. Benjamini and Y. Hochberg. Controlling the false discovery rate; a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society*, 1:289–300, 1995.

[93] P. Whittle. Estimation and information in stationary time series. *Arkiv för Matematik*, 2(5):423–434, 1953.

[94] S. Butterworth. On the theory of filter amplifiers. *Wireless Engineer*, 7:536–541, 1930.

[95] S. C. Olhede and A. T. Walden. Generalized morse wavelets. *Signal Processing, IEEE Transactions on*, 50(11):2661–2670, 2002.

[96] S. Vanhatalo and K. Kaila. Development of neonatal EEG activity: from phenomenology to physiology. *Semin Fetal Neonatal Med*, 11(6):471–478, Dec 2006.

[97] K. Palmu, N. Stevenson, S. Wikstrom, L. Hellstrom-Westas, S. Vanhatalo, and J. M. Palva. Optimization of an NLEO-based algorithm for automated detection of spontaneous activity transients in early preterm EEG. *Physiol Meas*, 31(11):85–93, Nov 2010.

[98] K. Palmu, S. Wikstrom, E. Hippelainen, G. Boylan, L. Hellstrom-Westas, and S. Vanhatalo. Detection of 'EEG bursts' in the early preterm EEG: visual vs. automated detection. *Clin Neurophysiol*, 121(7):1015–1022, Jul 2010.

[99] T. Gneiting, W. Kleiber, and M. Schlather. Matérn cross-covariance functions for multivariate random fields. *Journal of the American Statistical Association*, 105(491), 2010.

[100] A. C. Harvey. *Forecasting, structural time series models and the Kalman filter*. Cambridge university press, 1990.

[101] World Medical Association. World medical association declaration of helsinki: Ethical principles for medical research involving human subjects. *JAMA*, 310(20):2191–2194, 2013.

[102] P. M. Robinson. Gaussian semiparametric estimation of long range dependence. *The Annals of statistics*, pages 1630–1661, 1995.

[103] C. M. Hurvich and C.L. Tsai. Regression and time series model selection in small samples. *Biometrika*, 76(2):297–307, 1989.

[104] F. Wilcoxon. Individual comparisons by ranking methods. *Biometrics bulletin*, pages 80–83, 1945.

[105] I. Jolliffe. *Principal component analysis*. Wiley Online Library, 2005.

[106] A. N. Kolmogorov. *Sulla determinazione empirica di una legge di distribuzione*. na, 1933.

[107] N. Smirnov. Table for estimating the goodness of fit of empirical distributions. *The Annals of Mathematical Statistics*, pages 279–281, 1948.

[108] K. J. Anand, V. Coskun, K. V. Thrivikraman, C. B. Nemeroff, and P. M. Plotsky. Long-term behavioral effects of repetitive pain in neonatal rat pups. *Physiol. Behav.*, 66(4):627–637, Jun 1999.

[109] D. R. Brillinger. *Time Series: Data Analysis and Theory*. Society for Industrial and Applied Mathematics, 2001.

[110] G. C. Reinsel. *Elements of multivariate time series analysis*. Springer Series in Statistics. Springer, 2003.

[111] D. B Percival. Simulating gaussian random processes with specified spectra. *Computing Science and Statistics*, 24:534–538, 1992.

[112] L. Rankine, N. Stevenson, M. Mesbah, and B. Boashash. A nonstationary model of newborn EEG. *IEEE Trans Biomed Eng*, 54(1):19–28, Jan 2007.

[113] N. J. Stevenson, J. M. O'Toole, L. J. Rankine, G. B. Boylan, and B. Boashash. A nonparametric feature for neonatal EEG seizure detection based on a representation of pseudo-periodicity. *Med Eng Phys*, 34(4):437–446, May 2012.

[114] S. Faul, G. Gregorcic, G. Boylan, W. Marnane, G. Lightbody, and S. Connolly. Gaussian process modeling of EEG for the detection of neonatal seizures. *IEEE Trans Biomed Eng*, 54(12):2151–2162, Dec 2007.

[115] S. Z Li and S. Singh. *Markov random field modeling in image analysis*, volume 26. Springer, 2009.

[116] 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*, 145(3):997–1006, Mar 2007.

[117] M. T. Colonnese, A. Kaminska, M. Minlebaev, M. Milh, B. Bloem, S. Lescure, G. Moriette, C. Chiron, Y. Ben-Ari, and R. Khazipov. A conserved switch in sensory processing prepares developing neocortex for vision. *Neuron*, 67(3):480–498, Aug 2010.

[118] S. B. Nagaraj, N. J. Stevenson, W. P. Marnane, G. B. Boylan, and G. Lightbody. Robustness of time frequency distribution based features for automated neonatal EEG seizure detection. *Conf Proc IEEE Eng Med Biol Soc*, 2014:2829–2832, Aug 2014.

[119] M. I. Garrido, J. M. Kilner, S. J. Kiebel, K. E. Stephan, and K. J. Friston. Dynamic causal modelling of evoked potentials: a reproducibility study. *Neuroimage*, 36(3):571–580, Jul 2007.

[120] B. R. Greene, G. B. Boylan, R. B. Reilly, P. de Chazal, and S. Connolly. Combination of EEG and ECG for improved automatic neonatal seizure detection. *Clin Neurophysiol*, 118(6):1348–1359, Jun 2007.

[121] B. Telenczuk, S. N. Baker, A. V. Herz, and G. Curio. High-frequency EEG covaries with spike burst patterns detected in cortical neurons. *J. Neurophysiol.*, 105(6):2951–2959, Jun 2011.

[122] N. X. Tritsch, E. Yi, J. E. Gale, E. Glowatzki, and D. E. Bergles. The origin of spontaneous activity in the developing auditory system. *Nature*, 450(7166):50–55, Nov 2007.

[123] F. Z Roueff and R. von Sachs. Locally stationary long memory estimation. *Stochastic Processes and their Applications*, 121(4):813 – 844, 2011.