

**Assessing Variability of EEG and  
ECG/HRV Time Series Signals using a  
Variety of Non-Linear Methods**

Submitted to the University of Hertfordshire in  
partial fulfilment of the requirements for the degree  
of

*Doctor of Philosophy*

**Ronakben P. Bhavsar**

Department of Computer Science

October 2019

This thesis is dedicated to

. . . to my beloved son, **Dhven Vyas** for his endless love, support, encouragement, giving me strength, and allowing me to stay away from him to pursue my goals.

. . . to my beloved husband, **Mr. Chiragkumar Vyas**, for his love and presence that always gives me strength and support to face the challenges, and urges me to strive to achieve my goals in life. He is a constant source of inspiration in my life.

-Ronakben Bhavsar

## Acknowledgements

I take this opportunity to express profound gratitude and deep regards to my supervisor, Dr. Na Helian, for her exemplary guidance, stimulating suggestions, monitoring and constant encouragement throughout the course of the project. With her calm, patient and attentive approach, I was encouraged to give my best efforts in this thesis. Also, I thank her for her constant availability to discuss any doubts related to the project.

I would like to acknowledge with much appreciation the crucial role of Dr. Yi Sun, and Dr. Neil Davey, my second supervisors, who helped me understand the various concepts giving me a clear view of the project. I would like to thank them for the timely guidance provided with respect to the analysis. I am grateful to Mr. David Mayor for his insights to the data provided and suggestions related to the research for the project, and Mr. Tony Steffert for giving me the training to collect the EEG and ECG recordings. I was able to get knowledge of EEG and ECG inside out with their support and advice.

Ramon Y. Cajal, the father of modern neurobiology, almost a hundred years ago wrote: "more than once I was hopelessly discouraged about my ability to pursue science". Such times have been aplenty in my career and in the development not only of this thesis, but also in the investigation of a novel area of research that is different from traditional approaches. In these times, there is one certainty: that solace is needed around you to overcome even the highest of obstacles. This is why I am eternally grateful for my cornerstone, my husband Chiragkumar, my son Dhven, my In-laws Jayantilal and Ramaben, and my parents Dipaben and Prafulkumar, for their love, support, and constant encouragement. I would like to thank almighty, my husband and son for their constant support, motivation and sacrifices while I couldn't be with them. My parents for blessing me with this wonderful life and positive attitude. My friends for their encouragement and confidence in me achieving my goals. A special thanks to my friends – Helen Partou, Zaheed Mahmood, and Weam Binjumah for their endless support and motivation which helped me in pursuit of my goals for this project.

## Abstract

Time series signals, such as Electroencephalogram (EEG) and Electrocardiogram (ECG) represent the complex dynamic behaviours of biological systems. The analysis of these signals using variety of nonlinear methods is essential for understanding variability within EEG and ECG, which potentially could help unveiling hidden patterns related to underlying physiological mechanisms. EEG is a time varying signal, and electrodes for recording EEG at different positions on the scalp give different time varying signals. There might be correlation between these signals. It is important to know the correlation between EEG signals because it might tell whether or not brain activities from different areas are related. EEG and ECG might be related to each other because both of them are generated from one co-ordinately working body. Investigating this relationship is of interest because it may reveal information about the correlation between EEG and ECG signals.

This thesis is about assessing variability of time series data, EEG and ECG, using variety of nonlinear measures. Although other research has looked into the correlation between EEGs using a limited number of electrodes and a limited number of combinations of electrode pairs, no research has investigated the correlation between EEG signals and distance between electrodes. Furthermore, no one has compared the correlation performance for participants with and without medical conditions. In my research, I have filled up these gaps by using a full range of electrodes and all possible combinations of electrode pairs analysed in Time Domain (TD). Cross-Correlation method is calculated on the processed EEG signals for different number unique electrode pairs from each datasets. In order to obtain the distance in centimetres (cm) between electrodes, a measuring tape was used. For most of our participants the head circumference range was 54-58cm, for which a medium-sized I have discovered that the correlation between EEG signals measured through electrodes is linearly dependent on the physical distance (straight-line) distance between them for datasets without medical condition, but not for datasets with medical conditions.

Some research has investigated correlation between EEG and Heart Rate Variability (HRV) within limited brain areas and demonstrated the existence of correlation between EEG and HRV. But no research has indicated whether or not the correlation changes with brain area. Although Wavelet Transformations (WT) have been performed on time series data including EEG and HRV signals to extract certain features respectively by other research, so far correlation between WT signals of EEG and HRV has not been analysed. My research covers these gaps by conducting a thorough investigation of all electrodes on the human scalp in Frequency Domain (FD) as well as TD. For the reason of different sample rates of EEG and HRV, two different approaches (named as Method 1 and Method 2) are utilised to segment EEG signals and to calculate Pearson's Correlation Coefficient for each of the EEG frequencies with each of the HRV frequencies in FD. I have demonstrated that EEG at the front area of the brain has a stronger correlation with HRV than that at the other area in a frequency domain. These findings are independent of both participants and brain hemispheres.

Sample Entropy (SE) is used to predict complexity of time series data. Recent research has proposed new calculation methods for SE, aiming to improve the accuracy. To my knowledge, no one has attempted to reduce the computational time of SE calculation. I have developed a new calculation method for time series complexity which could improve computational time significantly in the context of calculating a correlation between EEG and HRV. The results have a parsimonious outcome of SE calculation by exploiting a new method of SE implementation. In addition, it is found that the electrical activity in the frontal lobe of the brain appears to be correlated with the HRV in a time domain.

Time series analysis method has been utilised to study complex systems that appear ubiquitous in nature, but limited to certain dynamic systems (e.g. analysing variables affecting stock values). In this thesis, I have also investigated the nature of the dynamic system of HRV. I have disclosed that Embedding Dimension could unveil two variables that determined HRV.

**Keywords:** Time series Signals, EEG, ECG, HRV, Cross-Correlation (CC), Pearson Correlation Coefficient (PCC), Wavelet Transform (WT), Sample Entropy, Embedding Dimension (ED), False Nearest Neighbours (FNN).

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Research Motivation . . . . .	2
1.2	Research Questions . . . . .	3
1.3	Research Contribution . . . . .	4
1.4	Publications on This Thesis . . . . .	5
1.5	Structure of this Thesis . . . . .	5
1.6	Terminology Abbreviations . . . . .	6
<b>2</b>	<b>Literature Review: Assessing Variability of Time Series Data EEG and ECG</b>	<b>8</b>
2.1	Introduction . . . . .	8
2.1.1	Dynamical, Nonlinear and Chaos System . . . . .	8
2.1.2	Time Series Data . . . . .	12
2.1.2.1	Length of the Time Series . . . . .	14
2.1.2.2	Sampling Frequency . . . . .	15
2.1.2.3	Noise . . . . .	15
2.1.2.4	Filtering/Smoothing . . . . .	16
2.1.2.5	Stationarity . . . . .	18
2.1.3	Time Series Data Analysis Methods . . . . .	19
2.2	Electroencephalogram (EEG) . . . . .	22
2.2.1	The Nervous System . . . . .	22
2.2.2	EEG Recordings . . . . .	23
2.2.3	Types of EEG Frequencies . . . . .	24
2.2.4	EEG Artefacts . . . . .	25
2.2.5	Research on EEG . . . . .	26
2.3	Electrocardiogram (ECG) and Heart Rate Variability (HRV) . . . . .	31
2.3.1	Electrocardiography (ECG) . . . . .	31
2.3.2	Heart Rate Variability (HRV) . . . . .	32
2.3.2.1	Types of HRV frequencies . . . . .	34

2.3.2.2	HRV Artefacts . . . . .	35
2.3.3	Research on ECG and HRV . . . . .	36
2.4	Correlation Between EEG and HRV Signals . . . . .	38
2.4.1	Significance of the Wavelet Transformation on EEG and HRV Signals	41
2.5	Summary . . . . .	43
<b>3</b>	<b>Data Description</b>	<b>45</b>
3.1	Introduction . . . . .	45
3.2	Self Recording Dataset . . . . .	45
3.2.1	Recording Process . . . . .	46
3.2.2	Dataset 1 . . . . .	48
3.3	Other Datasets . . . . .	49
3.3.1	Dataset 2 . . . . .	50
3.3.2	Dataset 3 . . . . .	50
3.3.3	Dataset 4 . . . . .	53
3.3.4	Dataset 5 . . . . .	53
3.3.5	Dataset 6 . . . . .	54
3.3.6	Dataset 7 . . . . .	54
3.3.7	Dataset 8 . . . . .	55
3.3.8	Dataset 9 . . . . .	55
3.3.9	Dataset 10 . . . . .	56
3.4	Summary . . . . .	56
<b>4</b>	<b>Time Series Data Preprocessing</b>	<b>59</b>
4.1	Introduction . . . . .	59
4.1.1	Data Preprocessing and Its Importance . . . . .	59
4.1.2	Preprocessing Techniques for Time Series Data . . . . .	60
4.1.3	Preprocessing Techniques for EEG and HRV Time Series Data . . . .	62
4.2	Independent Component Analysis (ICA) . . . . .	63
4.2.1	An Example of ICA Calculation . . . . .	68
4.3	Fast Fourier Transform . . . . .	71
4.3.1	An Example of FFT Calculation . . . . .	73
4.4	Wavelet Transform . . . . .	77
4.4.1	Haar Wavelet . . . . .	78
4.4.1.1	An Example of Haar Wavelet Calculation . . . . .	79
4.4.2	Daubechies Wavelet . . . . .	81
4.4.2.1	An Example of Daubechies Wavelet Calculation . . . . .	84

4.5	Summary . . . . .	86
<b>5</b>	<b>Methods for Analysing EEG and HRV Time Series Data</b>	<b>88</b>
5.1	Introduction . . . . .	88
5.2	Approximate Entropy . . . . .	90
5.2.1	An Example of Approximate Entropy Calculation . . . . .	91
5.3	Sample Entropy . . . . .	94
5.3.1	An Example of Sample Entropy Calculation . . . . .	95
5.4	Pearson’s Correlation Coefficient . . . . .	97
5.4.1	An Example of Pearson’s Correlation Coefficient . . . . .	98
5.5	Cross-Correlation . . . . .	99
5.5.1	An Example of Cross-Correlation . . . . .	100
5.6	Embedding Dimension . . . . .	105
5.6.1	An Illustration of Embedding Dimension . . . . .	106
5.6.2	An Example of Embedding Dimension Calculation . . . . .	108
5.6.3	An Example of Lorenz Attractor . . . . .	110
5.7	Conclusion . . . . .	112
<b>6</b>	<b>The Correlation between EEG Signals Measured at Different Positions on Scalp Varying with Distance</b>	<b>114</b>
6.1	Introduction . . . . .	114
6.2	Dataset Information . . . . .	116
6.3	Experiments and Results . . . . .	116
6.3.1	Experiments . . . . .	116
6.3.2	Results . . . . .	116
6.4	Conclusion . . . . .	122
<b>7</b>	<b>Correlation Analysis between EEG and HRV Time series data</b>	<b>124</b>
7.1	Introduction . . . . .	124
7.1.1	Dataset Information . . . . .	125
7.2	Experiments . . . . .	126
7.2.1	Time Domain (TD) Analysis . . . . .	127
7.2.2	Frequency Domain (FD) Analysis . . . . .	127
7.2.2.1	Method 1 . . . . .	127
7.2.2.2	Method 2 . . . . .	128
7.3	Experimental Results and Discussion . . . . .	129
7.3.1	TD Analysis . . . . .	129

7.3.2	FD Analysis . . . . .	134
7.4	Conclusion . . . . .	141
<b>8</b>	<b>Efficient Methods for Calculating Sample Entropy in Time Series Data Analysis</b>	<b>145</b>
8.1	Introduction . . . . .	145
8.1.1	Dataset Information . . . . .	146
8.2	Sample Entropy and Proposed Implementation . . . . .	146
8.2.1	Efficient and Parsimonious way for Sample Entropy Calculation . . . . .	146
8.3	Experimental Results . . . . .	147
8.3.1	Experiments using three proposed SE Calculation Methods . . . . .	148
8.3.2	Experiment 4 . . . . .	150
8.4	Discussion and Conclusion . . . . .	151
<b>9</b>	<b>Heart Rate Variability Time Series Analysis using Embedding Dimension</b>	<b>153</b>
9.1	Introduction . . . . .	153
9.1.1	Dataset Information . . . . .	153
9.2	Experiments and Results . . . . .	154
9.2.1	HRV Analysis using ED . . . . .	154
9.2.2	HRV Prediction using Linear Regression . . . . .	158
9.3	Discussion and Conclusion . . . . .	160
<b>10</b>	<b>Conclusion and Future Work</b>	<b>161</b>
10.1	Summary of Chapters . . . . .	161
10.2	Contribution to knowledge . . . . .	164
10.3	Future Work . . . . .	166
<b>A</b>	<b>Dataset 1- Recording Photos</b>	<b>168</b>
<b>B</b>	<b>Cross-correlation performance of Electrodes Pair against Time lags and Distance</b>	<b>174</b>
B.1	Cross-correlation performance of Electrodes Pair against Time Lags . . . . .	174
B.1.1	Dataset 1 . . . . .	174
B.1.2	Dataset 4 . . . . .	180
B.2	Cross-Correlation performance of Electrodes Pair against Distance between them . . . . .	184
B.2.1	Dataset 1 . . . . .	184
B.2.2	Dataset 4 . . . . .	187

B.2.3	Dataset 5 . . . . .	189
B.2.4	Dataset 6 . . . . .	192
B.2.5	Dataset 7 . . . . .	195
B.2.6	Dataset 8 . . . . .	197
<b>C</b>	<b>Additional SE performance of SE-Method 2 and SE-Method 3</b>	<b>200</b>
C.1	Slot 1 Result . . . . .	200
C.1.1	SE-Method 2 . . . . .	200
C.1.2	SE-Method 3 . . . . .	202
C.2	Slot 5 Result . . . . .	203
C.2.1	SE-Method 1 . . . . .	203
C.2.2	SE-Method 2 . . . . .	203
C.2.3	SE-Method 3 . . . . .	205
<b>D</b>	<b>Additional ED Results for Chapter 9</b>	<b>206</b>
<b>E</b>	<b>Publications</b>	<b>211</b>
<b>F</b>	<b>Publications</b>	<b>245</b>
	<b>Bibliography</b>	<b>274</b>

# List of Figures

2.1	Not all nonlinear systems are chaotic, but all chaotic systems are nonlinear. Nonlinear systems are in turn a particular form of dynamical systems (Redrawn from (James and Walker, 1999)). . . . .	9
2.2	Phase space plots provide a means to geometrically show chaotic behaviour which may not be readily apparent within a time series James and Walker (1999). Left side of the figure shows the signal, and right side is the phase plane plot (Embedding Dimension), where (a) Periodic signal, (b) Chaotic signal, and (c) Random signal with white noise. . . . .	11
2.3	The length of Pinocchio’s nose: (a) continuous growth and (b) growth using discrete measurements. . . . .	13
2.4	Plot of 20 seconds ECG data from a participant under acupuncture treatment. The data were gathered using sampling frequency of 256Hz. . . . .	14
2.5	Plot of 10 seconds EEG data from a participant watching a short documentary film. The data were gathered using sampling frequency of 500Hz. . . . .	14
2.6	Time series data (left) and corresponding power spectra (right) for a (a) sine wave of 5 Hz, (b) a sine wave of 20 Hz, (c) a sine wave of 50 Hz, and (d) the sum of these three wave functions. . . . .	16
2.7	Time series data (left) and corresponding power spectra (right) for a (a) sine wave of 20 Hz, (b) white noise, and (c) the sum of the sine wave and white noise. . . . .	17
2.8	Time series left and power spectrum right: (a) unfiltered, (b) filtered using a cutoff frequency of 33 Hz, (c) filtered using a cutoff frequency of 15 Hz, and (d) filtered using a cutoff frequency of 5 Hz. The filter used was a “brick wall” filter (very high-order Butterworth filter), and cutoff frequency is indicated by vertical bar (James and Walker, 1999). . . . .	18

2.9	White noise time series data showing (a) stationarity, (b) nonstationarity due to change in mean, and (c) nonstationarity due to change in variance (James and Walker, 1999). . . . .	20
2.10	The international 10%-20% system seen from A (left side of the head) and B (above the head). The letter F, T, C, P, O, A, Fp and Pg stands for frontal, temporal, central, parietal, occipital, earlobes, frontal polar, and nasopharyngeal, respectively. The figure is obtained from (Klem et al., 1999).	24
2.11	Traced Frequencies from EEG. The Amplitude (Power) of the signals increases as the frequencies decreases (Gastaut, 1952). . . . .	25
2.12	Cardiac Artefact presents in the Data Grouiller et al. (2007) . . . . .	27
2.13	Electrode Artefact showing noise caused by lead movement Grouiller et al. (2007) . . . . .	27
2.14	Ocular Artefacts: (a) Eye blinks with rapid movement of the eyes both up and down, and (b) Lateral Eye Movement . . . . .	28
2.15	Muscle Artefacts: (a) Swallow effect, and (b) Jaw movement. . . . .	29
2.16	Determination of Inter Beat Interval (IBI). Simulated ECG containing three beats with arbitrary units of time and amplitude. Time intervals corresponding to the IBI are indicated by IBI(1) and IBI(2). ECG morphology is shown by five characteristic waves P, Q, R, S, and T. . . . .	34
2.17	HRV pattern, showing the contribution of Sympathetic and Parasympathetic Nervous Activity (Steffert and Mayor). . . . .	35
2.18	Power Spectrum of HRV showing presence of traced HRV frequencies . . .	36
2.19	Traced Frequencies from HRV. The Amplitude (Power) of the signals increases as the frequencies decreases (Lin et al., 2010). . . . .	36
2.20	Presence of extra beat and premature ventricular contractions (PVC) artefact in HRV. The highlighted in red circles are the extra beat . . . . .	37
3.1	Transcutaneous Electrical Acupoint Stimulation (TEAS), a safe and standardized acupuncture technique, which does not require needles insertion (V.Kaye, 1998-2019) . . . . .	46
3.2	Visit to Acupuncture Clinic, showing myself as a participant, along with Dr. Tony Steffert (EEG Specialist and Researcher) and Dr. Na Helian (My Principal Supervisor). . . . .	47
3.3	Dataset 1- Time line diagram showing slot's information . . . . .	49
3.4	Dataset 2- Time line diagram showing slot's information . . . . .	51
3.5	Electro acupuncture (EA) technique (Dr.Evans, 1998-2019). . . . .	51

3.6	Manual Acupuncture (MA) technique (Dr.Xie, 1998-2019) . . . . .	52
3.7	Dataset 3- Time line diagram showing slot's information . . . . .	52
3.8	Dataset 4- Time line diagram showing slot's information . . . . .	53
4.1	Illustration of BSS problem that Independent Component Analysis (ICA) addresses using the popular cocktail-party problem. The goal is to recover the individual signals $S1(t)$ and $S2(t)$ from the mixtures signals $X1(t)$ and $X2(t)$ measured from Microphone 1 and 2, respectively. . . . .	64
4.2	Illustration of Independent Component Analysis (ICA) showing two stage process: Do stage (left hand side of the figure), and Undo stage (right hand side of the figure). . . . .	66
4.3	ICA example, a) showing two signals in blue colour $S1$ (top left) and $S2$ (top Right), b) showing two mixed sources in red colour $X1$ (middle panel on left) and $X2$ (middle panel on right), with a linear mixture of sources $S1$ and $S2$ , and c) showing uncovering original activation in green colour of two signals $S1$ (bottom left) and $S2$ (bottom right) after ICA. . . . .	70
4.4	An example of 4 data point FFT algorithm structure, using a decomposition into half-size FFTs. $X_E[k]$ is representing even samples, and $X_O[k]$ is representing Odd samples. . . . .	73
4.5	An example of 4 point FFT algorithm structure, using a decomposition into half-size FFTs of odd and even sequences. $x_E[k]$ is representing even samples, and $x_O[k]$ is representing Odd samples. . . . .	74
4.6	The sequence in TD $X[n]$ (top in blue color) and in FD $x[k]$ (bottom in red colour). . . . .	75
4.7	An example of FFT performance of signals S (Bottom panel- as shown in colour green). . . . .	76
4.8	Three level sub-band decomposition of discrete wavelet transform (DWT) implementation; $g[n]$ is the low-pass filter, and $h[n]$ is the high-pass filter. $g[n]$ is the approximation part of the signal containing high scale and low frequencies, and $h[n]$ is the detail part of the signal containing low scale and high frequencies (Akay, 1997). . . . .	78
4.9	Scaling (a) and Wavelets (b) functions being considered for Haar Wavelet (Mutihac, 2006). . . . .	80
4.10	Example of Haar Wavelet Transform: The top panel shows original signal in magenta colour, the middle panel shows approximation part of WT in red colour, and the bottom panel shows detail part of WT in blue colour. . .	81

4.11	Example of 1-level Haar Wavelet Transform (WT) on EEG signal: The top panel shows EEG signal, the middle panel (approximation part of WT signal) and the bottom panel(detailed part of WT signal). . . . .	82
4.12	Scaling (a) and Wavelets (b) functions being considered for Daubechies Wavelet (Mutihac, 2006). . . . .	84
4.13	Example of Daubechies Wavelet Transform: The top panel shows the original signal in magenta colour, the middle panel shows approximation part of WT in red colour, and the bottom panel shows detailed part of WT in blue color.. . . .	86
4.14	Example of 1-level Daubechies Wavelet Transform (WT) on EEG Signal: The top panel shows EEG signal, the middle panel (approximation part of WT signal) and the bottom panel(detailed part of WT signal). . . . .	87
5.1	Comparison for matches at each length of sequence of $m$ , for the input time series $x(n)$ . Where, Length of the sequence is shown as red, and point sequences are shown in black at $m= 1;2;$ and 3. . . . .	92
5.2	Comparison for matches at each length of sequence of $m$ , for the input time series $x(n)$ . Where, Length of the sequence is shown as red, and point sequences are shown in black at $m = 0;1;$ and 2. . . . .	96
5.3	Correlation at Lag of 0. . . . .	101
5.4	Correlation at Lag of 1. . . . .	102
5.5	Correlation at Lag of 2. . . . .	103
5.6	Correlation at Lag of -1. . . . .	104
5.7	Geometric explanation of the <i>False Nearest Neighbours</i> Method (Kennel et al., 1992): (a) In one-dimensional, the nearest neighbour for Red is Green, (b) In two-dimensional, the nearest neighbour for Red is Blue, and not Green (This means Green was a false nearest neighbour in one-dimensional), and (c) In three-dimensional, the Blue is still the nearest neighbour for Red, so they are real nearest neighbour. . . . .	107
5.8	Reconstruction of the state-space with a embedding dimension of 3, from a scalar time series represented by a sequence of $x_1, x_2, \dots, x_{21}$ , considering Time Lag (t) = 1. . . . .	109
5.9	The percentage of <i>False Nearest Neighbour</i> example. The percentage of <i>False Nearest Neighbour</i> hits zero when the dimension is three. . . . .	110
5.10	A visualisation of the Lorenz attractor in 3-dimensional phase space $x(t), y(t), z(t)$ (Frank et al., 2001). . . . .	111

5.11	A visualisation of the Lorenz attractor in Three-dimensional phase space $x(t), y(t), z(t)$ with Noise (Frank et al., 2001). . . . .	111
5.12	The percentage of <i>False Nearest Neighbours</i> in the Lorenz data set (Frank et al., 2001). . . . .	112
6.1	Cross-Correlation at all possible Lags for all electrode pairs for Electrode Fp1. Blue colour shows the result of one of the participants, and Orange colour shows the average of 15 participants result from Dataset 1 . . . . .	117
6.2	Dataset 1 - Positive Cross-Correlation at Lag 0. . . . .	118
6.3	Dataset 1 - Negative Cross-Correlation at Lag 0. . . . .	118
6.4	Cross-Correlation between electrodes at varying distance on Dataset 1. . . . .	119
6.5	Cross-Correlation between electrodes at varying distance on Dataset 4. . . . .	119
6.6	Cross-Correlation between electrodes at varying distance on Dataset 5. . . . .	120
6.7	Cross-Correlation between electrodes at varying distance on Dataset 6 (Autism). . . . .	120
6.8	Cross-Correlation between electrodes at varying distance on Dataset 7 (Epilepsy). . . . .	121
6.9	Cross-Correlation between electrodes at varying distance on Dataset 8 (Seizure). . . . .	121
7.1	Method 1 - experiment steps of the correlation performance. . . . .	129
7.2	Method 2- experiment steps of the correlation performance. . . . .	130
7.3	WT EEG signals showing: a) unfiltered EEG signal in blue colour, b) Low-pass filtered EEG shown in red colour (Approximation part, showing trend of the EEG signal), c) High-pass filtered EEG shown in green colour (Detailed part, showing fluctuation of the EEG signal), and d) WT EEG signal with low-passed filter. . . . .	131
7.4	Dataset 1 - Time domain (TD) analysis of correlation performance between each of the EEG electrodes with HRV for each 5-minute slot. . . . .	132
7.5	Dataset 2 - Time domain (TD) analysis of correlation performance between each of the EEG electrodes with HRV for each 5-minute slot. . . . .	133
7.6	Method 1 - Best Electrodes Correlation Performance, highlighted in yellow colour: (a) Experiment 1 - Dataset 1 Correlation performance on pre-processed HRV and EEG, (b) Experiment 3 - Dataset 1 Correlation performance on pre-processed HRV and WT signals of EEG ,(c) Experiment 1 - Dataset 2 Correlation performance on pre-processed HRV and EEG, (d) Experiment 3 - Dataset 2 Correlation performance on pre-processed HRV and WT signals of EEG. . . . .	135

7.7	Method 2 - Best Electrodes Correlation Performance, highlighted in yellow colour: (a) Experiment 1 - Dataset 1 Correlation performance on pre-processed HRV and EEG, (b) Experiment 3 - Dataset 1 Correlation performance on pre-processed HRV and WT signals of EEG ,(c) Experiment 1 - Dataset 2 Correlation performance on pre-processed HRV and EEG, (d) Experiment 3 - Dataset 2 Correlation performance on pre-processed HRV and WT signals of EEG. . . . .	136
7.8	Key neuronal projections that maintain alertness, and possibly the path from the cardiovascular centre to the frontal lobe of the brain's communication. The figure is obtained from (Saper et al., 2005). . . . .	137
8.1	An example of how the SE can be calculated efficiently. . . . .	147
8.2	Comparison of original SE calculation and proposed SE methods, showing average results for all electrodes are shown for the individual participants : (a) Method 1-Shortened Neighbours comparison, (b) Method 2- Using Moving Window of 2 seconds of data, and (c) Method 3- Calculating SE from Mean values of each 2 seconds of data. Red= Original SE performance, Blue= Proposed Method's SE Performance. . . . .	149
8.3	Electrode Ranking based on correlation performance between SE values of EEG and HRV, showing best three performing electrodes across participants, highlighted in yellow colour: (a) Ranking based on the original approach for SE calculation, and (b) Ranking based on the new approach (SE-Method 3) for SE calculation. . . . .	152
9.1	Dataset 2- Embedding Dimension Result: Two participants (Participant 1 and 2) ED result from Dataset 2. . . . .	155
9.2	Dataset 3 - Embedding Dimension Result: Two participant's (Participant 3 and 4) ED result from Dataset 3. . . . .	156
9.3	Dataset 9 - Embedding Dimension Result: (a) Two Participants (Participant 3 and 4) ED result for HRV without medical condition, (b) Two Participants (Participant 1 and 3) ED result for HRV with a medical condition Congestive Heart Failure (CHF). . . . .	157
9.4	Embedding Dimension Result: Two participants (Participant 2 and participant 4 (5 measured Electrodes, as shown in different colour)) from Dataset 10. . . . .	158

9.5	Window Size selection for the HRV prediction, where red coloured data points are not considered, and the yellow highlighted data points are repeated in the window size of 1 and 2. . . . .	159
9.6	Relative error of the predictor for varying window size for HRV Time series.	159
A.1	Location of the Recording at the University of Hertfordshire Premises. . . .	170
A.2	Device used to control Transcutaneous Electroacupuncture Stimulation (TEAS) Stimulation Frequency. . . . .	171
A.3	ECG electrodes attached on both wrists, TEA electrodes attached on both hands, and sensor attached on the indexed finger. . . . .	171
A.4	Tie leads together just behind cap, keeping them loose. . . . .	172
A.5	Explaining one of the participants the next step of getting a good connection between the surface of the scalp and EEG electrodes. Explaining (black = no connection vs white = very good connection or yellow = good connection colour) what colours represented on the screen in front of them.	172
A.6	Adjustment between the gel and the hair is needed to make the connection between the electrodes and the surface of the scalp. Therefore, rubbing the gel, and checking the connection on the laptop screen (This process can be seen by participants also). . . . .	173
A.7	Impedance screen showing all 21 electrodes before them, being connected on the surface of the scalp (Black colour = No connection). . . . .	173
B.1	Cross-correlation at all possible Lags for all electrode pairs for Electrode Fp2. Blue color shows the result for one of the participant, and Orange color shows the average of 15 participants result from Dataset 1 . . . . .	175
B.2	Cross-Correlation at all possible Lags for all electrode pairs for Electrode T3. Blue color shows the result for one of the participant, and Orange color shows the average of 15 participants result from Dataset 1 . . . . .	176
B.3	Cross-Correlation at all possible Lags for all electrode pairs for Electrode T4. Blue color shows the result for one of the participant, and Orange color shows the average of 15 participants result from Dataset 1 . . . . .	177
B.4	Cross-Correlation at all possible Lags for all electrode pairs for Electrode O1. Blue color shows the result for one of the participant, and Orange color shows the average of 15 participants result from Dataset 1 . . . . .	178
B.5	Cross-Correlation at all possible Lags for all electrode pairs for Electrode O2. Blue color shows the result for one of the participant, and Orange color shows the average of 15 participants result from Dataset 1 . . . . .	179

B.6	Cross-Correlation at all possible Lags for all electrode pairs for Electrode F8. Blue color shows the result for one of the participant, and Orange color shows the average of 15 participants result from Dataset 4 . . . . .	181
B.7	Cross-Correlation at all possible Lags for all electrode pairs for Electrode P3. Blue color shows the result for one of the participant, and Orange color shows the average of 15 participants result from Dataset 4 . . . . .	182
B.8	Cross-Correlation at all possible Lags for all electrode pairs for Electrode P4. Blue color shows the result for one of the participant, and Orange color shows the average of 15 participants result from Dataset 4 . . . . .	183
B.9	Cross-Correlation between electrodes at varying distance on Dataset 1 for: (a) Electrode Fp2, (b) Electrode F7, (c) Electrode F3, (d) Electrode Fz, (e) Electrode F4, (f) Electrode F8, (g) Electrode T3, and (h) Electrode C3 . . .	185
B.10	Cross-Correlation between electrodes at varying distance on Dataset 1 for: (a) Electrode Cz, (b) Electrode C4, (c) Electrode T4, (d) Electrode T5, (e) Electrode P3, (f) Electrode Pz, (g) Electrode P4, and (h) Electrode T6 . . .	186
B.11	Cross-Correlation between electrodes at varying distance on Dataset 1 for: (a) Electrode O1, and (b) Electrode O2. . . . .	187
B.12	Cross-Correlation between electrodes at varying distance on Dataset 4 for: (a) Electrode F3, (b) Electrode Fz, (c) Electrode F4, (d) Electrode F8, (e) Electrode T5, (f) Electrode P3, (g) Electrode Pz, (h) Electrode P4, and (i) Electrode T6 . . . . .	188
B.13	Cross-Correlation between electrodes at varying distance on Dataset 5 for: (a) Electrode Fp2, (b) Electrode F7, (c) Electrode F3, (d) Electrode Fz, (e) Electrode F4, (f) Electrode F8, (g) Electrode C3, and (h) Electrode Cz. . .	190
B.14	Cross-Correlation between electrodes at varying distance on Dataset 5 for: (a) Electrode C4, (b) Electrode P3, (c) Electrode Pz, (d) Electrode P4, (e) Electrode O1, and (f) Electrode O2. . . . .	191
B.15	Cross-Correlation between electrodes at varying distance on Dataset 6 for: (a) Electrode Fp2, (b) Electrode F7, (c) Electrode F3, (d) Electrode Fz, (e) Electrode F4, (f) Electrode F8, (g) Electrode T3, and (h) Electrode C3 . . .	193
B.16	Cross-Correlation between electrodes at varying distance on Dataset 6 for: (a) Electrode Cz, (b) Electrode C4, (c) Electrode T4, (d) Electrode T5, (e) Electrode P3, (f) Electrode Pz, (g) Electrode P4, (h) Electrode T6, (i) Electrode O1, and (j) Electrode O. . . . .	194

B.17	Cross-Correlation between electrodes at varying distance on Dataset 7 for: (a) Electrode Fp2, (b) Electrode F7, (c) Electrode F3, (d) Electrode Fz, (e) Electrode F4, (f) Electrode F8, (g) Electrode T3, and (h) Electrode C3 . . . . .	196
B.18	Cross-Correlation between electrodes at varying distance on Dataset 8 for: (a) Electrode Fp2, (b) Electrode F7, (c) Electrode F3, (d) Electrode Fz, (e) Electrode F4, (f) Electrode F8, (g) Electrode T3, and (h) Electrode C3 . . . . .	198
B.19	Cross-Correlation between electrodes at varying distance on Dataset 8 for: (a) Electrode Cz, (b) Electrode C4, (c) Electrode T4, (d) Electrode T5, (e) Electrode P3, (f) Electrode Pz, (g) Electrode P4, (h) Electrode T6, (i) Electrode O1, and (j) Electrode O. . . . .	199
C.1	SE-Method 2 - Comparison of original SE calculation and proposed SE methods, showing average results for all electrodes are shown for the indi- vidual participants: SE-Method 2 on moving windows for: (a) 10 Seconds, (b) 20 Seconds, (c) 30 Seconds, (d) 40 Seconds, (e) 50 Seconds, (f) 60 Seconds, (g) 70 Seconds, (h) 80 Seconds, and (i) 90 Seconds. . . . .	201
C.2	Slot 1 - Comparison of original SE calculation and proposed SE methods, showing average results for all electrodes are shown for the individual par- ticipants: SE-Method 3 on moving windows for: (a) 0.06 Seconds, (b) 0.12 Seconds, (c) 0.25 Seconds, (d) 0.55 Seconds, (e) 1 Seconds, (f) 3 Seconds, and (g) 4 Seconds. . . . .	202
C.3	SE-Method 1 - Comparison of original SE calculation and proposed SE methods, showing average results for all electrodes are shown for the indi- vidual participants: SE-Method 1-Shortened Neighbours comparison. . . . .	203
C.4	SE-Method 2 - Comparison of original SE calculation and proposed SE methods, showing average results for all electrodes are shown for the indi- vidual participants: SE-Method 2 on moving windows for: (a) 2 Seconds, (b) 10 Seconds, (c) 20 Seconds, (d) 30 Seconds, (e) 40 Seconds, (f) 50 Sec- onds, (g) 60 Seconds, (h) 70 Seconds, (i) 80 Seconds, and (j) 90 Seconds. . . . . .	204
C.5	Slot 5 - Comparison of original SE calculation and proposed SE methods, showing average results for all electrodes are shown for the individual par- ticipants: SE-Method 3 on moving windows for: (a) 0.06 Seconds, (b) 0.12 Seconds, (c) 0.25 Seconds, (d) 0.55 Seconds, (e) 1 Seconds, (f) 2 Seconds (g) 3 Seconds, and (h) 4 Seconds. . . . .	205

D.1	Dataset 4- Embedding Dimension Result: Five participant's (Participants 3, 4, 5, 6 and 7) ED result from Dataset 4. . . . .	207
D.2	Dataset 5- Embedding Dimension Result: Ten participant's (Participants 1, 2, 5, 6, 7, 8, 9, 10, 11 and 12) ED result from Dataset 4. . . . .	208
D.3	Dataset 9 - Embedding Dimension Result: (a) Three Participant's (Participant 1, 2 and 5) ED result for HRV without medical condition, (b) Three Participant's (Participant 2, 4 and 5) ED result for HRV with medical condition Congestive Heart Failure (CHF). . . . .	209
D.4	Dataset 10- Embedding Dimension Result: Three participant's (Participants 1, 3 and 5 ) ED result from Dataset 10 showing 5 measured Electrodes, as shown in different colour). . . . .	210

# List of Tables

1.1	Words and Expressions . . . . .	7
1.2	Methodology . . . . .	7
2.1	Summary of Some of the Research work for Correlation Between EEG Signals. . . . .	32
2.2	Summary of Correlation Research on EEG and HRV. . . . .	41
2.3	Summary of Research on Well known Wavelet Transformation Methods for EEG and HRV. . . . .	41
3.1	Summary of Datasets and detail of their utilisation. . . . .	58
4.1	Complex multiplication required for DFT and FFT for different value of $N$ . . . . .	72
5.1	Probability of Sequence matches obtained at $m = 2$ . . . . .	92
5.2	Probability of Sequence matches obtained at $m = 3$ . . . . .	93
5.3	Point sequences at $m = 1; 2$ ; and $3$ along with the count of match obtained for $A^m(r)$ and $B^m(r)$ . Here, X represents that the point sequence was not considered for $B(m)$ . Columns $A(m)$ and $B(m)$ indicates count for the total number of matches obtained . . . . .	97
5.4	Pearson's Correlation Calculation. . . . .	98
5.5	Cross-correlation Calculation for corresponding lag . . . . .	105
6.1	Summary of Work in this Chapter . . . . .	115
7.1	Summary of the Research on the correlation of EEG and HRV. . . . .	125
7.2	Dataset 1- Method 1 - Averaged participants' correlation performance: <b>(a)</b> Level 0 (No WT), <b>(b)</b> Level 1 WT, <b>(c)</b> Level 2 WT, <b>(d)</b> Level 3 WT) . . . . .	138
7.3	Dataset 2- Method 1 - Averaged participants' correlation performance: <b>(a)</b> Level 0 (No WT), <b>(b)</b> Level 1 WT, <b>(c)</b> Level 2 WT, <b>(d)</b> Level 3 WT) . . . . .	139
7.4	Dataset 1- Method 2 - Averaged participants' correlation performance: <b>(a)</b> Level 0 (No WT), <b>(b)</b> Level 1 WT, <b>(c)</b> Level 2 WT, <b>(d)</b> Level 3 WT) . . . . .	140

7.5	Dataset 2- Method 2 - Averaged participants' correlation performance: (a) Level 0 (No WT), (b) Level 1 WT, (c) Level 2 WT, (d) Level 3 WT	141
7.6	Method 1 - Heat Map Results of Averaged participants' correlation performance: (a) Dataset 1, (b) Dataset 2. Colour coding is from Red to Dark Blue, Red=Strongest, Dark-Blue=Weakest)	142
7.7	Method 2 - Heat Map Results of Averaged participants' correlation performance: (a) Dataset 1, (b) Dataset 2. Colour coding is from Red to Dark Blue, Red=Strongest, Dark-Blue=Weakest)	143
8.1	Computation time for the SE calculation using the Original approach and the three efficient methods.	150
F.1	Dataset 1- Method 1 - Averaged participants' correlation performance: (a) Level 0 (No WT), (b) Level 1 WT, (c) Level 2 WT, (d) Level 3 WT	270
F.2	Dataset 2- Method 1 - Averaged participants' correlation performance: (a) Level 0 (No WT), (b) Level 1 WT, (c) Level 2 WT, (d) Level 3 WT	271
F.3	Dataset 1- Method 2 - Averaged participants' correlation performance: (a) Level 0 (No WT), (b) Level 1 WT, (c) Level 2 WT, (d) Level 3 WT	272
F.4	Dataset 2- Method 2 - Averaged participants' correlation performance: (a) Level 0 (No WT), (b) Level 1 WT, (c) Level 2 WT, (d) Level 3 WT	273

# Chapter 1

## Introduction

In recent years, time series data analysis has been widely used to study complex behaviours and different structures of biological systems (Al-Angari and Sahakian, 2007). Time series data Electroencephalogram (EEG) and Electrocardiogram (ECG) are recordings of biological systems and are commonly used for checking human medical conditions. Time series analysis has been proven to be a robust approach for the assessment of different biological systems because it can unveil hidden patterns related to underlying physiological mechanisms (Alcaraz and Rieta, 2010) (Richman and Moorman, 2000).

Electroencephalogram (EEG) is the electric potentials on the surface of the human scalp, which can reflect different brain activities at different areas of the brain. The nervous system of the human brain contains about 100 billion neurons (nerve cells), and each neuron communicates with approximately 7000 others (Quiroga, 1998). While you think, dream, see, and sense, your brain is constantly active: absorbing information, compacting and re-connecting information, and integrating everything into a consistent experience. EEG signals change with people's behaviours (Bob et al., 2010) (Jeong et al., 2015) and medical conditions (Na et al., 2002). Analysing these brain activities recorded through EEG signals should help to understand the functionalities of the brain. This could be beneficial for the people working in neurology and medical area.

Electrocardiogram (ECG) is a recording of the heart's rhythm and electrical activity. ECG can reflect a variety of intertwined and complex chemical, electrical, and mechanical processes present in the heart (Burch and DePasquale, 1990). It conveys a great deal of valuable diagnostic information describing not only the function of the heart but also other systems, such as blood circulation and nervous systems. Analysing ECG could help check whether or not a heart is functioning well, for example, to identify people at risk for a cardiovascular disease (Cooper et al., 2007).

Analysing EEG and ECG data could give a better understanding of the phenomenon of interest. However, the issues with the data, such as noise and artefacts, may prevent this.

Therefore, data preprocessing could be beneficial for data analysis performance because it sustains important information, and removes noise and artefacts. For given EEG and ECG data, appropriate data analysis techniques should be utilised in order to achieve good results. The following two paragraphs will briefly describe what those two approaches could achieve, respectively.

Data preprocessing can be utilised to eliminate noise and artefacts in order to make useful components of data eminent(Beckmann et al., 2005). This is because the power of noise and artefacts is much stronger than real signals in EEG(Proakis et al., 1992). Hence, without data preprocessing, the analysis of data might be misleading. In most applications, there might be a need for more than one form of data preprocessing techniques. Identifying appropriate types of data preprocessing might be a crucial task for a given data EEG.

Data analysis is the basis for investigations in many fields, from science to engineering, and from management to process control. Data analysis comprises a wide variety of methods that allow extracting various characteristics of data (Kantz and Schreiber, 2004). EEG and ECG data which contain nonstationary and unpredictable signals are generated by a highly complex system. Data analysis could quantify distinct states of the brain and heart, which helps diagnose medical conditions (da Silva et al., 1994), (Ku'nel and Dolce, 1975), (Silva, 1987).

## **1.1 Research Motivation**

The human brain is very complex to be understood because each of the billions of neurons can communicate with many others. A huge amount of research has been carried out on EEG analysis to understand human brain activities. It has been suggested by (Niedermeyer and da Silva, 2005) that various characteristics of EEG signals are representative of distinct states of brain activities. Researchers (Na et al., 2002), (Li et al., 2013) have demonstrated that brain activities might be similar within the same (local) brain region, but different from other regions (globally). Those researches imply that EEG might change gradually from one region to another. But, to my knowledge, I have not seen anybody showing how EEG signal (measured through electrode) changes with electrode location. This motivated me to think whether or not there is any variation in correlation value of EEG signals with the distance between them. Talking about a distance between electrodes, there could be three possible definitions: a) straight-line distance, b) geodesic distance and c) travelled distance between neurons. Therefore, investigating whether EEG changes with distance or not, and if so with which distance definition will form one of my research questions.

EEG and ECG time series signals represent complex dynamic behaviours of biological systems and might be related to each other because both of them are generated from one co-ordinately working body. Investigating this relationship is of interest because it may reveal information about the correlation between EEG and ECG signals. In recent years, the correlation between EEG and ECG during sleep (Miyashita et al., 2003), (Yang et al., 2002), (Ako et al., 2003), (Jurysta et al., 2003), (Abdullah et al., 2010), (Chua et al., 2012), (Berg et al., 2005) and during meditation (Takahashi et al., 2005), (Kim et al., 2013) have been analysed. The results from these studies have shown a correlation between EEG and ECG, but each of the studies only focused on part(s) of the brain, rather than the whole brain. Therefore, the results are not comprehensive. (Na et al., 2002) indicates that EEG from the left side of the brain is correlated with ECG, and (Bob et al., 2010) demonstrated the correlation was from the right side of the brain. These findings have made me curious to find out if there is any particular area within the brain having a stronger correlation between EEG and ECG.

Time series could be very complex because it might contain regular signals, as well as disordered and random signals of a dynamic system. In practice, it is difficult to predict the complex time series. Analysing complexity of time series might provide insights into time series data (Abásolo et al., 2006), (Ramdani et al., 2009) so that the state of the system could be predicated. For example, the complexity of the heart and brain data sometimes can predict heart attack and mental medical conditions (Kantz et al., 2012). But calculating complexity might be very time consuming, particularly while the number of data points is large. I, therefore, would like to explore more on this.

A time series is a sequence of one visible signal taken in time from a dynamic system. There might be many underlying variables determining the visible signal. In practice, it can be difficult to know what variables determine the behaviour of the dynamic system. For example, stock data could be affected by many interacting variables, such as economic data, exchange rates and so on. Time series signal ECG might be the results of the interaction of many underlying variables (Chun-Hua and Xin-Bao, 2004) and finding those variables contributing to ECG might be important to understand the behaviour of the dynamic system. It will be interesting to find out whether or not the actual underlying variables determining ECG could be unveiled.

## 1.2 Research Questions

1. How does the correlation between EEG signals measured through electrodes vary with the physical distance (straight-line distance) between them?

2. Is there any particular area within the brain having a stronger correlation between EEG and ECG?
3. Can the efficiency (computational time) for the calculation method for time series complexity be improved in the context of calculating a correlation between EEG and ECG?
4. Could underlying variables determining ECG be unveiled?

## 1.3 Research Contribution

My contributions are:

- **Discovered that the correlation between EEG signals measured through electrodes is linearly dependent on the physical (straight-line distance) distance between them.**

Although other research has looked into the correlation between EEGs using a limited number of electrodes and a limited number of combinations of electrode pairs, no research has investigated the correlation between EEG signals and distance between electrodes. My research filled up this gap by using a full range of electrodes and all possible combinations of electrode pairs.

- **Demonstrated that EEG at the front area of the brain has a stronger correlation with Heart rate variability (HRV) than the other area.**

HRV is the physiological phenomenon of variation in time interval between heartbeats, which is retrieved from ECG. Some research has investigated correlation between EEG and HRV limited to certain brain areas and demonstrated the existence of correlation between EEG and HRV. But no research has indicated whether or not the correlation changes with brain area. My research covers this gap by conducting a thorough investigation of all electrodes on the human scalp.

- **Designed a new calculation method for time series complexity which could improve computational time significantly in the context of calculating a correlation between EEG and HRV.**

Recent research has proposed new calculation methods for time series complexity, but their focus was on improving accuracy. No one has attempted to reduce the computational time of it. My application required a fast calculation method, therefore, I

designed three calculation methods for time series complexity. Experimental results show that one of the methods can improve efficiency dramatically.

- **Two variables determining ECG (HRV) were unveiled.**

Time series analysis method has been utilised to study complex systems that appear ubiquitous in nature, but limited to certain dynamic systems (e.g. analysing variables affecting stock values). No literature has investigated the nature of the underlying dynamic system of HRV. My research highlights this matter by analysing actual variables determining HRV. The results strongly suggest that the Autonomic Nervous System driving the heart is a two-dimensional dynamic system.

## 1.4 Publications on This Thesis

During my PhD study, I have published five conference papers: three papers are published in the conference journal, and two in conference proceedings. For all details about the five papers, please see Appendix E. The lists of all five publications are shown below:

1. The Correlation between EEG Signals Measured at Different Positions on Scalp Varying with Distance by Ronakben Bhavsar, Yi Sun, Na Helian, Neil Davey, David Mayor, and Tony Steffert - presented at 8th International Conference of Biological Inspired Cognitive Architectures (BICA), 1-6 August 2017, Moscow, Russia (I have received an *Outstanding Research Award* by BICA Society).
2. An Investigation of How Wavelet Transform Can Affect the Correlation Performance of Biomedical Signals- The Correlation of EEG and HRV Frequency Bands in the Frontal Lobe of the Brain by Ronakben Bhavsar, Neil Davey, Yi Sun, and Na Helian - presented at the 11th Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2018), 19-22 January 2018, Funchal, Portugal
3. Efficient Methods for Calculating Sample Entropy in Time Series Data Analysis by Ronakben Bhavsar, Na Helian, Neil Davey, Yi Sun, Tony Steffert, and David Mayor - presented at 8th International Conference of Biological Inspired Cognitive Architectures (BICA), 22-26 August 2018, Prague, Czech Republic.
4. Time Series Analysis using Embedding Dimension on Heart Rate Variability by Ronakben Bhavsar, Neil Davey, Na Helian, Yi Sun, Tony Steffert, and David Mayor - presented at 8th International Conference of Biological Inspired Cognitive Architectures (BICA), 22-26 August 2018, Prague, Czech Republic.

5. The Correlation between EEG Signals Measured at Different Positions on Scalp Varying with Distance for Datasets With and Without Medical Conditions by Ronakben Bhavsar, Yi Sun, Na Helian, Neil Davey, David Mayor and Tony Steffert - presented at Engineering and Computer Science Conference, 17th April 2019, University of Hertfordshire, UK (I have received an *Highest Recommended Research* by Engineering and Computer Science Society, UH).

## 1.5 Structure of this Thesis

The structure of the rest of the thesis is as follows:

**Chapter 2** gives the concept of time series data derived from nonlinear systems. This is because EEG and ECG are time series data and they form the datasets of this research. Then, a literature review about the analysis of EEG, ECG and their correlation is presented.

**Chapter 3** provides the information of the datasets utilised in this research: the first-hand dataset (which I have recorded in person) and second-hand datasets (obtained from the Internet and that were given to me).

**Chapter 4** explains the importance of time series data preprocessing. Then, commonly used preprocessing methods for EEG and ECG data are described.

**Chapter 5** describes the nonlinear analysis methods used for time series data analysis. The explanation of each method is provided in this chapter.

**Chapter 6** is the first result chapter of the thesis, in which the correlation between EEG signals measured through electrodes and the distance between electrode is presented using a time series data analysis method.

**Chapter 7** is the second result chapter of the thesis, in which the correlation between EEG and HRV is shown utilising various time series data preprocessing and time series analysis methods.

**Chapter 8** is the third result chapter of the thesis, in which designs for calculation methods of time series complexity are given. The aim is to improve efficiency in the context of calculating a correlation between EEG and HRV.

**Chapter 9** is the fourth (final) result chapter, in which the actual underlying variables that determine HRV are unveiled using a time series analysis method.

**Chapter 10** summarises all the findings in this thesis. Besides, some future work is indicated.

## 1.6 Terminology Abbreviations

The terminology and expressions used frequently in the time series data studies and this thesis are listed in Table 1.1, and computational methods used in this thesis are listed in Table 1.2.

Table 1.1: Words and Expressions

<b>Index</b>	<b>Terminology</b>	<b>Abbreviation</b>
1	Electroencephalogram	EEG
2	Electrocardiogram	ECG
3	Heart Rate Variability	HRV
4	Heart Rate	HR
5	Inter Beat Interval	IBI
6	Autonomic Nervous System	ANS
7	Peripheral Nervous System	PNS
8	Very Low Frequency	VLF
9	Low Frequency	LF
10	High Frequency	HF
11	Time Domain	TD
12	Frequency Domain	FD
13	Transcutaneous Electro Acupuncture Stimulation	TEAS
14	Electro Acupuncture	EA
15	Manual Acupuncture	MA

Table 1.2: Methodology

<b>Index</b>	<b>Terminology</b>	<b>Abbreviation</b>
1	Approximate Entropy	AE
2	Sample Entropy	SE
3	Pearson Correlation Coefficient	PCC
4	Cross-Correlation	CC
5	Embedding Dimension	ED
6	False Nearest Neighbours	FNN
7	Pearson Correlation Coefficient	PCC
8	Fast Furious Transform	FFT
9	Independent Component Analysis	ICA
10	Wavelet Transform	WT
11	Discreet Wavelet Transform	DWT
12	Continuous Wavelet Transform	CWT

# Chapter 2

## Literature Review: Assessing Variability of Time Series Data EEG and ECG

### 2.1 Introduction

This chapter sets the stage to the relevance of time series analysis to the scientific study of complex systems. The key concepts that have motivated the development and application of nonlinear methods, including the notion of a dynamical system, concepts such as chaos, and nonlinear dynamics. Descriptions of nonlinear time series data analysis on EEG and ECG is provided, followed by their relative importance in time series analysis.

#### 2.1.1 Dynamical, Nonlinear and Chaos System

A dynamical system is a system which changes in time - what changes is the state of the system. Studies of dynamical systems have led to the understanding of important concepts for biologists (Mpitsos and Soinila, 1992). The way a set of functions can describe the state of a dynamical system change over time citepjames1999primer. We may define these functions either in continuous time or discrete time by differential equations. Within the realm of dynamical systems, many will exhibit nonlinear characteristics. Among those nonlinear systems, exists a subset of chaotic systems, as shown in Figure 2.1. Chaos (a state of disorder and confusion) often thought in terms of noise within a system that is unpredictable. However, its technical meaning is random and unpredicted phenomenon. Chaotic behaviour exists in many simple systems such as the movement of a ferromagnetic beam buckled between two magnets under the effect of sinusoidal oscillations (Moon and Holmes, 1979), a double pendulum (Richter and Scholz, 1984), or a dripping tap (Shaw, 1984). Dynamical systems have been used in a variety of applications, including human motion (action) modelling (Al-Angari and Sahakian, 2007), Bissacco et al. (2001), (Bregler, 1997), (Wang et al., 2007) and dynamic textures (Chan and Vasconcelos, 2007),

(Doretto et al., 2003), (Ghanem and Ahuja, 2007), (Liu et al., 2006), (Yuan et al., 2004), (Wang and Zhu, 2002), (Schödl et al., 2000). Most of these approaches utilised linear dynamical systems, while others use nonlinear dynamical systems. Time series (detail about time series data is available in section 2.1.2) modelling and prediction has been an active area of research because of the wide variety of applications in the financial market, weather, biology, etc. More sophisticated approaches rely on nonlinear modelling (Casdagli, 1989) and state space projection of the time series (Ralaivola et al., 2004). It is necessary to understand the space that it occupies to understand the underlying properties of a dynamical system. For example, Electrocardiogram (ECG) data, might result from underlying variables from the Autonomous Nervous System (ANS).

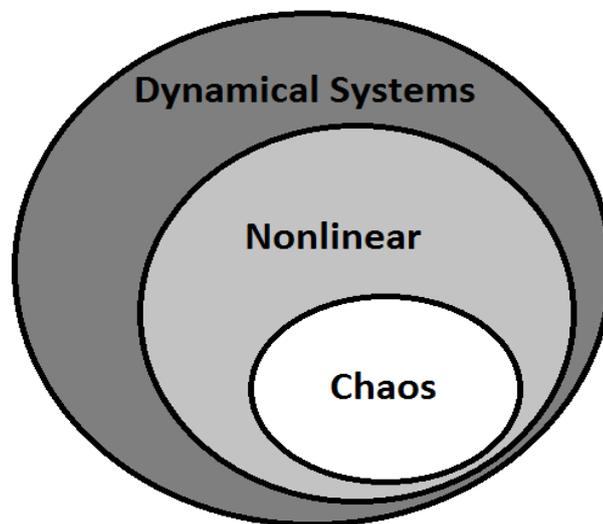


Figure 2.1: Not all nonlinear systems are chaotic, but all chaotic systems are nonlinear. Nonlinear systems are in turn a particular form of dynamical systems (Redrawn from (James and Walker, 1999)).

It is important to know that not all nonlinear dynamical systems are chaotic, but all chaotic systems are nonlinear (Hilborn et al., 2000). A nonlinear system is defined as a system that does not have a linear origin. This includes a system that may contain nonlinearity, but the underlying dynamics are linear. In this case, the presence of nonlinearity is caused by some measurement distortion, but it is originally generated by a linear stochastic process. There are two major approaches to identify the evidence of nonlinearity in a time series. The first approach involves the direct application of nonlinear methods (Kaplan and Glass, 1995), (Mitra et al., 1997), while the second approach involves the application of surrogate methods (Stergiou, 2004), (Breakspear and Terry, 2002), (Dingwell and Cusumano, 2000), (Paluš, 1996). Methods commonly used for the first approach include

the application of the correlation dimension or the largest Lyapunov exponent. The correlation dimension is a measure of self-similarity of a time series, while the largest Lyapunov exponent quantifies the exponential rate of divergence of nearby trajectories in the state space. The use of other nonlinear methods besides correlation dimension and the largest Lyapunov exponent are also limited in terms of detecting nonlinearity in a time series, since the probability distributions of those methods of time series with finite data length are unknown (Paluš, 1995), (Pompe, 1993), (Prichard and Theiler, 1995). Thus, applications of these nonlinear measures alone in detecting nonlinearity, possibly chaotic behaviour in the system have been shown to be difficult (Miller et al., 2006), (Schreiber and Schmitz, 2000), (Theiler and Rapp, 1996). Methods commonly used for the second approach include surrogate methods. Surrogate methods were originally developed to prevent misdiagnoses of random stochastic processes from being characterised as chaotic dynamical processes or vice versa (Stergiou, 2004), (Theiler et al., 1992), (Theiler and Rapp, 1996). They take a form of hypothesis testing to determine whether a given time series is consistent with a specific null hypothesis. Surrogate method has been applied to identify the evidence of nonlinearity in many biological systems such as postural control, ECG, EEG, gait mechanics, and so forth (Breakspear and Terry, 2002), (Stergiou, 2004), (Paluš, 1996), (Miller et al., 2006), (Acharya et al., 2005), (Buzzi et al., 2003), (Chang et al., 1994), (Cignetti et al., 2009), (Collins and De Luca, 1995), (Costa et al., 2014), (Ehlers et al., 1998), (Govindan et al., 1998), (Ivanov et al., 1996), (Janjarasjitt et al., 2008), (Kunhimangalam et al., 2008), (Ladislao and Fioretti, 2007), (Martinerie et al., 1998), (Nurujjaman et al., 2009), (Porta et al., 2006), (Rieke et al., 2003), (Rombouts et al., 1995), (Stam et al., 1997), (Zhao et al., 2008). Applying surrogate methods to nonstationary time series can lead to problems regarding the proper interpretation of results (Breakspear and Terry, 2002), (Paluš, 1996). Breakspear and Terry noted this problem in their study of electroencephalographic (EEG) data (Breakspear and Terry, 2002). Citeppeng1995quantification in the analysis of heart rate variability also highlighted the problem of nonstationary. Specifically, nonstationary makes it difficult to determine whether the structure of the time series results from the dynamics of the system or from changes in the external environment.

A chaotic system is known to be nonlinear. The theory of chaos has been applied to many fields of biological and nonbiological analysis. It has been used in systems ranging from psychology (Ayers, 1997) hydrology (Sivakumar, 2000) and to analyse the financial markets (Peters et al., 1994). When a dynamical system displays sensitivity to its initial conditions, which leads to irregularity, it can be termed chaotic (Kaplan and Glass, 2012). Such a system while appearing irregular is actually deterministic (Kaplan and Glass, 2012); however, it is impossible to make long-term predictions for such a system. (Kaplan and

Glass, 2012) add further to the definition of chaos, noting the same state is never repeated and that it is bounded. Representation of system behaviour in phase space provides a powerful basis for both visualizing and quantifying the dynamics of both nonlinear and chaotic systems (Figure 2.2). The phase space provides the underlying variables of the system. Figure 2.2 shows 3-dimensional representation of periodic (Figure 2.2a), chaotic (Figure 2.2b) and random signals (Figure 2.2c). The concept of a phase space representation rather than a time or frequency domain approach is the hallmark of nonlinear dynamical time series analysis (Kantz and Schreiber, 2004). Understanding if a system is chaotic may provide important information concerning whether the system is deterministic and the feasibility of making a longer-term prediction about future states of the system.

For nonlinear time series EEG and ECG analysis, my focus is on characterising nonlinear dynamics by applying nonlinear tools. Nonlinearity is considered as one of the key features of EEG and ECG time series that exhibit chaos, which has been shown to have a potential link with overall health of the biological system (Amato, 1992), (Buchman et al., 2001), (Cavanaugh et al., 2009), (Garfinkel et al., 1992), (Goldstein et al., 1998), (Orsucci, 2006), (Slutzky et al., 2001), (Wagner et al., 1996). Therefore, to detect chaos in EEG and ECG time series, identifying nonlinearity in the system is essential. In recent years, several methods have been proposed to compute dynamical parameters from EEG and ECG time series. Such parameters are extracted by using the information dimension, entropy, Lyapunov exponents, correlation dimensions, embedding dimensions, recurrence plot, and so on. In all cases, it is assumed that the EEG and ECG time series are obtained from an autonomous dynamical system. It is also assumed that EEG and ECG times series are much longer than the characteristics times of the dynamical system.

### **2.1.2 Time Series Data**

Time series could be a list of numbers, assumed to measure some process sequentially in time (Stergiou, 2004). Mathematicians have a more formal definition, a set or a sequence of observations, with each one recorded at specific times, or at least sequentially (Brockwell et al., 2002), (Box et al., 2015). Time series could be created from multiple sources for research to understand various behaviours. For example, social scientists could collect graduation rates, physiologists record heart rates and brain waves, economists study consumer spending, and climatologists examine weather patterns. Time series inherently possess dependence between adjacent observations. This dependence is of interest because it reveals information about the source producing the behaviour (Stergiou, 2016). Time series analysis is essential for assessing variability because time series analysis reveals how the system evolves.

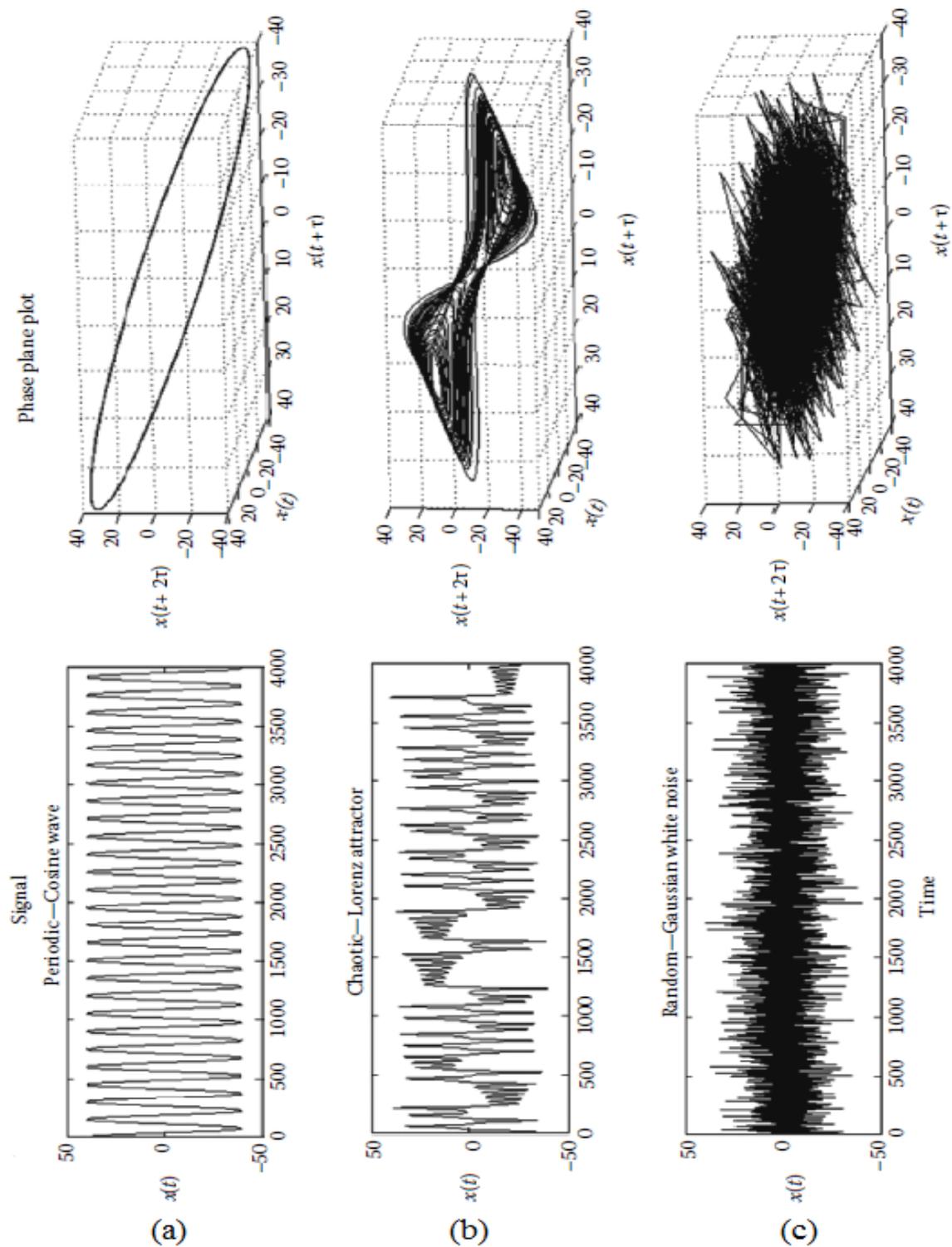


Figure 2.2: Phase space plots provide a means to geometrically show chaotic behaviour which may not be readily apparent within a time series James and Walker (1999). Left side of the figure shows the signal, and right side is the phase plane plot (Embedding Dimension), where (a) Periodic signal, (b) Chaotic signal, and (c) Random signal with white noise.

The term ‘continuous data’ or ‘continuous time series’ is described in various ways. According to Brockwell and Davis, continuous series are those in which observations are recorded continuously for a specific amount of time (Brockwell et al., 2002). A continuous measure is one that evolves continuously. The outcome is, thus, a function of time, for example,  $x = f(t)$ . Consider the children’s story character Pinocchio; his nose grows when he tells lies. We can create a function to describe the length of Pinocchio’s nose. Let us say that his nose grows 0.1 cm, multiplied by the time in minutes, raised to the third power:  $L = 0.1 \times t^3$  (James and Walker, 1999). When plotted, this appears, as shown in Figure 2.3a. Discrete time series is also described as a sampling of continuous time series at certain intervals (Box et al., 2015). Discrete time series is a series in which observations made in discrete sets, such as a specific, fixed time intervals (Brockwell et al., 2002). If the length of Pinocchio’s nose calculated (and grows) at discrete intervals, say at the end of each minute, then the length grows, as shown in Figure 2.3b.

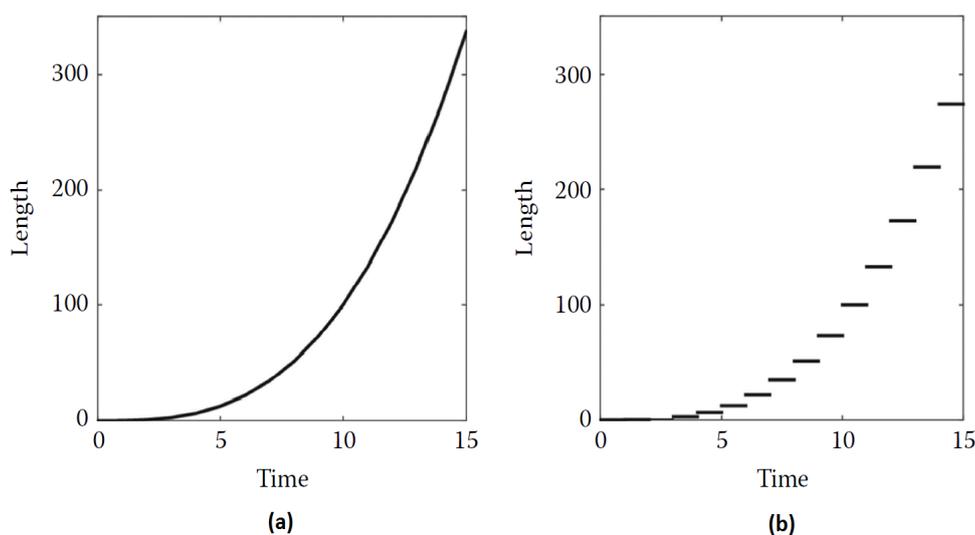


Figure 2.3: The length of Pinocchio’s nose: (a) continuous growth and (b) growth using discrete measurements.

Time series analysis is useful in many applications whenever you are trying to describe the change of a system with time; that is, the dynamics of the system. For example, a psychologist has subjects filling out a survey once per day for 4 weeks, scoring the answers, and creating a time series that has 28 points (Fredrickson and Losada, 2005). Astronomers have observed sunspot activity annually since 1700 by counting sunspots visible on the sun - creating a time series of length 314 (James and Walker, 1999). Physiologists analysed blood samples for luteinising hormone every 10 min for 4 days - creating a time series 577 points long (Liou et al., 2014). Neurologists analyse brain signals for changes

in brain activity while participants are under acupuncture treatments (see Figure 2.4), or watching a documentary film or short videos (Koelstra et al., 2012) (see Figure 2.5). Scientists, engineers, and clinicians from many disciplines use time series analysis to provide an understanding of the dynamics of whatever system they are studying. While all these different applications may seem entirely unrelated, the methodologies of the analysis overlap considerably.

The time series which are analysed using techniques described in my research are discrete time series with observations made at equal intervals (Brockwell et al., 2002), (Box et al., 2015). Although time series that look continuous, in Figure 2.4 and 2.5, they are sampled at specific time intervals.

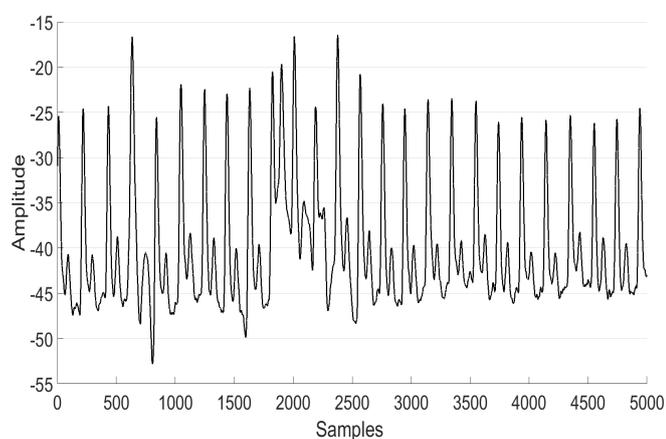


Figure 2.4: Plot of 20 seconds ECG data from a participant under acupuncture treatment. The data were gathered using sampling frequency of 256Hz.

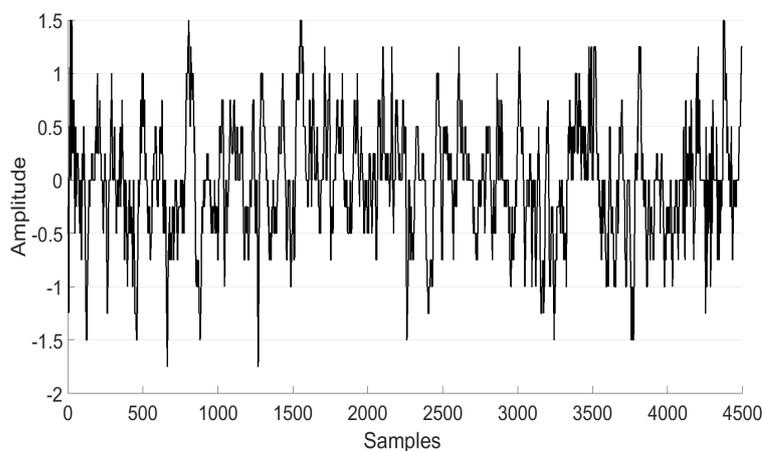


Figure 2.5: Plot of 10 seconds EEG data from a participant watching a short documentary film. The data were gathered using sampling frequency of 500Hz.

Linear time series analysis methods assume observations are independent of history (Cryer and Kellet, 1991). However, continuous observations of nonlinear time series, are rarely independent of each other (Abarbanel et al., 1993). This chapter acknowledges this and approaches EEG and ECG from this perspective that seems mostly lost in the current literature. Therefore, the essential features discussed here are entered on nonlinear algorithmic methods to investigate time series data EEG and ECG. However, regardless of the investigative approach, it is always essential to consider several fundamental issues of time series analysis. These include a length of the time series, sampling frequency, noise, filtering and smoothing, and stationary. This chapter details what makes up time series data such as Electroencephalogram (EEG) and Electrocardiogram (ECG), and describes specific considerations that should be kept in mind when working with these data.

### **2.1.2.1 Length of the Time Series**

The length of the time series could be seen as a major limitation for the utilisation of a certain analyses of time series data. In particular, for nonlinear analysis, the mathematicians who have derived the formulas suggest that a certain number of data points are critical for performing the analysis. The problem is not that the nonlinear calculation with the fewer data points cannot be done, but the problem is whether the answer received from using a shorter time series is really an accurate characterisation of the dynamics of the system (Warner, 1998).

As an example, consider the Electrocardiogram (ECG) data. It takes 1 minute to gather heart rate of an individual from ECG data. If you only record 30 seconds ECG, you could not discover what would be the heart rate of an individual. If you collect 2 minutes of ECG, you will observe heart rate per minute from ECG, but you would not know that if the same pattern will be repeated. Therefore, The longer the ECG we have, the better our ability to characterise the dynamics (Shaffer and Ginsberg, 2017). So how long is long enough for the time series? This depends on which analytical technique you want to use — some require more data than others.

### **2.1.2.2 Sampling Frequency**

The sampling frequency is a critical consideration when dealing with time series data. It is a measure of how often you gain a data sample, and thus, sampling frequency multiplied by the total time that you sampled, gives the number of data points in time series (Scargle, 1982). The sampling frequency needs to be high enough to capture the dynamics of the quick changes in the system. So going up to a sampling frequency of about five times faster frequency is a good rule of thumb for periodic data (Theiler et al., 1992).

To show how different frequencies contribute to a signal, consider: a sine wave of 5 Hz, a sine wave of 20 Hz, a sine wave of 50 Hz, and the sum of these three wave functions. The peak in the power spectrum corresponds to the frequency of the wave function in the time series (Figure 2.6a through c), and the time series that is a sum of three wave functions has three peaks, corresponding to the three frequencies which were added together (Figure 2.6d). The reason for wanting to divide data into a sine function is because then you can see which frequencies are contributing the most to the data just by examining the peak positions. If the peak corresponding to 50 Hz is very high, then you likely have a 50 Hz component in the signal. If the 50 Hz component of your signal is the highest frequency that is significant, then we need a sampling frequency of at least 100 Hz ( $2 \times 50\text{Hz} = 100\text{Hz}$ ) to see it, but something more like 250 Hz ( $5 \times 50\text{Hz} = 250\text{Hz}$ ) would be best to define it better. For Electroencephalogram (EEG) data, the highest frequencies that occur at gamma rhythm with 50 Hz. Thus, a 250 Hz ( $5 \times 50$ ) sampling rate should be satisfactory; however, in reality, neurologists usually sample at 5 to 10 times the highest frequency in the signal (that is between 250Hz-500Hz for EEG). If data are under-sampled, it does not capture the entire signal. If data are over-sampled, it could introduce more measurement noise (Antoniol and Tonella, 1997).

### **2.1.2.3 Noise**

In any experimental measurement, there are always concerns about measurement error or contamination of what you are trying to measure with other information that you are not trying to measure. The problem with noise in the nonlinear dynamics analysis is that you are trying to detect the dynamics of the system of interest, and a signal with unknown dynamics may contaminate the experimental data. As with any experiment, anything that can optimise the signal-to-noise ratio is beneficial. The example of a power spectrum of a time series without noise and with noise is shown in Figure 2.7a, and 2.7c, respectively.

Another example would be for EEG, for example, If I want to record EEG signals with eyes open, what needs to be recorded, is the brain's electrical activity. But during the recording process, noise from a known source (eye blinks, eye movements, muscle movements, etc.) or an unknown source (cable movement, equipment, etc.) can be found with EEG.

### **2.1.2.4 Filtering/Smoothing**

Experimental noise is a problem in time series analysis. As highlighted in the previous subsection, interference from biological signals, movement artefacts or high-frequency noise contaminate the dynamics of our signal of interest. One approach commonly used to deal

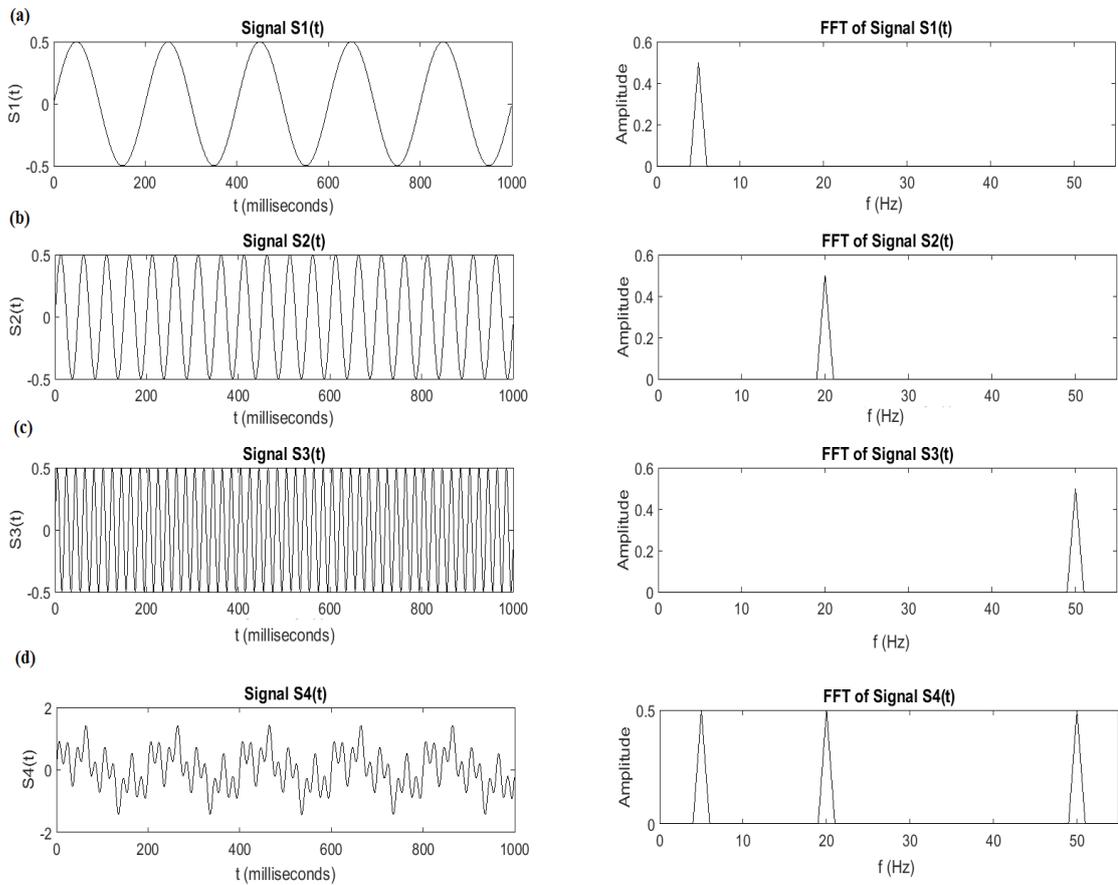


Figure 2.6: Time series data (left) and corresponding power spectra (right) for a (a) sine wave of 5 Hz, (b) a sine wave of 20 Hz, (c) a sine wave of 50 Hz, and (d) the sum of these three wave functions.

with these issues in time series analysis is filtering/smoothing the data (Woltring, 1985), (Busby and Trujillo, 1988), (Vaughan et al., 1999), (Giakas, 2004). The more filtering that is performed, or the more frequencies that are removed, the smoother the signal will be.

There are different methods that can be implemented to filter the data correctly. Two of the most common implementations of this technique are the Butterworth filter, the critically damped filter, and the Jackson filter (Smith, 2002). The algorithms will use the selected cut-off and contain different bands depending on what type of data is being smoothed. One can select whether to remove frequencies above a certain cutoff frequency, called a “low-pass” filter, because it allows lower frequencies to pass through the filter. A “high-pass” filter would block low frequencies and allow high frequencies to pass through. A “band-pass” or “notch-pass” filter passes through frequencies in an intermediate range while rejecting higher and lower frequencies. Because random noise is a high-frequency component of the measured signal, a low-pass filter is used to remove it. We must carefully select the cutoff

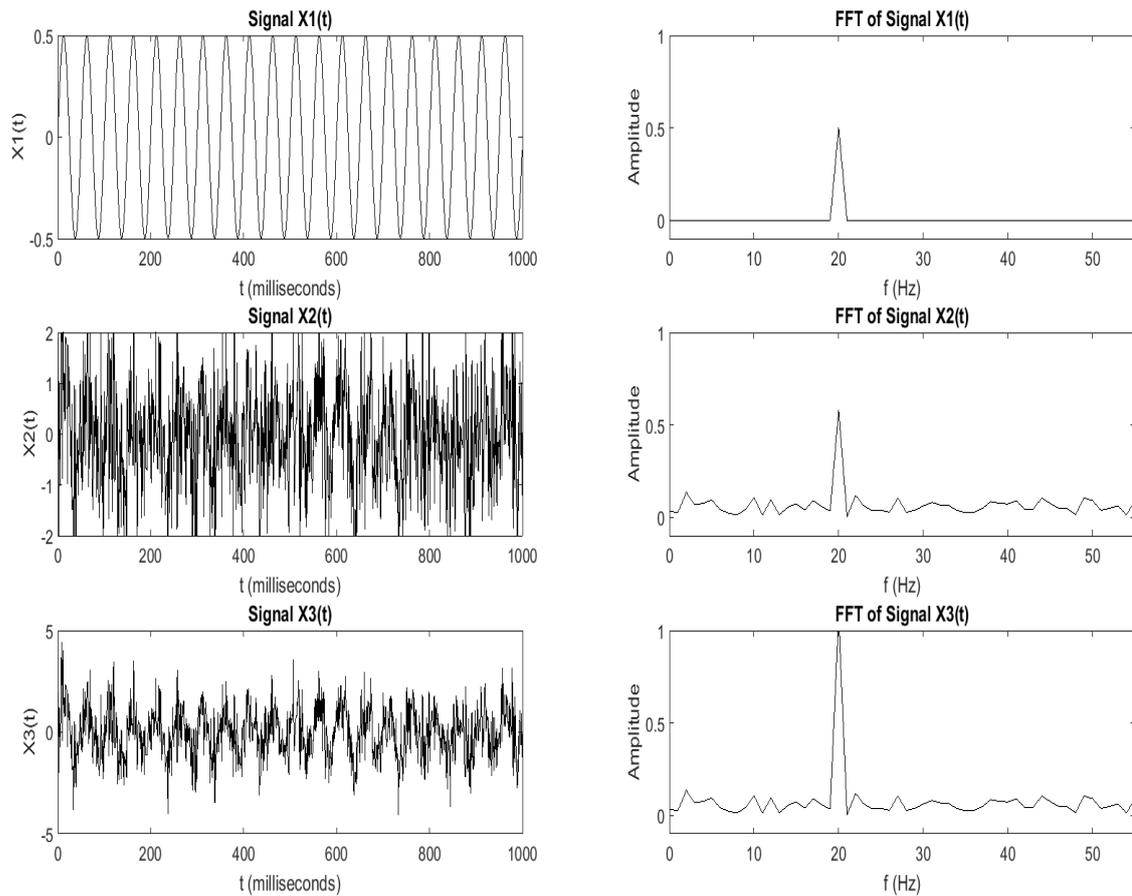


Figure 2.7: Time series data (left) and corresponding power spectra (right) for a (a) sine wave of 20 Hz, (b) white noise, and (c) the sum of the sine wave and white noise.

frequency to remove noise without removing the signal of interest.

The time series and its power spectrum of filtered and unfiltered time series are shown in Figure 2.8. Figure 2.8b shows that a cutoff frequency above the three features of interest leaves them intact, but if the cutoff frequency is too low (Figure 2.8c and d), the peaks of interest are removed. It is important to select a cutoff frequency that will preserve most of the data of interest.

Filtering data is a very common method of data manipulation for many types of linear analyses. The problem is that most of the filtering techniques are based on statistically preserving the linear features of the data. There is no reason to believe that the nonlinear dynamics of the time series would still be intact after filtering, and in fact, one would expect that these methods would be counterproductive for nonlinear analysis Rapp (1994), Theiler et al. (1992).

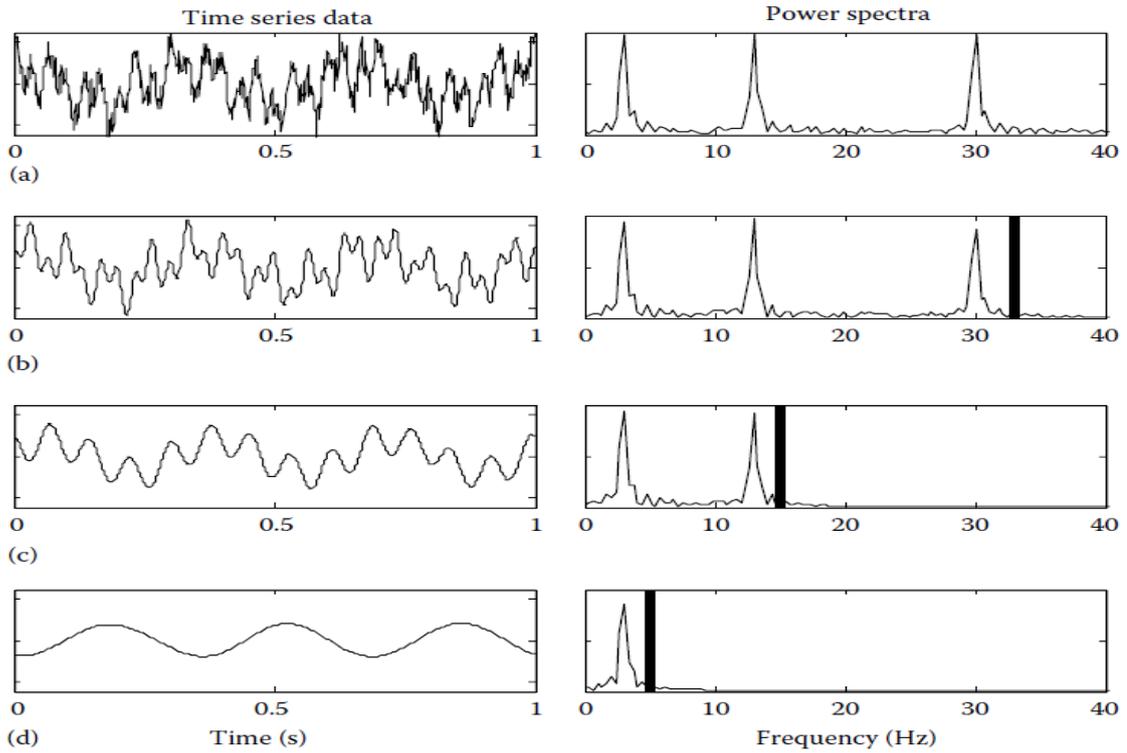


Figure 2.8: Time series left and power spectrum right: (a) unfiltered, (b) filtered using a cutoff frequency of 33 Hz, (c) filtered using a cutoff frequency of 15 Hz, and (d) filtered using a cutoff frequency of 5 Hz. The filter used was a “brick wall” filter (very high-order Butterworth filter), and cutoff frequency is indicated by vertical bar (James and Walker, 1999).

### 2.1.2.5 Stationarity

The concept of stationary is the requirement that there is a statistical similarity of successive parts of a time series. It shows that the mean and the variance should not change as a function of time in the time series (James and Walker, 1999). The stationary issue is perplexing because the exact nature of the stationarity required for nonlinear analysis is not clear. The stationarity requirement for the experimental data comes from an assumption commonly made by mathematicians in the derivation of the mathematical algorithms that are used for the nonlinear analysis. For example, Pincus assumes stationarity in the algorithm’s derivation for approximate entropy (Pincus, 1991), and (Wolf et al., 1985). The stationarity assumed in the algorithm’s derivation for maximum Lyapunov exponent (Wolf et al., 1985). Some authors have attempted to quantify nonstationarity of a time series as a useful measure (Rieke et al., 2003), (Cao et al., 2004), (Mäkinen et al., 2005), (Gourévitch and Eggermont, 2007), (Tong et al., 2007). For example, some debate about global climate change centres on the stationarity, or lack thereof, of climate-related variables (Kärner,

2002). Nonstationary may be inherent to biological systems and should be embraced and studied, and not be considered as a limitation.

A commonly used technique to remove nonstationary of time series data is to differentiate the data (Chatfield, 2016). Differentiating is subtracting values between two data points to create a new point in a differenced series. Another technique that can address nonstationary is de-trending. Briefly, de-trending occurs before the application of a non-linear algorithm, usually as a first step in the calculation process. In Figure 2.9a, there is a stationary white noise time series, in Figure 2.9b, the time series is not stationary because the mean is different in the first half compared to the second half, and in Figure 2.9c, the time series is not stationary because the variance is different in the first half compared to second half.

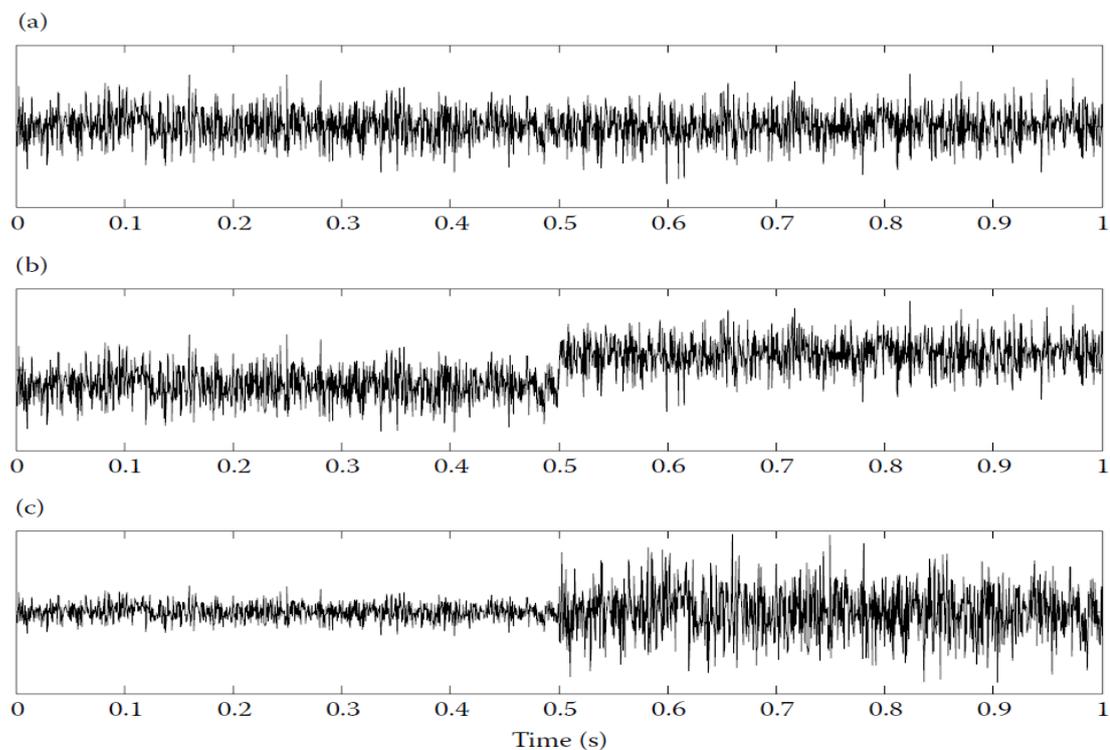


Figure 2.9: White noise time series data showing (a) stationarity, (b) nonstationarity due to change in mean, and (c) nonstationarity due to change in variance (James and Walker, 1999).

### 2.1.3 Time Series Data Analysis Methods

There are many methods being considered for time series data analysis depending on the discipline of the time series data has been produced. There are regression methods, for

example, for forecasting prediction (Hurvich and Tsai, 1989), which I have not considered in my work. I describe some of the popular and most suitable methods used for time series data analysis in this section.

The concept of entropy was first developed in classical thermodynamics, where it grew out of the work by (Carnot, 1824) on steam engines, to develop an understanding of the limits of a mechanical work such engines could produce that. The term entropy was introduced to the vocabulary of classical thermodynamics by (Clausius, 1867). Entropy is defined as losing information in a time series or signal. It is based on what you know about the current state of a time series or signal, how well can you predict the next state of the system? If a system has a very low entropy, the next state of the system is very predictable. However, high entropy would show a higher level of uncertainty in what the next state will be. The aim to compute entropy of time series data, is to determine the sensitivity changes. There are a number of different algorithms that have been used to estimate the entropy of a time series. Historically, the most popular was Approximate Entropy (AE), so it is discussed in detail in Chapter 3. In addition to being popular, other techniques such as Sample Entropy (SE) is built upon the AE algorithm, so understanding AE in some detail is worthwhile. This does not mean that AE is the best entropy measure to apply to all data. There are several items that I have carefully considered before choosing the correct entropy measure to answer my research questions. Considering the limitation of AE with bias, relative inconsistency, and depending on the sample length (Alcaraz and Rieta, 2010), SE would be most suitable for time series data like EEG and HRV. SE has been used widely to investigate various biological conditions in the human body, such as arrhythmia studied through ECG (Alcaraz and Rieta, 2010), Alzheimer's patients' EEG background activity (Abásolo et al., 2006), analysing human postural sway data (Ramdani et al., 2009) and studying HRV in the case of obstructive sleep apnoea syndrome (Al-Angari and Sahakian, 2007). SE is also used to detect the termination of a particular medical condition like seizures Yoo et al. (2012) and to test the effect of a therapy like ketogenic diet used for controlling intractable seizures (Takahashi et al., 2010). These studies have concluded that SE is a robust quantifier of complexity, which offers an accurate nonlinear metric for quantification (Alcaraz and Rieta, 2010). It gives a good dynamical signature and is a helpful tool that provides insights into various biological time series (Abásolo et al., 2006),(Ramdani et al., 2009).

A time series can be used to reconstruct the attractor of the underlying dynamic process. State-space reconstruction of a time series is a powerful approach for the analysis of the complex, nonlinear systems that appear ubiquitous in the natural and human world. This is a very important step in identifying the structural characteristics of a time series. Embedding is a transformation of a single sequence in time into a higher dimensional space.

This transformation provides increased information that increases the uniqueness of the signal being analysed. The goal of an embedding Dimension (ED) is to create a state-space where the structure of a system is embedded. If the dimension of the space that contains the true structure of the time series can be identified correctly, then it is possible to observe the actual variable of the time series. The dimension of the space that contains the true structure of this system is called the embedding dimension, and it is the minimum number of variables required to form a valid state-space from a given time series. The technique False Nearest Neighbour (FNN) is the most commonly used method for finding an ED (Abarbanel et al., 1993), (Kennel et al., 1992), (Stergiou, 2004). This approach is based on eliminating false projections that can occur when the dimension is not large enough to unfold the dynamics of the attractor. FNN discusses in detail in Chapter 3. ED would be suitable to find the underlying variables for time series data such as Heart Rate Variability (HRV) since many underlying variables might drive HRV.

The Pearson's correlation coefficient (PCC) measures how closely two different time series are related to each other with the same sequence length and linear dependency. The correlation coefficient ranges between 1 (when the matching entities are the same) and  $-1$  (when the matching entities are inverses of each other). A value of zero shows no relationship existing between the signals (Benesty et al., 2009). Cross-correlation (CC) measures how closely two different time series are related to each other taking time lag into consideration, at the same or different time. CC can be performed to analyse the time delay between two related time series. For example, CC has been successfully applied in analysing EEG signals in the Time Domain (TD) (Bob et al., 2010), as well as Frequency Domain (FD) (Li et al., 2013). This method can be used to determine the relationship between activities in global and local areas, and also among the different local areas of the human brain.

The typical approach to analysing data is to describe the data in terms of how they change over time, which is known as the TD. An alternative approach arises in the form of data analysis in the FD. This analysis presents data as a function of the frequencies in the signal rather than a function of amplitudes. FD analysis is used extensively to provide additional insights into health and pathological movement (Giakas, 2004), (Giakas et al., 1996), (Giakas and Baltzopoulos, 1997), (Stergiou et al., 2002), (Wurdeman et al., 2011), and (McGrath et al., 2012). To decipher frequencies in time series data, one method is spectral analysis, which entails breaking down the biological signal into simple signals. Spectral analysis is a numerical technique to write data as the sum of multiple discrete sine and cosine functions of different frequencies. There are different frequency transforms available, but the most commonly used transform is the Fourier transform. This transform

uses sums of sine and cosine functions to represent the more complex function. There are several software programs that will allow to calculate the power spectrum of your data (Percival et al., 1993), (Stoica and Moses, 1997), (Huang et al., 1998), (Beard, 2013), and (Prabhu, 2013). To find the insights into time series data, EEG and HRV, FFT would be more suitable because it can tell us the functionality of the brain, and could differentiate functionality for participants with and without medical conditions.

FFT can transform TD to FD and show us the presence of frequencies in the signal. However, FFT loses the time information and shows only frequency information. Whereas, Wavelet Transform (WT) can keep both time and frequency information. The WT of the signal can be thought of as an extension of the classic Fourier transform (FT) - it works on a multi-scale basis, instead of working on a single scale (Time or Frequency) as FT. This is achieved by decomposition of the signal over dilated (scale) and translated (time) version of the wavelet. So, for spectral analysis, WT is more suitable than FFT (Akin, 2002). For example, various methods based on Discrete Wavelet Transform (DWT) reported for removing noise by using WT for EEG (Faust et al., 2015), and ECG data (Sudarshan et al., 2017).

## **2.2 Electroencephalogram (EEG)**

Electroencephalogram (EEG) signals provide a measure of brain nerve cell electrophysiological activity accessible on the surface of the scalp Lewis et al. (1988), thus provide information about different brain activity. The electrical activity of the brain is recorded via electrodes attached to the surface of the scalp. The EEG signals vary, depending on the location of the electrodes on the scalp. To analyse these signals, it is essential to understand the basics of the nervous system, EEG recordings, different EEG frequencies, and artefacts.

### **2.2.1 The Nervous System**

The nervous system of a human brain contains about 100 billion neurons (nerve cells), and each neuron communicates to approximately 7000 others (Quiroga, 1998). When the body responds an action from an outside world, it sends messages via spinal cords to the brain. Then the parts of neuron cell called *dendrites* and *axons* between one neuron and other neurons receive and transmits messages to each other across a synapse as known as synaptic transmission Siegel and Sapru (2006). This synaptic transmission allows two cells transferring information and passing the message back to different parts of the body along the spinal cord for the reaction. The transmission and reception of messages provide the standard human functions, such as social behaviour, personality, movement, cognition, feeling,

thinking and perception (McAllister et al., 1995). The complex electrical signals generated in brain control all body activities and understanding, and these allow us to identify and create links between certain brain activities and disease states which are very useful in clinical areas and scientist (Quiroga, 1998). The first human brain measuring recorded by Hans Berger, the German psychiatrist in 1924. He also gave the name EEG for the electrical impulses of brain activities. The process of generating the electrical signal among membranes while two neurons transmit a message to each other is called Electroencephalogram (EEG) (Luck, 2014).

### **2.2.2 EEG Recordings**

The electrical activities in the brain recorded via measurement electrodes attached to the surface of the scalp to capture the voltage from billion active neurons lying under the skull (Subasi, 2005). Recording EEG requires several channels of electrode placement. There are various electrode placement systems used, including internal 10%-20% system with 21 channels, 10%-5% system with 128 channels and 32 channels (Jurcak et al., 2007). The commonly used system is 10%-20% electrode placement system which contains 21 positions (19 positions around the brain, and 2 reference electrode positions on each ear) and can be recorded by the following processes: apply the gel on the scalp to maximise skin contact and aid a low-resistance recording through the scalp. The electrodes then detect the electrical signals generated by different parts of the brain. The different voltage between sets of electrodes measured by a voltmeter (subsequently picked up by an amplifier) to magnify the signal, is then processed by a computerised system (Tatum IV et al., 2008). Recording voltage after a while results as EEG (Chall and Mirsky, 1978).

Following the standard 10%-20% system, as shown in Figure 2.10, 21 electrodes (Fp1, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, O2, A1 and A2) for EEG recording were used. The values of 10%-20% shown refer to the distances between adjacent electrodes: either 10% or 20% of the total front-to-back or right-to-left distance over the skull. The front-to-back distance is based on the measurement from the Nasion (a point between forehead and nose) to Inion (the lowest point of the skull from the back of the head showed by a prominent bump), and right-to-left distance is based on the measurement between the left and right pre-auricular points. Measuring the length of the skull is of paramount importance for choosing and positioning right sized EEG cap. In this 10%-20% system, most electrodes usually are labelled with a letter and a number; a 'number' referring to the position of the electrode on the scalp, and a 'letter' corresponding part of the cerebral cortex lying beneath. The letters used are F, T, P, O and C which represents Frontal lobes, Temporal lobes, Parietal lobes, Occipital lobes and Central electrodes (which overlie

the central sulcus or Rolandic Fissure of the brain) respectively. In some cases, lower case letters “p” and “z” follow the original letter of the part corresponding to the cerebral cortex. For example, Fp1 and Fp2 refer to electrodes placed just above the eyes over the “frontal poles” of the frontal lobes referred to as Frontopolar (Fp) electrodes. The letter “z” indicates the midline; therefore, Fz would imply Frontopolar electrodes lying over the midline and Fz, Cz and Pz imply Frontal lobe, central and parietal lobe electrodes lying over the midline.

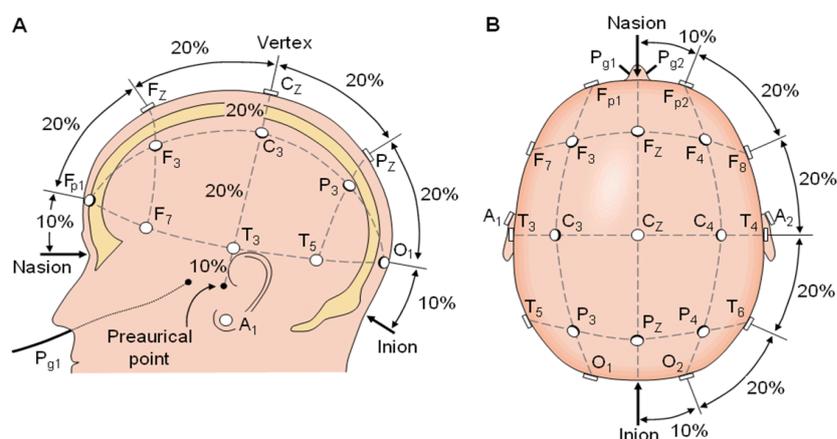


Figure 2.10: The international 10%-20% system seen from A (left side of the head) and B (above the head). The letter F, T, C, P, O, A, Fp and Pg stands for frontal, temporal, central, parietal, occipital, earlobes, frontal polar, and nasopharyngeal, respectively. The figure is obtained from (Klem et al., 1999).

### 2.2.3 Types of EEG Frequencies

EEG recordings reflect different brain activities from different areas of the brain. The presence of five different brain activities is usually found with five different EEG frequencies: Delta (0.5 Hz - 4 Hz), Theta (4 Hz - 7.5 Hz), Alpha (7.5 Hz - 13 Hz), Beta (13 Hz - 30 Hz), and Gamma (30 Hz - 50 Hz). The presence of these activities can be different for different participants. Figure 2.11 shows the different brain wave that could be recorded by an EEG.

Gamma ( $\gamma$ ) frequency band ranges from  $30\text{Hz} - 50\text{Hz}$ . It is the highest frequency wave, which indicates fast activities. It also presents when objects, sounds and tactile sensations are matched, such as in short-term memory (Teplan, 2002). This wave reflects the mechanism of consciousness when persons notice different combined senses, for example, sound and sight. Beta ( $\beta$ ) frequency band ranges from  $14\text{Hz} - 30\text{Hz}$ . It mostly presents on Front-central areas of the head, especially with electrode Fz. This brain wave activity appears in vigorous activity. It is the dominant rhythm of alert, active thinking, active attention,

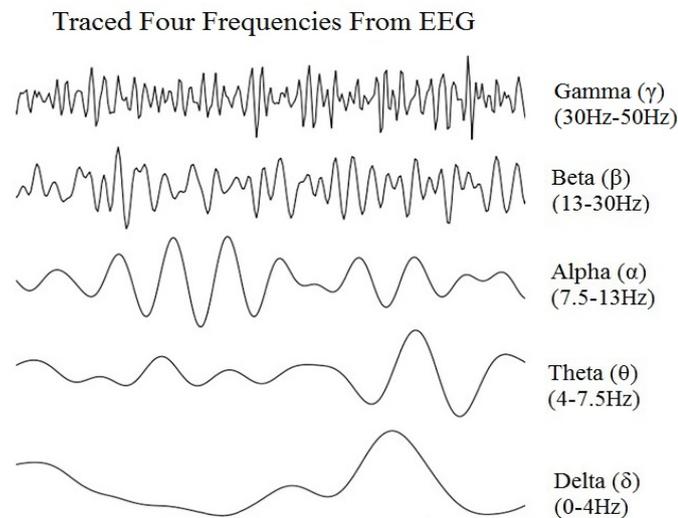


Figure 2.11: Traced Frequencies from EEG. The Amplitude (Power) of the signals increases as the frequencies decreases (Gastaut, 1952).

focuses on an outside world, concentration, fear, or stress. It has found that the presence of Beta is higher with open eyes (Quiroga, 1998).

Alpha ( $\alpha$ ) frequency band ranges from  $7.5Hz - 14Hz$ . It presents during relaxation, inattention, empty mind, light trances, and so on. Alpha mostly appears in the Posterior and Frontal areas, such as electrodes P3, Pz, P4, and electrodes F3, Fz, and F4. The effect of the Alpha wave reduces when an individual becomes aroused by certain other brain activities, such as thinking, calculating, super learning, and hearing strange sounds. Alpha wave increases when eyes closed, but decreased when eyes are open. (Teplan, 2002).

Theta ( $\theta$ ) frequency band ranges from  $4Hz - 7.5Hz$ . It is often known as slow activity in children, due to the density of its neural layers, the hippocampus generates some of the largest EEG signals and often continuing for many seconds. Theta also appears during dreaming sleep or in waking state of adult (Teplan, 2002).

Delta ( $\delta$ ) frequency band ranges from  $0.5Hz - 4Hz$ . It presents with the lowest frequency and high amplitude. The Delta frequency appears in deep, dreamless sleep and in profound, transcendental meditation where awareness is fully detached. Sometimes it appears in waking state such as during quiet or drowsiness (Teplan, 2002). It also may be easily confused with the artefact signals of the muscle near the skin surface, such as the muscle of the neck and jaw.

Figure 2.6 shows the traced frequencies from EEG signals.

## 2.2.4 EEG Artefacts

Although EEG records cerebral activity, it also records electrical activities arising from sites other than a brain. These recorded non-cerebral activities known as artefacts and divided into physiologic and extra-physiologic artefacts. Physiologic artefacts get generated from participants. They arise from sources other than the brain (that is another part of the body). Extra-physiologic artefacts arise from outside the body (types of equipment, environment). There are a few types of artefacts commonly found with EEG: 1) Cardiac artefacts, 2) Electrode artefacts, 3) Ocular artefacts, and 4) Muscle artefacts (Teplan, 2002). Cardiac artefacts (electrical or mechanical) are most easily identifiable by their synchronisation with complexes in the ECG channel. The presence of cardiac artefact mainly depends on the referential montages (placement of electrodes) of 10%-20% because of their greater inter-electrode distances. Montages with an average reference have a minimal cardiac artefact, and with bipolar montages, the artefact occurs with maximum amplitude and clearest QRS morphology over the temporal regions Carmichael et al. (2012). Electrode artefacts mostly occur due to poor connection of electrodes or electrode lead movement, leading to sharp and slow wanes of varying morphology and amplitude (Teplan, 2002).

The ocular artefact is the most common one in the EEG; it can occur during eye blinks, eye flutter, lateral eye movement, roving eye. Roving eye movements, occur with drowsiness and are an involuntary and repeated horizontal ocular movement. The movements have a relatively constant period and show a phase reversal because of the eyes' dipoles. The field around the right front-temporal electrodes becomes positive, and the left front-temporal electrodes become negative, with gaze (Vigário, 1997). Blinking produces an ocular artefact because of the rapid movement of the eyes, both up and down and appears in EEG with a field that does not extend beyond the frontal region. Muscle artefacts can occur during the chew/swallow, talking, and movement of any facial muscle. Movement during the recording of an EEG may produce an artefact through both the electrical fields generated by muscle and through a movement effect on the electrode contacts and their leads (O'Regan et al., 2010).

The visual appearance of the four types of artefacts described earlier, such as Cardiac artefacts, Lead Movements, Ocular artefacts, and Muscle movements, can found in Figure 2.12-2.15, respectively. However, artefacts caused by Cardiac and cable movement have not found with the datasets utilised in this research. Therefore, results shown in Figures 2.12 and 2.13, are not for the datasets I have used- information about the datasets utilised in my research described in Chapter 4.



Figure 2.12: Cardiac Artefact presents in the Data Grouiller et al. (2007)

## 2.2.5 Research on EEG

In recent years, research has been done on EEG signals to analyse the conditions of the brain, not only by comparing brain activities between different part of the brain but also by finding differences between participants with and without medical conditions. For example, (Na et al., 2002) examined information transmission between different cortical areas in participants with medical condition schizophrenia and Alzheimer. EEG signals have been investigated by (Jeong et al., 2015) to distinguish the EEG signals of participants without any medical conditions, and participants with conditions, such as Parkinson's related dementia and Alzheimer. (Kannathal et al., 2005) and (Sun et al., 2018) has studied the complexity of the EEG signals from participants with Epilepsy and participants without any medical condition. (Oberman et al., 2005) and (Rajaguru and Prabhakar, 2017) investigated whether individuals with Autism Spectrum Disorder (ASD) showed dysfunction in the mirror neuron system, given their behavioural impairments in understanding and responding appropriately to others' behaviours. (Choi and Cho, 2019) investigated the effect of electroacupuncture at different frequencies and intensities via EEG, and found decreased in absolute power for Theta band. Furthermore, EEG is used in the acupuncture research area to explore the effect of acupuncture stimulation in the central nervous system (Rastiti

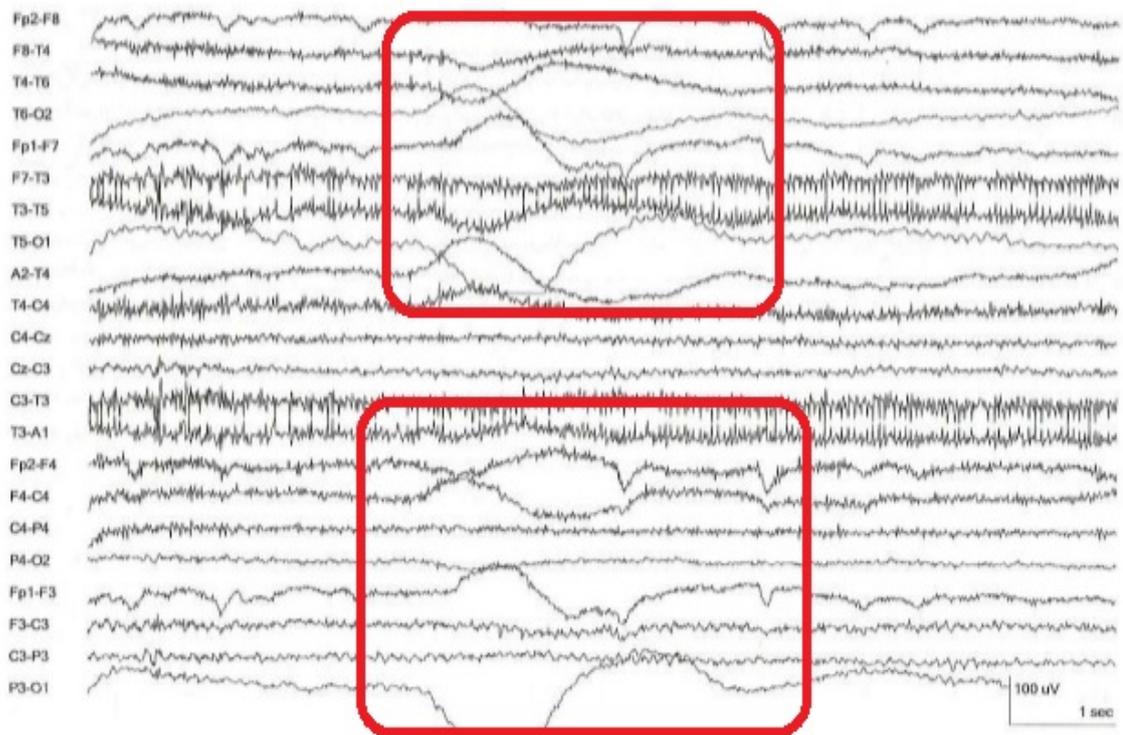
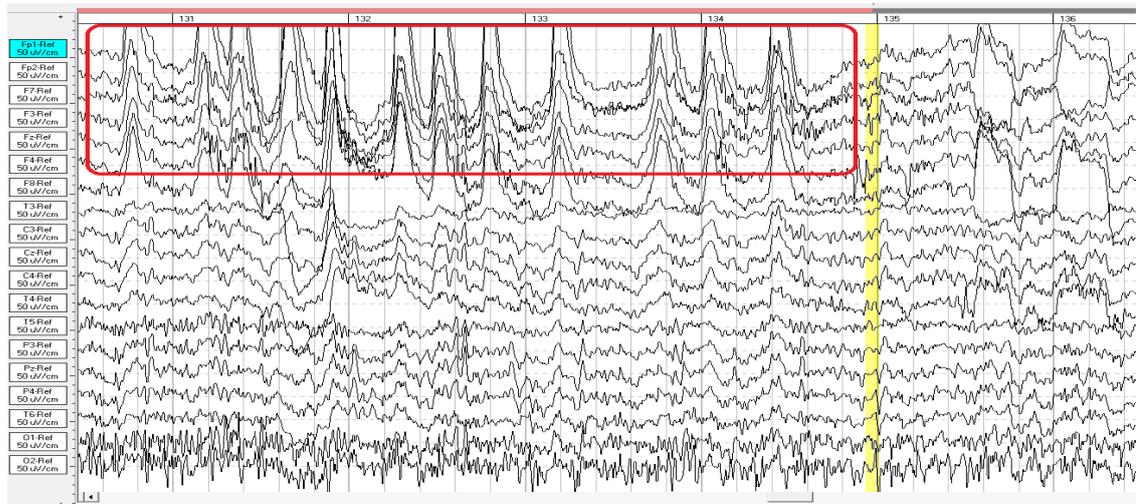


Figure 2.13: Electrode Artefact showing noise caused by lead movement Grouiller et al. (2007)

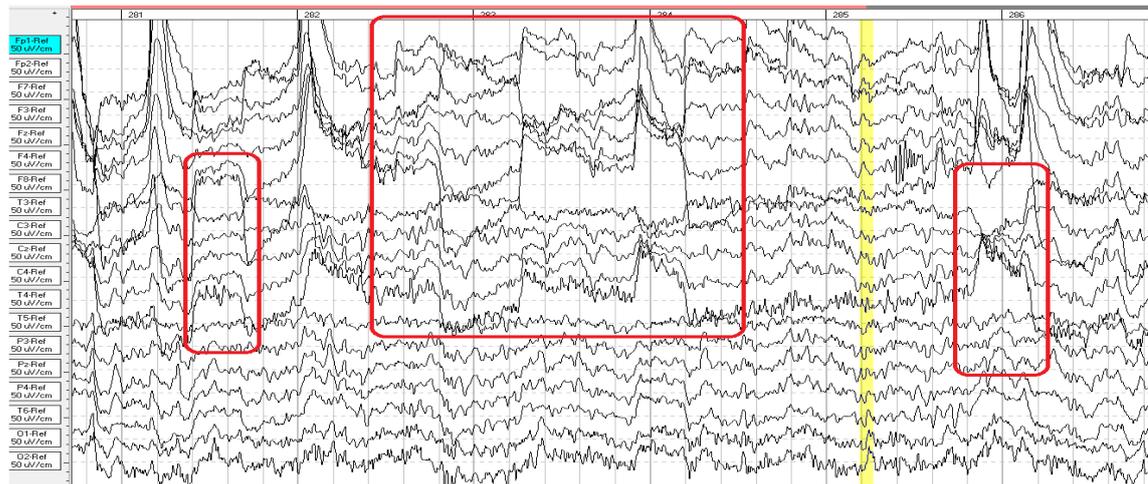
et al., 2018).

Various characteristics of EEG signals represent distinct states of brain activity (Niedermeyer and da Silva, 2005). We can quantify these distinct states using linear or non-linear measures. Previous research has showed a correlation between EEG signals (or brain activity) from different part of the brain (Left and right brain hemisphere) (Bob et al., 2010), (Na et al., 2002), (Jeong et al., 2015), (Bilucaglia et al., 2019), (Abdulla et al., 2019). A high correlation between the signals from different electrodes indicates similar brain activity, and a low correlation suggests that the brain activity at the different measurement sites is relatively independent. Researchers (Na et al., 2002), (Li et al., 2013), (Hevia-Orozco et al., 2017), (Almanza-Sepúlveda et al., 2018), and (Gartstein et al., 2020) have expressed that brain activities within the same (local) region might be similar, but that they might be different among non-identical regions (globally). One question my research address is whether the activities of the two brain hemispheres are similar.

So far, various numbers of electrodes and combinations of electrode pairs have been used to analyse EEG signals. The combinations of electrode pair depend on the total number of electrodes. For example, if there are 19 electrodes, then the number of different potential electrode pairs is 171. According to recent research on EEG signal analysis, elec-



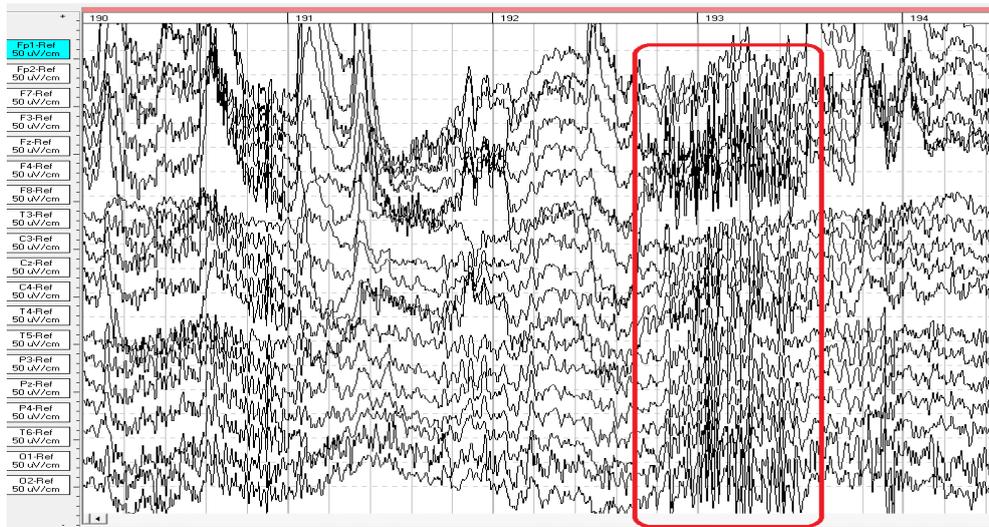
(a)



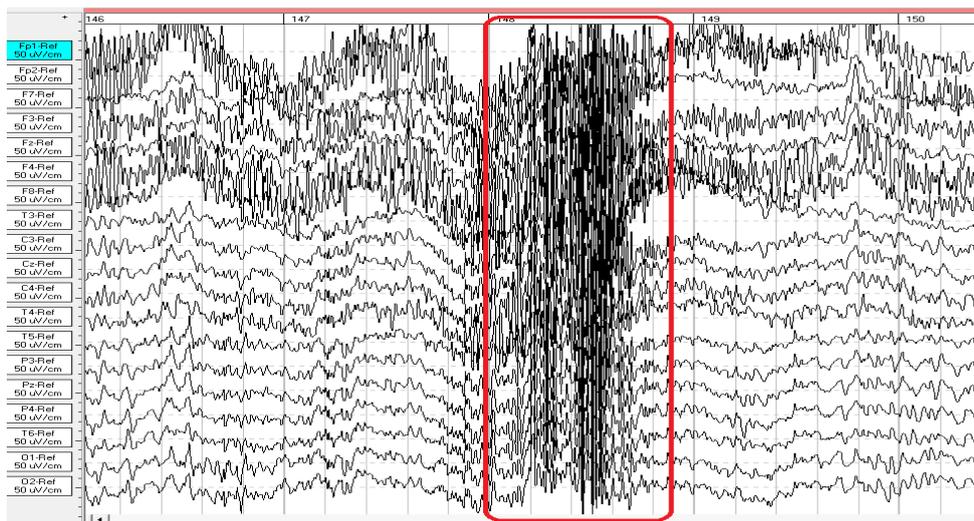
(b)

Figure 2.14: Ocular Artefacts: (a) Eye blinks with rapid movement of the eyes both up and down, and (b) Lateral Eye Movement

trodes from the central part of the brain deserve the best consideration, possibly because of the minimum noise found in the recorded signals (Klein and Thorne, 2006). Consideration of the central part of the brain was one reason I found papers in which they analysed EEG signals using only limited numbers of electrodes and combinations of electrode pairs. For example, the cortical EEG, recorded from 6 electrodes, during performance of working memory (WM) tasks in each trimester of pregnancy has been studied. Their result suggested characteristic patterns of EEG synchronization between the prefrontal and parietal cortices during performance of both WM tasks in each trimester of pregnancy (Almanza-Sepúlveda et al., 2018). Furthermore, the EEG correlation during a social decision making



(a)



(b)

Figure 2.15: Muscle Artefacts: (a) Swallow effect, and (b) Jaw movement.

task (Ultimatum Game) in a group of institutionalized (INST) adolescents with a never institutionalized group (NINST), with 6 electrodes (Hevia-Orozco et al., 2017). Their results have suggested significant changes in the EEG signals of electrodes from the right side of the brain when compared to those on the left side. Na et al. (Na et al., 2002), examined 16 electrodes with 38 pairs of electrodes within the right hemisphere and within the left hemisphere. Their results showed less complex EEG activity in the left temporal regions. Bob et al. (Bob et al., 2010), inspected 8 electrode and 16 electrode pairs to examine the relation between EEG activity with the Dissociative Experiences Scale (DES)

in paranoid schizophrenia patients. Their results demonstrated a significant correlation of DES in 9 EEG electrode pairs. Similar electrode pair effects have been found by Cuevas et al. (Cuevas and Bell, 2011), who studied 8 electrodes and 16 electrode pairs in their investigation of patterns with EEG signals of children's developing brains. Their results suggested an age-related increase in EEG power for 9 electrode pairs. Li et al. (Li et al., 2013), examined 16 electrodes and 4 electrode pairs and proposed more significant changes in the EEG signals of electrodes from the right side of the brain when compared to those on the left side.

In recent research, the correlation between EEG signals has been analysed in the Frequency Domain (FD) using various methods, such as Mutual information, Coherence analysis, Wavelet coherence, Correlation coefficient, Auto-correlation and Cross-Correlation. Mutual information has been utilised to examine information transmission between different cortical areas in subjects with both schizophrenia and Alzheimer's diseases (Na et al., 2002). This research found lower mutual information between EEG signals of subjects with these conditions when compared to EEG signals without a medical condition. Coherence analysis has been applied to study brain interactions between EEG signals (Nolte et al., 2004), showing a significant correlation in EEG Beta ( $\beta$ ) frequency range between the left and right motor areas of the human brain of participants without any medical conditions. Wavelet coherence has been applied to distinguish between EEG signals of participants without medical conditions, and EEG signals of participants with medical conditions, such as Parkinson's related dementia and Alzheimer's diseases (Jeong et al., 2015). Correlation coefficient has been utilised to discover changes in EEG signals and autonomic nervous activity, and the association of these with personality traits (Takahashi et al., 2005), with an increase in EEG theta ( $\theta$ ) power and EEG alpha ( $\alpha$ ) power predominantly in the frontal area. Cross-Correlation has been utilised to study the degree of association between activities in symmetrical (left and right) parts of the brain (Li et al., 2013), (Hevia-Orozco et al., 2017), and (Almanza-Sepúlveda et al., 2018) with the indication of stronger correlation in the delta ( $\delta$ ) frequency range on the right side of a brain than the left.

In my research, I have also used EEG data to differentiate the brain activity of participants with and without any medical condition, like recent research as shown earlier. From the research above, I have found that the focus of the analysis of EEG was on either left or right brain hemisphere when comparing the EEG activity from different part of the brain. However, in my research, I have focused on the combination of EEG electrode pairs from and between both brain hemispheres. In addition, to my knowledge, they have conducted limited research to analyse EEG signals in the Time Domain (TD) with all combination of electrode pairs. TD analysis is used to analyse a signal in its actual state, which is the

earliest and direct way of analysing EEG signals - it is utilised to analyse changes in EEG signals, such as power (or amplitude) over time. Previous research, as described earlier, focused on FD. There is a necessity to conduct TD analysis on EEG. Therefore, it is essential to perform a comparative analysis and an interpretation of EEG signals in the TD, not just the FD. The summary of the research work reported in this section shown in Table 2.1.

Table 2.1: Summary of Some of the Research work for Correlation Between EEG Signals.

Reference	Electrodes	Pairs	TD	FD	CC	Other Methods
(Lewis et al., 1988)	8	28 pairs	✓	-	-	MI
(Klem et al., 1999)	19	38 pairs	-	✓	-	CA
(Na et al., 2002)	16	120 pairs	✓	-	-	MI
(Nolte et al., 2004)	122	No pairs	-	✓	-	CA
(Niedermeyer and da Silva, 2005)	6	No pairs	-	✓	-	CC
(Takahashi et al., 2005)	1-Cz	No pairs	✓	-	✓	-
(Klein and Thorne, 2006)	21	No pairs	-	✓	-	SC
(Fields, 2008)	106	No pairs	-	✓	✓	CC
(Bob et al., 2010)	8	16 pairs	✓	-	✓	-
(Cuevas and Bell, 2011)	16	8 pairs	-	✓	-	CA
(Li et al., 2013)	16	4 pairs	-	✓	✓	-
(Jeong et al., 2015)	19	56 pairs	✓	✓	-	WC
(Hevia-Orozco et al., 2017)	6	No pairs	-	✓	-	PCC
(Almanza-Sepúlveda et al., 2018)	6	No pairs	-	✓	-	rEEG

MI = Mutual Information, CA = Coherence Analysis, CC = Correlation Coefficient, SC = Spearman's Correlation, WC = Wavelet Coherence, PCC= Pearson's correlation Coefficient.

## 2.3 Electrocardiogram (ECG) and Heart Rate Variability (HRV)

### 2.3.1 Electrocardiography (ECG)

Electrocardiography (ECG) signals reflect activities of heart muscles. They are related to a variety of intertwined and complex chemical, electrical, and mechanical processes present in the heart. They convey a great deal of valuable diagnostic information not only describing the functioning of heart but also other systems such as circulation or nervous systems. The ECG signal has been a subject of studies for over 100 years. The English physiologist August Waller has realised the first recording of electrical activities of the heart in 1887, who used surface electrodes placed on the skin and connected to the capillary electrometer. He was the first to call the registered signal ECG (Burch and DePasquale, 1990). Nevertheless, W.Eintheoven is regarded to be the father of electrocardiography who

in 1902 recorded the first ECG with the use of a string galvanometer (Einthoven, 1912). Furthermore, ECG records electrical activities of the heart, where each beat of the heart started by an electric signal from the heart vagus. ECG signals might contain a plethora of information with and without pathological physiology of the heart and its health.

### **2.3.2 Heart Rate Variability (HRV)**

Heart Rate Variability (HRV) can be extracted from ECG. HRV is the estimation of Neuro-cardiac function that reflects heart-brain interactions and autonomic nervous system dynamics (McCraty et al., 2001). The measurement of HRV is a valuable investigative tool in clinical cardiology as it gives a primary method to test the physiological state of the heart directly. Many neurological and psychological investigations have used HRV to assess the effects of stress, emotion, and work on the autonomic nervous system (Malik and Camm, 1990). The heart rate and rhythm are mainly under the control of the Autonomic Nervous System (ANS), which is part of the Peripheral Nervous System (PNS). PNS act as a control system functioning mostly below the level of consciousness to control physical functions. ANS contains two primary components: the Sympathetic and Parasympathetic Nervous system. Both the sympathetic and parasympathetic nervous systems innervate the heart. The parasympathetic nervous system functions in regulating heart rate through the vagus nerve, with increased vagal activity producing a slowing of heart rate. The sympathetic nervous system has an excitatory influence on heart rate and contractility (Robinson et al., 1966). The study of HRV series helps to determine the interaction between sympathetic and parasympathetic activity (Karmakar et al., 2011).

HRV is widely adopted to measure the heart function, which can help identify patients at risk for a cardiovascular event or death (Cooper et al., 2007). HRV analysis is not only a non-invasive tool for assessing cardiovascular system function but also serves as a useful index for evaluating the function of the autonomic nervous system (ANS) in regulating human organs and muscles (Lin and Hu, 2007). HRV testing is concerned about an analysis of the changes in heart rhythm. This analysis helps us evaluate the quality of the autonomic heart control system. The autonomic nervous system controls the functioning of the heart comprises sympathetic and parasympathetic systems. These two systems work in two opposite directions regarding the frequency of heart action, controlling it depending upon the existing situation. The heart rhythm is increased when there is an increased activity of the sympathetic system. When the activity of the parasympathetic system increases, the heart rhythm decreases. In a healthy individual, there are some fluctuations of the rhythm within the bounds of the adaptation. However, under certain stimulus (say, alcohol, nicotine, some medications) and because of some diseases (for instance, failure of kidneys), the change of

the heart rhythm is made difficult, and it significantly reduces the variability (Acharya et al., 2006). The analysis of the variability of heart rhythm reduces to the analysis of changes in the length of RR distances for QRS complexes of sinus origin. An essential component of the analysis is a precise localisation of the peaks of the R waves. The analysis of RR distances (RR periods) is done in a time domain, frequency domain by utilising methods of time-frequency analysis or other nonlinear methods (Acharya et al., 2006), (Sörnmo and Laguna, 2005).

The R waves of the ECG detected through a QRS complex detection algorithm, e.g. Pan-Tomkins QRS detection algorithm (Pan and Tompkins, 1985), after the ECG recording is digitised. Pan-Tomkins QRS detection algorithm detects R waves and finds the difference between consecutive R peaks, which is known as HRV. R waves typically have the largest amplitudes compared to surrounding P, Q, S, and T waveforms, as shown in Figure 2.16. Thus a beat-to-beat interval can be defined as the time difference between consecutive R peaks (RR Interval). RR intervals originate from normal sinus rhythms, sometimes referred to as normal-to-normal (NN) or Inter-Beat Interval (IBI) intervals. Thus, standard nomenclature of "NN" can be used in place of IBI or RR to indicate IBI's containing no ectopic intervals. Figure 2.16 shows a hypothetical ECG and how IBI's are determined based on R waves. IBI (1) and IBI (2) represent the first and second data point of the IBI time series signal. Therefore, the HRV series consists of RR intervals representing the fluctuations in the interval between heartbeats (Brennan et al., 2001).

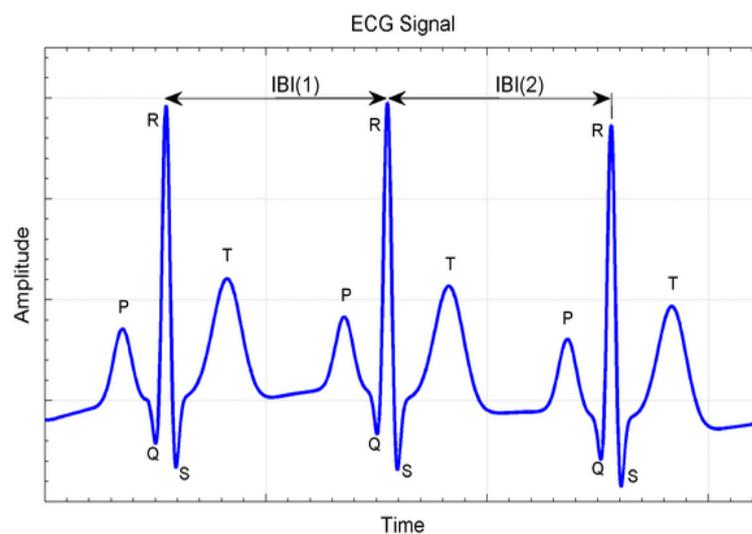


Figure 2.16: Determination of Inter Beat Interval (IBI). Simulated ECG containing three beats with arbitrary units of time and amplitude. Time intervals corresponding to the IBI are indicated by IBI(1) and IBI(2). ECG morphology is shown by five characteristic waves P, Q, R, S, and T.

### 2.3.2.1 Types of HRV frequencies

The HRV frequencies help to determine the interaction of sympathetic and parasympathetic activities. These frequencies can be useful for understanding the state of the heart of a person. There are three main frequencies in HRV: Very Low Frequency (VLF) ranging from 0-0.04 Hz, Low Frequency (LF) ranging from 0.04-0.15 Hz, and High Frequency (HF) ranging from 0.15-4 Hz.

The presence of VLF found if a person is angry, stressed or with anxiety (Shaffer et al., 2014). This frequency results from the sympathetic nervous activity of the ANS (Petretta et al., 1997). The presence of HF found if the person is feeling relaxed, calmed or with appreciation. This frequency results from the parasympathetic nervous activity of the ANS (Petretta et al., 1997). Figure 2.17 shows how the brain and heart interact with each other, and how HRV pattern being dominant by ANS (Sympathetic and Parasympathetic). The cerebral cortex of the brain is the body's ultimate control and information processing (McCraty et al., 2009). The usual Heart-Brain communication path is through the spinal cord. Figure 2.17 shows that the brain and heart communicate through 'Medulla' (cardiovascular centre placed in medulla controls the heart beating), which is part of the brain stem.

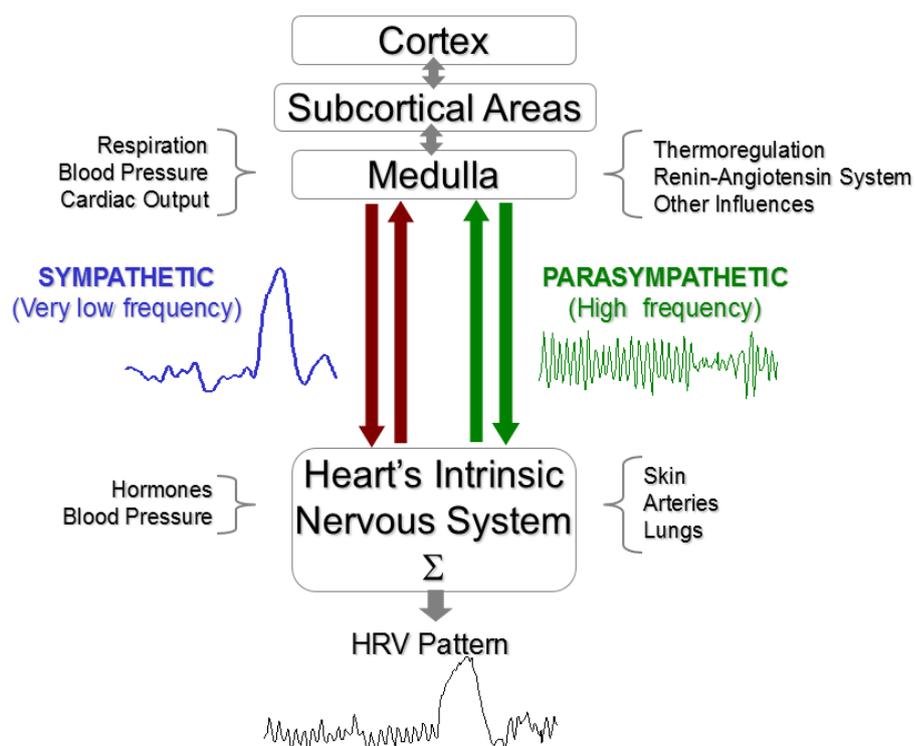


Figure 2.17: HRV pattern, showing the contribution of Sympathetic and Parasympathetic Nervous Activity (Steffert and Mayor).

Spectral analysis for one HRV shown in Figure 2.18, as an example, where the presence of all three HRV frequencies shown. Besides, the detail about traced frequencies of HRV signals are shown in Figure 2.19.

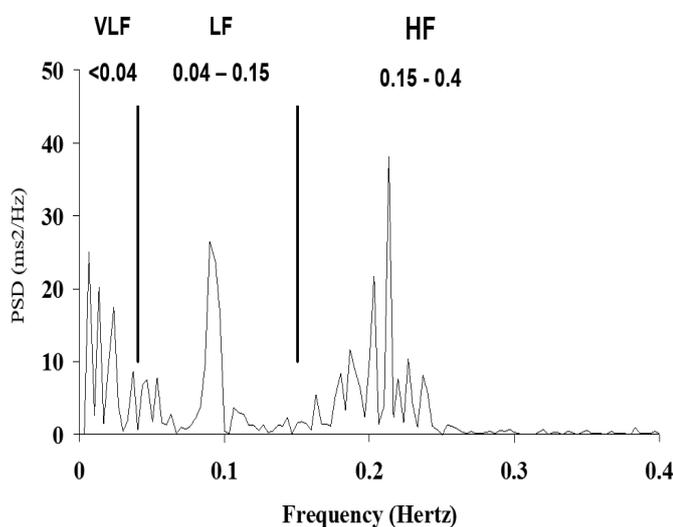


Figure 2.18: Power Spectrum of HRV showing presence of traced HRV frequencies

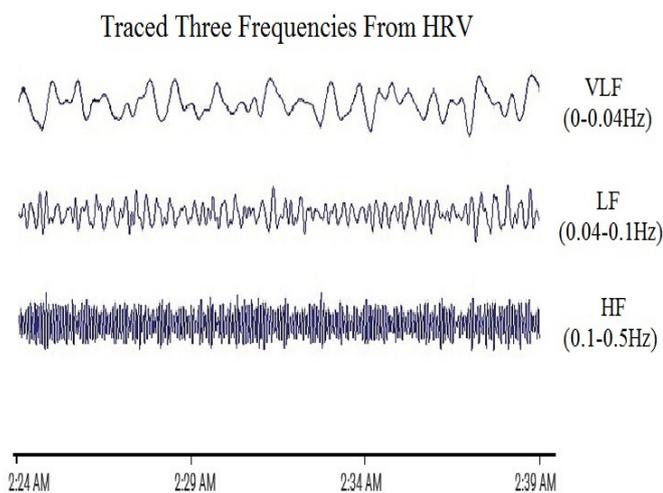


Figure 2.19: Traced Frequencies from HRV. The Amplitude (Power) of the signals increases as the frequencies decreases (Lin et al., 2010).

### 2.3.2.2 HRV Artefacts

Artefacts in IBI time series can cause significant distortion of HRV analysis results, and thus, all artefacts should be either corrected or excluded from analysis as recommended

in (Electrophysiology, 1996). Typical artefacts include missing, extra or misaligned beat detections and ectopic beats, such as premature ventricular contractions (PVC) or other arrhythmias. The procedure of removing these artefacts is shown in Chapter 7.

Figure 2.20 shows the presence of an extra beat type of artefacts in HRV as an example.

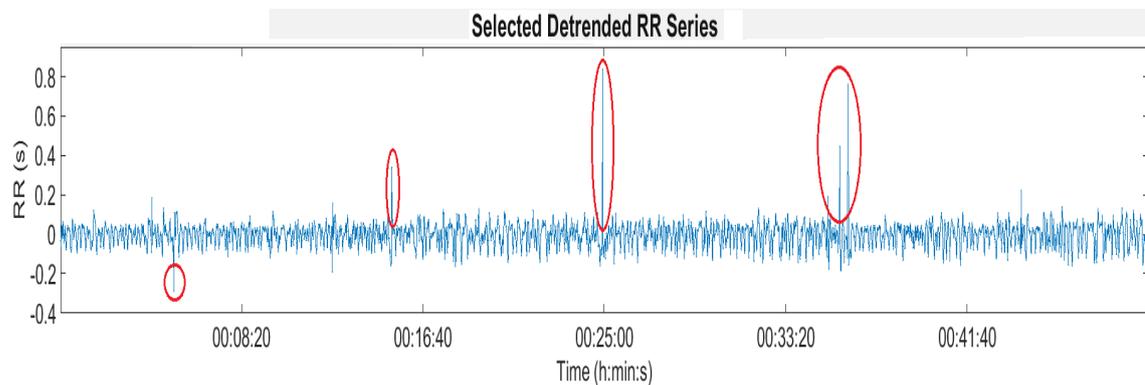


Figure 2.20: Presence of extra beat and premature ventricular contractions (PVC) artefact in HRV. The highlighted in red circles are the extra beat

### 2.3.3 Research on ECG and HRV

ECG analysis is a very useful diagnostic tool for clinical purposes. The ECG varies from person to person due to the differences in position, size, and anatomy of the heart, age, sex, relative body weight, chest configuration, and various other factors (Simon and Eswaran, 1997). Biel et al. (Biel et al., 1999), (Biel et al., 2001) showed that it is possible to identify individuals based on an ECG signal. (Shen et al., 2002) has used ECG as a new biometric for human identity verification in which, they have demonstrated successfully that it is possible to identify a specific person from a group of candidates using a one-lead ECG (Shen et al., 2002). There are numbers of study devoted to the analysis and classification of ECG signals by using combinations of syntactic and probabilistic methods (Horowitz, 1975), (Udupa and Murthy, 1980), (Papakonstantinou and Gritzali, 1981), (Skordalakis, 1986), (Trahanias and Skordalakis, 1990). neural networks were used to arrhythmia analysis (Wang et al., 2001), determination of P waves (de Azevedo Botter et al., 2001)), and classification of heart evolution with the use of second order cumulants (Osowski and Linh, 2001). They also used fuzzy systems in the analysis of ECG and EEG signals (Moon et al., 2002). Fuzzy rule-based systems were used in the detection of arrhythmias (Chowdhury and Ludeman, 1994) and construction of ECG signals (Wang et al., 1991). Many algorithms for automatic detection and classification of ECG heartbeat patterns have been

presented in the literature including signal processing techniques such as frequency analysis (Minami et al., 1999), wavelet transform (Shyu et al., 2004), (Ince et al., 2008)), and filter banks (Afonso et al., 1999), statistical (Willems and Lesaffre, 1987) and heuristic approaches (Talmon, 1983), hidden Markov models (Coast et al., 1990), support vector machines (Osowski et al., 2004), artificial neural networks (ANNs) (Hu et al., 1993), and mixture-of-experts method (Hu et al., 1997).

Some research supports the idea of reporting an inverse relationship between inflammatory cytokines and vagally-mediated HRV using both short-term (less than 1 hour) (Soares-Miranda et al., 2012), (Young et al., 2014), and long-term (greater than 1 hour) (Araújo et al., 2006), (Janszky et al., 2004) recordings of HRV. Prospective studies have also found similar results, showing vagally-mediated HRV to predict negatively inflammation four years into the future (Jarczok et al., 2014). In the biomedical setting, they often use HRV metrics for risk stratification, where clinical endpoints (e.g., myocardial infarction) across a range of chronic health conditions may be forecaster by earlier measurements of HRV. Low HRV, for example, is associated with mortality in participants with coronary artery disease (Huikuri and Stein, 2013), (Martin et al., 1987), chronic heart failure (Nolan et al., 1998), and among those with a history of myocardial infarction (Bigger Jr et al., 1992), (Bigger Jr et al., 1988), (Buccelletti et al., 2009), (Camm et al., 2004), (Kleiger et al., 1987). Beyond mortality, hypertension (Singh et al., 1998), end-stage renal disease (Brotman et al., 2010), and diabetes (Schroeder et al., 2005) are also associated with low HRV. Notably, higher HRV does not always signal apparent protection, as high HRV confers risk for atrioventricular (AV) block, sick sinus syndrome, and atrial fibrillation (Vikman et al., 2003). Besides clinical applications, HRV metrics are often employed to better understand the peripheral physiological correlates of complex brain and behavioural processes, such as emotion and its regulation (Graziano and Derefinko, 2013), (Rottenberg et al., 2007) and executive cognitive functioning (Thayer and Lane, 2000), (Thayer et al., 2009), possibly by reflecting the functionality of higher brain systems, such as the pre-frontal cortex (Beauchaine and Thayer, 2015). The predictive utility of pretreatment HRV has been investigated for outcomes of antidepressant medication in major depressive disorder (MDD), with pretreatment anxious depression as a hypothesised moderator of HRV effects (Kircanski et al., 2019). The increased resting heart rate (HR) in heart transplant patients is associated with enhanced metabolic demand, the potential for fatigue, and lower quality of life. (Moreira et al., 2019) presented that transcutaneous electrical acupoint stimulation (TEAS) can modulate autonomic balance and reduce resting HR in heart transplant patients. The effect of acupuncture was studied through HRV to investigate the possible role

of the autonomic nervous system in mediating, and showed an increase in the LF/HF ratio (indicating greater sympathetic activity) (Chang et al., 2010).

The main goal of the use of HRV in my applications is to draw more specific inferences about the autonomic nervous system (ANS) activity that is enabled by the heart period. Changes in heart period, cannot be necessarily interpreted as reflecting symmetric, but opposite, changes in cardiac sympathetic and parasympathetic (that is vagal) control (Berntson et al., 1991), (Berntson et al., 1994), (Gianaros and Quigley, 2001). Instead, heart period and evoked changes in heart period are ambiguous with respect to their autonomic origins. Furthermore, the utilization of HRV is to understand the complex brain employing HRV metrics.

## **2.4 Correlation Between EEG and HRV Signals**

A series of data points in time order, or time series, provides a view of a signal as it evolves, in Time Domain (TD). TD analysis is used to analyse the signal in its actual state - in biomedical signals, the power (or amplitude) over time is utilised to analyse changes. In parallel, the frequencies present in the signal investigated, and such an analysis take place in the Frequency Domain (FD). FD analysis is used to identify frequencies present in the signal.

Recent studies seem to indicate that executive function: processes that control and regulate thought and action (e.g., suppressing habitual responses) (Friedman et al., 2006) are correlated with brain activity in the resting state (Mennes et al., 2010), (Takeuchi et al., 2011), (Mackey et al., 2013), (Martínez et al., 2013), (Thompson et al., 2016), (Fox et al., 2005), (Fox et al., 2006), (Seeley et al., 2007). The two brain networks in the resting state has been identified by (Fox et al., 2006). One network comprises regions that are routinely positively correlated with cognitive task performance, and the other includes regions that are routinely negatively correlated. The presence of significant positive/negative correlations between a cerebral region and a task across participants suggests that at least some part of the cerebral response induced by a particular task is intrinsically represented in the brain (Mennes et al., 2010). Similarly, when variability is disrupted, the brain has little capacity to adapt to environmental conditions, resulting in neuropathological diseases such as epilepsy and attention-deficit/hyperactivity disorder (Mizuno et al., 2010), (Catarino et al., 2011), (Vakorin et al., 2011), (Ramon and Holmes, 2013), (Alba et al., 2016), (Chen et al., 2017).

Lately, relations between HRV and the endogenous dynamic of brain regions involved in autonomic control and emotional regulation during the resting state have been explored.

These studies showed that high- and low-frequency components of HRV are strongly coupled with functional connectivity (Chang et al., 2013), (Jennings et al., 2016), (Sakaki et al., 2016). However, these studies have not addressed the relationship between the variability of functional connectivity and HRV or whether both factors might predict the outcomes of cognitive tasks. (Palva et al., 2013) correlated the variability of functional connectivity using magneto-encephalography (MEG) and HRV during a stimulus detection task and during the resting-state period. Strong correlations were found between neuronal oscillations and task performance during the task and during the resting-state period. These results suggest that the variability of functional connectivity in the resting state is not specific to the task but is related to the performance of cognitive tasks. This study found that HRV in both task and rest conditions predicted task performance. Normally, any EEG measure is estimated during the resting state by averaging its values in a certain number of EEG segments. This procedure assumes that resting-state EEG remains static during the recording period. Even in this case, recent evidence (Kitzbichler et al., 2009), (Botcharova et al., 2014) suggests that brain synchronisation assessed from neurophysiological signals is not constant, but during this time, it presents significant variability, which is disrupted in neuropathological conditions (Ramon and Holmes, 2013), (Alba et al., 2016). The aim of this study was to determine whether there is any functional connectivity of EEG with HRV in the resting state.

In recent research, the correlation between the EEG and HRV signals have been analysed in FD using well-known methods. There are various reasons people have investigated the relationship between these signals. For example, to find changes in EEG activity and autonomic nervous system during sleep (Miyashita et al., 2003). This research found a striking correlation between the frequencies of the EEG and sympathetic activity (Low Frequency (LF)) of HRV. The relationship between the depth of sleep and the changes in autonomic nervous system have been explored (Yang et al., 2002), (Ako et al., 2003), (Jurysta et al., 2003), (Abdullah et al., 2010), (Chua et al., 2012) and (Moeynoi and Kitjaidure, 2017). Most of these studies suggest a negative correlation between the Delta frequency range of EEG with LF of HRV, and some of these studies demonstrate a significant correlation between the Delta frequency range of EEG with parasympathetic activity (High Frequency (HF)) of HRV. (Berg et al., 2005) has investigated the sleepiness and drowsiness of drivers by analysing the correlation between EEG and HRV. A strong correlation between Delta, Alpha, and Theta frequency range of EEG with LF of HRV was found, respectively, which suggested that HRV as an indicator of sleepiness. The changes in EEG and HRV during meditation have also analysed in (Takahashi et al., 2005). Their results suggest a significant correlation between Alpha and Theta frequency range of EEG with

LF and HF of HRV during meditation, respectively. investigated the EEG and ECG features that can reveal status change during cycling exercise (Jao et al., 2017). Their result suggests that the Cardiac Stress Index and high alpha in EEG electrode C4 are the most suitable signals in ECG and EEG for predicting status changes during exercise

The relationship between EEG and HRV has been analysed in different conditions. For example, the correlation analysis of EEG signals and ECG signals between epileptic and normal population was carried out based on the improved synchronization algorithm IRC (Sun et al., 2018). This research has found that the correlation between EEG signals and ECG signals in the left forehead and left anterior temporal regions of normal and epileptic populations has been distinguished using the IRC algorithm. Furthermore, the impacts of the driving duration and circadian rhythm on the vigilance level of the drivers has been examined and the correlations between the vigilance and driving performance has been identified (Wang et al., 2019). (Doufesh et al., 2018) investigated the correlations between alpha electroencephalography (EEG) and other physiological parameters during Muslim prayer utilizing the self organizing map (SOM). Their result indicated that alpha power of EEG showed significant positive correlation in the occipital and parietal electrodes with the normalized unit of high-frequency (HF) power of HRV. The effect of fast cable car ascent on both the autonomic and central nervous system has been analysed (Edlinger and Guger, 2006) and suggested a positive correlation of Alpha, and Beta frequency range of EEG with LF of HRV. The effect of acupuncture on EEG and HRV has also been analysed (Sakai et al., 2007) and found a negative correlation between power in all EEG frequency bands with the LF/HF ratio of HRV. An investigation has been done to assess the relationship between cerebral cortices with peripheral cardiac autonomic (PAN) in uremic and healthy controls (Liou et al., 2014). Their study suggests the correlation between Delta / Theta ratio frequency range of EEG with LF of HRV, and between the Beta frequency range of the EEG with HF of the HRV. The relationship between HRV and Rolandic mu rhythm in relaxed condition of resting state has been observed (Triggiani et al., 2016) and observed a negative correlation between the Beta frequency range of EEG with LF of HRV. (Prinsloo et al., 2013) analysed the effect of HRV biofeedback (concentrative meditation) on EEG, and suggested a correlation between HRV biofeedback and EEG- significantly for Theta, and Beta frequency range of the EEG.

As discussed, the correlation between EEG and HRV signals has analysed in FD. Table 2.2, indicates that the Pearson correlation coefficient (PCC) is a well-known method for the FD analysis. I also aim to use well-known method PCC for the correlation performance of EEG and HRV in my research. However, the focus is not only in FD as shown in current research, but both FD and TD. I have found little information on the use 19 EEG electrodes

to analyse correlation performance between EEG and HRV in FD. Therefore, my prime focus for EEG and HRV correlation is to conduct analysis in both domains (TD and FD) including all 19 EEG electrode data (please refer to Chapter 7).

Table 2.2: Summary of Correlation Research on EEG and HRV.

Reference	TD	FD	PCC	Other Method	EEG Electrodes
(Miyashita et al., 2003)	-	✓	✓	-	4
(Yang et al., 2002)	-	✓	✓	-	2
(Ako et al., 2003)	-	✓	✓	-	1
(Jurysta et al., 2003)	-	✓	-	Coherency Analysis	3
(Takahashi et al., 2005)	-	✓	✓	-	6
(Edlinger and Guger, 2006)	-	✓	✓	-	2
(Berg et al., 2005)	-	✓	✓	-	2
(Sakai et al., 2007)	-	✓	✓	-	19
(Abdullah et al., 2010)	-	✓	-	Cross-Correlation	1
(Chua et al., 2012)	-	✓	-	✓	4
(Kim et al., 2013)	-	✓	-	Coherency Analysis	19
(Prinsloo et al., 2013)	✓	-	✓	-	3
(Liou et al., 2014)	-	✓	✓	-	19
(Triggiani et al., 2016)	-	✓	✓	-	19
(Jao et al., 2017)	-	✓	-	Poincare Analysis	4
(Sun et al., 2018)	✓	-	-	IRC	16

Table 2.3: Summary of Research on Well known Wavelet Transformation Methods for EEG and HRV.

Reference	EEG	ECG/HRV	TD	FD	Preprocessing Method
(Kutlu and Kuntalp, 2012)	-	✓	✓	-	DWT-Daubechies Wavelet
(Thomas et al., 2015)	-	✓	✓	-	DWT-Daubechies Wavelet
(Sudarshan et al., 2017)	-	✓	✓	-	DWT-Daubechies Wavelet
(Acharya et al., 2017)	-	✓	-	✓	DWT-Daubechies Wavelet
(Dolatabadi et al., 2017)	-	✓	✓	✓	PCA
(Kumari et al., 2014)	✓	-	✓	✓	DWT-Daubechies Wavelet
(Mumtaz et al., 2017)	✓	-	✓	✓	DWT-Daubechies Wavelet
(Kevric and Subasi, 2017)	✓	-	-	✓	DWT-Daubechies Wavelet
(Faust et al., 2015)	✓	-	✓	-	DWT-Daubechies Wavelet

## 2.4.1 Significance of the Wavelet Transformation on EEG and HRV Signals

Wavelet Transformations (WT) can be performed on time series data, including EEG and ECG signals to extract features. For EEG signals, various methods based on Discrete

Wavelet Transform (DWT) reported for removing noise by using WT. (Faust et al., 2015) presented a review of wavelet methods on continuous wavelet transform (CWT) and DWT for computer-aided seizure detection and epilepsy diagnosis. Their findings show that more scientific work carried out using the DWT methods than the CWT because EEG are discrete time series signals. (Kevric and Subasi, 2017) used Daubechies wavelet for the decomposition of the EEG signals for a classification task. (Mumtaz et al., 2017) also utilised Daubechies wavelet extracting features from frontal and temporal EEG data to improve the quality of life for major depressive disorder (MDD) patients. Their results suggested that the low frequencies of the EEG signals, such as delta and theta, may predict antidepressant's treatment outcome for MDD patients. (Kumari and Vaish, 2014) examined brain-wave energy features extracted by using Daubechies wavelet to differentiate one person from another. Their results found to differentiate a person from another person showing the energy distribution over the sub-bands of EEG signal corresponding to a delta, theta, alpha, beta and gamma waves.

For the HRV signal, various methods based on DWT reported for removing noise by using WT. (Sudarshan et al., 2017), extracted alarming features from the ECG using Daubechies wavelet to detect Congestive Heart Failure (CHF). Their results obtained accurate detection of CHF using only 2 seconds of ECG signal. (Acharya et al., 2017), developed an automated diagnostic system for the detection of Coronary Artery Disease (CAD) and Myocardial Infarction (MI) using three noises removing methods such as DWT, EMD and Discrete Cosine Transform (DCT). Their results indicated that detection of CAD is done more accurately with extracted features. (Thomas et al., 2015) proposed a novel technique using DWT based *Daubechies wavelet* for the automatic classification of cardiac arrhythmias, and indicated that the novel technique of feature extraction which utilised DWT based *db wavelet* for automatic heartbeat recognition, achieve better accuracy than the standard DWT method. (Kutlu and Kuntalp, 2012). Their results obtained success in discriminating five different ECG beats.

Based on the recent research, as shown in Table 2.3 on WT, it is straightforward to conclude that the DWT based methods are well known for EEG and ECG feature extraction and analysis. Among the DWT-based methods, *Daubechies wavelet* method has been considered by the researchers. Daubechies wavelet is the most popular wavelet family used for texture feature analysis, because of its support abilities with orthogonal support: the inverse of wavelet transform is the adjoint (a complex number with real part and an imaginary part equal in magnitude but opposite in sign) of the wavelet transform, and compact: the wavelet has a finite non-zero length which will result in increasing the probability of capturing events in short time instances. The Daubechies wavelet uses overlapping windows,

so the results reflect all changes between pixel intensities. The Daubechies D4 transform, for example, has four wavelet and scaling coefficients. The sum of the scaling function coefficients is also one; thus, the calculation is averaging over four adjacent data points. Although Wavelet Transformations (WT) have been performed on time series data including EEG and ECG signals, so far the correlation between WT signals has not been analysed (please refer to Chapter 7 for the correlation performance of EEG and HRV with and without WT signals.)

## 2.5 Summary

An underlying assumption of the most time series analysis is that EEG and ECG time series inherently possess dependence between adjacent observations. This dependence is of interest because it reveals information about the source producing the behaviour. In this way, time series analysis is essential for understanding variability within EEG and HRV, because time series analysis of EEG and HRV reveals how the system evolves. EEG and HRV signals represent complex dynamic behaviours of the biological system. Therefore, these endogenous electrical brain and heart signals need to be analysed further to understand the variation in a biological system.

SE has been used widely to investigate various biological conditions in the human body such as through ECG (Alcaraz and Rieta, 2010), HRV (Al-Angari and Sahakian, 2007), and EEG (Abásolo et al., 2006) determining sensitive changes. These studies have concluded that SE is a robust quantifier of complexity, which offers an accurate nonlinear metric for quantification (Alcaraz and Rieta, 2010). It gives a good dynamical signature and is a helpful tool that provides insights into various biological time series (Abásolo et al., 2006), (Ramdani et al., 2009), (Ramdani et al., 2009), (Al-Angari and Sahakian, 2007), (Yoo et al., 2012), (Takahashi et al., 2010). Therefore, SE is considered as an effective method for investigating different time series data.

A time series can be used to reconstruct the attractor of the underlying dynamic process. State-space reconstruction of a time series is a powerful approach for the analysis of the complex, nonlinear systems that appear ubiquitous in the natural and human world. The goal of an embedding Dimension (ED) is to create a state-space where the structure of a system is embedded. The technique False Nearest Neighbour (FNN) is the most commonly used method for finding an ED (Abarbanel et al., 1993), (Kennel et al., 1992), (Stergiou, 2004). I discuss FNN in Chapter 3. ED would be suitable to find the underlying variables for time series data such as Heart Rate Variability (HRV) since HRV might be driven by many underlying variables.

In recent research, the correlation between EEG signals has been analysed in the Frequency Domain (FD) using various methods, such as Mutual information (Na et al., 2002), Coherence analysis (Nolte et al., 2004), Wavelet coherence (Jeong et al., 2015), Correlation coefficient (Takahashi et al., 2005), Auto-correlation and Cross-correlation (Li et al., 2013). The focus of the analysis of EEG was on either left or right brain hemisphere when comparing the EEG activity from different part of the brain (Lewis et al., 1988), (Klem et al., 1999), (Na et al., 2002), (Jeong et al., 2015), (Li et al., 2013), (Cuevas and Bell, 2011), (Hevia-Orozco et al., 2017), (Almanza-Sepúlveda et al., 2018). In addition, to my knowledge, limited research has been conducted to analyse EEG signals in the Time Domain (TD) (Lewis et al., 1988), (Na et al., 2002), (Takahashi et al., 2005), (Bob et al., 2010), (Jeong et al., 2015), (Hevia-Orozco et al., 2017), (Almanza-Sepúlveda et al., 2018). In Chapter 6, I have covered the research gaps found to show the correlation between EEG signals for participants with and without medical conditions.

The recent research on the correlation between EEG and HRV has focused on Fourier analysis of the frequencies presents in these signals, to analyse their functionalities under certain conditions and to check whether these functionalities are related to each other. Research (Kim et al., 2013), (Chua et al., 2012), (Abdullah et al., 2009), (Sakai et al., 2007), (Berg et al., 2005), (Edlinger and Guger, 2006), (Sun et al., 2018), (Wang et al., 2019), (Doufesh et al., 2018), for example, suggesting the correlation between spectral bands of EEG and HRV has been conducted to assess the interaction between them, and remarkable correlation has been found. The WT of the signals is an important method not only to analyse EEG and HRV signals individually but also to analyse the correlation between them. According to recent research (Thomas and Moni, 2016), (Chandra et al., 2017), (Mirsadeghi et al., 2016), (Mporas et al., 2015), (Valderrama et al., 2012), (Nasehi and Pourghassem, 2011), (Cvetkovic et al., 2008), WT has been used to analyse either EEG or ECG signal, but the correlation between these transformed signals has not yet been conducted. In Chapter 7, I have shown both the correlation between WT signals and without WT signals.

# Chapter 3

## Data Description

### 3.1 Introduction

In this chapter, I am going to describe datasets used in my research. The chapter consists of two parts: the description of a dataset recorded by myself, and the description of datasets provided to me by David Mayor and Dr. Tony Steffert or obtained from the Internet. For EEG, all of these datasets follow 10%-20% electrode placement system, as shown (see Figure 2.7) and discussed in Chapter 2. In this chapter, some datasets contain both EEG and ECG signals, and some only EEG or ECG. However, what I meant by dataset is that it contains either both or one (EEG/ECG or EEG and ECG) signals.

### 3.2 Self Recording Dataset

This non-invasive dataset was collected for a joint study, conducted by a team of three people: myself, David Mayor <sup>1</sup> and Dr Tony Steffert <sup>2</sup>, to explore the effects of Electroacupuncture (EA). Transcutaneous Electrical Acupoint Stimulation (TEAS) is a type of EA. This EA is a treatment used in a wide variety of medical conditions and has beneficial effects on both cerebral and cardiac functions. The self-recording dataset was collected by using TEAS. TEAS is a safe, standardised acupuncture technique in which there is no needle insertion (V.Kaye, 1998-2019). As shown in Figure 3.1, it involves applying cutaneous electrical stimulation by placing electrodes at classical Chinese acupoints. A TEAS system consists of an electrical power unit connected by wires to one or more pairs of electrodes. I attach these electrodes to the participant's skin. When the power unit is switched on, a mild electrical current travels through the electrodes into the body. Participants may feel

---

<sup>1</sup><https://www.welwynacupuncture.co.uk/>

<sup>2</sup><http://qeeg.co.uk/>

localised tingling or warmth during treatment. A session typically lasts from five to thirty minutes, and treatments may be applied as often as needed.

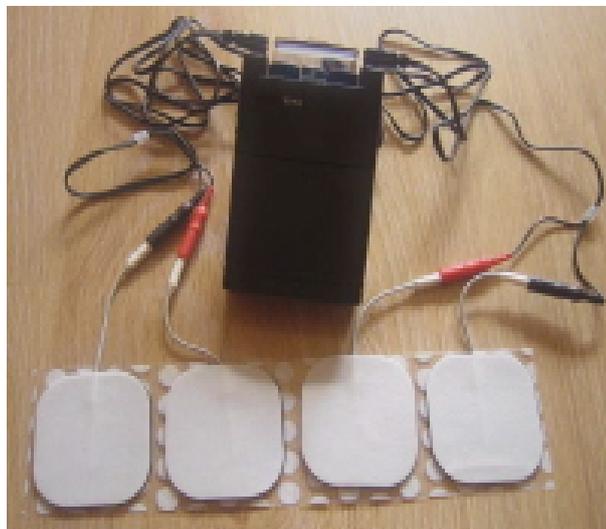


Figure 3.1: Transcutaneous Electrical Acupoint Stimulation (TEAS), a safe and standardized acupuncture technique, which does not require needles insertion (V.Kaye, 1998-2019)

The main aim of recording this dataset was to determine whether the electrical activity of the cerebral cortex (electroencephalograph, EEG) and heart (heart rate variability, HRV) change in different ways in response to different frequencies of TEAS applied to the hands of human participants.

### 3.2.1 Recording Process

For this study, I have been trained by an EEG specialist Dr Tony Steffert<sup>3</sup> working in this area for more than 20 years. This training took nearly two months to complete, including the understanding of the types of equipment being used, learning new software WinEEG and Biotrace to collect EEG and ECG data, record data, how to prepare myself before recordings, what precautions needed while recording, how to identify artefacts visually appearing on the system during recording.

In order to understand the recording process and to know a participants' feeling, I had been as a participant before recording this dataset.

The recording of this dataset was conducted in the premises of the University of Hertfordshire, approved by Health and Human Science (HHS) Ethics Committee with HHS Protocol Number HSK/SF/UH/00076. Once the ethical approval was granted, prospective participants have called and given opportunities to ask questions. During this conversation,

---

<sup>3</sup><http://qeeg.co.uk/>

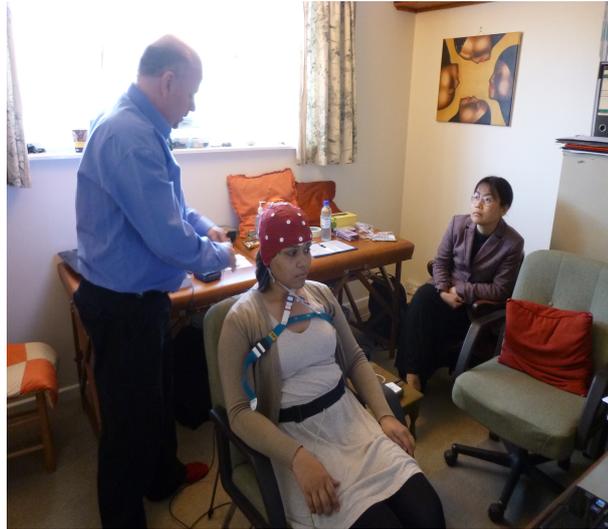


Figure 3.2: Visit to Acupuncture Clinic, showing myself as a participant, along with Dr. Tony Steffert (EEG Specialist and Researcher) and Dr. Na Helian (My Principal Supervisor).

they were asked to give their verbal agreements to proceed and to complete a 10-minute online questionnaire before attending their first sessions which were about a week later. There were certain requirements needed from participants before the recording. They are shown as follows:

1. To remove any bangles or bracelets and to keep still during measurements.
2. To wear comfortable ordinary clothing: not too tight anywhere.
3. To abstain from consuming caffeine, nicotine, alcohol or a heavy meal for at least two hours before attending for a session.
4. To wash hair not more than 12 hours before a session, and not to wear any hair products (such as conditioner) during the sessions.
5. To wear glasses rather than contact lenses during the EEG recording.
6. To avoid any strenuous activity for two hours before a session.
7. To bring with a list of any medication currently taking (including non-prescription drugs, nutritional supplements or herbal products).

People who had suffered a severe head injury in the past, suffered from epilepsy or diabetes, had cancer, wore an implanted electronic device, or dependent on psychoactive

medication were not eligible to take part in this study. Nor was anyone who was pregnant, or who had a condition in which they impaired peripheral circulation (such as Raynaud's syndrome), or who had any shoulder, arm or hand injury.

The steps used for recording, and some photos from the recording process of this dataset are available in Appendix A.

### 3.2.2 Dataset 1

This dataset consists of EEG and ECG recordings from 15 participants without any known medical conditions. The dataset was obtained for ten 5-minute time sequential slots in a single session in a relaxed state with eyes open. EEG and ECG recording were made simultaneously.

For EEG data, the sampling rate used was 250Hz (because of the storage rate of the Mitsar device (Mitsar-EEG-BT) used for EEG recording was 250Hz), and the reference was linked to ear electrodes (A1 and A2). 19 electrodes (Fp1, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, and O2) were used. A mild skin abrasive gel was used to abrade the scalp lightly and then an electrically conductive gel was applied to their scalp at the electrodes (both gels have been washed out of their hair easily).

For ECG data, the sampling rate was 256Hz (because of the storage rate of the Biotrace device used for ECG recording was 256Hz). The measurement of ECG involved detecting the heart electrical signals via two electrodes to record the electrical activity of the heart over time- one electrode was positioned on the volar surface of each forearm, with an additional electrode as a ground on the dominant side. However, during the data preprocessing steps we have encountered ECG data was very noisy, and we could not remove the noise (obtained from a faulty device) from the signals. Therefore, for this dataset, I have used Blood Volume Pulse (BVP) data, which was measured using the sensor attached to the participant's finger. According to (Kushki et al., 2011), Heart Rate Variability (HRV) can be estimated using the ECG or BVP. Chapter 2 describes the detail about HRV.

Each participant attended for a single session 85-minutes (50-minute recording, and 35-minute set-up time) sessions (in randomised, counterbalanced order), at intervals of 1-2 weeks. In each session, TEAS were applied to a different combination of points on the hands, at different frequencies by David Mayor <sup>4</sup> (Who is an acupuncturist, and has over 25 years experience in using TEAS). The intensity of stimulation was always adapted to what participants found comfortable.

---

<sup>4</sup><https://www.welwynacupuncture.co.uk/>

At the beginning of each recording, participants were required to sit in a comfortable chair with forearms supported and asked to complete two short questionnaires about medical conditions, medications, recent intake of food, drink, how well they slept, and their concern to use their data in research (10-minute) and a quick measurement scale of current feelings (5-minute). After this, the EEG electrode cap was positioned on the participant's head. Other sensors were attached to arms and fingers to record heart rate, skin blood flow, and temperature. Following an initial 5-minute baseline recording (with no acupuncture stimulation), TEAS were applied on both hands for 5-minute at one acupuncture stimulation frequency (2.5 Hz/10 Hz/80 Hz) in a randomised order for each participant, followed by a 10-minute break during which monitoring was continued. Halfway through this break, I asked participants to complete the same measurement scale as before and had the opportunity to move around to make themselves comfortable. Another 5-minute of stimulation at a different frequency was followed, again with a 10-minute break afterward during which they completed the measurement scale. This process was repeated for the third time, and then the various electrodes and sensors were removed. Finally, participants were asked to complete the initial questionnaire again, and any other feedback provided was recorded at this time as well.

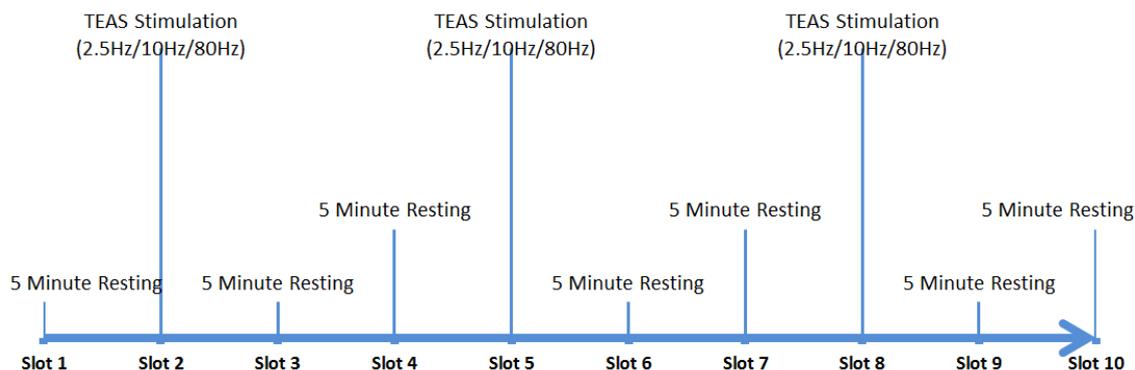


Figure 3.3: Dataset 1- Time line diagram showing slot's information

### 3.3 Other Datasets

In this section, I am going to give information for the other nine datasets: Datasets 2-10; Datasets 2, 3, 4 and 6 belong to EEG researcher Dr. Tony Steffert <sup>5</sup> and Acupuncturist David Mayor <sup>6</sup>(Steffert, 2018), (Steffert and Mayor, 2013), (Steffert and Mayor), (Mayor

<sup>5</sup><http://qeeg.co.uk/>

<sup>6</sup><https://www.welwynacupuncture.co.uk/>

and Steffert, 2016), (Mayor and Steffert, 2013), (Mayor and Steffert, 2016), (MAYoR), (Mayor and Steffert), who gave me their permission to use these datasets (Datasets 2, 3, 4, and 6) in my research; and Datasets 5, 7, 8, 9, and 10 was obtained from Internet.

### **3.3.1 Dataset 2**

This dataset consists of EEG and ECG recordings from 7 participants with no prior history of neurological or psychiatric disorders, learning disabilities, drug abuse, or chronic illness. Participants were asked to refrain from drinking caffeine or alcohol during the 12 hours prior to the recording sessions, and to arrive with clean, dry hair. The main aim of recording this dataset was to determine whether the electrical activity of the cerebral cortex (EEG) and heart (HRV) change in different ways in response to different frequencies of TEAS method applied to four different body locations (Left Hand, Below Left Knee, Right Hand, and Below Right Knee). These data were obtained over ten 5-minute slots with eyes open, including resting state data in the first and the last slot. The EEG and ECG recording were made simultaneously.

For EEG data, the sampling rate used was 250Hz, and the reference was linked to ear electrodes. 19 electrodes (Fp1, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, and O2) for EEG recording were used.

For ECG data, the sampling rate was 256Hz. Measurement of heart rate variability (HRV) involves detecting the heart's electrical signals via two electrodes to record the electrical activity of the heart over time; one electrode was positioned on the volar surface of each forearm, with an additional electrode as the ground on the dominant side.

The EEG and ECG recording are 50-minute long, dividing data into ten 5-minute slots: three baseline slots and six acupuncture stimulation slots. The stimulation parameters (e.g., body location) are kept constant within each intervention but varies between interventions. Each participant visited twice, during which the TEAS stimulation of either 2.5Hz or 10Hz is applied (randomised order used) at four different body locations (Slot 3 to 8) with eyes closed. The baseline measurements are slots 1, 2, and 9.

### **3.3.2 Dataset 3**

This dataset consists of EEG and ECG data of 12 participants with no prior history of neurological or psychiatric disorders, learning disabilities, drug abuse, or chronic illness. Participants were asked to refrain from drinking caffeine or alcohol during the 12 hours prior to the recording sessions, and to arrive with clean, dry hair. This dataset is derived using both Electroacupuncture (EA) and Manual Acupuncture (MA) method of acupuncture

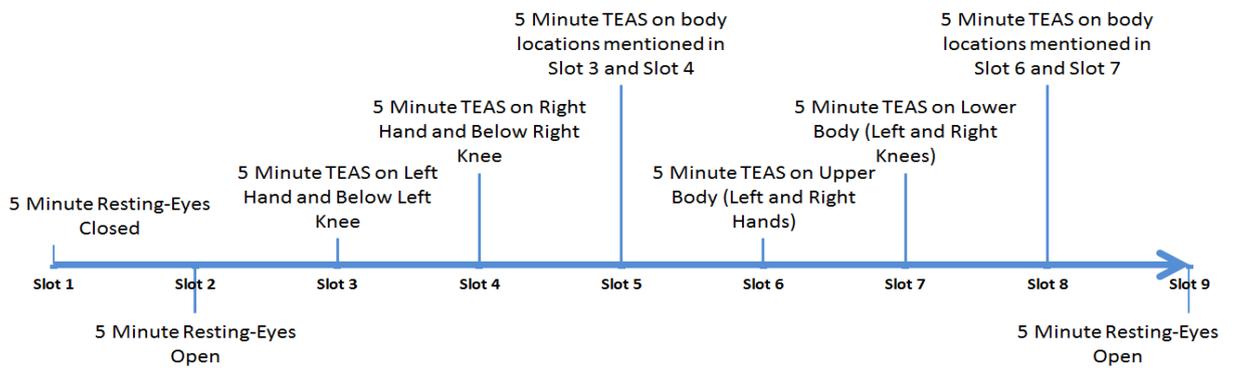


Figure 3.4: Dataset 2- Time line diagram showing slot's information

in turn. The main aim of recording this dataset was to determine whether the electrical activity of EEG and heart HRV change in different ways in response to different frequencies of EA and MA methods applied. EA is an acupuncture method, with needles being inserted at specific points on the body (Dr.Evans, 1998-2019), as shown in Figure 3.5. The needles are then connected to a device that generates continuous electric pulses. These devices are used to adjust the frequency and intensity of the impulse being delivered, depending on the condition being treated. EA uses pairs of needles so that the impulses can pass from one needle to the other. Manual Acupuncture (MA) is an acupuncture method, similar to EA, needles are inserted at specific points on the body (Dr.Xie, 1998-2019). As shown in Figure 3.6, instead of passing electric pulses through the needle, needles are twisted by or otherwise manipulated by an acupuncturist for MA.

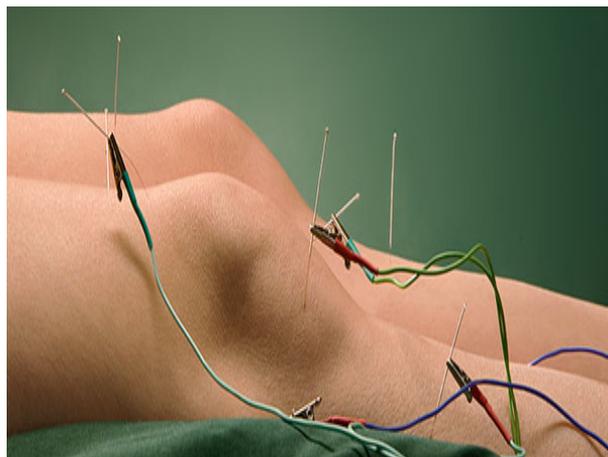


Figure 3.5: Electro acupuncture (EA) technique (Dr.Evans, 1998-2019).

For EEG data, the sampling rate used was 250Hz, and the reference was linked to ear electrodes (A1 and A2). 19 electrodes (Fp1, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, and O2) for EEG recording were used.



Figure 3.6: Manual Acupuncture (MA) technique (Dr.Xie, 1998-2019)

For ECG data, the sampling rate was 256Hz. Measurement of heart rate variability (HRV) involves detecting the heart's electrical signals via two electrodes to record the electrical activity of the heart over time; one electrode was positioned on the volar surface of each forearm, with an additional electrode as the ground on the dominant side.

All participants attended four visits, during each visit stimulation performed at four different locations (in randomised order): Right (Below Right Knee and Right Hand), Left (below Left Knee and Left Hand), Upper Body (Right and Left Hands) and Lower Body (below Left and Right Knees). EEG and ECG monitoring was carried out in eight 5-minute sequential slots with stimulation at a single location: EA stimulation of 2.5Hz, 10Hz, 20Hz, and 80Hz is applied (Slot 3 to 6), MA stimulation applied in two slots (Slot 2 and Slot 7), and baseline measurements are slots 1 and slot 8.

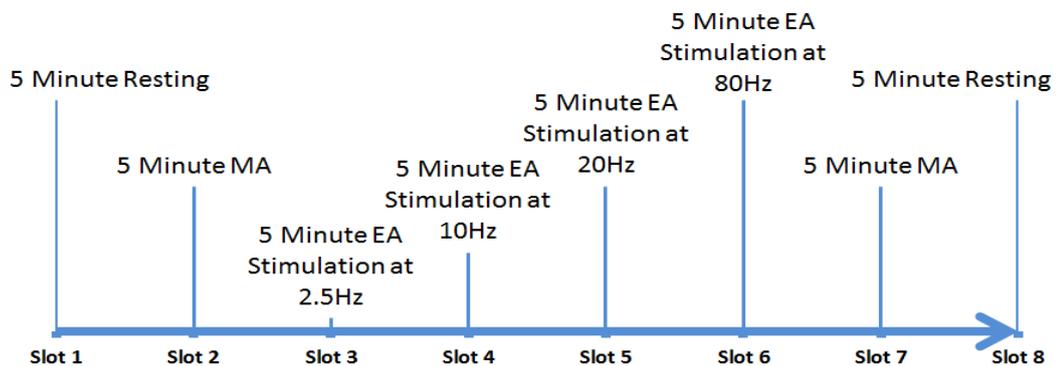


Figure 3.7: Dataset 3- Time line diagram showing slot's information

### 3.3.3 Dataset 4

This dataset consists of scalp EEG recordings from 20 participants, while they watched a short documentary movie. All participants were healthy with no prior history of neurological or psychiatric disorders, learning disabilities, drug abuse, or chronic illness. Participants were asked to refrain from drinking caffeine or alcohol during the 12 hours prior to the recording sessions, and to arrive with clean, dry hair. This dataset was obtained with four slots (three 3-minute slots, and one 60-minute slot) in a relaxed state. The main aim of recording this dataset was to determine whether EEG activity change with different emotions shown in the film.

The sampling rate of this dataset was 500Hz, and the reference was linked to ear electrodes (A1 and A2). 10 electrodes (F7, F3, Fz, F4, F8, T5, P3, Pz, P4, and T6) were used.

The EEG recording is 69-minutes long, which is divided into four slots: three baseline slots and one film-watching slot. Each participant visited once, during which 1-hour long documentary film was played. The recordings of 20 participants took 4 days to complete—recording two sessions a day with each session collecting EEG recordings of three participants simultaneously. The baseline measurements are slots 1, 2, 4, and 60-minute slot watching a film is slot 3.

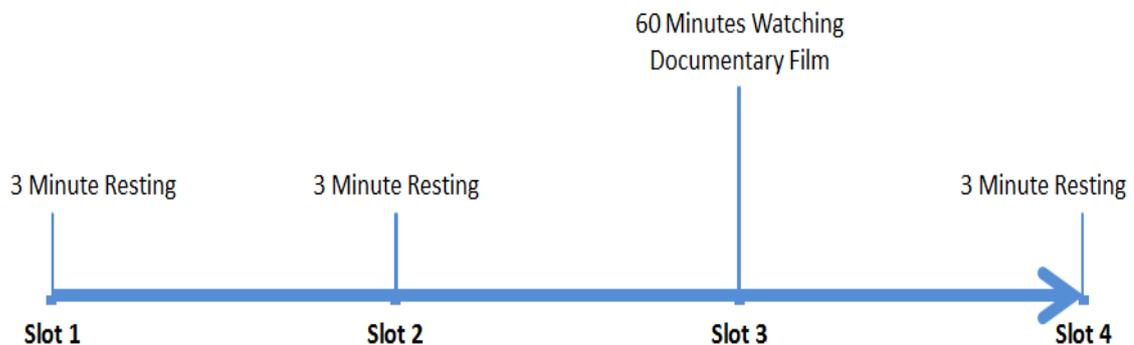


Figure 3.8: Dataset 4- Time line diagram showing slot's information

### 3.3.4 Dataset 5

This dataset consists recordings of EEG signals for the analysis of human affective states (Koelstra et al., 2012). 32 participant's EEG signals and peripheral physiological signals were recorded while the participants watched forty 1-minute long excerpts of music videos. Participants rated each video in terms of the levels of arousal, valence, like/dislike, dominance, and familiarity. For 22 of the 32 participants, the frontal face video was also

recorded. In my thesis, I have only utilised EEG signals of 32 participants from this dataset. The main aim of this dataset was to detect emotional cues occurring during human-computer interaction and synthesizing emotional responses.

For EEG data, the sampling rate used was 512Hz, and the reference was ear electrodes. This dataset contains 48 recorded electrodes (32 EEG electrodes, 12 peripheral electrodes, 3 unused electrodes, and 1 status electrode). As they used the standard EEG electrode names (according to the 10%-20% system) I was able to identify 15 electrodes (from the standard EEG electrodes name from 10%-20% system) detail based on the information provides (Koelstra et al., 2012). These are the 15 electrodes (Fp1, Fp2, F7, F3, Fz, F4, F8, C3, Cz, C4, P3, Pz, P4, O1, and O2).

Furthermore, The data were recorded in two separate locations. Participants 1-22 were recorded in Twente and participants 23-32 in Geneva. The EEG electrodes' numbering was slightly different. Based on dataset information provided on <sup>7</sup>, I could identify the difference and used the appropriate channels accordingly. More information about this dataset can be found on the link above, and in (Koelstra et al., 2012), where they utilised this dataset for the first time.

### **3.3.5 Dataset 6**

This dataset consists of EEG recordings from 13 participants with medical condition Autism. Autism is a mental condition, and presented from early childhood, is characterized by great difficulty in communicating and forming relationships with other people and in using language and abstract concepts (Frith, 2003). This dataset was recorded with a 5-minute time slot in a relaxed state with eyes opened.

The sampling rate of this dataset was 250Hz, and the reference was linked to ear electrodes (A1 and A2). 19 electrodes (Fp1, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, and O2) were used.

### **3.3.6 Dataset 7**

This dataset consists of EEG recording from 5 participants with medical condition Epilepsy. Epilepsy is a neurological disorder marked by sudden recurrent episodes of sensory disturbance, loss of consciousness, associated with abnormal electrical activity in the brain (Lennox, 1960). This dataset was obtained from machine learning repository (Dheeru and Karra Taniskidou, 2017), and belongs to (Andrzejak et al., 2001). This dataset was recorded with a 5-minute time slot. The aim of this dataset was to analyse sets of EEG time series:

---

<sup>7</sup><https://www.eecs.qmul.ac.uk/mmv/datasets/deap/readme.html>

surface EEG recordings from healthy participants with eyes closed and eyes open, and intracranial EEG recordings from epileptic patients during the seizure free interval from within and from outside the seizure generating area as well as intracranial EEG recordings of epileptic seizures (Andrzejak et al., 2001).

The sampling rate of this dataset was 173Hz, and the reference was linked to ear electrodes (A1 and A2). An EEG cap of 100 Electrode had been used, out of which 19 electrodes (Fp1, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, and O2) following the standard 10%-20% system (Klem et al., 1999), were selected in my research. More Information about this dataset is available at <sup>8</sup>.

### **3.3.7 Dataset 8**

This dataset contains EEG recording from 22 participants with medical condition Seizure (Goldberger et al., 2000). A Seizure is a sudden, uncontrolled electrical disturbance in the brain. It can cause changes in behaviour, movements or feelings, and in the levels of consciousness. If a participant has two or more seizures or a tendency to have recurrent seizures, then they might have epilepsy (MayoClinic, 1998-2018). This database, collected at the Children’s Hospital Boston, consists of EEG recordings from pediatric participants (5 males, ages 3–22; and 17 females, ages 1.5–19) with intractable seizures. Participants were monitored for up to several days following withdrawal of anti-seizure medication in order to characterize their seizures and assess their candidacy for surgical intervention (Goldberger et al., 2000). This dataset consists of twelve 5-minute time slots.

The sampling rate used was 256Hz, and the reference was linked to ear electrodes (A1 and A2). 19 electrodes (Fp1, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, and O2) were used. This dataset was obtained from PhysioNet Database, more information about this dataset is available at <sup>9</sup>.

### **3.3.8 Dataset 9**

This dataset contains ECG recordings from 10 participants; including 5 participants without any medical condition; and 5 participants with medical condition Congestive Heart Failure (CHF). The aim of this dataset was to determine whether normal heart rate is chaotic and to identify the difference between participants with CHF and without CHF. CHF occurs when the heart cannot pump sufficient, to maintain blood flow to meet the needs of the body.

---

<sup>8</sup><http://epileptologie-bonn.de/cms/upload/workgroup/lehnertz/eegdata.html>

<sup>9</sup><https://archive.physionet.org/pn6/chbmit/>

Symptoms commonly include shortness of breath, tiredness and leg swelling. This dataset was obtained from PhysioNet Database <sup>10</sup>, and belongs to (Goldberger et al., 2000).

Each ECG time series is about 24-hours long (roughly 100,000 intervals). One ECG electrode (placed on the left chest) was used to record the ECG signals. The sampling rate used for ECG data is 256 Hz. More information about this dataset can be found on <sup>11</sup>.

### 3.3.9 Dataset 10

This data set consists of ECG recordings from 5 women in labour, between 38 and 41 weeks of gestation, with cardiac condition Fetal echo-cardiography (Fetal ECG). Fetal ECG is the name of the test used to diagnose cardiac conditions in the fetal stage. Cardiac defects are amongst the most common congenital disabilities. The diagnosis in the fetal stage is important because it might provide an opportunity to plan and manage a baby when the baby is born. The main aim of this dataset was to assess reliability of indirect abdominal electrocardiography as an alternative to the commonly used Doppler ultrasound monitoring technique. I have obtained this dataset from PhysioNet Database (Goldberger et al., 2000), (Matonia et al., 2006).

The sampling rate used for ECG data was 1000 Hz (1 kHz). Each recording comprises four different signals acquired from the maternal abdomen and the reference direct fetal electrocardiogram registered from the fetal head. Four ECG electrodes were placed around the navel, and a reference electrode placed above the pubic symphysis and a common mode reference electrode (with active-ground signal) placed on the left leg. In all cases, the scalp electrode was placed for a clinical indication, and all women consented to participate in this study. More information about this dataset can be found at <sup>12</sup>

## 3.4 Summary

In this chapter, I have presented the description of datasets I have used in my research, dividing them into two parts: Self Recording Dataset, Other Datasets. I will use these datasets in the chapters, as shown in Table 3.1.

The differences among these datasets, other than participants and the number of electrodes are:

---

<sup>10</sup><https://archive.physionet.org/challenge/chaos/>

<sup>11</sup><https://physionet.org/challenge/chaos/>

<sup>12</sup><https://physionet.org/physiobank/database/adfecgdb/?C=D;O=A>

1. Body location where TEAS stimulation has been performed. For example, for Dataset 2, four different body location (Left Hand, Below Left Knee, Right Hand, and Below Right Knee) has been used to perform TEAS stimulation, and for Dataset 1, and Dataset 3 only one body location (Dominant Hand),
2. Stimulation technique used. For Example, Dataset 1, Dataset 2 used TEAS, whereas Dataset 3 used EA and MA both, the rest of the datasets did not use any stimulation,
3. Total time length and individual slot length is also different for each of them,
4. Medical conditions: Datasets 1-5 contains participants without any medical conditions, whereas Datasets 6-10 contains participants with medical conditions, such as Autism, Epilepsy, Seizure, Congestive Heart failure, and fetal ECG, respectively.

Last and most important is the utilisation of these datasets in individual Chapters in my thesis, as summarised in Table 3.1. For Example, Dataset 1 is used in Chapter 6, 7, and 8. Dataset 4, Dataset 5, and Datasets 6-8 are used in Chapter 6. Dataset 2 is used in Chapter 7 and 9. Finally, Dataset 3, and Datasets 9-10 are used in Chapter 9.

Table 3.1: Summary of Datasets and detail of their utilisation.

Label	Dataset 1	Dataset 2	Dataset 3	Dataset 4	Dataset 5	Dataset 6	Dataset 7	Dataset 8	Dataset 9	Dataset 10
Number of Participants	15	7	12	20	32	13	5	22	10	5
EEG-Electrodes	19	19	19	10	15	19	19	19	None	None
EEG-Sampling Rate	250 Hz	250 Hz	250 Hz	500 Hz	512 Hz	250 Hz	173 Hz	256 Hz	None	None
EEG-Electrode	1	1	1	None	None	None	None	None	1	5
EEG-Sampling Rate	256 Hz	256 Hz	256 Hz	None	None	None	None	None	256 Hz	1000 Hz
Stimulation Technique	TEAS	TEAS	EA and MA	None	None	None	None	None	None	None
Stimulation Locations	1	4	1	None	None	None	None	None	None	None
Total Time Length	50 minutes	45 minutes	40 minutes	70 minutes	40 minutes	NA	NA	1 Hour	24 Hours	NA
Slot Time Length	5 minutes	3 and 5 minutes	40 minutes	5 minutes	5 minutes	NA	NA	NA	NA	NA
Medial Condition	None	None	None	None	None	Autism	Epilepsy	Seizure	CHF	Fetal ECG
Utilised in Chapter	Chapters 6,7, and 8	Chapters 7,9	Chapters 9	Chapters 6	Chapters 6	Chapter 6	Chapter 6	Chapter 6	Chapter 9	Chapter 9

NOTE: NA means Not Available.

# Chapter 4

## Time Series Data Preprocessing

### 4.1 Introduction

Data preprocessing is to keep valuable information about the raw data excluding noise. The effect of how well data are processed can be found in the analysis process. Without preprocessing, the analysis of the data won't be accurate and might be misleading. Therefore, in this chapter the focus is to describe what is data preprocessing, why is it important, preprocessing techniques used for time series data, and in particular for EEG and HRV time series data are discussed. Furthermore, well-known data preprocessing techniques for time series data, such as Independent Component Analysis (ICA), Fast Fourier Analysis (FFT), and Wavelet Transform (WT), are described, and illustrated along with an example. The aim of this chapter is to provide thorough overview of data preprocessing techniques with concrete examples, which could be very helpful to understand how the different techniques work, how can I justify the preprocessing outcome obtained for my dataset and how can I be sure to conclude my research contributions.

#### 4.1.1 Data Preprocessing and Its Importance

Data analysis is the basis for investigations in many fields of knowledge, from science to engineering and from management to process control. Data on a particular topic acquired in the form of symbolic and numeric attributes. The source of these data varies from human beings to sensors with different degrees of complexity and reliability. Analysis of these data gives a better understanding of the phenomenon of interest. The main objective of any data analysis is, therefore, to discover knowledge used to solve problems or make decisions. However, problems with data may prevent this. Therefore, data preprocessing analyse data intelligently. Data preprocessing might be a time-consuming task, and may be performed on the data to understand the nature of the data by performing a more meaningful data

analysis, and extracting more meaningful knowledge. In most applications, there is a need for more than one form of data preprocessing. Identifying types of data preprocessing are, therefore, a crucial task.

Noise in data can be attributed to several sources, noise added by amplifiers and signal conditioning circuitry, jittering in the sampling device, non-linearities and quantisation noise in the analogue-to-digital (A/D) converter, extraneous noise picked up from the environment (Rubel et al., 1991), and data transmissions between channels and sensor thresholds (upper/lower).

When data exhibits too much variation or non-stationary behaviour, the use of time series models may provide a reliable approach. Variation is present in data from most application domains. For example, the variation could be because of: (i) process equipment, (ii) raw material used, (iii) process environment, (iv) human operating procedures and (v) individual decisions or process plans (Yarling, 1993). In most process monitoring applications, time series means transforming data into a static collection of features that represent a view of the operation.

#### **4.1.2 Preprocessing Techniques for Time Series Data**

Moving Average (MA) smooths data by replacing each data point with the average of the neighbouring data points. The method is based on the idea that large irregular component will exert a smaller effect with its immediate neighbours when averaged. MA smooths the data and makes it easier to spot a trend, removing short term noise from a data set (Shumway and Stoffer, 2011). MA can be used for time series data such as economical, signal processing and financial data.

Regression Analysis (RA) is the process of estimating the relationships among variables (a dependent variable and one or more independent variables), which is used for prediction and forecasting (Shumway and Stoffer, 2011). RA helps to understand how the typical value of the dependent variable changes when any one of the independent variables varies, while the other independent variables are held fixed. The theory associated with linear regression is well understood and allows for the construction of different types of easily interpretable statistical intervals for predictions, calibrations, and optimizations.

Singular Spectrum Analysis (SSA) is a non-parametric spectral estimation method. SSA combines elements of classical time series analysis, multivariate statistics, multivariate geometry, dynamical systems and signal processing (Elsner and Tsonis, 2013). SSA, aiming to answer following: what time series components can be separated by SSA, and how to choose the window width and make proper grouping for extraction of a desirable

component (Golyandina et al., 2001). SSA can be used for classical time series analysis, multivariate statistics (observation and analysis of more than one outcome variable), multivariate geometry, dynamical systems and signal processing data.

Morphological Component Analysis (MCA) is the method of decomposing a signal into its components. This method is based on the assumption that every signal component has a different shape that enables its reconstruction using sparse representation. Each component is sparsely represented by different bases (Discrete Cosine Transform (DCT), wavelet and Dirac basis). MCA separates features that present different morphological aspects in an image, and it can be deemed as a fast and simple basis pursuit (Gao et al., 2010), in which 1) its dictionary is a concatenation associated with a fast transformation and 2) constraints can be easily imposed on decomposed components. MCA can be applied to single-channel EEG signals to remove artefacts (Yong et al., 2009) even though it has been shown to have interesting applications in image inpainting (Elad et al., 2005) and Magnetoencephalography (MEG) signal decomposition (Ozkurt et al., 2007).

Principal Component Analysis (PCA) transforms a set of multivariate data with correlated components into a set of uncorrelated components by finding the orthogonal direction of the largest variance in the EEG signals. PCA can be used for image processing, biological data, telecommunication, audio processing, etc. PCA finds orthogonal directions of greatest variance in the data. For example, a method to remove eye artefacts from multi-channel EEG It explains the maximum amount of variance with few components (Duszak and Koczkodaj). Independent Component Analysis (ICA) observes that random data are linearly transformed into components that are maximally independent of each other and simultaneously have interesting distributions (Comon, 1994). ICA is an application that can be found in many areas, such as audio processing, biomedical signal processing, image processing, telecommunications, and econometrics. ICA can be formulated as the estimation of a latent variable model. This method can be used to detect and remove a wide variety of artefacts (including eye blinks, muscle noise, heart signal, and line noise) from spontaneous EEG data.

Fast Fourier Transform (FFT) is among the most common methods of noise modelling for data preprocessing. Classic Fourier transform/spectral analyses signals in terms of frequency components among the whole spatial domain, which loses time localisation. The Fourier transform is therefore appropriate for long time periodic signals. Short time window Fourier transform uses a set of window functions to restrict transform length and can be used to provide better time localisation. The popularity of FFT is evidenced by the wide variety of application areas. In addition to conventional radar, communications, sonar, and speech signal processing application, the current field of FFT usage include biomedical

engineering, imaging, analysis of stock market data, spectroscopy, metallurgical analysis, nonlinear system analysis, mechanical analysis, and so on (Brigham and Brigham, 1988).

The Wavelet Transform (WT) is similar to the Fourier transform (or much more to the windowed Fourier transform) with a completely different merit function. The main difference is this: Fourier transform decomposes the signal into sines and cosines, i.e. the functions localised in Fourier space; in contrary the wavelet transform uses functions that are localised in both the real and Fourier space (Bultheel et al., 1995). WT is a mathematical tool that can be used for extracting information from a variety of data forms, such as image and audio signals (Lee and Lim, 2012). The theory of wavelet is utilised as an essential technique in specialised research in electronics, mechanics, computers, communications, medicine, biology, astronomy and so on (Addison, 2017). It is important in the deconstruction of non-stationary and/or non-periodic signals. Therefore, it has been used in different places in signal processing to compress and de-noise data (Sifuzzaman et al., 2009).

For time series data such as EEG and HRV, certain preprocessing techniques work better than others, and they are described in the following section.

#### **4.1.3 Preprocessing Techniques for EEG and HRV Time Series Data**

In the previous section general preprocessing techniques for time series, in general, are described. Some preprocessing technique suits to analyse EEG and ECG signals, but not all of them. For example, MCA is the best technique when analysing images, MA and RA are perfect for economic data and when prediction are needed. ICA, FFT and WT would be most appropriate techniques for EEG and ECG, not just because of their extensive use by researchers, but because of their ability to analyse signals in both domains time and frequency.

The main goal of identifying PCA is to select proper attributes. However, due to some drawbacks with PCA, such as it cannot completely separate eye movements from EEG signals, especially when they have comparable amplitudes. It does not recover original signals because it only uses covariance. Whereas ICA is computationally efficient than PCA (Comon, 1994), it can simultaneously separate the EEG and artefacts into independent components without relying on the availability of reference artefacts. It avoids the problem of mutual contamination between EEG and Electrooculography (EOG) channels that could not be solved with filters, regression and PCA. ICA is able to separate superimposed signals into components having different statistical characteristics. It is applicable to removal of the wide variety of EEG artefacts, and it can find original co-ordinate (Beckmann et al., 2005), (Makeig et al., 1996).

Fourier/spectral analysis of a signal involves decomposition of the signal into its frequency (sinusoidal) components. In other words, the original signal can be separated into its sub-spectral components by using spectral analysis methods (Akin et al., 2000). Among spectral analysis techniques, FFT is considered being the best transformation between time and frequency domains because of it being time shift invariant (Proakis et al., 1992).

The fact is that EEG signals, are nonstationary, and by the use of the FFT, small changes may not be realised and the analysis may change depending on the length of data. So, for spectral analysis, it can be said that WT is more suitable than FFT (Akin, 2002). The reason for this success depends on the scaling and the shifting properties of the mother wavelet (Lee and Yamamoto, 1994). Another advantage of the WT is a 3 dimensional representation of signals as amplitude, frequency, and time.

For EEG signals, various methods based on discrete wavelet transform (DWT) are reported for the pre-processing (Faust et al., 2015), (Kevric and Subasi, 2017), (Mumtaz et al., 2017), (Kumari and Vaish, 2014), (Kumari et al., 2014). Their results recommended that more scientific work has been carried out using the DWT methods than the Continuous Wavelet Transform (CWT) because EEG and ECG are discrete signals. These research suggests, that using WT, it is possible to differentiate a person from another person, based on the coefficient of variation over each brain region such as frontal, cerebral, parietal and occipital. For ECG signal, various feature extraction methods based on DWT are reported (Sudarshan et al., 2017), (Acharya et al., 2017), (Dolatabadi et al., 2017), (Thomas et al., 2015), (Kutlu and Kuntalp, 2012). WT has been successfully utilised to extract features from ECG data, For example, for accurate detection of congestive heart failure (CHF) (İşler and Kuntalp, 2007), coronary artery disease (CAD) (Giri et al., 2013) and myocardial infarction (MI) (Jayachandran et al., 2010), cardiac arrhythmias (Khadra et al., 1997). Their results suggest, that the detection of heart disorders can be done more accurately with DWT.

In summary, the different types of preprocessing techniques which have been used for EEG and ECG are *Fast Fourier Analysis (FFT)*, *Independent Component Analysis (ICA)*, and *Discrete Wavelet Transform (DWT)*. In this chapter, the explanation of these methods, how these methods work and why they are best for preprocessing EEG and ECG times series are available. In addition, it also describes each method along with an illustrative example.

## 4.2 Independent Component Analysis (ICA)

Independent Component Analysis (ICA) is a blind source separation (BSS) method that finds statistically independent and non-Gaussian components of multivariate data. The target of ICA is to separate non-Gaussian signals. The concept of the ICA was first proposed by (Jutten and Herault, 1991). In the early 1990s (Tong et al., 1991b), and (Tong et al., 1991a) conducted a more systematic study of the uncertainty of the BSS problem solution, that is identifiable. The basic of ICA for solving BSS problems were defined by (Comon, 1994), assuming that the components recovered from the mixed signals using appropriate linear transformations are statistically independent of each other.

ICA is mainly used to solve BSS problems of finding signals from observations (more than 1 unknown source). In the cases where the source signals and the mixed signals are unknown, assuming that the source signals are statistically independent ensures that ICA can separate the source signals from the mixed signals well. In practical problems, the assumption of independence is reasonable, therefore, ICA can be applied in many fields. The reason for using ICA in this work is to remove artefacts from EEG. There are different types of artefacts presenting EEG signals. For example, muscle movement includes: chew, swallow, facial muscle, and so on. These artefacts usually contain high amplitude and very fast activities. For example, muscle artefacts contain a greater number of amplitudes. For facial muscle and scalp movements, the amplitudes should be higher than of clinical EEG and too fast to be visually estimated. These artefacts most commonly occur in frontal and temporal electrodes. For chew and swallow movement, the artefacts have a wide field with maximal amplitude frontally and comprise isolated slow waves, Delta frequency range activity, or more typically slowing with faster frequencies. Therefore, the ICA focusses to remove the artefacts from signals.

To illustrate the fundamental principle of ICA, let's consider a cocktail party example, where two groups are talking, and their conversations are being recorded using two microphones. Let's consider signal  $S1(t)$  and  $S2(t)$  are being recorded from two microphones  $X1(t)$  and  $X2(t)$ . However, during these recordings, only the mixing of these two signals can be recorded (for example from Microphone 1: a bit of signal  $S1(t)$  and a bit of signal  $S2(t)$ ).

Two recordings from microphone  $X1(t)$  and  $X2(t)$  are as below:

$$\begin{aligned}X1(t) &= a_{11} \cdot S1(t) + a_{12} \cdot S2(t); \\X2(t) &= a_{21} \cdot S1(t) + a_{22} \cdot S2(t); \end{aligned}$$

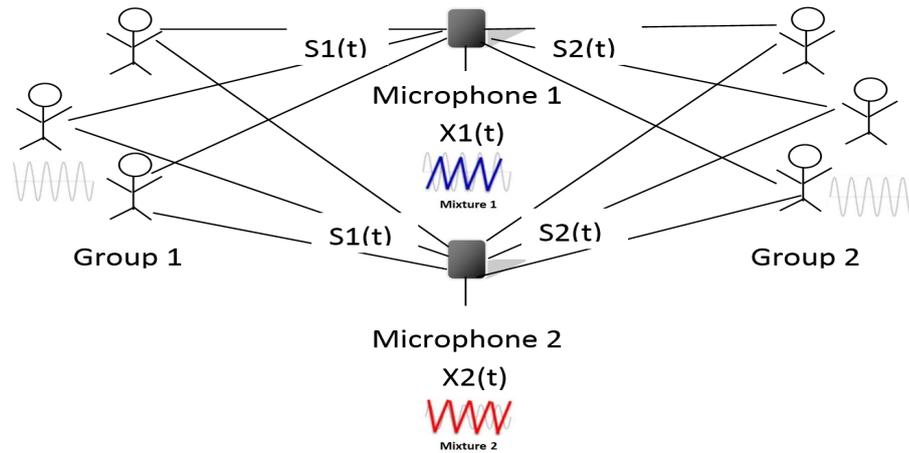


Figure 4.1: Illustration of BSS problem that Independent Component Analysis (ICA) addresses using the popular cocktail-party problem. The goal is to recover the individual signals  $S1(t)$  and  $S2(t)$  from the mixtures signals  $X1(t)$  and  $X2(t)$  measured from Microphone 1 and 2, respectively.

Where,  $a_{ij}$  is mixing coefficient.

Now the question is can we, with the two measurements, separate these two signals  $S1(t)$  and  $S2(t)$ . As shown in Figure 4.1, signals  $X1(t)$  and  $X2(t)$  are linear combinations of signals  $S1(t)$  and  $S2(t)$ . Now, what we have is  $X1(t)$  and  $X2(t)$ , but what we want is  $S1(t)$  and  $S2(t)$ . In practice, we do not really know what  $a_{ij}$  coefficients are. The technique needed here is not how to get  $a_{ij}$ , but to get approximate  $a_{ij}$ . Once we know the approximate mixing coefficient matrix  $A$ , then we can easily get our signals  $S1(t)$  and  $S2(t)$ . The determination of matrix  $A$  can be dependent on the actual location of the microphones.

Apart from the cocktail party problem, ICA application includes Radar detection, EEG analysis, and so on. In Radar detection, we aim at detecting the returning signal which we sent out bounces on an aircraft. The problem occurs when more than one aircraft multi-returning signals. In this situation, we have to figure out how to pick up the target aircraft from all the objects in the air. For EEG, each electrode measures the brain activities collectively. In practice, we need to separate different activity signals that are going around the brain.

The formal framework for ICA is, given  $N$  distinct linear combinations of  $N$  signals, determine the original  $N$  signal. More general settings for the cocktail party problem can be described here:

$$X = A \cdot S$$

Where,  $X$  is the measurement used,  $S$  is the signal,  $A$  is the coefficient matrix:

$$X = \begin{pmatrix} X_1 \\ X_2 \\ \cdot \\ \cdot \\ X_n \end{pmatrix} \quad S = \begin{pmatrix} S_1 \\ S_2 \\ \cdot \\ \cdot \\ S_n \end{pmatrix} \quad A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

$U$  and  $V$  are unitary matrix (unitary matrix is the one, if its conjugate transpose is equal to its inverse  $A^* = A^{-1}$ , please refer (Periwal and Shevitz, 1990) for more detail) to make rotation, and  $\Sigma$  is a stretching matrix (Sparse matrix method that makes matrices sparser by making them larger (Grcar, 2012), (Pissanetzky, 1984)) to stretch according to principle components (set of values of linearly uncorrelated variables (Johnson and Bjordal, 2011)) (Comon, 1994). In the context of ICA, according to  $X = A \cdot S$ , this will become  $X = (U \cdot \Sigma \cdot V^T) \cdot S$ , where  $U \cdot \Sigma \cdot V$  is the singular value decomposition matrix  $A$ ,  $^T$  is the transpose, and  $S$  is the signal. Once we know  $U$ ,  $V$  and  $\Sigma$ , the process for ICA becomes easy, because we just need to get our independent signals considering  $S = (U \cdot \Sigma \cdot V^T)^{-1} \cdot X$ , which becomes  $S = A^{-1} \cdot X$ .

The task is to recover the source signals given no advance knowledge of the nature of the sources or of the mixing process. To do this, it is necessary to find a square matrix,  $A$ , specifying a filter that linearly inverts the mixing process. The key assumption used to identify source signals from measures is that sources,  $S_i$ , are statistically independent.

Let's use Figure 4.2 to illustrate the fundamental concept of ICA:

The problem here is we don't know the values for matrix  $A$ . All we are given is the mixed unitary matrix  $U$  resulted from rotating  $V$ , stretching  $\Sigma$  of the signal  $S$ . Once, we have mixed unitary matrix  $U$ , all we need to do is get the maximum variance along the direction to find the angle  $\Theta$  (between 0 to 90 degrees) of the data. Once, we have found  $\Theta$ , we can rotate it back to  $U^T$ , and  $\Sigma^{-1}$ , calculating the variance. The most tricky part is how to separate the probability distribution (p).

The problem with recovering the source signals is that we don't know the values of matrix  $A$ , all we need to consider is the mixed matrix  $U$  for ICA. Once, we have mixed matrix  $U$ , using PCA first get the maximum variance in the data to find the  $\Theta$  of the data.

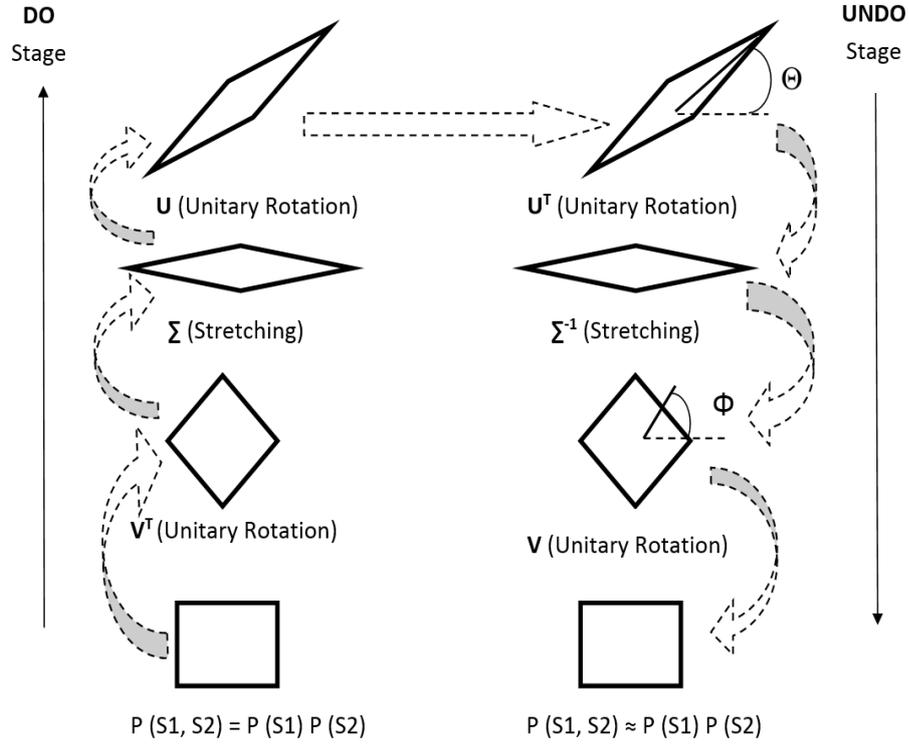


Figure 4.2: Illustration of Independent Component Analysis (ICA) showing two stage process: Do stage (left hand side of the figure), and Undo stage (right hand side of the figure).

Once, we have found  $\Theta$ , we can rotate it back to  $U^T$ , and  $\Sigma^{-1}$  calculating the variance, by using ICA.

As shown in Figure 4.2, we need four steps to separate signals  $S1$  and  $S2$  from the mixture of sources independently. The assumption is that the "DO" stage, as shown, is we already know, and the "UNDO" stage is the one that we will need to get through ICA.

1. Calculate angle  $\Theta$  to get a maximum variance.

$$\Theta_0 = \frac{1}{2} \tan^{-1} \frac{-2 \sum U1U2}{\sum (U2^2 - U1^2)} \quad (4.1)$$

$$U^T = \begin{pmatrix} \cos \Theta_0 & \sin \Theta_0 \\ -\sin \Theta_0 & \cos \Theta_0 \end{pmatrix} \quad (4.2)$$

where,  $U1$ , and  $U2$  are the unitary rotation vectors for signal  $S1$  and  $S2$ , resulted after subtraction of the mean for  $U1$ , and  $U2$ , respectively. Therefore,  $U1 = U1 - \text{Mean}(U1)$ , and  $U2 = U2 - \text{Mean}(U2)$

2. Undo Scaling of Singular Values.

$$\Sigma = \begin{pmatrix} \frac{1}{\sqrt{var1}} & 0 \\ 0 & \frac{1}{\sqrt{var2}} \end{pmatrix} \quad (4.3)$$

where,

$$\begin{aligned} var1 &= \sum((U1 \times \cos(\Theta_0) + U2 \times \sin(\Theta_0))^2) \\ var2 &= \sum((U1 \times \cos(\Theta_0 - \frac{\pi}{2}) + U2 \times \sin(\Theta_0))^2) \end{aligned}$$

where,  $var1$  is the variance against principal component angel  $\Theta_0$ , and  $var2$  is the variance against orthogonal angel for second principal component direction.

3. Undo last unitary transformation calculating fourth-moment Kurtosis to make probability density separable.  $\Phi_0$  calculation is based on the data after applying  $U^*$ , and  $\Sigma^{-1}$ .

$$V = \begin{pmatrix} \cos(\Phi_0) & \sin(\Phi_0) \\ -\sin(\Phi_0) & \cos(\Phi_0) \end{pmatrix} \quad (4.4)$$

where,

$$\Phi_0 = 0.25 \times \tan[-\sum(2 \times U1Sigma^3 \times U2Sigma - 2 \times U2Sigma^3 \times U1Sigma) \times \sum(3 \times U1Sigma^2 \times U2Sigma^2 - 0.5 \times U1Sigma^4 - 0.5 \times U2Sigma^4)]$$

$$\begin{aligned} U1Sigma &= \Sigma(1,1) \times (U^T(1,1) \times X1 + U^T(1,1) \times X2); \\ U2Sigma &= \Sigma(2,2) \times (U^T(2,1) \times X1 + U^T(2,2) \times X2); \end{aligned}$$

where,  $U1Sigma$  and  $U2Sigma$  are the mixed vectors, as a result of unitary rotation  $U^T$ , stretching  $\Sigma^{-1}$ , and mixed signals  $X1$ ,  $X2$ . The  $U1Sigma$  and  $U2Sigma$  are considered as an input to  $V$  for the unitary rotation again.

4. Once, these three steps are measured, we need to combine all steps, following Eq.(4.4):

$$S1_{ICA} = V(1,1) \times U1Sigma + V(1,2) \times U2Sigma \quad (4.5)$$

$$S2_{ICA} = V(2,1) \times U2Sigma + V(2,2) \times U1Sigma \quad (4.6)$$

### 4.2.1 An Example of ICA Calculation

In order to understand how *ICA* works, let us mix and then separate two sources. First, let's define the time series of 2 independent sources  $S1$ (*topleft*) and  $S2$ (*topright*) in Figure 4.3, where,

$$S1 = [0, -0.66, -0.99, -0.81, -0.22, 0.47, 0.94, 0.92, 0.44, -0.26]$$

$$S2 = [-0.95, 0.30, 0.60 - 0.99, 0.51, 0.41, -0.98, 0.70, 0.18, -0.91]$$

Then we linearly mix these two sources. For this let us define matrix  $AA$  for the mixing values to be used in the signals  $S1$  and  $S2$  to create mixed signals  $X1$  and  $X2$ .

$$AA = \begin{pmatrix} 0.75 & 0.15 \\ 0.85 & 0.25 \end{pmatrix}$$

The target is to create mixed signals  $X1$  and  $X2$ , where mixture will be the result of both signals  $S1$  and  $S2$ . For example,  $X1$  will contains 75 percentages of  $S1$  and 15 percentage of  $S2$ . Whereas,  $X2$  will contain 85 percentages of  $S1$  and 25 percentage of  $S2$ . Adding these values of mixing matrix  $AA$  matrix, mixed signals  $X1$  and  $X2$  will be as follow:

$$\begin{aligned} X1 &= AA(1,1) \times S1 + AA(1,2) \times S2 \\ &= 0.75 \times S1 + 0.15 \times S2 \\ &= [-0.14, -0.45, -0.65, -0.76, -0.09, 0.42, 0.56, 0.80, 0.36, -0.33] \end{aligned}$$

$$\begin{aligned} X2 &= AA(2,1) \times S1 + AA(2,2) \times S2 \\ &= 0.85 \times S1 + 0.25 \times S2 \\ &= [-0.24, -0.49, -0.69, -0.95, -0.06, 0.51, 0.55, 0.96, 0.43, -0.45] \end{aligned}$$

The mixed signals  $X1$  and  $X2$  are shown Figure 4.3. The middle panel on the left is  $X1$ , middle panel on the right is  $X2$ . Finally, input these two signals into the ICA algorithm which is able to uncover the original activation of  $S1$  and  $S2$ , as shown in Figure 4.3 bottom left and bottom right, respectively. In order to get source signals  $S1$  and  $S2$  separating from the mixture  $X1$  and  $X2$  for the result shown in Figure 4.3, we will need to follow four steps described earlier. They are as follows:

1. According to Eq. (4.1 and 4.2):

$$\Theta_0 = 0.71, \text{ and}$$

$$U^T = \begin{pmatrix} 0.76 & 0.65 \\ -0.65 & 0.76 \end{pmatrix}$$

2. According to Eq. (4.3):

$var1 = 5.99$  (the variance against principal component angel  $\Theta_0$ ) , and  $var2 = 5.17$  (the variance against orthogonal angel for second principal component direction).

$$\text{Therefore, } \Sigma = \begin{pmatrix} 0.41 & 0 \\ 0 & 0.44 \end{pmatrix}$$

3. According to Eq. (4.4):

$$\Phi_0 = 0.17.$$

$$\text{Therefore, } V = \begin{pmatrix} 0.99 & 0.17 \\ -0.17 & 0.99 \end{pmatrix}$$

4. According to Eq. (4.5 and 4.6):

$$S1_{ICA} = [-0.11, -0.27, -0.39, -0.50, -0.04, 0.27, 0.32, 0.51, 0.23, -0.23];$$

$$S2_{ICA} = [-0.02, 0.01, 0.02, -0.01, 0.01, 0.00, -0.03, 0.01, 0.00, -0.02];$$

The result shows that information of mixed signal  $X1$  and  $X2$  are lost from original signals  $S1$  and  $S2$  after applying ICA. The resulted signals  $S1_{ICA}$ , and  $S2_{ICA}$ , obtained using ICA are not the same as  $S1$  and  $S2$ , but they are approximated by ICA. The aim of ICA is to approximate the probability distribution of signals  $S1$  and  $S2$ , as seen earlier in the discussion of ICA.

The software *Mitsar/WinEEG* is used for removing artefacts visually from EEG using method ICA in my research.

### 4.3 Fast Fourier Transform

Fourier Transform (FT) allows us to examine signal from the perspective of both Time Domain (TD) and Frequency Domain (FD). FT transforms the TD signal to FD which allow the possibility to see the frequency components of a signal. The essence of the FT of a signal is to decompose a signal into different frequencies of sinusoids. If these sinusoids sums to the original signal at TD, then we have constructed the FT of a signal. A signal

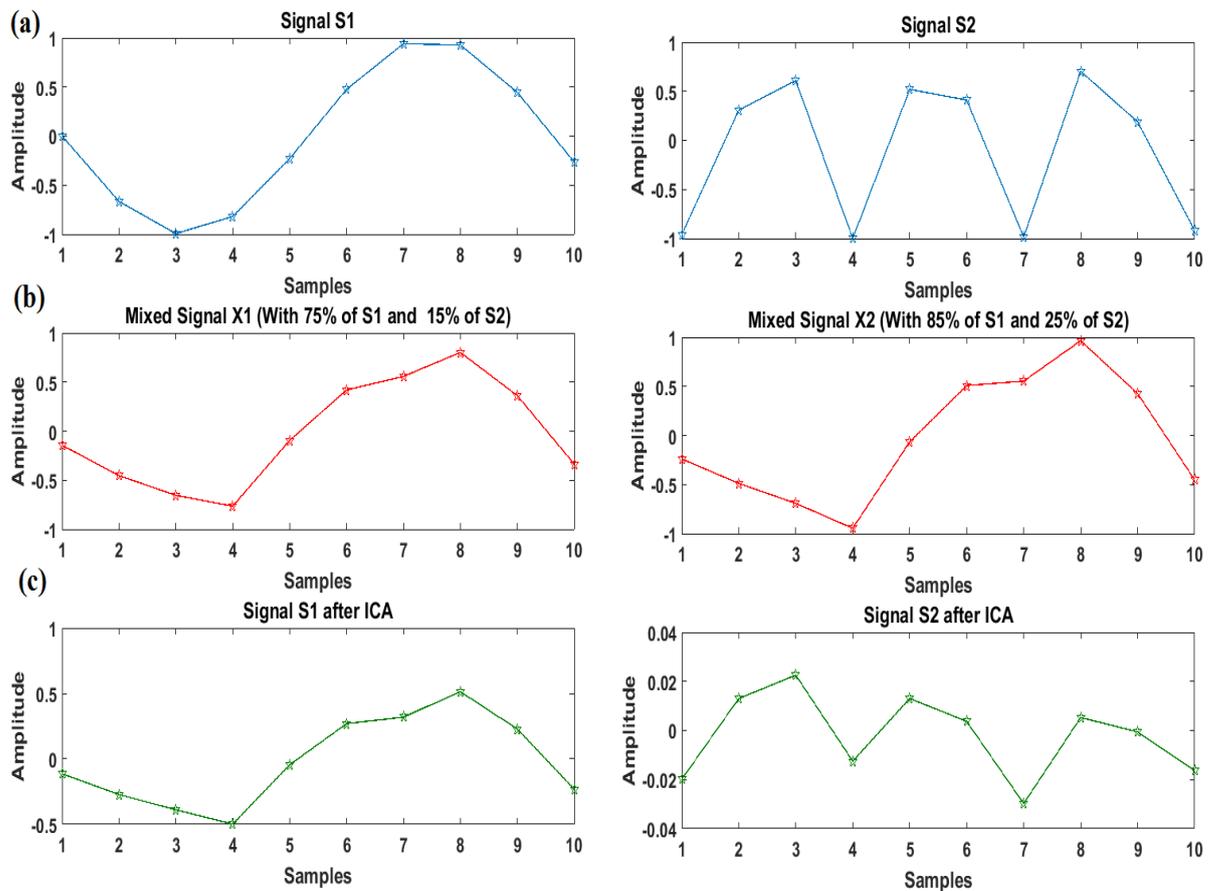


Figure 4.3: ICA example, a) showing two signals in blue colour S1(top left) and S2(top Right), b) showing two mixed sources in red colour X1 (middle panel on left) and X2 (middle panel on right), with a linear mixture of sources S1 and S2, and c) showing uncovering original activation in green colour of two signals S1 (bottom left) and S2 (bottom right) after ICA.

in real life, EEG signal, for example, consists of a bunch of sine waves presenting in TD (Burrus and Parks, 1991). This kind of signal presenting in TD is difficult to identify the component by looking at them. The best way to identify specific components is to transform the signal from time to frequency domain. Therefore, it is essential to convert these signals from a time domain to the frequency domain, to find out the presence of different frequency range within time series signals, such as EEG and ECG.

Discrete Fourier Transform (DFT) is widely used to transfer digital data to understandable data (in terms of frequencies of the signals). DFT provides both convert and reverse functions. Converting function is to convert the TD to FD, and reversing function is to revert it back to TD (Duhamel and Vetterli, 1990). However, DFT requires a complex computation which spends a long time to calculate which is the biggest problem with it. It

requires complex multiplication for the number of samples  $N$  ( $N \cdot N = N^2$ ), therefore; it is computationally expensive. The formula for DFT is given below:

$$x[k] = \sum_{n=0}^{N-1} x[n] \cdot W_N^{nk} \quad (4.7)$$

Where,  $W_N^{nk} = -j2\pi k \frac{n}{N}$  is the constant stored in a table precomputed,  $x[k]$  is sampled in frequency,  $x[n]$  is sampled in time,  $n$  is time index,  $k$  is frequency index, and  $N$  is the number of samples.

Fast Fourier Transform (FFT) algorithm is developed from Discrete Fourier Transform (DFT), to reduce the computation time (Cooley and Tukey, 1965). FFT processing faster than DFT, which can be beneficial for a huge data set. FFT is mostly used in digital signal processing to provide a frequency spectrum analysis. FFT is a computational algorithm that reduces the computing time of DFT Eq. (4.7) to a time proportional to  $N \log_2 N$  (Duhamel and Vetterli, 1990). For example, let's have a look at a different value for  $N$ , as shown in Table 4.1 below:

Table 4.1: Complex multiplication required for DFT and FFT for different value of  $N$ .

Length of $N$	For DFT $N^2$	For FFT $N \log_2 N$
1000	$10^6$	$10^4$
$10^6$	$10^{12}$	$20 \times 10^6$
$10^9$	$10^{18}$	$30 \times 10^9$

As shown in Table 4.1, for  $N = 10^9$ , DFT will calculate  $N \cdot N = N10^{18}$  samples. If each operation took a nanosecond (NS) for the complex multiplications, then  $N10^{18}$  NS turn out to be 31.2 years. In contrast, with FFT we can cut this down to  $N \log_2 N = 30 \times 10^9$  NS taking only 30 seconds for the operation of multiplying  $N$ . The general idea of FFT is to divide the data into the sequence as the even and odd index subsequence to compute the DFT, assuming  $N$  must be a power of 2, for being faster. The formula for the FFT is given below:

$$x[k] = \sum_{n \text{ Even}} x[n] \cdot W_N^{nk} + \sum_{n \text{ Odd}} x[n] \cdot W_N^{nk} \quad (4.8)$$

where, even and odd indices are  $n = 2r$ , and  $n = 2r + 1$ , respectively.  $r = 0, 1, 2, \dots, \frac{N}{2} - 1$ .

$$x[k] = \sum_{r=0}^{\frac{N}{2}-1} x[2r] \cdot W_N^{kr} + W_N^k \sum_{r=0}^{\frac{N}{2}-1} x[2r+1] \cdot W_N^{kr} \quad (4.9)$$

Where,  
the first part of the addition is the  $\frac{N}{2}$  DFT of even samples  $X_E[k]$ , the second part of the addition is the  $\frac{N}{2}$  DFT of Odd samples  $X_O[k]$ ,  $W_N^k$  is the phase factor.

Therefore,

$$x[k] = X_E[k] + W_N^k \cdot X_O[k] \quad (4.10)$$

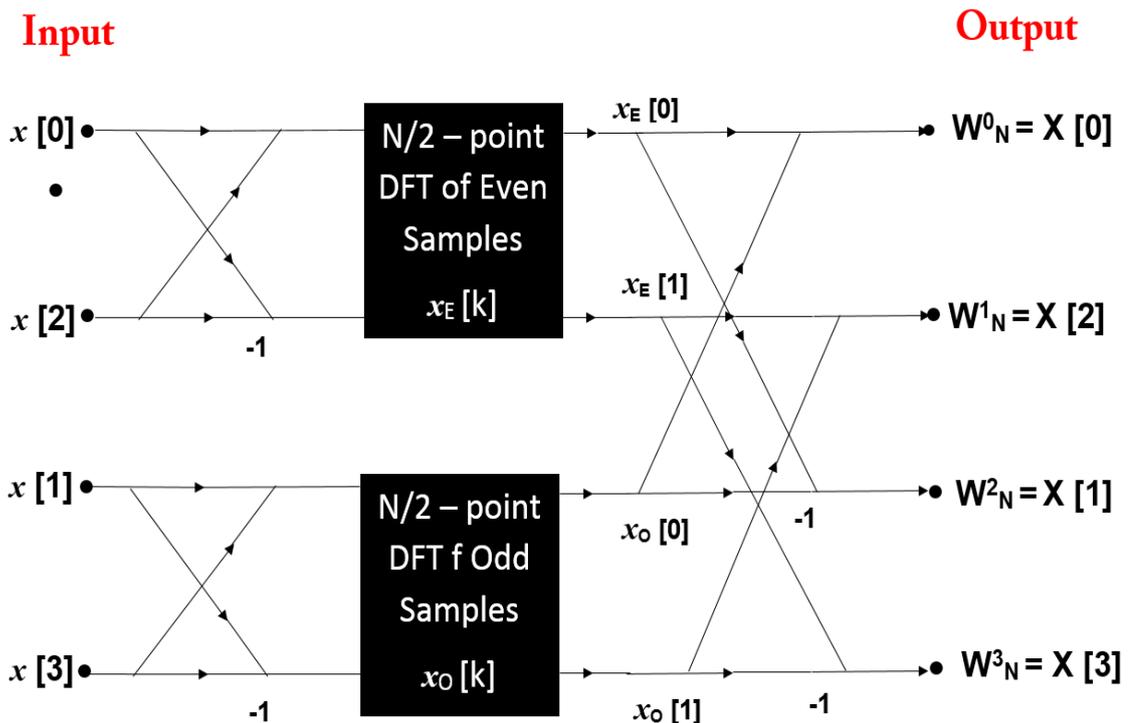


Figure 4.4: An example of 4 data point FFT algorithm structure, using a decomposition into half-size FFTs.  $X_E[k]$  is representing even samples, and  $X_O[k]$  is representing Odd samples.

Figure 4.4 shows the graphical presentation of how the FFT works, where the even indexes are shown as  $x[0], x[2]$ , and odd indexes as  $x[1], x[3]$ . The output of these sequences through  $\frac{N}{2}$  DFT will be  $x_E[0]$  and  $x_E[1]$  for the FFT on even sequence, and  $x_O[0]$  and  $x_O[1]$  for the FFT on odd sequence. Combining these two  $\frac{N}{2}$  DFT's of even and odd sequences can give us the DFT of the overall sequence. The arrow indicates multipliers.

1. To calculate even sequence from the Eq 4.10:

$$x_E[0] = x[0] + x[2]$$

$$x_E[1] = x[0] + W_N^k \times x[2]$$

2. To calculate odd sequence from the Eq 4.10:

$$x_O[0] = x[1] + x[3]$$

$$x_O[1] = x[1] + W_N^k \times x[3]$$

3. To calculate DFT of the overall sequence with Eq 4.10:

$$W_4^0 = x_E[0] + x_O[0]$$

$$W_4^1 = x_E[0] + W_N^k \times x_O[0]$$

$$W_4^2 = x_E[1] + x_O[1]$$

$$W_4^3 = x_E[1] + W_N^k \times x_O[1]$$

### 4.3.1 An Example of FFT Calculation

In order to understand how *Fast Fourier Transform* work, let us consider the following example:

Find the FFT  $x[k]$  of the sequence [11,22,33,44], where  $N = 4$ , and  $k = 0, 1, 2, 3$ . Based on Eq. (4.8) and Figure 4.4, the result of the FFT calculation is shown in Figure 4.5.

To calculate the DFT of the overall sequence, we need to first calculate FFT on even and odd samples, and then combining these two  $\frac{N}{2}$  DFT's of even and odd sequences, as seen earlier in Figure 4.4. The steps to calculate the overall DFT for the sequence [11,22,33,44], for  $N = 4$  we will need to compute the DFT into 2 stages ( $\log_2 N = 2$ ), and total complex multiplication we will need be ( $N \log_2 N = 4$ ). Here, the complex multiplication  $W_N^k = -1$ . The calculation for FFT following Eq 4.10 is as follows:

1. To calculate even sequence from the Eq 4.10:

$$x_E[0] = x[0] + x[2] = 11 + (33) = 44$$

$$x_E[1] = x[0] + W_N^k \times x[2] = 11 + (-1) \times 33 = -22$$

2. To calculate odd sequence from the Eq 4.10:

$$x_O[0] = x[1] + x[3] = 22 + 44 = 66$$

$$x_O[1] = x[1] + W_N^k \times x[3] = 22 + (-1) \times 44 = -22$$

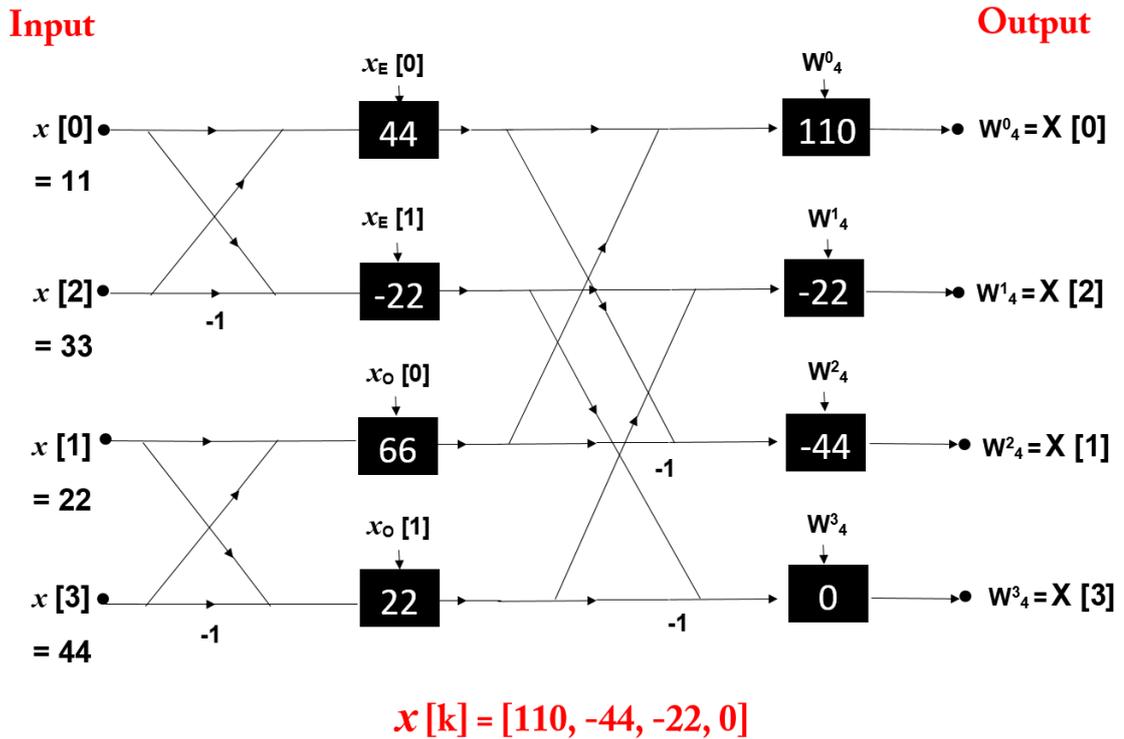


Figure 4.5: An example of 4 point FFT algorithm structure, using a decomposition into half-size FFTs of odd and even sequences.  $x_E[k]$  is representing even samples, and  $x_O[k]$  is representing Odd samples.

3. To calculate DFT of the overall sequence with Eq 4.10:

$$W_4^0 = x_E[0] + x_O[0] = 44 + 66 = 110$$

$$W_4^1 = x_E[0] + W_N^k \times x_O[0] = 44 + (-1) \times 66 = -22$$

$$W_4^2 = x_E[1] + x_O[1] = -22 - 22 = -44$$

$$W_4^3 = x_E[1] + W_N^k \times x_O[1] = -22 + (-1) \times -22 = 0$$

Based on the example the output  $x[k]$  for the sequence  $[11, 22, 33, 44]$  is  $W_N^n = [110, -22, -44, 0]$ . However, putting them in the right order for  $k = 0, 1, 2, 3$ , the output  $x[k] = [110, -44, -22, 0]$ . The sequence in TD  $X[n]$  and in FD  $x[k]$  are plotted in Figure 4.6.

The example we have seen earlier is the illustration of how to calculate the FFT. Let us now see another example of how the FFT can transform TD to FD and display the frequencies present in the signal. To illustrate this, I will show an FFT example on a sinusoid signal using MATLAB. Let us create a signal containing a 50 Hz sinusoid of amplitude 0.5 and a 120 Hz sinusoid of amplitude 1. The parameters needed to create this signal are Sampling frequency  $F_s = 1000$ , Sampling period  $T = 1/F_s$ , Length of signal  $L = 1000$ , and Time vector  $t = (0 : L - 1) \times T$ .  $S = 0.5 \times \sin(2 \times \pi \times 50 \times t) + \sin(2 \times \pi \times 120 \times t)$

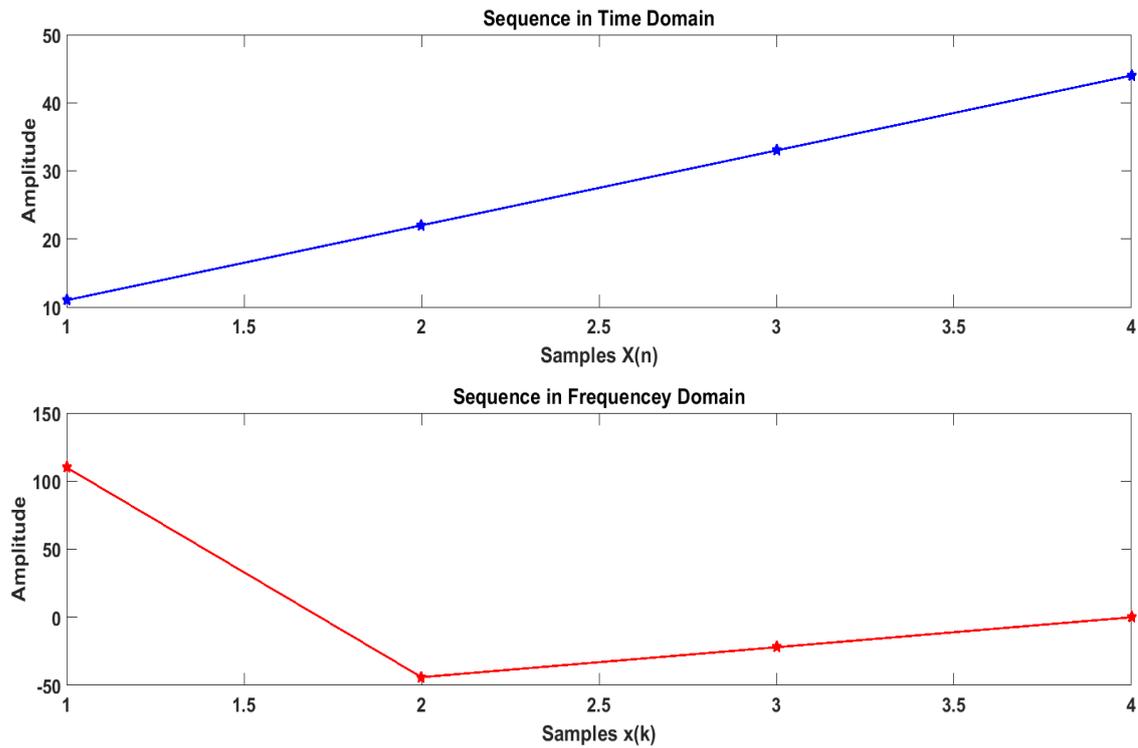


Figure 4.6: The sequence in TD  $X[n]$  (top in blue color) and in FD  $x[k]$  (bottom in red colour).

In Figure 4.7, the top panel shows a sinusoidal signal (in blue), and the bottom (in green) shows the FFT of the Signal  $S(t)$ .

As seen in Figure 4.7, the presence of frequencies and their amplitude values can be clearly visible by transforming the TD signal to FD using FFT. Therefore, FFT is used in this research, to gather information of frequency ranges from EEG and ECG signals.

A MATLAB function, which is  $fft()$ , was used to find the frequency components of EEG and HRV. It takes four parameters as follows:

$$FFT = fft(r, NFFT)/L;$$

Where, Fast Fourier Transform  $fft$  was used,  $r$  is the input signal,  $NFFT$  is used for making FFT faster, assuming  $N$  must be the power of 2 from the original signal length  $L$ . This pads the signal  $r$  with trailing zeros to improve the performance of the FFT.

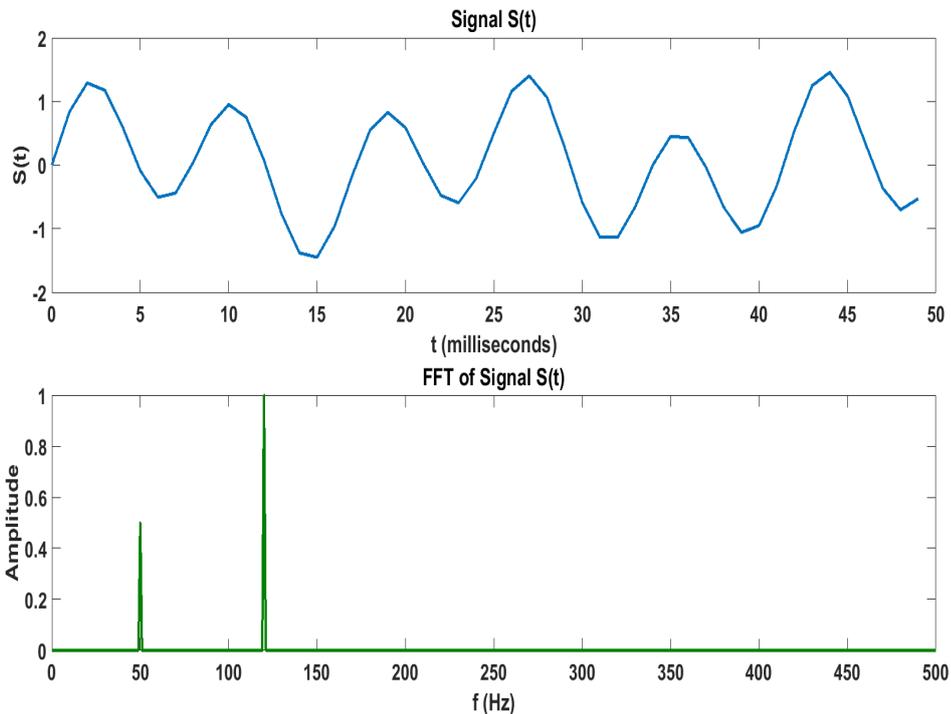


Figure 4.7: An example of FFT performance of signals S (Bottom panel- as shown in colour green).

## 4.4 Wavelet Transform

FFT can transform TD to FD and show us the presence of frequencies in the signal. However, FFT loses the time information and shows only frequency information. Whereas, Wavelet Transform (WT) can keep both time and frequency information. The WT of the signal can be thought of as an extension of the classical Fourier transform (FT) - it works on a multi-scale basis, instead of working on a single scale (Time or Frequency) as FT. This is achieved by decomposition of the signal over dilated (scale) and translated (time) version of the wavelet. So, for spectral analysis, WT is more suitable than FFT (Akin, 2002). WT is designed to address the problem of signals with nonstationary. It includes representation of time function in terms of simple blocks, termed wavelets. These blocks are derived from a signal generating a function called the *mother wavelet* by translation and dilation operations. Dilation, also known as scaling, compresses or stretches the mother wavelet and translation shifts it along the time axis (Daubechies, 1990), (Akay, 1997), (Unser and Aldroubi, 1996). The WT can be categorised into continuous and discrete. Continuous wavelet transform (CWT), implies that the scaling and translation parameters change continuously, and thus, represent considerable effort and a vast amount of data calculation for

every possible scale. In addition, time series signals EEG and ECG are discrete signals. Therefore, I have considered a discrete wavelet transform (DWT).

When undertaking the WT, two components are extracted from the signal: the approximation and detail components. The *approximation* sub-signal is defined as the high scale, the low-frequency component of the original signal. It is also referred to as a smoothed signal. The *detail* sub-signal is the low scale, the high-frequency component of the original signal. An input signal is decomposed by using a low-pass filter and high-pass filter followed by down-sampling in each stage. The output of the first stage high-pass filter gives the detail coefficient  $d_m^1$ , whereas the low-pass filter gives the approximation coefficient  $a_m^1$ , where  $m$  is the scaling with the support of time-units (translation in time by an even number of time-units). The decomposition of a signal up to three scales is shown in Figure 4.8.

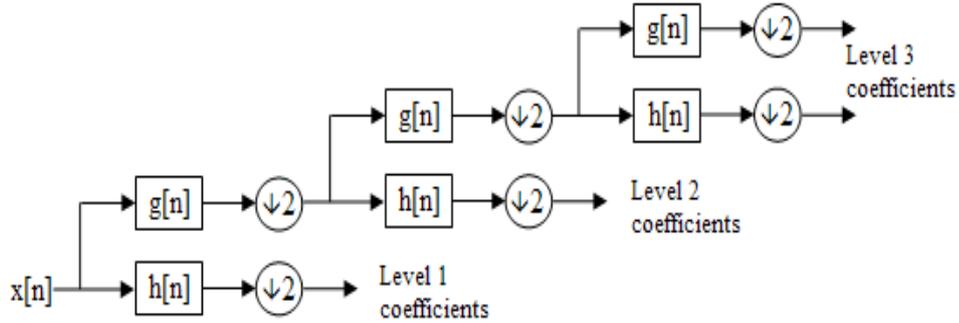


Figure 4.8: Three level sub-band decomposition of discrete wavelet transform (DWT) implementation;  $g[n]$  is the low-pass filter, and  $h[n]$  is the high-pass filter.  $g[n]$  is the approximation part of the signal containing high scale and low frequencies, and  $h[n]$  is the detail part of the signal containing low scale and high frequencies (Akay, 1997).

At each level in the above diagram, DWT decomposes the signal into low and high frequencies. The approximation part of a signal contains high scale and low frequencies, and detail part of a signal contains low scale and high frequencies. The output of low pass and high-pass filters for an input signal  $x[n]$  is given by:

$$y_{low} = \sum_n x[n] \cdot g[-n + 2k] \quad (4.11)$$

$$y_{high} = \sum_n x[n] \cdot h[-n + 2k] \quad (4.12)$$

#### 4.4.1 Haar Wavelet

In this section, I start with the simplest wavelet transform: Haar wavelet transform, which can be used for signal decomposition (James and Walker, 1999). I have used extensively

haar wavelets as examples in teaching because of its simplicity. In fact, it is the simplest wavelet and has been a prototype for all other types of wavelet transforms (Malik and Verma, 2012). Suppose a signal  $x$  consists of  $N$  elements in TD:

$$x = \{x_i\}_{i=1}^N \quad (4.13)$$

Like all wavelet transforms, the Haar transform decomposes a signal into two sub-signals of half its length. One sub-signal is a running average or *approximation*; the other sub-signal is a running difference or *detail*. The *approximation* part captures the trend of the signal, and the *detail* part keeps the fluctuations of each feature. The precise formula for the value of  $a_m^1$  and  $d_m^1$  for  $m = 1, 2, 3, \dots, N/2$  are defined as follows:

$$a_m = \frac{x_{2m-1} + x_{2m}}{\sqrt{2}} \quad (4.14)$$

$$d_m = \frac{x_{2m-1} - x_{2m}}{\sqrt{2}} \quad (4.15)$$

It can be seen from Eq.(4.14), that the *approximation* part consists of a set of average values for each pair in the original signal, divided by the square root of 2. Eq.(4.15) shows that the *detail* part captures the difference of each pair in the original signal. Note that an important property of the Haar transform is that it conserves the energies of signals. The energy of the signal  $E$  is defined as follows:

$$E = \int x^2 dt, \quad (4.16)$$

In the discrete case, where the energy is defined as the sum of the squares of signal values, that is:

$$E = \sum_{i=1}^N x_i^2 \quad (4.17)$$

The energy of the *approximation* part accounts for a large percentage of the total energy.  $\sqrt{2}$  is used in Eq. (4.14) and (4.15) so that the Haar transform conserves the energy of a signal (James and Walker, 1999).

The Haar transform can be performed at multiple levels. At the top level, that is level one (level-1), a signal is transformed to two sub-signals  $a_m^1$  (see Eq. (4.14) and  $d_m^1$  (see Eq. (4.15)). The second level is then carried out by computing a second trend  $a_m^2$  and a second fluctuation  $d_m^2$  for the  $a_m^1$  only. That is, we can continue recursively with the same process to work on the next level, where the signal is always the approximation part obtained from the preceding level.

Figure 4.9 shows an example of the scaling and the wavelet function of Haar wavelet transforms. As we can see from Figure 4.9, the Haar wavelets have two constant scaling and wavelet numbers for the approximation and detail parts. Haar wavelet is discontinuous and like a step function. The transformation with Haar wavelet is a simple wavelet that involves averaging and differencing terms, storing detailed coefficients, eliminating data, and reconstructing the matrix in a way that the resultant matrix is similar to the initial matrix. It is not well adapted to approximate smooth functions.

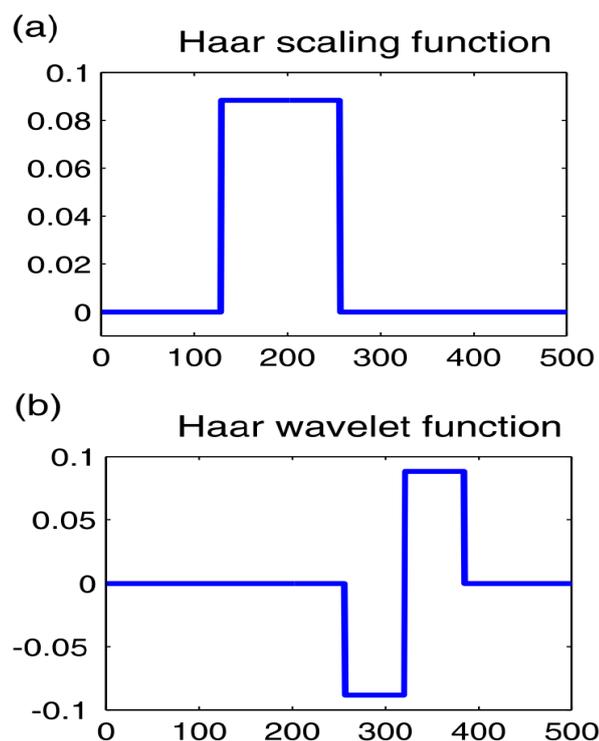


Figure 4.9: Scaling (a) and Wavelets (b) functions being considered for Haar Wavelet (Mutihac, 2006).

#### 4.4.1.1 An Example of Haar Wavelet Calculation

In this section, how to calculate Haar Wavelet up to 2-levels are explained using a simple time series, as an example. Let the input time series be  $x = [4, 6, 10, 12, 8, 6, 5, 5]$ , consisting of eight values. The first *approximation*  $a_m^1$ , and first *detail*  $d_m^1$  of the signal  $x$  are computed as follows:

$$a_m^1 = \left[ \frac{4+6}{\sqrt{2}}, \frac{10+12}{\sqrt{2}}, \frac{8+6}{\sqrt{2}}, \frac{5+5}{\sqrt{2}} \right] = [7.07, 15.55, 9.89, 7.07]$$

$$d_m^1 = \left[ \frac{4-6}{\sqrt{2}}, \frac{10-12}{\sqrt{2}}, \frac{8-6}{\sqrt{2}}, \frac{5-5}{\sqrt{2}} \right] = [-1.41, -1.41, 1.41, 0]$$

The second *approximation*, and *detail* parts of the signal  $x$  are computed by decomposing the first approximation part  $a^1$ . The second approximation  $a^2$ , and second detail  $d^2$  of the signal  $x$  are computed as follows:

$$a_m^2 = \left[ \frac{7.07+15.55}{\sqrt{2}}, \frac{9.89+7.07}{\sqrt{2}} \right] = [16, 12]$$

$$d_m^2 = \left[ \frac{7.07-15.55}{\sqrt{2}}, \frac{9.89-7.07}{\sqrt{2}} \right] = [-6, 2]$$

Figure 4.10 shown the original time series  $x = [4, 6, 10, 12, 8, 6, 5, 5]$ , the first level approximation part, and the first level detail part, to check the performance of Haar Wavelet transformation.

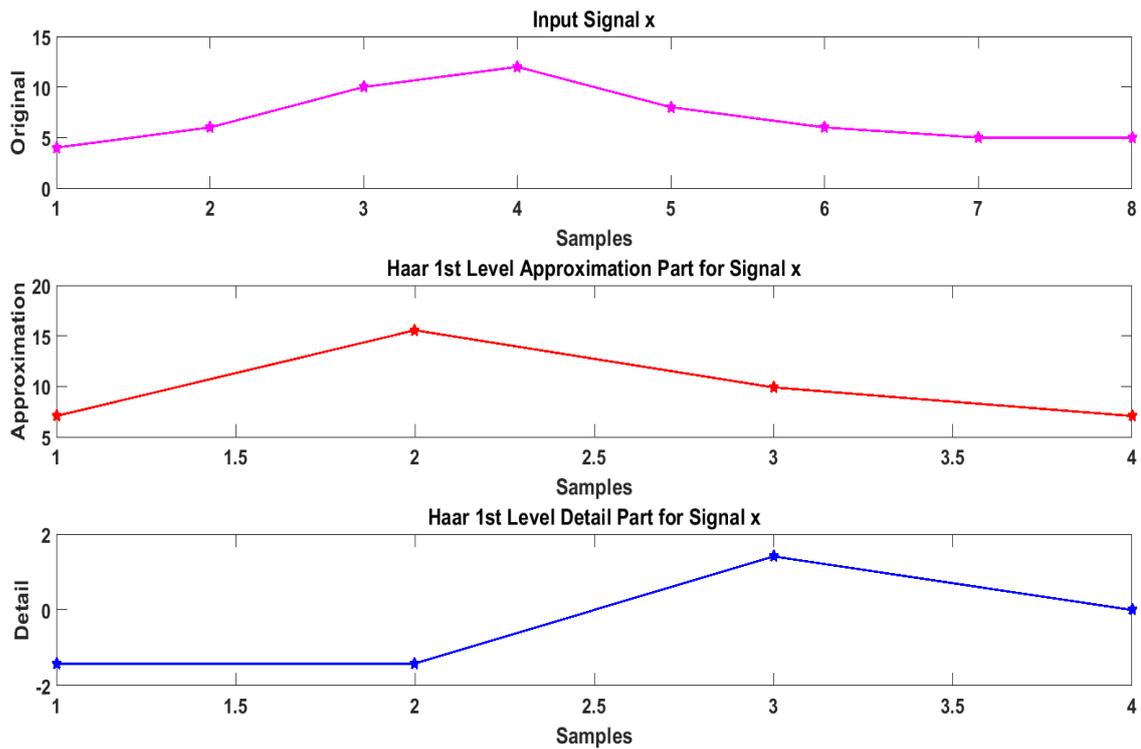


Figure 4.10: Example of Haar Wavelet Transform: The top panel shows original signal in magenta colour, the middle panel shows approximation part of WT in red colour, and the bottom panel shows detail part of WT in blue colour.

How Haar works on EEG signal is displayed in Figure 4.11, the top panel shows the EEG signal. The panels in the middle (approximation part of WT signal) and the bottom (detailed part of WT signal) show the 1-level Haar transform of the original signal at the

top. We can see it that the approximation part of the signal (in the middle panel) becomes smoother than the original noisy one.

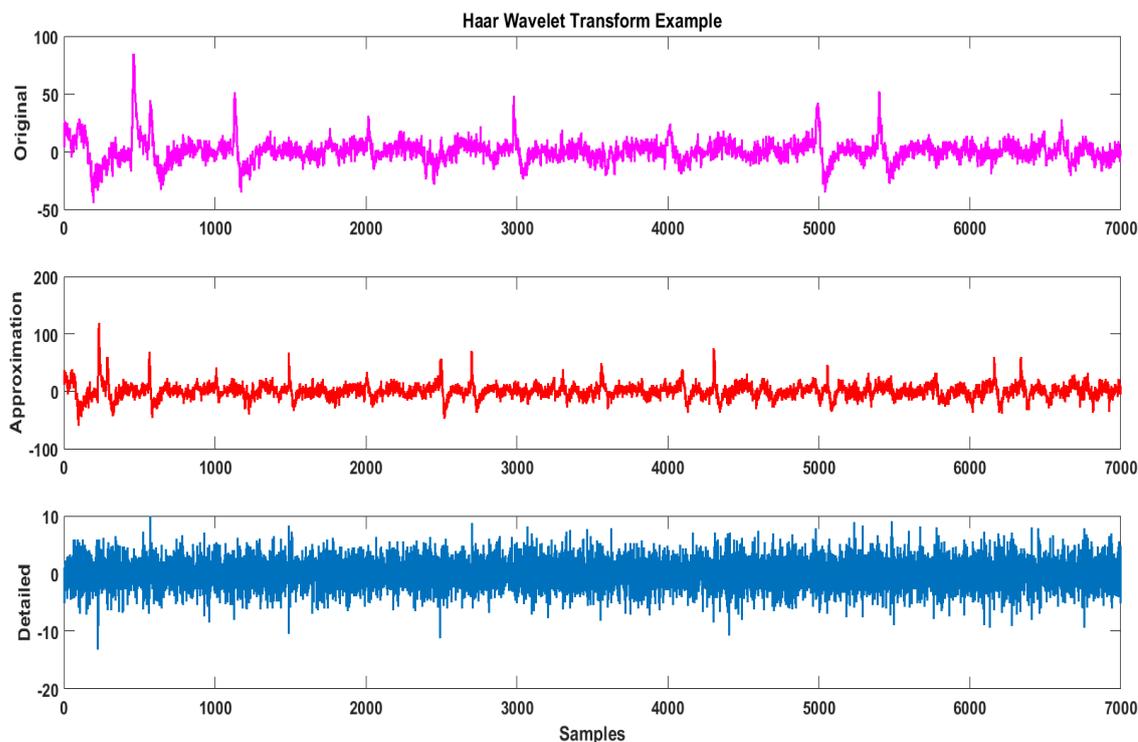


Figure 4.11: Example of 1-level Haar Wavelet Transform (WT) on EEG signal: The top panel shows EEG signal, the middle panel (approximation part of WT signal) and the bottom panel(detailed part of WT signal).

#### 4.4.2 Daubechies Wavelet

Apart from the Haar transform, there are many other different types of wavelet transforms. For example, another widely used type is Daubechies wavelets. Daubechies wavelet transform has been successfully applied in many engineering related works (Amaratunga et al., 1994). Similar to the Haar wavelets, it decomposes signals into the *approximation* and *detail* parts, and preserves the energy of each signal. However, Daubechies uses more data points to compute both the *approximation* and *detail*, rather than using just pairs of data as was done in the Haar transform. Moreover, unlike the Haar wavelets using  $\sqrt{2}$ , Daubechies uses different numbers to multiply with data points so that the signal's energy can be kept (Daubechies, 1990).

There are many Daubechies transforms, but they are all very similar. In this section, I am concentrating on the simplest and commonly used for time series signals, the *Daub4*

wavelet transform (Daubechies, 1990). The *Daub4* wavelet transform is defined in essentially the same way as the Haar wavelet transform.

By using Daub4 transform, the signal can be decomposed into the *approximation* and *detail*. The precise formula for the value of  $a_m^1$  and  $d_m^1$  for window  $m = 1, 2, 3, \dots, N/2$  are defined as follows:

$$a_m^1 = x \cdot V_m^1 \quad (4.18)$$

$$d_m^1 = x \cdot W_m^1 \quad (4.19)$$

Where,  $x$  is the signal,  $a_m$ , and  $d_m$  are the scalar product of  $x$  with a 1-level *scaling signal*  $V_m^1$ , and *wavelet*  $W_m^1$ , respectively.

The only difference between the Haar and Daub4 wavelet transform lies in the way that the *scaling signals* and *wavelets* are defined. Therefore, let's discuss the *scaling signals* first. Let the *scaling numbers*  $p_1, p_2, p_3, p_4$  be defined by

$$p_1 = \frac{1 + \sqrt{3}}{4\sqrt{2}}, p_2 = \frac{3 + \sqrt{3}}{4\sqrt{2}}, p_3 = \frac{3 - \sqrt{3}}{4\sqrt{2}}, p_4 = \frac{1 - \sqrt{3}}{4\sqrt{2}} \quad (4.20)$$

According to Eq.(4.20)  $p_1 = 1.61$ ,  $p_2 = 3.61$ ,  $p_3 = 2.38$ ,  $p_4 = 0.38$ . Using these *scaling numbers*, the 1-level Daub4 scaling signals are defined by:

$$\begin{aligned} V_1^1 &= (p_1, p_2, p_3, p_4, 0, 0, 0, 0) \\ V_2^1 &= (0, 0, p_1, p_2, p_3, p_4, 0, 0) \\ V_3^1 &= (0, 0, 0, 0, p_1, p_2, p_3, p_4) \\ V_{N/2}^1 &= (p_3, p_4, 0, 0, 0, 0, p_1, p_2) \end{aligned}$$

Each scaling signal has the support of just four-time units. Notice also that the second scaling signal  $V_2^1$  is just translation by two time-units of the first scaling signal  $V_1^1$ . Likewise, each subsequent scaling signal is a translation by the multiplication of two times-units of  $V_1^1$ . For  $V_{N/2}^1$ , we *wrap-around* to the start, since  $(p_1, p_2, p_3, p_4)$  has length 4, this would send  $p_3$ , and  $p_4$  beyond the length  $N$  of the signal  $x$ .

Using the natural basis of signals  $V_1^0, V_2^0, V_N^0$ , the first level Daub4 *scaling numbers* satisfy

$$V_m^1 = p_1 \cdot V_{2m-1}^0 + p_2 \cdot V_{2m}^0 + p_3 \cdot V_{2m+1}^0 + p_4 \cdot V_{2m+1}^0 \quad (4.21)$$

We now turn the discussion of the Daub4 wavelets. Let the *wavelet numbers*  $q_1, q_2, q_3, q_4$  be defined by

$$q_1 = \frac{1 - \sqrt{3}}{4\sqrt{2}}, q_2 = \frac{\sqrt{3} - 3}{4\sqrt{2}}, q_3 = \frac{3 + \sqrt{3}}{4\sqrt{2}}, q_4 = \frac{-1 - \sqrt{3}}{4\sqrt{2}} \quad (4.22)$$

Notice that the *wavelet numbers* in Eq. (4.22) are related to the *scaling numbers* in (4.20) by:  $q_1 = p_4, q_2 = -p_3, q_3 = p_2$ , and  $q_4 = -p_1$ . According to Eq.(4.20),  $q_1 = -0.12, q_2 = -0.22, q_3 = 0.84, q_4 = -0.48$ . Using these *wavelet numbers*, the 1-level Daub4 wavelets signals are defined by:

$$\begin{aligned} W_1^1 &= (q_1, q_2, q_3, q_4, 0, 0, 0, 0) \\ W_2^1 &= (0, 0, q_1, q_2, q_3, q_4, 0, 0) \\ W_3^1 &= (0, 0, 0, 0, q_1, q_2, q_3, q_4) \\ W_{N/2}^1 &= (q_3, q_4, 0, 0, 0, 0, q_1, q_2) \end{aligned}$$

These wavelets are all translates of  $W_1^1$ , with wraparound for the last wavelet. Each wavelet has the support of just four-time units. The first level Daub4 *wavelets* satisfy

$$W_m^1 = q_1 \cdot V_{2m-1}^0 + q_2 \cdot V_{2m}^0 + q_3 \cdot V_{2m+1}^0 + q_4 \cdot V_{2m+1}^0 \quad (4.23)$$

Figure 4.12 shows an example of the scaling and wavelet function of Daubechies (db4) wavelet transforms. As we can see from Figure 4.12, the Daubechies wavelets has four variable scaling and wavelet numbers for the approximation and detail parts. Daubechies wavelet (db) is a family of orthogonal wavelets and known as most popular wavelet among others. This mother wavelet is characterised by a maximal number of vanishing moments. It is well adapted to approximate smooth functions.

#### 4.4.2.1 An Example of Daubechies Wavelet Calculation

In this section, how to calculate Daubechies Wavelet at 1-level is explained using a simple time series, as an example. Let the input time series be the same as previously defined in her wavelet example,  $x = [4, 6, 10, 12, 8, 6, 5, 5]$ , consisting of eight values.

By examining Eq.(4.21), first level *approximation*  $a^1$  of the signal  $x$  is now computed as follows:

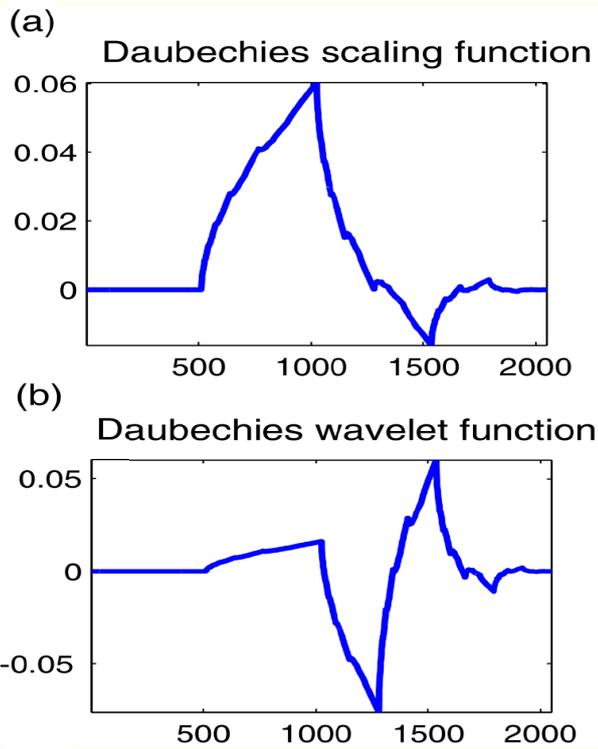


Figure 4.12: Scaling (a) and Wavelets (b) functions being considered for Daubechies Wavelet (Mutihac, 2006).

$$a_1^1 = [1.61 * 4 + 3.61 * 6 + 2.38 * 10 + 0.38 * 12] = [14.16]$$

$$a_2^1 = [1.61 * 10 + 3.61 * 12 + 2.38 * 8 + 0.38 * 6] = [20.22]$$

$$a_3^1 = [1.61 * 8 + 3.61 * 6 + 2.38 * 5 + 0.38 * 5] = [12.11]$$

$$a_4^1 = [1.61 * 5 + 3.61 * 5 + 2.38 * 4 + 0.38 * 6] = [9.5]$$

Therefore,

$$a_m^1 = [14.16, 20.22, 12.11, 9.5]$$

By examining Eq.(4.23), first level *detail*  $d^1$  of the signal  $x$  is now computed as follows:

$$d_1^1 = [(0.38 * 4 + -2.38 * 6 + 3.61 * 10 + -1.61 * 12) = [1]$$

$$d_2^1 = (0.38 * 10 + -2.38 * 12 + 3.61 * 8 + -1.61 * 6) = [-1.38]$$

$$d_3^1 = (0.38 * 8 + -2.38 * 6 + 3.61 * 5 + -1.61 * 5) = [-0.30]$$

$$d_4^1 = (0.38 * 5 + -2.38 * 5 + 3.61 * 4 + -1.61 * 6) = [-1.30]$$

Therefore,

$$d_m^1 = [1, -1.38, -0.30, -1.30]$$

Figure 4.13 shown the original time series  $x = [4, 6, 10, 12, 8, 6, 5, 5]$ , the first level approximation part, and the first level detail part, to check the performance of Daubechies Wavelet transformation.

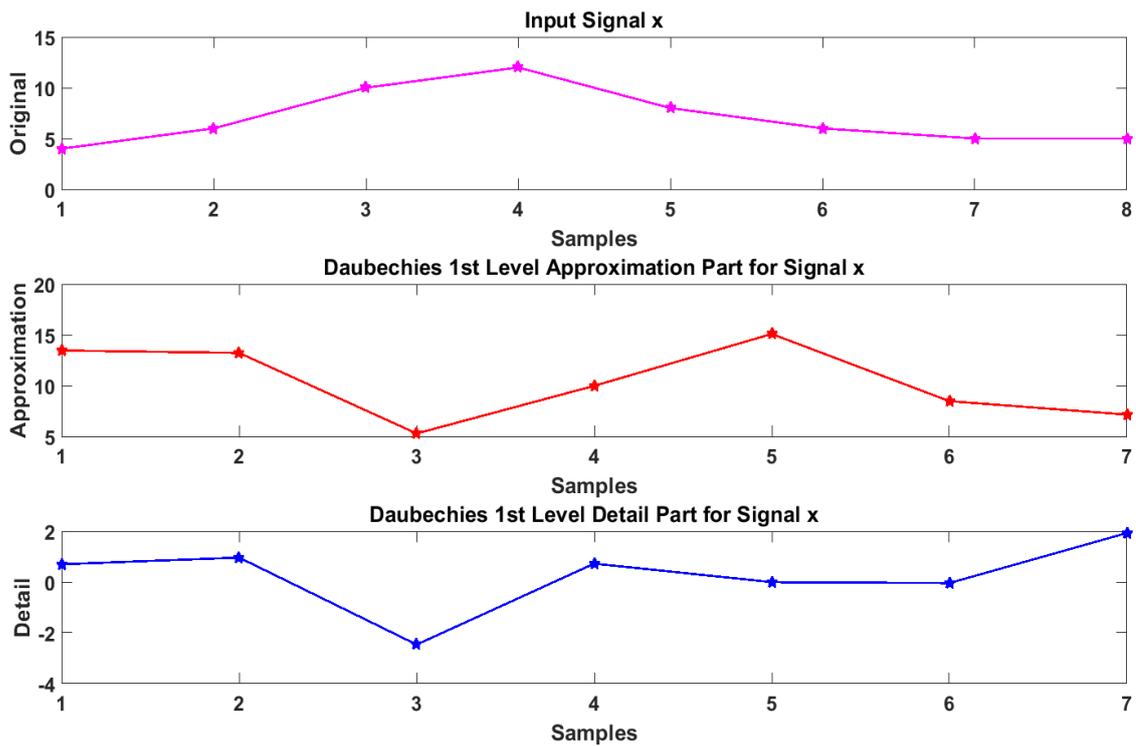


Figure 4.13: Example of Daubechies Wavelet Transform: The top panel shows the original signal in magenta colour, the middle panel shows approximation part of WT in red colour, and the bottom panel shows detailed part of WT in blue color..

How Daubechies works on EEG signal is displayed in Figure 4.14, the top panel shows the EEG signal. The panels in the middle (approximation part of WT signal) and the bottom (detailed part of WT signal) show the 1-level Daubechies transform of the original signal at

the top. We can see that the approximation part of the signal (in the middle panel) becomes smoother than the original noisy one.

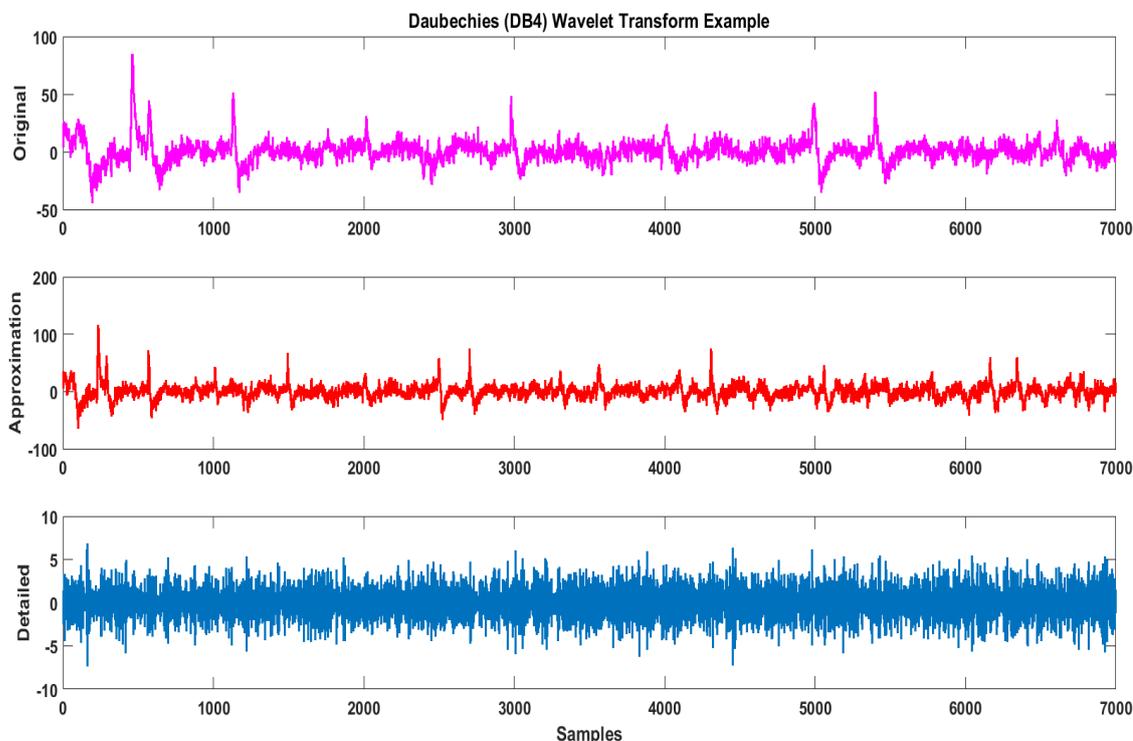


Figure 4.14: Example of 1-level Daubechies Wavelet Transform (WT) on EEG Signal: The top panel shows EEG signal, the middle panel (approximation part of WT signal) and the bottom panel(detailed part of WT signal).

Daubechies wavelet found to provide a better results than the Haar wavelet transform for the datasets used in this work, also it is best suited for EEG and ECG data based on the literature. Therefore, I have focused on the Daubechies wavelet transform.

A MATLAB function, which is `dddtree()`, was used to decompose the signals. It takes four parameters as follows:

$$wt = dddtree(typtree;x;level;df)$$

Where, discrete wavelet transform 'dwt' was used as a *typtree*, *x* is the input signal, *level* is the number of times the transform is applied, and *df* is the decomposition filters that is used by the wavelet transform, which are 'haar' and 'db4' in this research. The argument 'wfilters' was used for wavelet filters.

## 4.5 Summary

One of the most important problems in data preprocessing is how we know what valuable information exists in the raw data so that we can make sure it is preserved. This may depend upon our definition of data preprocessing. Some may argue that data preprocessing is not a completely "pre" process of data analysis. It needs feedback from the main data analysis process. After all, the ultimate judgement whether one has done a good job for data preprocessing is to see if the valuable information has been found in the later data analysis process. It is very clear from the research shown in this chapter, that key features of EEG and ECG signal can improve the analytical performance. Therefore, preprocessing of these signals is vital, before any further analysis on analysing them. In this Chapter, I have focused on the well-known data preprocessing techniques for time series data such as ICA, FFT, and WT.

# Chapter 5

## Methods for Analysing EEG and HRV Time Series Data

### 5.1 Introduction

A general aim of time series analysis is to understand patterns in data. The understanding of pattern in the time series data will influence the analysis that needs to be performed, such as nonlinear or linear system. Time series analysis includes forecasting, determining a transfer function for predictive purposes, describing relationships between related time series, studying the effects of interventions on the time series, developing control schemes, and other (Schiecke et al., 2016). Nonlinear time series analysis is a practical spin-off from *complex dynamical system theory* and chaos theory. It allows one to characterise dynamical systems in which nonlinearity gives rise to a complex temporal evolution. Importantly, this concept allows extracting information that cannot be resolved using classical linear techniques, such as power spectrum or spectral coherence.

Nonlinear time series dynamic analyses has been widely used to study the complex behaviours and different structures of biological systems (Al-Angari and Sahakian, 2007). Nonlinear dynamic analysis proves to be a robust approach for the assessment of different physiological time series because it can unveil hidden patterns related to underlying physiological mechanisms (Alcaraz and Rieta, 2010) (Richman and Moorman, 2000). The chaotic behaviours of a cardiac system and brain waves indicate nonlinearity (Abásolo et al., 2006). With the given nature of nonlinearity with Electroencephalogram (EEG) and Heart Rate Variability (HRV), it turns out to be appropriate for analysing these nonlinear time series data using well known nonlinear methods (Abásolo et al., 2006).

The framework of nonlinear time series analysis comprises a wide variety of methods that allow one to extract various characteristic features of a dynamical system underlying

some measured signal (Kantz and Schreiber, 2004). These methods include Correlation Dimension (CD) as an estimate of the number of independent degrees of freedom. Lyapunov Exponent (LE) as a measure for the divergence of similar system states in time. Prediction errors as detectors for the characteristic traits of deterministic dynamics. Detrended Fluctuation Analysis (DFA) to determine the statistical self-similarity. Approximate Entropy (AE), Sample Entropy (SE), and Multi-scale entropy (MSE) for measuring complexity of the signals. Embedding Dimension (ED), and Poincare plots (PP) as state-space reconstruction for finding nature of the underlying dynamic system, Pearson's Correlation Coefficient (PCC) and Cross-Correlation (CC) for finding similarity between time series signals, and so on. The aforementioned nonlinear time series measures are univariate, which means, they are applied to a single signal measured from individual dynamics. In contrast, bivariate measures are in analysing pairs of signals measured simultaneously from two dynamics. Such bivariate time series analysis measures aim to distinguish whether two dynamics are independent or interacting through some coupling. Some of these bivariate measures aim to indicate not only the strength, but also the direction of these couplings.

Entropy is to predict the next state of a system based on what is known about the current state of a time series. There are a number of different algorithms that have been used to estimate entropy of a time series, Historically, the most popular was *Approximate Entropy*, so it will be discussed in detail in this chapter. *Approximate Entropy* is popular, other methods, with *Sample Entropy* as one of them, are built upon the *Approximate Entropy* with the purpose of overcoming the drawbacks of *Approximate Entropy* such as bias, relative inconsistency, and dependency with the sample length. Therefore, *Sample Entropy* will be discussed in this chapter.

*Pearson's Correlation Coefficient* is used to find the linear dependency between time series data, whereas, *Cross-correlation* can find both linear and nonlinear dependency between time series data (Zou et al., 2003). Therefore, in this chapter I have used both *Pearson's Correlation Coefficient* and *Cross-Correlation* methods for analysing EEG and HRV. A time series can be used to reconstruct the attractor of the underlying dynamic process. State-space reconstruction of a time series is a powerful approach for the analysis of complex, nonlinear systems that appear ubiquitous in the nature and human world. Therefore, the method *Embedding Dimension* for state-space reconstruction is also considered in this chapter.

In summary, the different types of nonlinear methods for analysing EEG and HRV used in this Chapter are *Approximate Entropy (AE)*, *Sample Entropy (SE)*, *Pearson's Correlation Coefficient (PCC)*, *Cross-Correlation (CC)*, and *Embedding Dimension (ED)*. I have discussed the importance of linearity and nonlinearity along with their application in Chapter

2. This Chapter describes, the nonlinear methods used to analyse for assessing variability of time series data. The explanation of these methods, how these methods work, why these methods are chosen, and why they are best for analysing EEG and HRV time series are available in this Chapter. In addition, it also describes each method along with an example of each method in this chapter. The aim of this chapter is to provide thorough overview of data analysis techniques with concrete examples, which could be very helpful to understand how the different techniques work, how can I justify the analysis outcome obtained for my dataset and how can I be sure to conclude my research contributions (for example, how to improve computational time for time series complexity).

## 5.2 Approximate Entropy

*Approximate Entropy (AE)* is a regularity statistic that quantifies the unpredictability of fluctuations in a time series. It was introduced by (Pincus, 1991) to measure system complexity of a different time series. *AE* reflects the likelihood that the similar patterns of observations will not be followed by additional similar observations. A time series containing many repetitive patterns has a relatively small *AE*; a less predictable time series has a higher *AE*. *AE* was used in (Srinivasan et al., 2007) to detect epilepsy from EEG signals, by using an artificial neural network *Elman and probabilistic* neural networks.

$AE(m, r, N)$  is defined as approximately the negative average natural logarithm of the conditional probability for a data set of length  $N$ , where two sequences similar to  $m$  points remain similar, within a tolerance value  $r$  at the next  $m + 1$  point (Richman and Moorman, 2000),(Lake et al., 2002). Here, the parameter  $m$  is the length of the sub-sequences also known as the pattern length to compare, and  $r$  is the tolerance value for accepting matches,  $N$  represents the length of the time series (Richman and Moorman, 2000).  $AE(m, r, N)$  analyses the time series of similar epochs. It uses a template wise approach to find the conditional probability for matches (Richman and Moorman, 2000).

For this research, “Fast Approximate Entropy” code available at Mathworks (2016) by Kijoon Lee is used. The value of  $m$  is 2, and the value of  $r$  is  $0.2 \times SD$  ( $SD$  is the standard deviation of the time series) (Lake et al., 2002). The description of  $AE(m, r, N)$  implemented in this work following the one shown in (Manis, 2008), is:

$$C_i^m(r) = \frac{\sum_{j=1}^{N-m+1} f(i, j, m, r)}{N - m + 1}, \quad (5.1)$$

$C_i^m(r)$  represents the probability of the pattern of length  $m$  that resembles the patterns of similar length beginning at an interval  $i$ , for the match  $f$ . ( $C_i^m(r)$  = number of  $x_j$  such

that  $d[x_i, x_j] \leq r)/(N - m + 1)$ ;  $d$  represents the distance between  $x_i$  and  $x_j$ ).

$$f(i, j, m, r) = 1, \|\vec{x}_i - \vec{x}_j\|, \text{ is } m < r, \text{ otherwise}$$

$$f(i, j, m, r) = 0$$

Here,  $\vec{x}_i$ , and  $\vec{x}_j$  are vectors of the same size of  $m$ .

In order to calculate  $AE(m, r, N)$ , the following steps need to be taken:

1. Fix values for  $m$ ,  $r$ , and  $N$ .
2. For each  $m$ , define the sequence of vectors.
3. Use the sequence to calculate  $C_i^m(r)$  as the equation 5.1
4. Define

$$\Phi^m(r) = \frac{\log C_i^m(r)}{N - m + 1} \quad (5.2)$$

5. Finally, calculate difference between  $m$

$$AE(m, r, N) = \Phi^m(r) - \Phi^{m+1}(r) \quad (5.3)$$

### 5.2.1 An Example of Approximate Entropy Calculation

In this section, how to calculate AE is illustrated by using an example with a simple time series. In order to understand how *Approximate Entropy* work, let us consider the following example:

Let the input time series be  $x(n) = [0.01, 0.1, 0.11, 0.2, 0.01]$ . The standard Deviation of this series is 0.0796.

1.  $m = 2$ ,  $r = 0.2 \times SD$  (that is  $0.2 \times 0.0796 = 0.0159$ ), and  $N = 5$ .
2. Sequence of vectors for  $m = 2$  and  $m = 3$ , are listed shown in Tables 5.1 and 5.2, respectively.
3.  $C_i^m(r)$  calculation, as shown in Tables 5.1 and 5.2.

Figure 5.1 shows the comparison is considered for each length of sequence at  $m = 1$ ; 2; and 3 for the first point from the input time series  $x(n)$ . The probability of sequence matches for all points at  $m = 2$ , and  $m = 3$  are shown in Table 5.1, and 5.2, respectively.

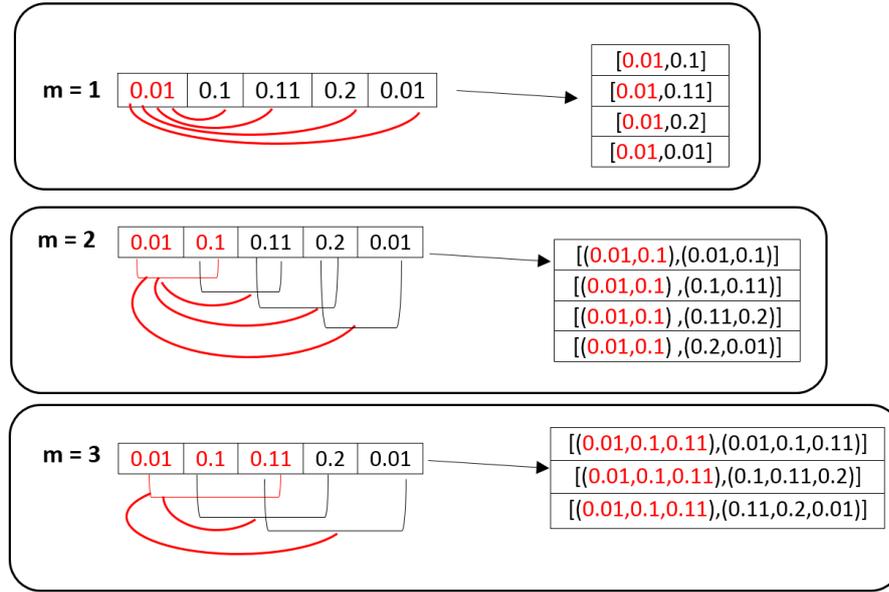


Figure 5.1: Comparison for matches at each length of sequence of  $m$ , for the input time series  $x(n)$ . Where, Length of the sequence is shown as red, and point sequences are shown in black at  $m= 1;2;$  and 3.

Table 5.1: Probability of Sequence matches obtained at  $m = 2$ .

Length of sequence	Point Sequences	Sequence Matches at $r = 0.0159(0.2 \times SD)$		
		Matches ( $f$ )	Total Matches	Probability of Matches $C_i^m(r)$
(0.01, 0.1)	(0.01,0.1)	(1,1)	1	1/4 = 0.25
	(0.1,0.11)	(0,0)		
	(0.11,0.2)	(0,0)		
	(0.2,0.01)	(0,0)		
(0.1, 0.11)	(0.01,0.1)	(1,1)	1	1/4 = 0.25
	(0.1,0.11)	(0,0)		
	(0.11,0.2)	(0,0)		
	(0.2,0.01)	(0,0)		
(0.11, 0.2)	(0.01,0.1)	(1,1)	1	1/4 = 0.25
	(0.1,0.11)	(0,0)		
	(0.11,0.2)	(0,0)		
	(0.2,0.01)	(0,0)		
(0.2, 0.01)	(0.01,0.1)	(1,1)	1	1/4 = 0.25
	(0.1,0.11)	(0,0)		
	(0.11,0.2)	(0,0)		
	(0.2,0.01)	(0,0)		

As shown in Table 5.1, to calculate the probability of matches  $C_i^m(r)$ , the number of matches obtained is counted as “1” for respective point sequences. The ”Total Match” column of the table indicates counting for the total number of matches ob-

Table 5.2: Probability of Sequence matches obtained at  $m = 3$ .

Length of sequence	Point Sequences	Sequence Matches at $r = 0.0159(0.2 \times SD)$		
		Matches ( $f$ )	Total Match	Probability of Matches $C_i^m(r)$
(0.01, 0.1, 0.11)	(0.01, 0.1, 0.11)	(1,1,1)	1	1/3 = 0.33
	(0.1, 0.11, 0.2)	(0,0,0)		
	(0.11, 0.2, 0.1)	(0,0,0)		
(0.1, 0.11, 0.2)	(0.01, 0.1, 0.11)	(1,1,1)	1	1/3 = 0.33
	(0.1, 0.11, 0.2)	(0,0,0)		
	(0.11, 0.2, 0.1)	(0,0,0)		
(0.11, 0.2, 0.01)	(0.01, 0.1, 0.11)	(1,1,1)	1	1/3 = 0.33
	(0.1, 0.11, 0.2)	(0,0,0)		
	(0.11, 0.2, 0.1)	(0,0,0)		

tained. For a particular length of sequence  $m$ , the point matches are obtained by calculating the absolute difference between the points in the sequences. The difference should be below the tolerance value  $r = 0.2$ . Considering the following sequence  $(x_k(i), x_k(j)) = [(0.01, 0.1), (0.1, 0.11)]$ , where  $i$  and  $j$  are the point sequence, and  $k$  is the index for these point sequences. To test the match,  $(|0.01 - 0.1|, |0.1 - 0.11|) = (-0.09, -0.01)$  is calculated. It can be observed that  $x_1(i)$  and  $x_1(j)$  (that is 0.01 and 0.1, and  $x_2(i)$  and  $x_2(j)$  (that is 0.1 and 0.11) both satisfies the condition because the absolute difference is greater than the tolerance value  $r$ . Since the point sequence is a complete match under the tolerance value  $r$ , this sequence is considered as a match, and shown as (1, 1).

Table 5.1 and 5.2 represents the point sequence match at a given length of sequence for  $m = 2$  and  $m = 3$  for the tolerance value  $r(0.2)$ , respectively. In third columns, "1" represents a match and "0" represent no match at tolerance value  $r$ .

### 3.1. Calculating $\Phi^m(r)$ using the equation 5.2

For  $m = 2$ , as shown in Table 5.1:

$$\begin{aligned} \Phi^2(0.2) &= [\log(0.25) + \log(0.25) + \log(0.25) + \log(0.25)] / (5 - 2 + 1) \\ &= (-5.5452) / 4 \\ &= -1.3863 \end{aligned}$$

For  $m = 3$ , as shown in Table 5.2:

$$\begin{aligned} \Phi^3(0.2) &= [\log(0.33) + \log(0.33) + \log(0.33)] / (5 - 3 + 1) \\ &= (-3.3260) / 3 \end{aligned}$$

$$= -1.1087$$

4. Calculating  $AE(m, r, N)$  using the equation 5.3:

$$\begin{aligned} AE(2, 0.0159) : \\ &= -1.3863 - (-1.1087) \\ &= -1.3863 + 1.1087 \\ &= -0.2776 \end{aligned}$$

In order to avoid the occurrence of  $\log 0$ , *Approximate Entropy* includes self-matches for each sequence (Richman and Moorman, 2000). Due to this, *Approximate Entropy* is said to have a strong bias (Lake et al., 2002). A low value of *Approximate Entropy* indicates many similar data points and a high degree of regularity in the time series (Richman and Moorman, 2000). For the given example, a value closer to 0 is obtained, which indicates that the data points are very frequent and has similar values under the given tolerance.

*Approximate Entropy* suffers a few shortcomings such as bias, relative inconsistency, and dependency with the sample length (Alcaraz and Rieta, 2010), (Richman and Moorman, 2000). To overcome these drawbacks with *Approximate Entropy*, a new complexity measure called *Sample Entropy (SE)* was proposed by (Richman and Moorman, 2000).

### 5.3 Sample Entropy

*Sample Entropy (SE)* has been used widely to investigate various biological conditions in the human body, like arrhythmia through ECG (Electrocardiogram) (Alcaraz and Rieta, 2010), EEG background activity with Alzheimer' (Abásolo et al., 2006), human postural sway data (Ramdani et al., 2009) and obstructive sleep apnoea syndrome (Al-Angari and Sahakian, 2007). *Sample Entropy* is also used to detect the termination of a particular medical condition like seizures (Yoo et al., 2012) and to test the effect of therapy like ketogenic diet used for controlling intractable seizures (Takahashi et al., 2010). These studies have concluded that *SE* is a robust quantifier of complexity, which offers an accurate nonlinear metric for quantification (Alcaraz and Rieta, 2010). It gives an excellent dynamical signature and is a helpful tool that provides insights into various biological time series (Abásolo et al., 2006),(Ramdani et al., 2009). Therefore, *Sample Entropy* has been considered as an effective method for investigating different types of time series data.

For a time series of length  $N$ , *SampleEntropy*( $m, r, N$ ) can be defined as the negative logarithm of conditional probability that two sequences are similar to  $m$  point (Lake et al.,

2002) within a tolerance value  $r$ , excluding any self-matches (Richman and Moorman, 2000). The equation can be represented according to (Richman and Moorman, 2000), as:

$$\text{Sample Entropy}(m, r, N) = -\ln\left(\frac{A}{B}\right), \quad (5.4)$$

where,  $m$  is the length of the sequences to be compared,  $r$  is the tolerance value for accepting matches,  $N$  is the length of the original data, and A and B are defined as follow:

$$A = \frac{[(N - m - 1)(N - m)]}{2} A^m(r), \quad (5.5)$$

$$B = \frac{[(N - m - 1)(N - m)]}{2} B^m(r), \quad (5.6)$$

where,  $A^m(r)$  is the probability that the two sequences match for  $m + 1$  points, and  $B^m(r)$  is the probability that the two sequences match for  $m$  points. Each *Sample Entropy* value indicates relative consistency with respect to any value of  $(m, r)$  (That is, if one  $m$  point has a lower *Sample Entropy* value, then it will be lower for any part of fixed  $m$  and  $r$  values (Lake et al., 2002). *Sample Entropy* is independent of the data length and shows an elimination of self-matching. In order to approximate the conditional probabilities of matches, *Sample Entropy* uses a point-wise approach (Richman and Moorman, 2000). Instead of finding differences between each point  $m$  with *Approximate Entropy*, point-wise approach of *Sample Entropy* calculates the probability for each point separately. Therefore, *Sample Entropy* can overcome the drawbacks of relative inconsistency, sample length dependency and biased with *Approximate Entropy*.

### 5.3.1 An Example of Sample Entropy Calculation

In this section, how to calculate *Sample Entropy* is explained using a simple time series, as an example. Let the input time series be  $x(N) = [0.1, 0.1, 0.2, 0.5, 0.22]$ , with  $m = 2, r = 0.2, N = 5$ .

The value of  $m$  specifies the length of the sequences to be considered for *Sample Entropy*. The default value of  $m$  is up to 2 (that is the maximum length of sequences considered is 2). The value of  $r$  represents the tolerance value below which a match is deemed. The input point sequence for  $A^m(r)$  is  $n$  points, while  $B^m(r)$  considers  $n - 1$  points of the input sequence. That means, the input point sequence is  $\{0.1, 0.1, 0.2, 0.5, 0.22\}$  and  $\{0.1, 0.1, 0.2, 0.5\}$ , for  $A^m(r)$  and  $B^m(r)$  respectively.

Figure 5.2 show the comparison considered for each length of sequence at  $m = 1; 2;$  and  $3$  for the first point from the input time series  $x(n)$ . The probability of sequence matches for all points at  $m = 1; 2;$  and  $3$ , are shown in Table 5.3.

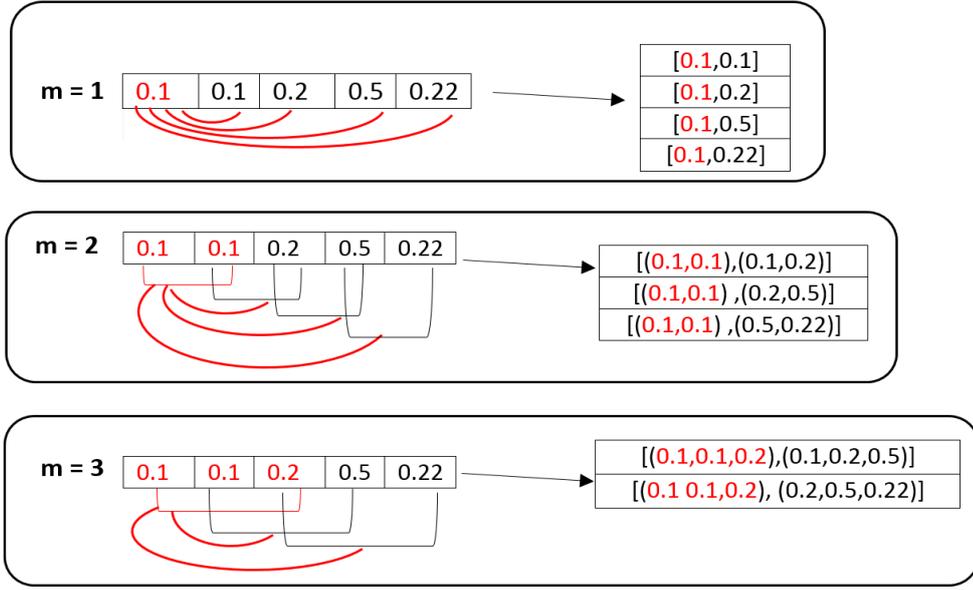


Figure 5.2: Comparison for matches at each length of sequence of  $m$ , for the input time series  $x(n)$ . Where, Length of the sequence is shown as red, and point sequences are shown in black at  $m = 0;1$ ; and 2.

As shown in Table 5.3, to calculate the probability for  $A$  and  $B$ , the number of matches obtained is counted as “1” for respective sequences of  $A^m(r)$  and  $B^m(r)$ .  $A(m)$  and  $B(m)$  indicates the count for the total number of matches obtained. For a particular length of sequence  $m$ , the point matches are obtained by calculating the absolute difference between the points in the sequences. The difference should be below the tolerance value  $r = 0.2$ . The calculation of similar segments can be summarized as:  $|x(i) - x(j)| < r$ . Considering the following sequence  $(x_k(i), x_k(j)) = [(0.1, 0.1), (0.2, 0.5)]$ , where  $i$  and  $j$  are the point sequence, and  $k$  is the index for these point sequences. To test the match,  $(|0.1 - 0.2|, |0.1 - 0.5|) = (0.1, 0.4)$  is calculated. It can be observed that  $x_1(i)$  and  $x_1(j)$  (that is 0.1 and 0.2) satisfy the condition, but  $x_2(i)$  and  $x_2(j)$  (that is 0.1 and 0.5) do not satisfy the condition because the absolute difference is greater than the tolerance value  $r$ . Since the point sequence is not a complete match under the tolerance value  $r$ , this sequence is not considered as a match.

Table 5.3 represents the point sequence match at a given length of sequence for  $m(1$  to 3) for the tolerance value  $r(0.2)$ . In third and fourth columns, “1” represents a match and “0” represents no match at tolerance value  $r$ .

According to equation (5.4),  $SE(m, r, n)$  value can be calculated as follows:

$$SE(1, 0.2, 5) = -\ln(A[1] \div ((n \times n - 1)/2)) = -\ln(6 \div 10) = 0.5108$$

Table 5.3: Point sequences at  $m = 1; 2;$  and 3 along with the count of match obtained for  $A^m(r)$  and  $B^m(r)$ . Here, X represents that the point sequence was not considered for  $B(m)$ . Columns  $A(m)$  and  $B(m)$  indicates count for the total number of matches obtained .

Length of sequence for $m$	Point Sequences	Point Matches at $r = 0.2$			
		$A^m(r)$	$B^m(r)$	$A(m)$	$B(m)$
$m = 1$	[0.1, 0.1]	1	1	6	3
	[0.1, 0.2]	1	1		
	[0.1, 0.5]	0	0		
	[0.1, 0.22]	1	X		
	[0.1, 0.2]	1	1		
	[0.1, 0.5]	0	0		
	[0.1, 0.22]	1	X		
	[0.2, 0.5]	0	0		
	[0.2, 0.22]	1	X		
	[0.5, 0.22]	0	X		
$m = 2$	[(0.1, 0.1),(0.1, 0.2)]	(1, 1)	(1, 1)	1	1
	[(0.1, 0.1),(0.2, 0.5)]	(1, 0)	(1, 0)		
	[(0.1, 0.1),(0.5, 0.22)]	(1, 0)	X		
	[(0.1, 0.2),(0.2, 0.5)]	(1, 0)	(1, 0)		
	[(0.1, 0.2),(0.5, 0.22)]	(1, 0)	X		
	[(0.2, 0.5),(0.5, 0.22)]	(0, 0)	(0, 0)		
$m = 3$	[(0.1, 0.1, 0.2),(0.1, 0.2, 0.5)]	(1, 1, 0)	(1, 1, 0)	0	0
	[(0.1, 0.1, 0.2),(0.2, 0.5, 0.22)]	(1, 0, 1)	X		
	[(0.1, 0.2, 0.5),(0.2, 0.5, 0.22)]	(1, 0, 0)	(1, 0, 0)		

$$SE(2, 0.2, 5) = -\ln(A[2] \div B[0]) = -\ln(1 \div 3) = 1.0986$$

$$SE(3, 0.2, 5) = -\ln(A[3] \div B[1]) = -\ln(0 \div 0) = NaN$$

From the *Sample Entropy* values obtained from the above examples, a low *Sample Entropy* value has been obtained at  $m = 1$ , and the *Sample Entropy* value increases with the increase of  $m$ . This increase indicates that for a longer point sequence, the similarity has decreased for this time series. A lower *Sample Entropy* value indicates a high degree of similarity in time series (Richman and Moorman, 2000).

## 5.4 Pearson's Correlation Coefficient

The *Pearson's Correlation Coefficient (PCC)* measures how closely two different time series are related to each other with the same sequence length and linear dependency. Corre-

lation coefficient ranges between 1 (when the matching entities are the same) and  $-1$  (when the matching entities are inverses of each other). A value of zero indicates no relationship existing between the signals. Let  $a$  and  $b$  be the time series signals. *Pearson's correlation coefficient* is defined as (Benesty et al., 2009):

$$\rho(a, b) = \frac{\sum_{i=1}^N ((a - \mu_a)(b - \mu_b))}{\sqrt{\sum_{i=1}^N (a - \mu_a)^2 \sum_{i=1}^N (b - \mu_b)^2}} \quad (5.7)$$

where,  $\rho(a, b)$  is the correlation coefficient between  $a$  and  $b$ ,  $N$  is the length of the two signals,  $i$  is the index of the signal points, and  $\mu_a$  and  $\mu_b$  are the mean of the signal  $a$  and  $b$ , respectively.

It has been found that the *Pearson's correlation coefficient* is the best-known method for performing correlation between time series data with the same length (Taylor, 1990).

#### 5.4.1 An Example of Pearson's Correlation Coefficient

In this section, I explain how the *Pearson's correlation coefficient (PCC)* is calculated in practice by giving an example. Suppose if the physical activity done by an individual is being affected by the number of hours slept. I believe when I sleep well, I can exercise more. To test my hypothesis, I have tracked 10 day record of physical activity I did each day, and the number of hours I slept the night previous. Let the input signals be:

$$a = [9, 8, 5, 8, 7, 7, 6, 4, 6, 5],$$

$$b = [60, 55, 25, 50, 40, 45, 35, 10, 30, 20],$$

$\mu_a = 6.50$ , and  $\mu_b = 37$ , are the mean for  $a$  and  $b$ , respectively.

Where  $a$  is the number of hours slept each day and  $b$  is the total minutes the physical activity done for the corresponding day. Table 5.4 shows the Pearson's correlation calculation with the given example.

1. Calculate  $(a - \mu_a)$ ,  $(b - \mu_b)$ ,  $((a - \mu_a) \times (b - \mu_b))$ ,  $(a - \mu_a)^2$ , and  $(b - \mu_b)^2$
2. Calculate  $\sum((a - \mu_a)(b - \mu_b)) = 225$ ,  
 $\sum(a - \mu_a)^2 = 22.5$ ,  
and  $\sum(b - \mu_b)^2 = 2310$
3. Calculate  $\sqrt{\sum(a - \mu_a)^2 \sum(b - \mu_b)^2} = \sqrt{(22.5 \times 2310)} = \sqrt{51975} = 227.98$
4. Finally, Perform *Pearson's correlation coefficient* as per equation (5.7).  
 $\rho(a, b) = 225 \div 227.98 = 0.9869$

Table 5.4: Pearson's Correlation Calculation.

$a$	$b$	$(a - \mu_a)$	$(b - \mu_b)$	$(a - \mu_a) \times (b - \mu_b)$	$(a - \mu_a)^2$	$(b - \mu_b)^2$	
1	9	60	2.5	23	57.5	6.25	529
2	8	55	1.5	18	27	2.25	324
3	5	25	-1.5	-12	18	2.25	144
4	8	50	1.5	13	19.5	2.25	169
5	7	40	0.5	3	1.5	0.25	9
6	7	45	0.5	8	4	0.25	64
7	6	35	-0.5	-2	1	0.25	4
8	4	10	-2.5	-27	67.5	6.25	729
9	6	30	-0.5	-7	3.5	0.25	49
10	5	20	-1.5	-17	25.5	2.25	289

The result obtained here 0.99 is a positive value which suggests that there is a strong positive correlation between my hours of sleep gotten the night before a physical activity. The degree of the positive correlation is 1 (when the matching entities are the same), 0 (when there is no match), and -1 (when the matching entities are opposite).

In order to find the correlation between different time series data, *Pearson's Correlation Coefficient* is an ideal measure. However, it requires the same length of input data for the comparison (Taylor, 1990). Time series data, EEG and HRV differ in the length based on their sampling rate, and *Pearson's Correlation Coefficient* might not be suitable to analyse these signals in their Time Domain. Therefore, *Cross-Correlation*, which does not require a similar length of input data, is also utilised to find the correlation, along with *Pearson's Correlation Coefficient*.

## 5.5 Cross-Correlation

*Cross-Correlation (CC)* can be performed to analyse the time delay between two related time series. It offers a valuable and sensitive method for investigating two time series signals, such as EEG, that are recorded at the same time from different electrodes with amplitudes. To analyse time series signals in time domain, *Cross-Correlation* stands out as the most appropriate correlation method, because of its ability to assess signal similarity at all possible time delays. *Cross-Correlation* has been successfully applied in analysing EEG signals in the time domain (Bob et al., 2010), as well as frequency domain (Li et al., 2013). This method can be used to determine the relationship between activities in global and local areas, and also among the different local areas of the human brain. Furthermore, *Cross-Correlation* has been utilized to study the degrees of association between activities

in symmetrical (left and right) parts of the brain (Li et al., 2013), and the results indicated that there was a stronger correlation in the delta  $\delta$  frequency range on the right side of a brain than on the left.

*Cross-Correlation* measures how closely two different time series are related to each other taking time lag into consideration, at the same or different time. In this research, to find the similarity between two time series signals, “*Normalized Cross-Correlation*” (Lewis, 1995) is used, to get ranges between 1 (when the matching signals are the same) and  $-1$  (when the matching signals are inverses of each other). A value of zero indicates no relationship existing between the signals. The *Normalized Cross-Correlation* for time sequence  $P_t$  and  $Q_t$  of signals  $P$  and  $Q$ , respectively, is defined as follows:

$$R_{xy}(T) = \frac{\frac{1}{N} \sum_{t=1}^{N-T} [(P_t)(Q_{t+T})]}{\sigma_P \sigma_Q}, \quad (5.8)$$

$T$  is the time lag at which the similarity between signals is investigated,  $N$  is the length of signals  $P$  and  $Q$ .  $\sigma_P$  is the standard deviation of  $P_t$  and  $\sigma_Q$  is the standard deviation of  $Q_t$ . Note that both the *Cross-Correlation* and *Normalized Cross-Correlation* can be evaluated for any length of  $P_t$  and  $Q_t$ , and they are not required to be the same (Lewis, 1995). If  $P_t$  and  $Q_t$  are of different length, then the shorter one is zero-padded (Buck et al., 2002).

### 5.5.1 An Example of Cross-Correlation

In this section, how the *Normalized Cross-Correlation* is calculated in practice, is explained using an example with a simple time series. Let the input time series be:

$P = [0.1, 0.2, -0.1, 4.1, -2, 1.5, -0.1]$ , and  $Q = [0.1, 4, -2.2, 3.6, 0.1, 0.1, 0.2]$ , with  $T = \pm 0 - 6$ , and  $N = 7$ .

Figure 5.3 illustrates the position of two signals  $P$  and  $Q$  at the time lag of 0, Figure 5.4 illustrates shifting position of signal  $Q$  at the positive time lag of 1, Figure 5.5 illustrates shifting position of signal  $Q$  at the positive time lag of 2, and Figure 5.6 illustrates shifting position of signal  $Q$  at the negative time lag of 1. When one of the signals (signal  $Q$ ) has finished shifting right and left, the correlation at each lag is then calculated. For example, the correlation at the lag of 0, 1 and  $-1$  is shown below, respectively:

1. Correlation at Lag of 0 is:

$$\begin{aligned} 1.1. \text{ For, } & \frac{1}{N} \sum_{t=1}^{N-T} [(P_t)(Q_{t+T})]: \\ & = (1/7) \times [(0.1)(0.1) + (0.2)(4) + (-0.1)(-2.2) + (4.1)(3.6) + (-2)(0.1) + \\ & \quad (1.5)(0.1) + (-0.1)(0.2)] \\ & = 0.1429 \times 15.72 \end{aligned}$$

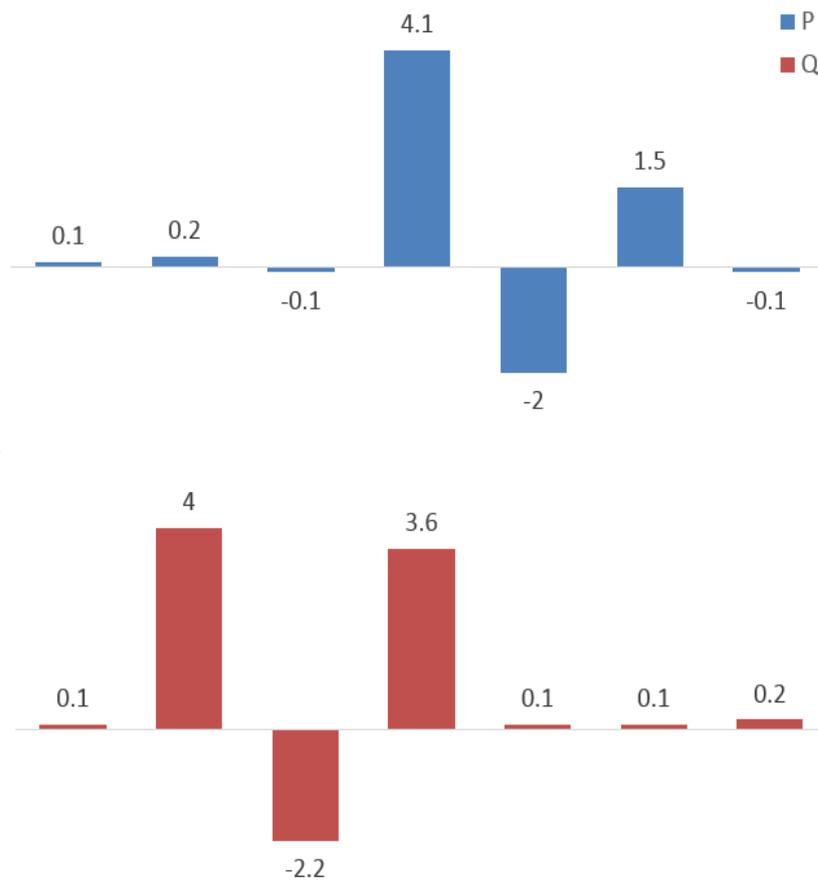


Figure 5.3: Correlation at Lag of 0.

$$= 2.2457$$

1.2. For,  $\sigma_P \sigma_Q$ :

$$= 1.8786 \times 2.1946$$

$$= 4.1228$$

1.3. As per the equation 5.8, *Normalized Cross-Correlation*:

$$= 2.2457/4.1228$$

$$= 0.5448$$

2. Correlation at Lag of 1 is :

2.1. For,  $\frac{1}{N} \sum_{t=1}^{N-T} [(P_t)(Q_{t+T})]$ :

$$= (1/6) \times [(0.2)(0.1) + (0.1)(4) + (4.1)(-2.2) + (-2)(3.6) + (1.5)(0.1) + (-0.1)(0.1)]$$

$$= 0.1667 \times -16.46$$

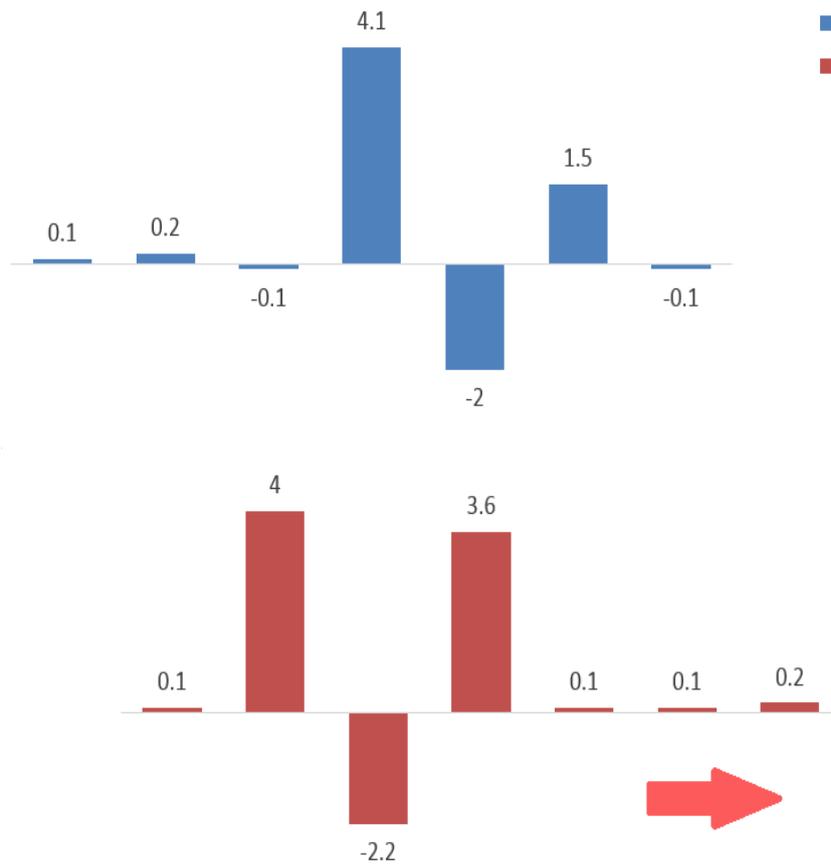


Figure 5.4: Correlation at Lag of 1.

$$= -2.7438$$

2.2. For,  $\sigma_P \sigma_Q$ :

$$= 2.0474 \times 2.3839$$

$$= 4.8808$$

2.3. As per the equation 5.8, *Normalized Cross-Correlation*:

$$= -2.7438 / 4.8808$$

$$= -0.5621$$

3. Correlation at Lag of 2 is :

3.1. For,  $\frac{1}{N} \sum_{t=1}^{N-T} [(P_t)(Q_{t+T})]$ :

$$= (1/5) \times [(-0.1)(0.1) + (4.1)(4) + (-2)(-2.2) + (1.5)(3.6) + (-0.1)(0.1)]$$

$$= 0.2000 \times 26.18$$

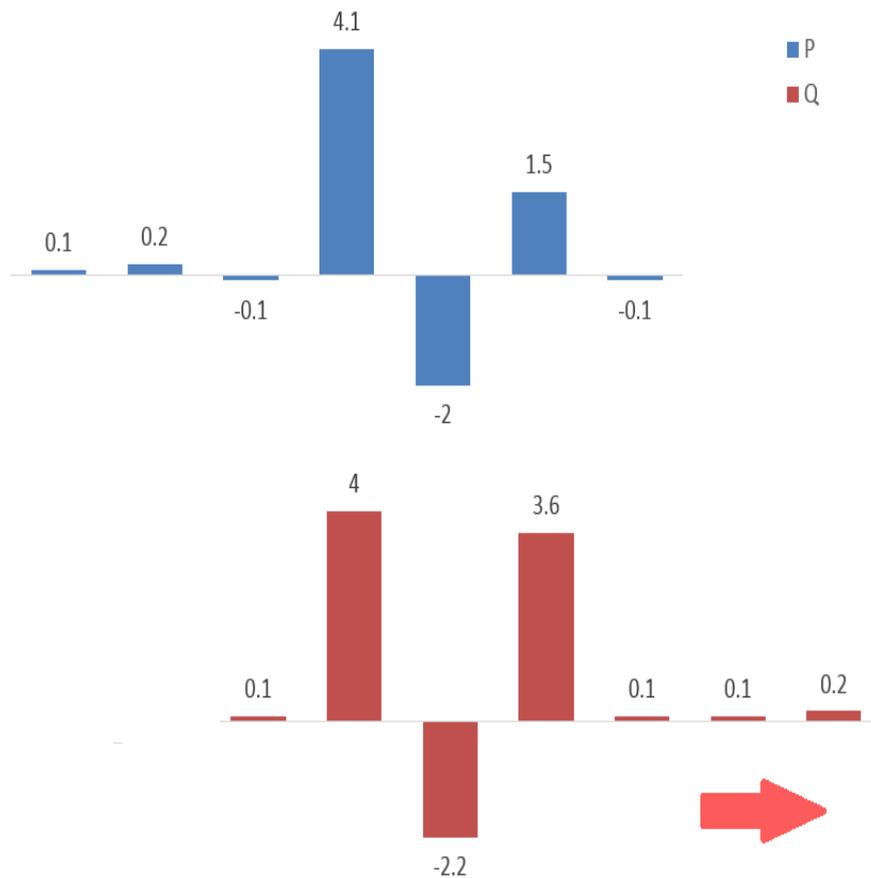


Figure 5.5: Correlation at Lag of 2.

$$= 5.2360$$

3.2. For,  $\sigma_P \sigma_Q$ :

$$= 2.2786 \times 2.6243$$

$$= 5.9797$$

3.3. As per the equation 3.6, *Normalized Cross-correlation*:

$$= 5.2360/5.9797$$

$$= 0.8756$$

4. Correlation at Lag of -1 is:

4.1. For,  $\frac{1}{N} \sum_{t=1}^{N-T} [(P_t)(Q_{t+T})]$ :

$$= (1/6) \times [(1.5)(0.2) + (0.1)(-2) + (4.1)(0.1) + (-0.1)(3.6) + (0.2)(-2.2) + (0.1)(4)]$$

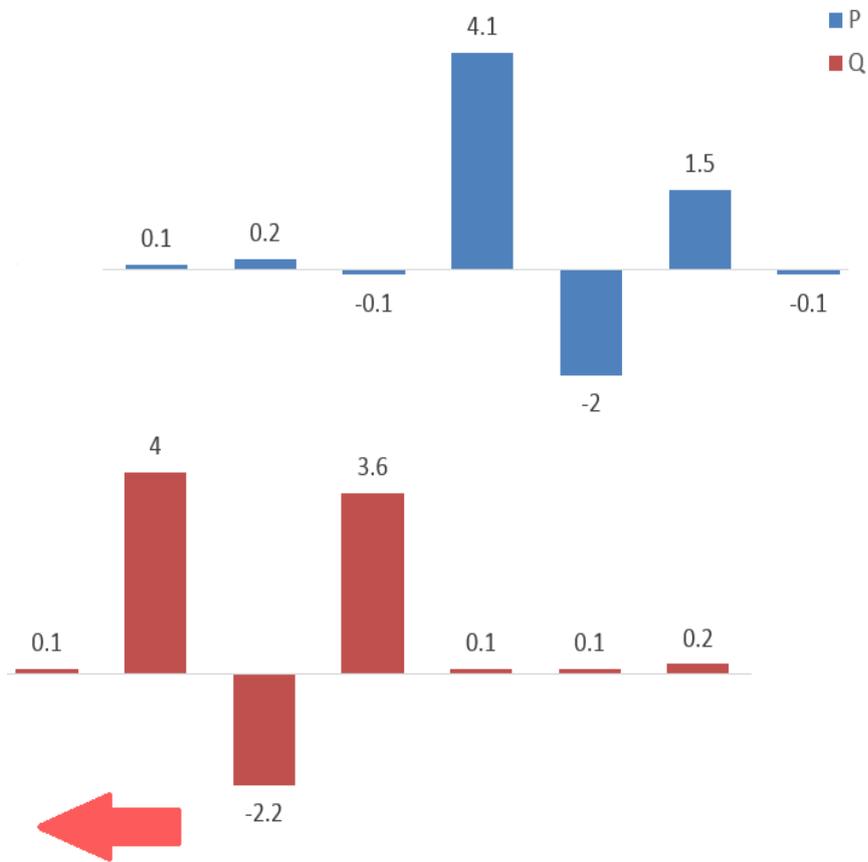


Figure 5.6: Correlation at Lag of -1.

$$= 0.1667 \times 0.11$$

$$= 0.0183$$

4.2. For,  $\sigma_P \sigma_Q$ :

$$= 2.0354 \times 2.3771$$

$$= 4.8383$$

4.3. As per the equation 5.8, *Normalized Cross-Correlation*:

$$= 0.0183 / 4.8383$$

$$= 0.003$$

The correlation is performed until  $N - 1$  lags for both positive and negative lags. Once, the correlation for each lag is performed, the correlation sequence of all lags will be:

The results in Table 5.5 show that the highest correlation is found at the positive lag of 2. This result suggests that the lag of 2 values between both signals (that is ignoring

Table 5.5: Cross-correlation Calculation for corresponding lag

Lags	Correlation Value
-6	0.0007
-5	0.0018
-4	0.0004
-3	0.0425
-2	0.0179
-1	0.0039
0	0.5448
1	-0.5621
2	0.8756
3	-0.4019
4	0.2151
5	-0.0089
6	-0.0004

first two values from  $P$ , and last two values from  $Q$ ) gives the highest positive correlation between these two signals.

## 5.6 Embedding Dimension

A time series, such as HRV, the measurement sequence of one or more visible variable of an underlying dynamic system, whose state changes with time. These time series could be the results of the interaction of many underlying variables. For example, a stock market is affected by many interacting factors, such as economic data, exchange rates and so on. In practice, it is difficult to know what variables determine the behaviour of the actual dynamic system. It is shown by (Takens, 1981) that, if only one scalar value can be measured from an active system, then the nature of the original multi variable dynamic system can be recaptured, by considering a sufficient number of consecutive values. In fact, (Takens, 1981) also mentioned, if the original dynamic system had a dimension of  $N$ , then an embedding of size  $2N$  will fully regain the original system. The size of this consideration is called the *Embedding Dimensions (ED)* (Abarbanel et al., 1993).

*Embedding Dimensions* is used to find out the nature of an underlying dynamical system. The method *False Nearest Neighbours (FNN)* is used to determine how many dimensions are sufficient to embed a particular time series (Kennel et al., 1992). The *False Nearest Neighbours* is designed to determine how many features are enough to present a specific time series (Kennel et al., 1992). The basic idea behind *False Nearest Neighbours* is that points in a state space should be close to each other because their dynamical state is

similar, not because of how they react in each dimension. In an embedding of dimension  $D$ , a series of scalar points has established as a lagged vector.

The *False Nearest Neighbour* method can be summarised as follows:

1. Find the nearest neighbour for each point in an embedding of dimension  $d$ ;
2. Find the percentage of those nearest neighbours which do not remain the nearest neighbour within an embedding of dimension  $d + 1$ , such points turns as false nearest neighbours;
3. Increase the embedding dimension until the number of false nearest neighbour is sufficiently small.

### 5.6.1 An Illustration of Embedding Dimension

To find the correct embedded dimensions,  $d$ , an incremental search, from  $n = 1$ , is performed. A set of time lagged vectors  $x_n(t)$ , for a given  $t$ , is formed, where  $n$  is the index for  $x$ , and  $t$  is the time. The nearest neighbour relation within the set of  $x_n(t)$ 's is then computed. When the correct value of  $d$  has been reached, the addition of an extra dimension to the embedding should not cause these nearest neighbours to spring apart. Any pair whose additional separation is of a high relative size is deemed *False Nearest Neighbours*. Specifically, if  $x_n(t)$  has *nearest neighbour*  $\tilde{x}_t$ , then the relative additional separation when the embedding dimension is incremented is given by (Abarbanel et al., 1993):

$$False\ Nearest\ Neighbours(x, d) = \left| \frac{d(x_t, \tilde{x}_t) - d(x_{t+1}, \tilde{x}_{t+1})}{d(x_t, \tilde{x}_t)} \right| > R_{tol}, \quad (5.9)$$

$$False\ Nearest\ Neighbours(x, d) = \left| \frac{d(x_t, \tilde{x}_t) - d(x_{t+1}, \tilde{x}_{t+1})}{\sigma} \right| > A_{tol}, \quad (5.10)$$

When the  $d$  value exceeds an absolute value, then  $x_t$  and  $\tilde{x}_t$  are denoted as FNN. In the equation 5.9,  $R_{tol}$  is the false neighbour Euclidean distance tolerance, the default value is  $R_{tol} = 15$ . If the ratio of the Euclidean distances between neighbour candidates in successive embedding dimensions exceeds  $R_{tol}$ , then those neighbours are declared as false neighbours. For example, if  $R_{tol} = 5$  neighbour candidates that are separated five times more so than in the previous embedding dimension are declared false neighbours, (Abarbanel et al., 1993). The equation 5.10 is the alternative equation to find a *False Nearest Neighbours* (Kennel et al., 1992), where  $A_{tol}$  arbitrary threshold value for a dataset with a short length, the default value is  $A_{tol} = 2$ .  $A_{tol}$  is a neighbour tolerance based on attractor

size. If the Euclidean distance between two neighbour candidates is  $A_{tol}$  times larger the estimated "size" of the attractor, then those neighbours are declared as false neighbours. To calculate closeness of neighbours, Euclidean Distance has been used (Kruskal, 1964). For example, a time series is a sequence of values  $x_n(t)$ , where  $x$  is the time series,  $n$  is the index for  $x$ , and  $t$  represent the time. Theoretically,  $x$  may be a value which varies continuously with  $t$ . An *Embedding Dimension* of 2 forms, vectors  $(x_0, x_1), (x_1, x_2)$  and so on. An *Embedding Dimension* of 3 forms the vectors  $(x_0, x_1, x_2), (x_1, x_2, x_3)$  and so on. Since these are numeric vectors, we can calculate the distance of any pair of these vectors. So for each vector in a given embedding, the nearest neighbour can be found. However, some of these nearest neighbours may be the false neighbour, in a sense that they are not the nearest neighbour in the embedding with one extra dimension (Frank et al., 2001). A geometric explanation of the concept is at the core of the *False Nearest Neighbours* technique is shown in Figure 5.7.

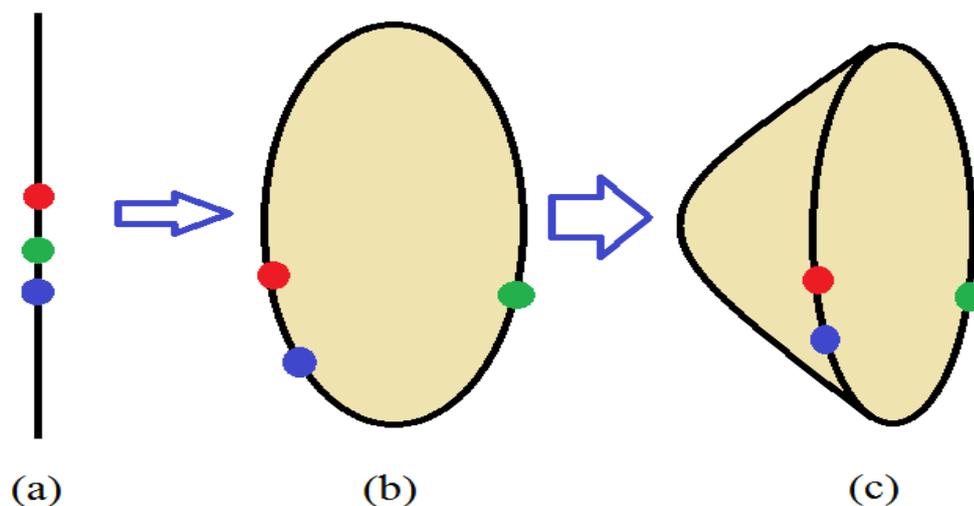


Figure 5.7: Geometric explanation of the *False Nearest Neighbours* Method (Kennel et al., 1992): (a) In one-dimensional, the nearest neighbour for Red is Green, (b) In two-dimensional, the nearest neighbour for Red is Blue, and not Green (This means Green was a false nearest neighbour in one-dimensional), and (c) In three-dimensional, the Blue is still the nearest neighbour for Red, so they are real nearest neighbour.

The line at the left represents a dimensional state space  $x_1$  (Red Point),  $x_2$  (Green Point), and  $x_3$  (Blue Point) and the nearest neighbour of the  $x_1$  (Red Point) is the  $x_2$  (Green Point). Next, the time series embedded in two-dimensional state space, represented by the oval in the middle of the picture. The  $x_1$  (Red Point) and  $x_2$  (Green point) are no longer near to each other. So,  $x_2$  (Green point) is labelled as a false nearest neighbour because it was only near to the  $x_1$  (Red Point) due to the projection of the time series onto the line. Next, the

nearest neighbour for each point in the two-dimensional state space found. Now the nearest neighbour to the  $x_1$  (Red Point) is the  $x_3$  (Blue Point). The time series is now embedded into a three-dimensional state space as represented in the rotated parabola at the right of the picture. The  $x_1$  (Red Point) and the  $x_3$  (Blue Point) are still near to each other, so  $x_3$  (Blue Point) is not a false nearest neighbour. This process continues until either there are no further false nearest neighbour or the data set becomes so sparse in a high dimensional space that no points can be considered being near neighbours, to begin with. The resulting percentage of *False Nearest Neighbours* for each *Embedding Dimension* is then plotted to create *False Nearest Neighbours* plot.

## 5.6.2 An Example of Embedding Dimension Calculation

In this section, the explanation of how the *Embedding Dimension* is calculated in practice is demonstrated by using an example. Let a time series be represented by a sequence of  $x_1, x_2, x_3, \dots, x_n$ , where  $x_1$ ,  $x_2$ ,  $x_3$ , and  $x_n$  represents first, second, and Nth data point, respectively. An actual construction of the state-space requires a great amount of data points, but for the sake of this calculation, let us use a sequence of observations that has 21 data points ( $n = 21$ ). Suppose we want to reconstruct the state-space with an embedding dimension of 3. We can do as described in the following example.

1. Gather input series of vectors to consider for the FNN calculation for each dimension.

As illustrated in Figure 5.8, first, take the data points from  $x_1$  to  $x_{19}$  and set this sequence, called  $x(t)$  aside. Next, take the data points from  $x_2$  to  $x_{20}$  and set this sequence, called  $x(t+1)$ , next to  $x(t)$ . Finally, we take data points from  $x_3$  to  $x_{21}$  and set this sequence, called  $x(t+2)$ , next to  $x(t+1)$ . Now, a matrix with 18 rows and 3 columns has been created.

The first row of this matrix is defined as vector  $\mathbf{v1}$ , the second row as vector  $\mathbf{v2}$ , and so on. Thus, 18 vectors have been created, with each vector having one element  $x(t)$ ,  $x(t+1)$ , and  $x(t+2)$ , respectively. Therefore,  $\mathbf{v1}$  contains  $x_1(t)$  that is 210,  $x_1(t+1)$  that is 214, and  $x_1(t+2)$  that is 202.  $\mathbf{v2}$  contains  $x_2(t)$  that is 214,  $x_2(t+1)$  that is 202, and  $x_2(t+2)$  that is 206 and so forth.

2. Calculate Euclidean Distance between each pair of vectors for each dimension.

In the two-dimensional space ( $m = 2$ ), there is a vector at time  $t$ ,  $v1 = [x_1(t), x_1(t+1)]$  (that is  $v1 = [210, 214]$ ), and its neighbour  $v2 = [x_2(t), x_2(t+1)]$  (that is  $v2 = [214, 202]$ ). In three-dimensional space ( $m = 3$ ), these vectors will be  $v1 = [x_1(t), x_1(t+1), x_1(t+2)]$

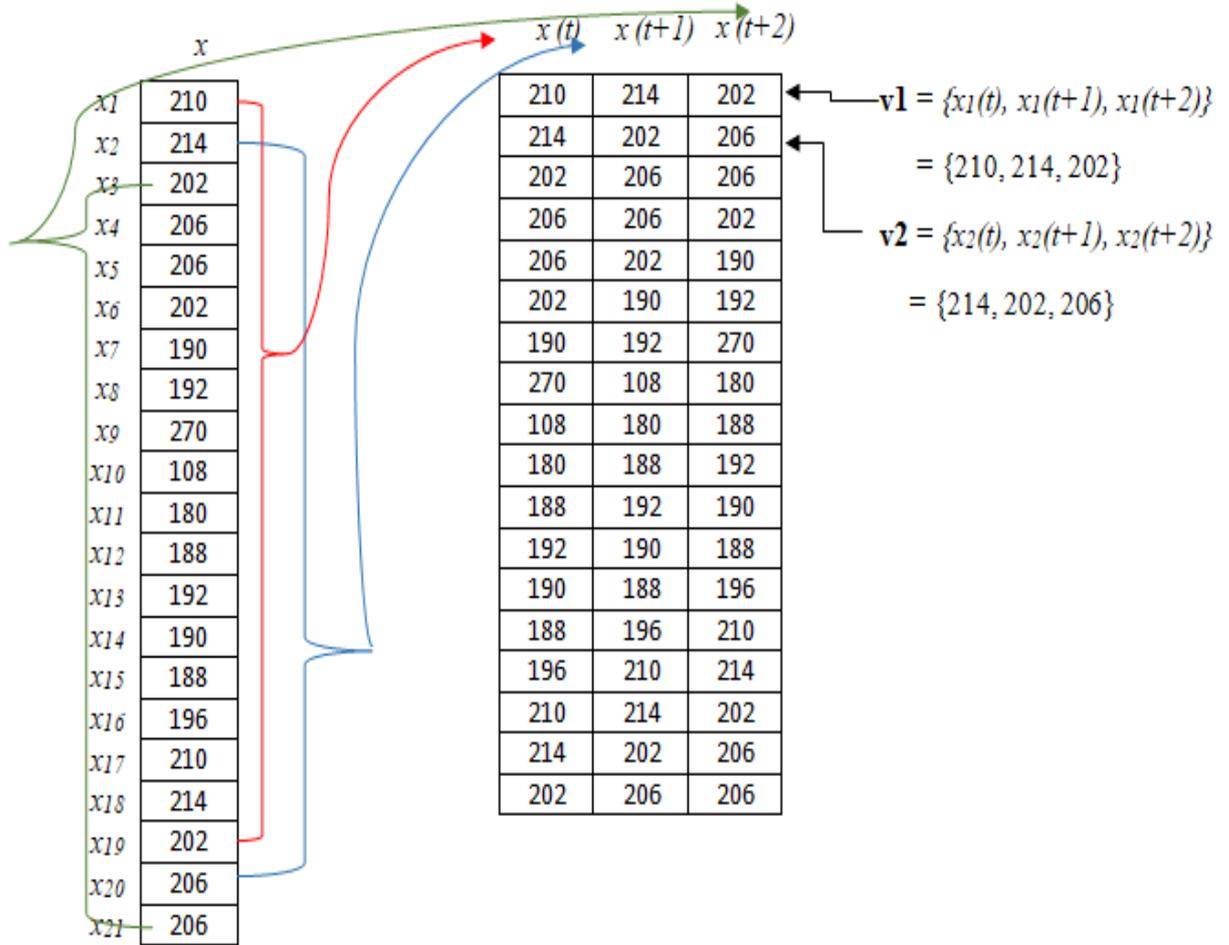


Figure 5.8: Reconstruction of the state-space with an embedding dimension of 3, from a scalar time series represented by a sequence of  $x_1, x_2, \dots, x_{21}$ , considering Time Lag ( $t$ ) = 1.

1),  $x_1(t+2)$ ] (that is  $v_1 = [210, 214, 202]$ ), and  $v_2 = [x_2(t), x_2(t+1), x_2(t+2)]$  (that is  $v_2 = [214, 202, 206]$ ).

2.1. The distance between the two vectors in the two-dimensional space is:

$$\begin{aligned}
 d(x_t - \tilde{x}_t) &= \sqrt{(210 - 214)^2 + (214 - 202)^2} \\
 &= \sqrt{-4 + 12} \\
 &= \sqrt{8} \\
 &= 2.8284
 \end{aligned}$$

2.2. The distance between the two vectors in the three-dimensional space is:

$$\begin{aligned}
 d(x_{t+1} - \tilde{x}_{t+1}) &= \sqrt{(210 - 214)^2 + (214 - 202)^2 + (202 - 206)^2} \\
 &= \sqrt{-4 + 12 + (-4)}
 \end{aligned}$$

$$= \sqrt{4}$$

$$= 2$$

2.3. Calculate  $FNN(x, \tau, m) = \left| \frac{d(x_t, \tilde{x}_t) - d(x_{t+1}, \tilde{x}_{t+1})}{d(x_t, \tilde{x}_t)} \right| > R_{tol}$

$$= (2.8284 - 2) / 2.8284$$

$$= 0.2929$$

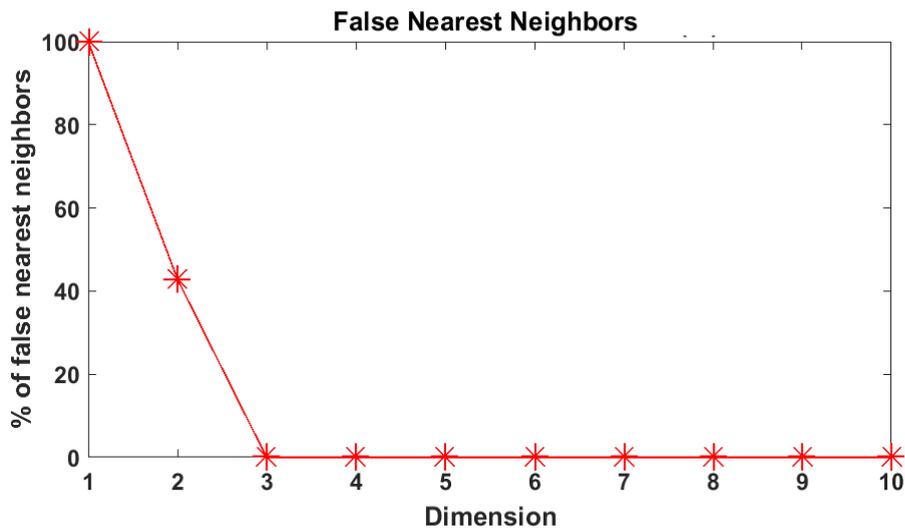


Figure 5.9: The percentage of *False Nearest Neighbour* example. The percentage of *False Nearest Neighbour* hits zero when the dimension is three.

The *False Nearest Neighbour* value obtained is less than the default threshold value  $R_{tol} = 15$ , when dimension is 3, as shown in Figure 5.8. Therefore, these two vectors are not *False Nearest Neighbour*. In this way, every point is examined to calculate how many nearest neighbours are false neighbours, and the percentage of *False Nearest Neighbour* to true nearest neighbour is computed at different dimensional space. The percentage of *False Nearest Neighbour* should drop at a higher-dimensional space as the dynamics of the system are being unfolded (Schiecke et al., 2016). The value of the dimension where the percentage of *False Nearest Neighbour* reaches 0 is considered as the dimension that is large enough to describe the dynamics of the system. That dimension is selected as the embedding dimension. In the case of the example above, the accurate dimension would be 3.

### 5.6.3 An Example of Lorenz Attractor

The well known Lorenz Attractor as shown in Figure 5.10 has three underlying cross-coupled variables. However, the attractor itself is almost two-dimensional. The minimum dimension of the attractor for the original Lorenz dataset, and new noisy version in which a significant amount of normal noise has then added. The visual effect of the noise is shown in Figures 5.10 and 5.11.

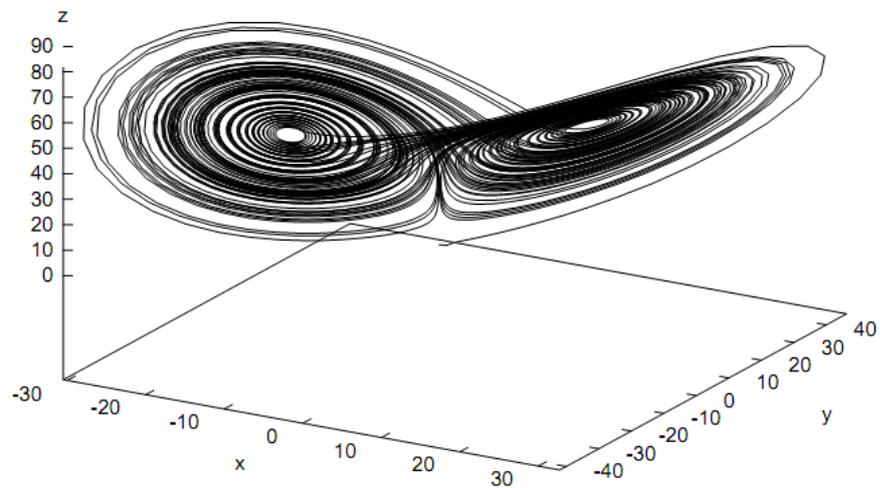


Figure 5.10: A visualisation of the Lorenz attractor in 3-dimensional phase space  $x(t)$ ,  $y(t),z(t)$ (Frank et al., 2001).

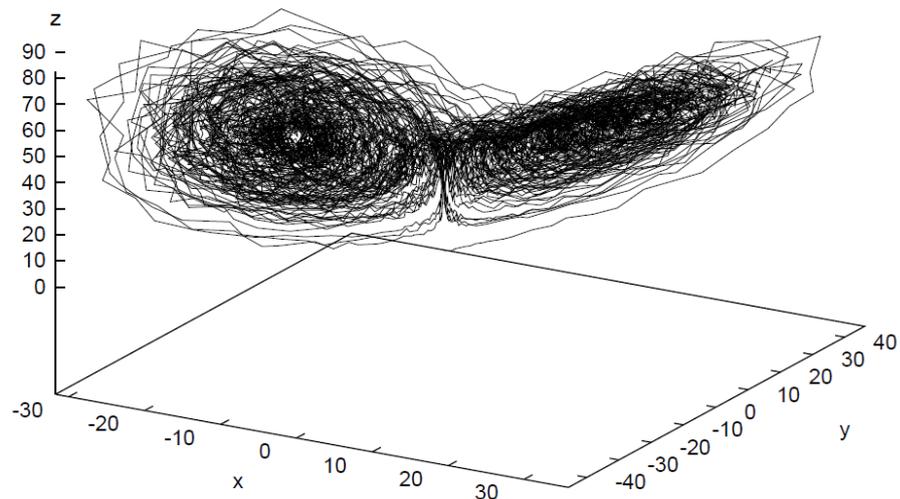


Figure 5.11: A visualisation of the Lorenz attractor in Three-dimensional phase space  $x(t)$ ,  $y(t),z(t)$  with Noise (Frank et al., 2001).

<i>Embedding Dimension</i>	<i>Percentage of False Nearest Neighbours</i>	
	<i>Clean Data</i>	<i>Noisy Data</i>
2	77%	77%
3	3.3%	3.3%
4	0.3%	0.3%
5	0.3%	0.3%

Figure 5.12: The percentage of *False Nearest Neighbours* in the Lorenz data set (Frank et al., 2001).

The false nearest neighbour results as shown in Figure 5.12, suggest that, in both cases, an embedding of 4 or 5 should be sufficient to represent the attractor. This result corresponds well with the theoretical upper bound of about 5, from the embedding theorem. The singular-value analysis shows the contrast between the clean and noisy data very clearly. For the noisy data, a “noise floor” for eigenvalues five onwards is apparent, but such a floor is, as expected, utterly absent from the clean data (Frank et al., 2001).

## 5.7 Conclusion

In this chapter, the explanations and examples of the methods used for nonlinear time series analysis have been discussed and illustrated. With the extensive use of the methods, such as *Approximate Entropy*, *Sample Entropy*, *Pearson’s Correlation Coefficient*, *Cross-Correlation*, and *Embedding Dimension (ED)* for the time series analysis, these measures prove to be appropriate to investigate the EEG and HRV time series.

*Approximate Entropy* was introduced (Pincus, 1991) to measure system complexity, which could be applied to different time series. However, due to few shortcomings with *Approximate Entropy*, such as bias, relative inconsistency and dependency with the sample length, a new complexity measure called *Sample Entropy* was proposed by (Richman and Moorman, 2000). Many studies such as (Alcaraz and Rieta, 2010), (Abásolo et al., 2006), (Ramdani et al., 2009), (Al-Angari and Sahakian, 2007), (Yoo et al., 2012), and (Takahashi et al., 2010) have concluded that *Sample Entropy* is a robust quantifier of complexity, which offers an accurate nonlinear metric for quantification (Alcaraz and Rieta, 2010). It gives a good dynamical signature and is a helpful tool that provides insights into various biological time series (Ramdani et al., 2009). Therefore, *Sample Entropy* has been considered as an effective method for investigating EEG and HRV time series data.

In order to find the correlation between different time series data, *Pearson’s Correlation Coefficient* is an ideal measure. However, it requires the same length of input data for the

comparison. Therefore, *Cross-Correlation* has been utilised to find the correlation between time series data, along with *Pearson's Correlation Coefficient*.

Reconstructing of a state-space is a very important procedure in terms of doing non-linear time series analysis. *Embedding Dimension* is required for the reconstruction of the state-space. There are several different algorithms to find an embedding dimension, but I have presented the most commonly used methods, *False Nearest Neighbour* method, to find an *Embedding Dimension*. The *Embedding Dimension* plays a vital role in nonlinear time series analysis(Chun-Hua and Xin-Bao, 2004). It has been successfully used in neural network approaches for time series prediction (Frank et al., 2001). They concluded that optimal performance could be achieved using the correct embedded dimensions. Furthermore, *Embedding Dimension* has been adopted by (Wendi et al., 2018), for recurrence plot generation of the reconstructed phase space to represent many real application scenarios when not all variables describe a system are available.

## Chapter 6

# The Correlation between EEG Signals Measured at Different Positions on Scalp Varying with Distance

*Results shown in Figures 6.3-6.7 are already published in the conference proceeding of BICA (9th International Conference Biological Inspired Cognitive Architecture, Moscow, Russia, August 2017).*

### 6.1 Introduction

Electroencephalogram (EEG) is a time varying signal, and different position of electrodes gives different time varying signals. There might be a correlation between these signals. It is likely that the correlation relates to the actual positions of the electrodes. This Chapter investigates the correlation of EEG signals in the TD using Cross-Correlation method introduced in Chapter 5. This analysis was carried out on a dataset with and without medical conditions.

Cross-Correlation can be used to analyse two related processes with time delay. In the present context, it offers a valuable and sensitive method for investigating EEG signals that are recorded at the same time from different electrodes. To analyse EEG signals in TD, CC method stands out as the most appropriate correlation method, because of its ability to assess signal similarity at all possible time delays. CC has been successfully applied in analysing EEG signals in the FD (Li et al., 2013), as well as TD (Bob et al., 2010). This method can be used to determine the relationship between activities in global and local areas, and also among the different local areas of the human brain.

The brief review of research on the correlation of EEG signals, as discussed in Chapter 2, indicated that investigations have been focused on the FD, and limited information has

been found in the correlation of EEG signals in the TD. In addition, this review has shown that the numbers of paired electrodes investigated, the number of datasets used, and use of CC for analysing EEG signals are limited. Instead, researchers' focus has been primarily on electrode combinations within the left and right brain hemispheres, respectively. It has been found from the recent research, that EEG signals with medical conditions tends to be similar when compared the EEG signals without any medical conditions from, and shows possible dysfunction. Huge amount of research is already conducted to understand correlation between EEG signals. Those researches imply that EEG might change gradually from one region to another. But, to my knowledge, I have not seen anybody showing how EEG signal (measured through electrode) changes with electrode location and the effect of the distance between two EEGs on the correlation performance is not considered. Talking about a distance between electrodes, there could be three possible definitions: a) straight-line distance, b) geodesic distance and c) travelled distance between neurons. This hypothesis is investigated in this chapter. The motivation is to find whether or not there is any variation in correlation value of EEG signals with distance between them. The following Table 6.1 shows the research gap I have addressed in this chapter.

Table 6.1: Summary of Work in this Chapter

<b>This Study</b>	<b>Electrodes</b>	<b>Paired Electrodes</b>	<b>TD</b>	<b>FD</b>	<b>CC Method</b>	<b>Other Methods</b>
Dataset 1	19	171 pairs	✓	-	✓	-
Dataset 4	10	45 pairs	✓	-	✓	-
Dataset 5	15	105 pairs	✓	-	✓	-
Dataset 6	19	171 pairs	✓	-	✓	-
Dataset 7	19	171 pairs	✓	-	✓	-
Dataset 8	19	171 pairs	✓	-	✓	-

To my knowledge, very limited work has been done on the correlation of EEG signals using all number of electrodes and their possible electrode pair combination as shown in Chapter 2. In this chapter, six datasets (3 datasets without any medical condition, and 3 datasets with a medical condition) are used to investigate the differences between EEG signals of subjects with and without medical conditions. Three datasets contain EEG signals of participants without any medical condition, and are named as Dataset 1, Dataset 4, Dataset 5, and three datasets contain EEG signals of participants with a specific medical condition and are named as Dataset 6, Dataset 7, and Dataset 8, as described in Chapter 3. Each dataset includes a different number of electrodes. Therefore, the number of unique electrode pairs to perform Cross-Correlation is different. From Datasets 1, 4-8 there are 171, 45, 105, 171, 171, and 171 pairs, respectively.

## 6.2 Dataset Information

Six datasets used in this study consist of different numbers of participants, electrodes and conditions. All of these datasets following the standard 10%-20% system (Klem et al., 1999), as shown in Chapter 2 (Figure 2.10). For more detail about the six datasets please refer to Chapter 3.

## 6.3 Experiments and Results

### 6.3.1 Experiments

The EEG signals were processed to remove artefacts, such as eye blinks, eye movements, jaw movements and muscle movements, by using Independent Component Analysis (ICA), as specified in Chapter 4. Cross-Correlation has been calculated on the processed EEG signals for the 171 (i.e.  $19 \times 18 / 2 = 171$ ) electrode pairs of Dataset 1, the 45 (i.e.  $10 \times 9 / 2 = 45$ ) electrode pairs of Dataset 4, the 105 (i.e.  $15 \times 14 / 2 = 105$ ) electrode pairs of Dataset 5, 171 (i.e.  $19 \times 18 / 2 = 171$ ) electrode pairs of Dataset 6, 171 (i.e.  $19 \times 18 / 2 = 171$ ) electrode pairs of Dataset 7, and 171 (i.e.  $19 \times 18 / 2 = 171$ ) electrode pairs of Dataset 8.

In order to obtain the distance in centimetres (cm) between electrodes, a measuring tape was used. For most of our participants from Datasets 1, 4, 5 and 6 the head circumference range was 54-58cm, for which a medium-sized 'electro-cap' is appropriate. According to (Song et al., 2015), the circumference of the medium-sized EEG cap is ideal for 64% of adults, both male and female. Therefore, I utilized a medium-sized EEG cap made of an elastic material which stretches according to the participant's head circumference, and measured distances using a straight line on the cap - not a curved line over the skull. Note that the distance between electrodes as shown in Figure 6.4-6.9 is a straight line distance between electrodes, not the distance as measured over the surface of the scalp.

### 6.3.2 Results

The maximum absolute correlation was found at lag 0. For example, Figure 6.2 shows the information for all electrode pairs for the Electrode Fp1 from Dataset 1, where all signals measured from the frontal lobe of the brain are positively correlated with electrode Fp1, and all signals measured from the back lobe of the brain are negatively correlated with electrode Fp1. The results of all other electrode pairs are available in Appendix B (Figure B.1-B.8). The other datasets present similar results as shown in Figure 6.1. Figure 6.2, and Figure 6.3 illustrate the closer look from Figure 6.1 for electrode pairs Fp1-Fp2, and Fp1-O1,

respectively, as an example. In Figures 6.2 and 6.3, the x-axis (horizontally) denotes the time lag where a lag of 1 corresponds to 4 milliseconds. Positive time lags (0 to 1000) and negative time lags (0 to -1000) indicate when the signal (the signal which being compared) shifts to the right and left the side of the reference signal Fp1, respectively. The y-axis (vertically) denote the Cross-Correlation value. The blue colour line is for an individual participant, and the red is for the average of all participants for each electrode pair.

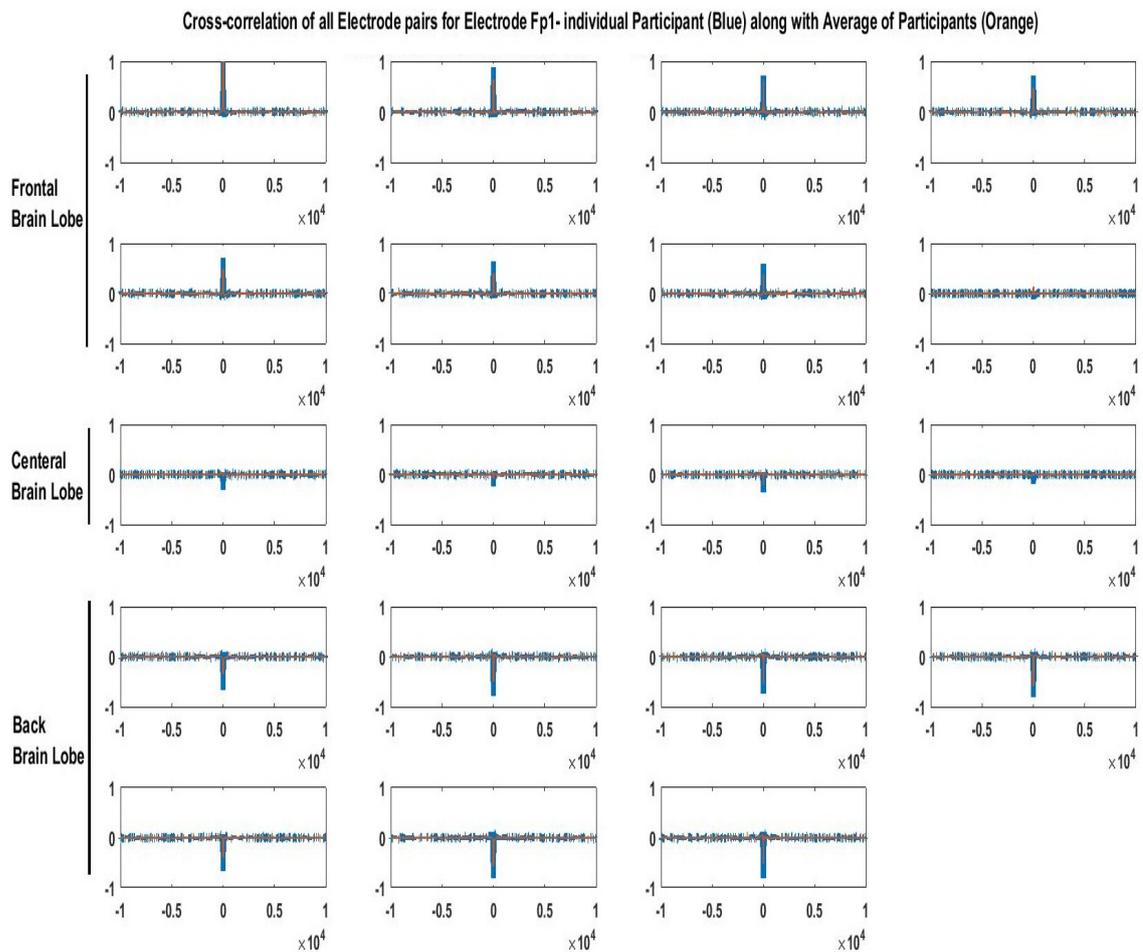


Figure 6.1: Cross-Correlation at all possible Lags for all electrode pairs for Electrode Fp1. Blue colour shows the result of one of the participants, and Orange colour shows the average of 15 participants result from Dataset 1

Electrode Fp1 from Dataset 1, Dataset 5, Dataset 6, Dataset 7, Dataset 8, and Electrode F7 from Dataset 4 have been chosen randomly from both hemispheres to show the Cross-Correlation performance. The result for other electrodes from these datasets are available in Appendix B (Figure B.9-B.23). The results show the averages for all participants. Figures 6.5, 6.6, 6.7, 6.8 show that there is an inverse linear relationship between Cross-Correlation

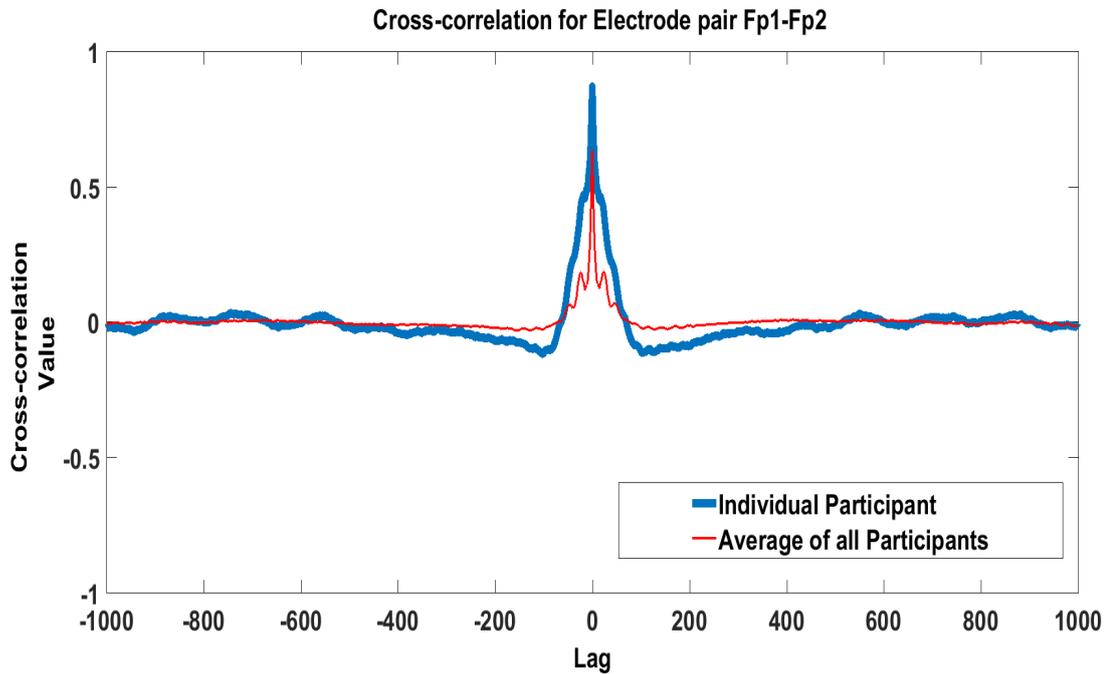


Figure 6.2: Dataset 1 - Positive Cross-Correlation at Lag 0.

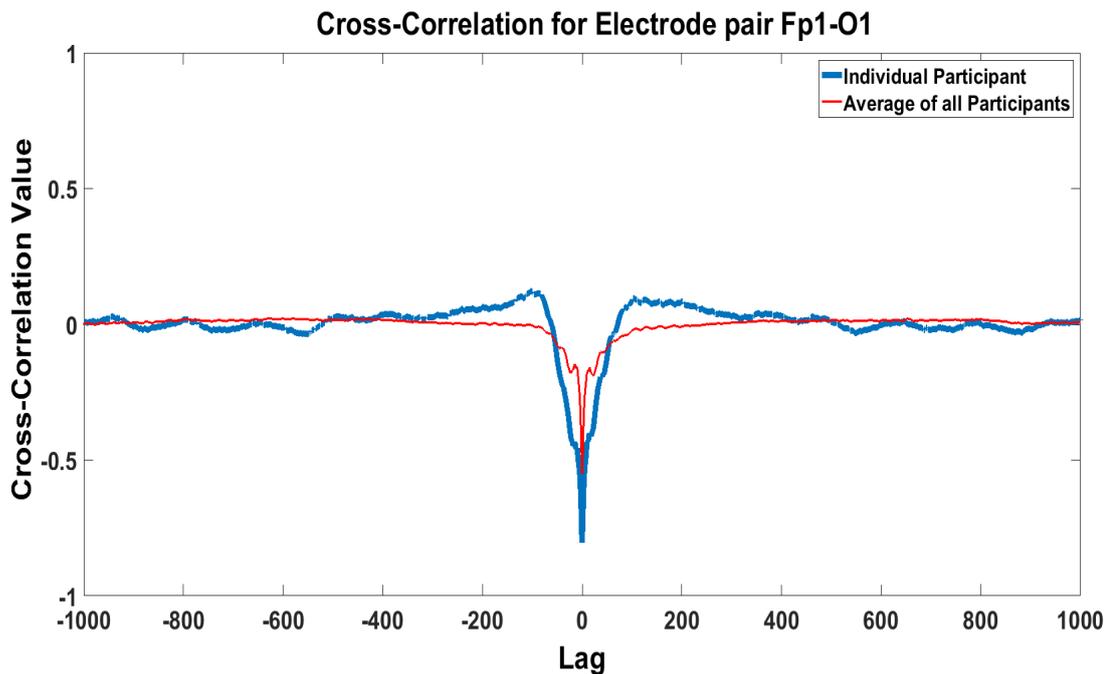


Figure 6.3: Dataset 1 - Negative Cross-Correlation at Lag 0.

and distance. Whereas, Figures 6.9, and 6.10 do not show such a linear dependency. The linear regression has been plotted to fit the data with a probability of  $p < 0.001$ . This indicates that the Cross-Correlation value decreases while the distance from Fp1 and F7

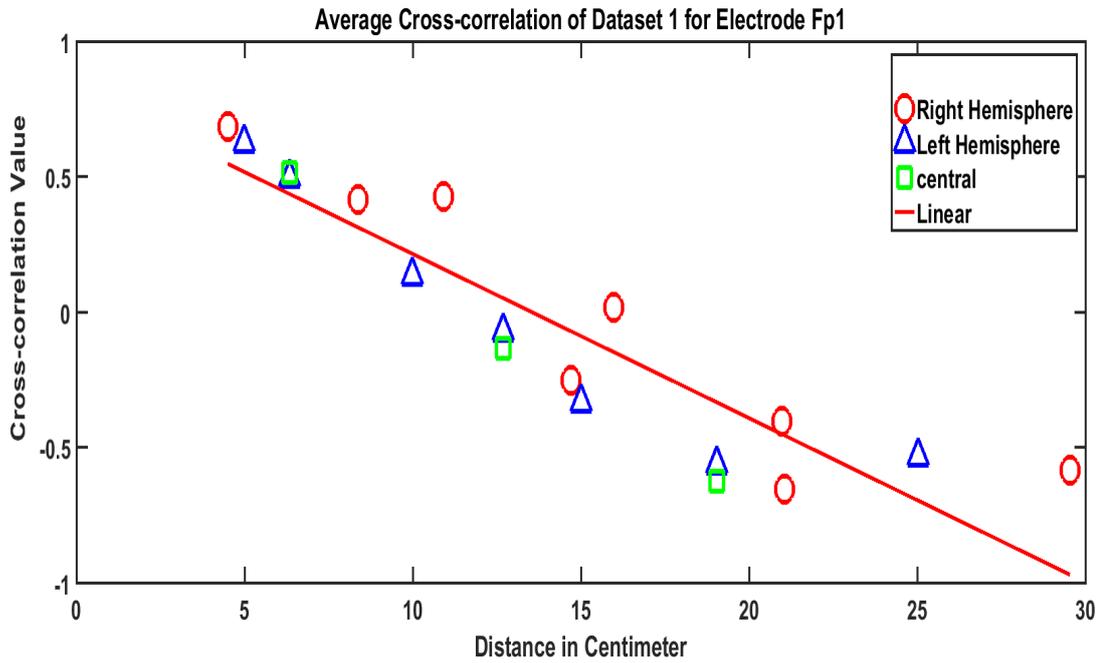


Figure 6.4: Cross-Correlation between electrodes at varying distance on Dataset 1.

increases, irrespective of brain hemisphere.

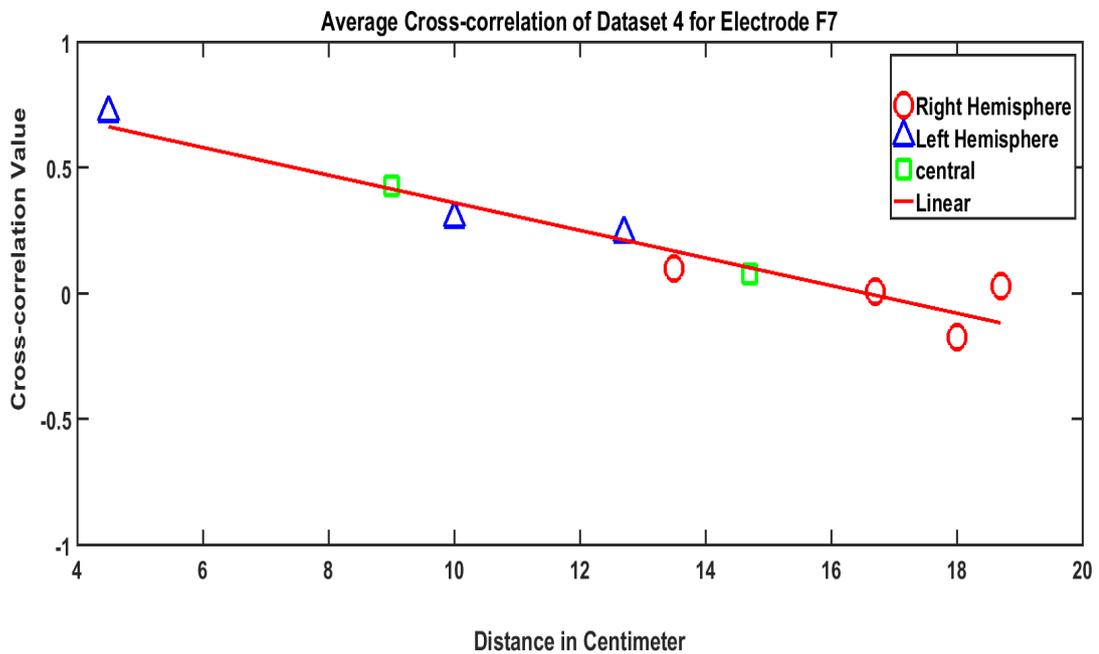


Figure 6.5: Cross-Correlation between electrodes at varying distance on Dataset 4.

The results of all other electrodes of all six datasets are similar to the ones shown in Figures 6.4, 6.5, 6.6, 6.7, 6.8, and 6.9. In order to evaluate linear dependency, the

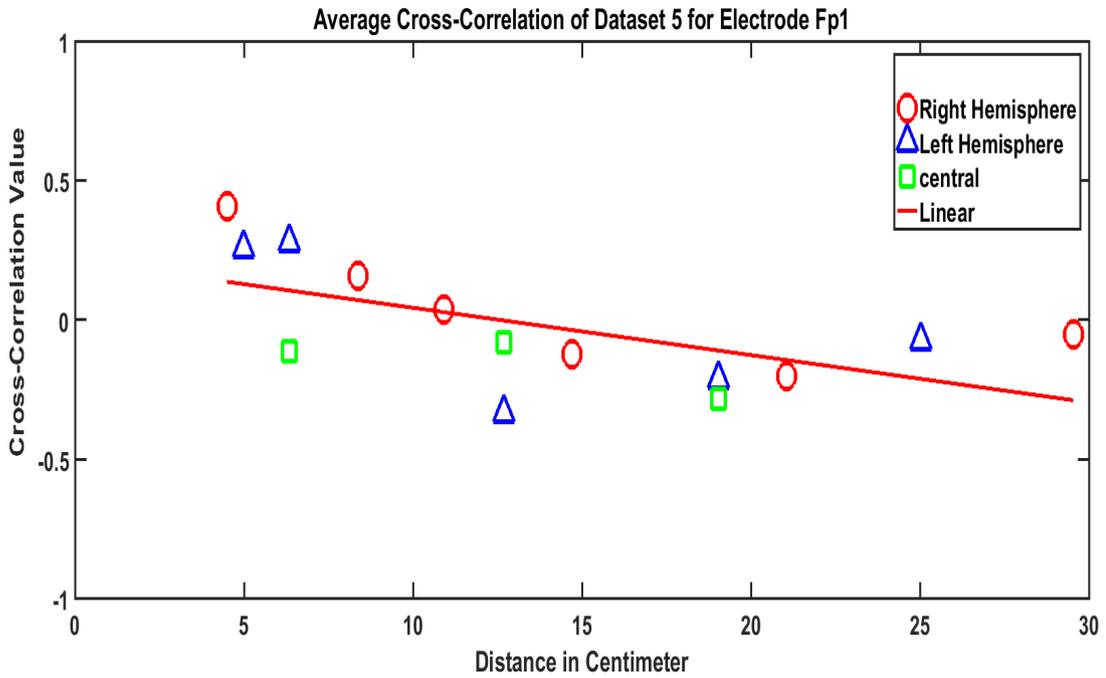


Figure 6.6: Cross-Correlation between electrodes at varying distance on Dataset 5.

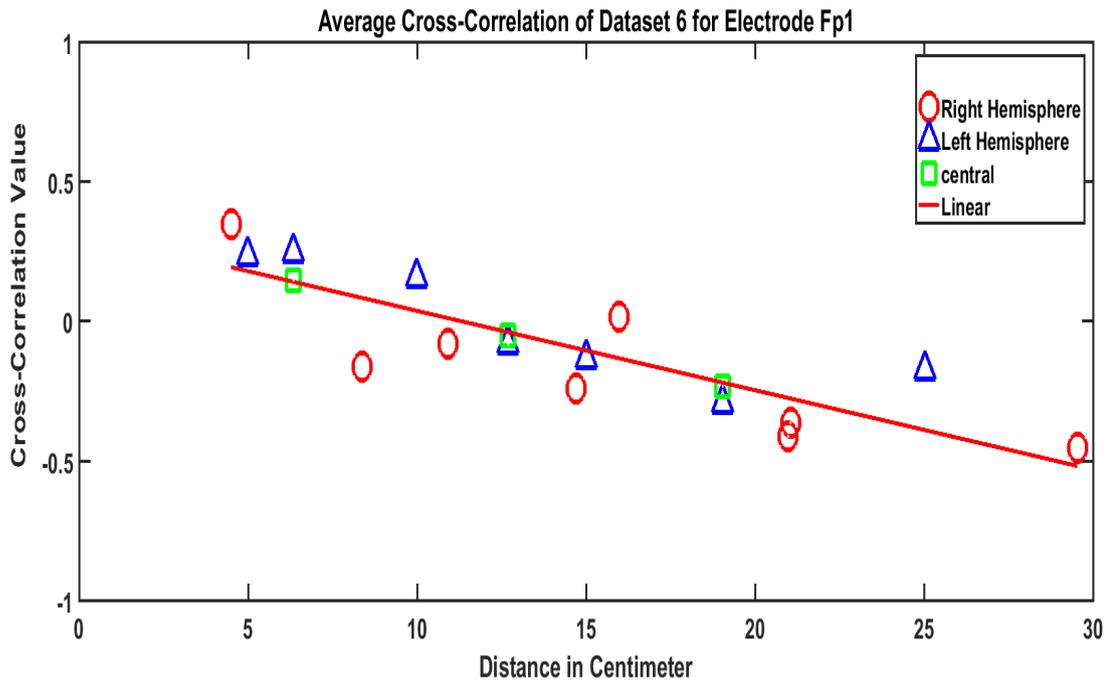


Figure 6.7: Cross-Correlation between electrodes at varying distance on Dataset 6 (Autism).

differences between each Cross-Correlation value and the corresponding value on the fitted linear function line were calculated. The results show that the difference between them is tiny (about 0.03, giving us 98% accuracy in the result obtained), which suggests that a

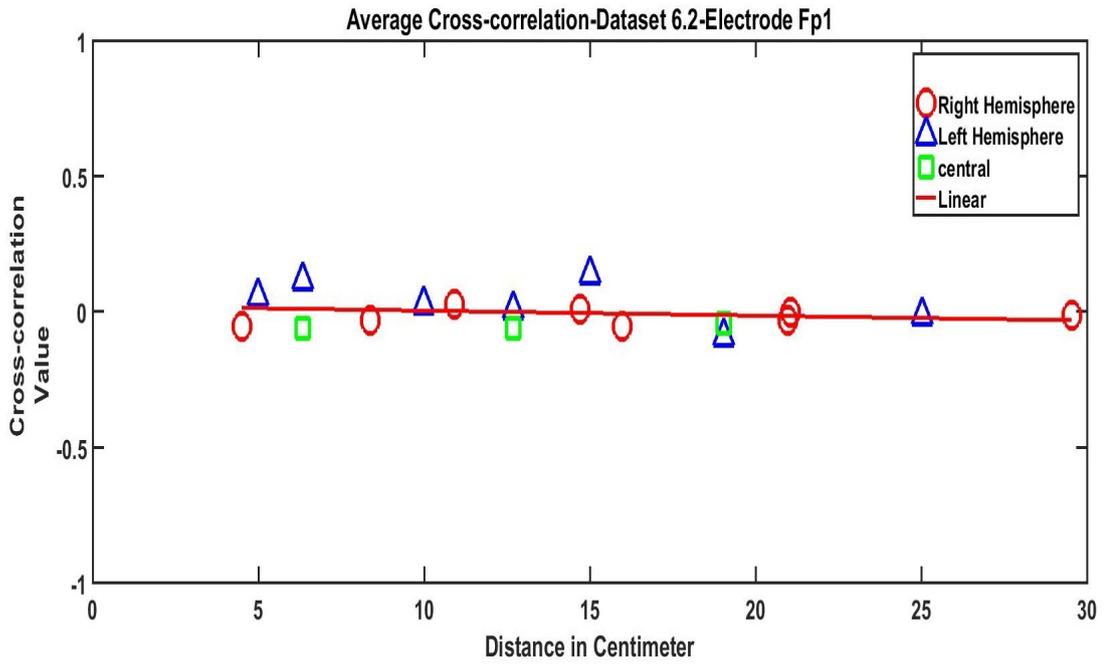


Figure 6.8: Cross-Correlation between electrodes at varying distance on Dataset 7 (Epilepsy).

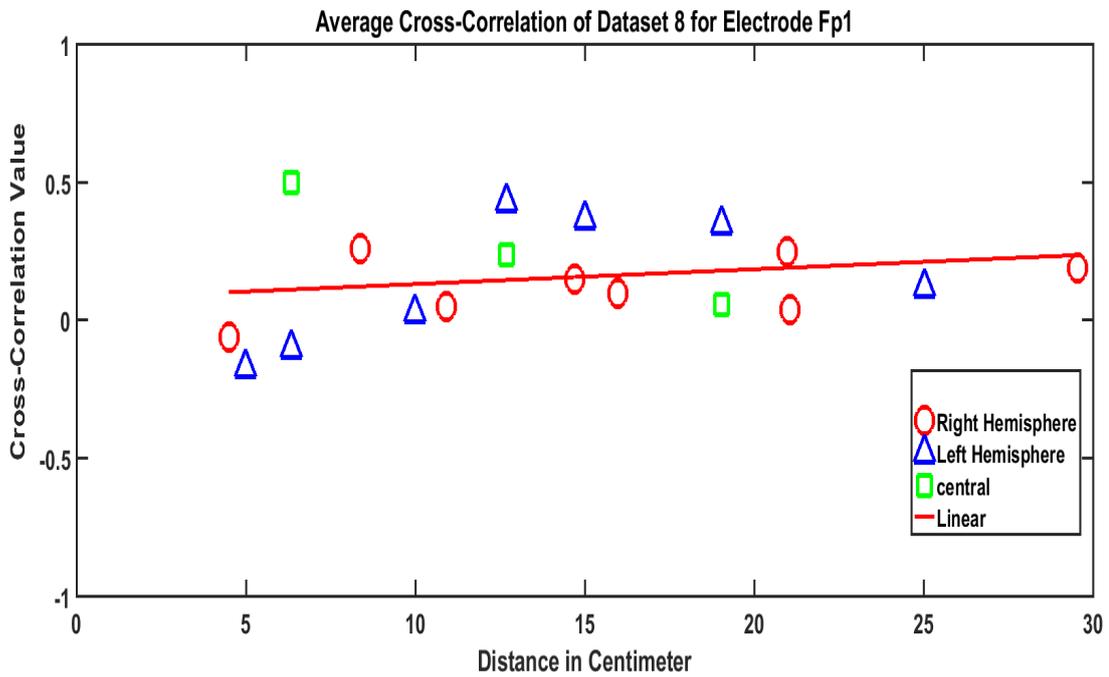


Figure 6.9: Cross-Correlation between electrodes at varying distance on Dataset 8 (Seizure).

close linear dependency exists in Figures 6.4, 6.5, 6.6, and 6.7. Whilst, Figures 6.8 and 6.9 do not show linear dependency with the distance.

## 6.4 Conclusion

The result suggest that, for dataset without any medical condition, the linear dependency with the physical distance measured at different position on the scalp is observed. However, for the dataset with medical conditions (excluding Autism dataset), such as Epilepsy and Epileptic Seizure, the linear dependency does not exist.

One of the main conclusions of this work is that electrical activity correlates linearly with distance within the brain, that is when the distance increases the correlation decreases. To my knowledge, previous research has not described this linear relationship in Time Domain (TD) (Li et al., 2013), (Jeong et al., 2015), (Bob et al., 2010), (Cuevas and Bell, 2011). The results from this research cover a gap in the research concerning the correlation of EEG signals in the TD using Cross-Correlation and combinations of all electrode pairs. It is important to note that the straight line distance between two electrodes is measured directly using a straight line on the cap, and not over the surface of the skull.

It is interesting to see that participants without medical condition shows the linear dependency with the physical distance on the scalp. However, participants with medical conditions such as Epilepsy and Epileptic Seizure, the linear dependency does not exist. It is important to note that data presented in Figure 6.8 and 6.9 have been collected during seizures. I have a lack of the expertise to find out the reason for this, but I think this might be of interest to the people working in the medical area.

The second conclusion from this work is that the correlation is independent of brain hemisphere for all six datasets. I do not know the exact reason for this independence of correlation result. However, it is well known that white matter in the brain could be significant in the transmission of electrical activity. Brain's white matter is a labyrinth underneath the outer gray matter, which is comprised of myelin-covered bundles of axons that connect billions of neurons and carry electrical signals between brain regions (Fields, 2008). According to the research (Durante et al., 2018), highly creative people have significantly more white matter connections between the right and left hemispheres of the brain. Connectomics research suggests that functional connectivity between brain regions may be linked to IQ (Men et al., 2014). Hopefully, future research will help us pinpoint a variety of ways to optimize functional connectivity and coordination between brain hemispheres to boost cognitive function and creative capacity for people from all walks of life across their lifespan. One issue I address here is how electrical activity can be communicated across the surface of the brain. My research focuses on evaluating the correlation of EEG signals between different brain regions. The aim is to determine the relationship between EEG

signals and electrode location on the scalp, and to check whether this relationship differs in the two brain hemispheres.

In summary, regardless of the anatomical substrates involved, the main finding is that the correlation between electrical activities in different parts of the brain is linearly related to the electrode distance between them for the participants without any medical condition across both hemispheres. Whereas, for participants with medical conditions such as Epilepsy and Epileptic Seizure, there is no such relationship found.

# Chapter 7

## Correlation Analysis between EEG and HRV Time series data

*Results shown in Figure 7.6, Table 7.2 and Table 7.3 have already been published in the conference proceedings of BIOSTEC (11th International Conference on Biomedical Engineering System and Technology, Funchal, Madeira, January 2018).*

### 7.1 Introduction

The correlation between biomedical signals, such as electroencephalograms (EEG) and electrocardiograms (ECG) time series signals, has been analysed using Pearson correlation coefficient method (Miyashita et al., 2003), (Yang et al., 2002), (Ako et al., 2003), (Jurysta et al., 2003), (Takahashi et al., 2005), (Edlinger and Guger, 2006), (Berg et al., 2005), (Sakai et al., 2007), (Abdullah et al., 2010), (Chua et al., 2012), (Kim et al., 2013), (Prinsloo et al., 2013), (Liou et al., 2014), (Triggiani et al., 2016). Although, Wavelet Transformations (WT) have been performed on time series data, including EEG and ECG signals, so far the correlation between signals preprocessed by WT has not been analysed. In this Chapter, the correlation between the EEG and ECG, with and without WT is analysed.

WT acts on the frequency and time of the recorded signals. Therefore, WT has been widely utilised for analysing time series including biomedical signals. The WT of a signal can be thought of as an extension of the classic Fourier transform (FT) - it works on multi-scale basis (Time and Frequency), instead of working on a single scale (Time or Frequency) as FT, and gives detailed and clear information of the signal. Recent research on the correlation between EEG and HRV, as seen in Chapter 2, has focused on Fourier analysis of the frequencies presented in these signals. These research to analyse brain's functionalities under certain conditions and to check whether these functionalities are related to each other.

The literature review in Chapter 2 shows the correlation between EEG and HRV has been analysed in FD using well-known method Pearson correlation coefficient (PCC). It is worth noting that, different numbers of EEG electrodes have been used to analyse the relationship between an EEG and HRV. The results from studies, as shown in Chapter 2, have shown a correlation between EEG and ECG, but each of the studies only focused on part(s) of the brain, rather than the whole brain. Therefore, the results are not comprehensive. For example, (Na et al., 2002) indicates that EEG from the left side of the brain is correlated with ECG, and (Bob et al., 2010) demonstrated the correlation was from the right side of the brain. To the best of my knowledge, very limited work has been done on the correlation between EEG and HRV, with and without WT signals, using 19 EEG electrodes. Moreover, no one has analysed these signals under the same condition (with TEAS acupuncture applied) that is utilised in this chapter. The aim of this chapter is to check if there is any particular area (not just one part) within the brain having a stronger correlation between EEG and ECG. Therefore, the correlation between EEG and HRV, with and without WT signals in FD using Pearson correlation coefficient considering all 19 EEG electrodes under the same condition are shown in this chapter. FD illustrates the functionality of EEG and HRV, comparing this functionality will be the base for the analysis. However, in this chapter the analysis of correlation performance between EEG and HRV in the TD is also presented. The performance of TD and FD is then analysed. Table 7.1 shows the information about the datasets utilised in this chapter.

Table 7.1: Summary of the Research on the correlation of EEG and HRV.

Datasets Utilised	TD	FD	PCC Method	Other Method	EEG Electrodes	Condition
Dataset 1	✓	✓	✓	-	19	TEAS
Dataset 2	-	✓	✓	-	19	TEAS

It is obvious from the research on WT as shown in Table 2.3 in Chapter 2, that extracting the key feature of EEG and ECG signals can improve the analytical performance. Therefore, it is interesting to analyse not just either EEG or ECG, but also the correlation between EEG and HRV. To my knowledge, I have yet to find the research on the correlation between wavelet transformed signals. In this work, we describe such an analysis.

### 7.1.1 Dataset Information

Two different datasets, Dataset 1 and Dataset 2, were obtained from Chapter 3, with each of them containing different numbers of participants, acupuncture stimulation location, and total time length, as shown in Chapter 3 (Table 3.1).

## 7.2 Experiments

The EEG signals were pre-processed to remove artefacts caused by the electrical activities in muscles, including eye, jaw and muscle movements using Independent Component Analysis (ICA), as explained in Chapter 4. It was straightforward to remove these using ICA (Hyvärinen and Oja, 2000).

The HRV signals were pre-processed to remove ectopic beats, extra beats, and premature ventricular contractions artefacts using *Kubios HRV standard* software (Tarvainen et al., 2014). This software considers the threshold-based correction method, in which the artefacts and ectopic beats are simply corrected by comparing every IBI value against a local average interval. The local average is obtained by median filtering the IBI time series, and thus, the local average is not affected by single outliers in IBI time series. If an IBI differs from the local average more than a specified threshold value (threshold in seconds), the interval is identified as an artefact and is marked for correction. We can select the value for threshold from:

1. Very low: 0.45 seconds
2. Low: 0.35 seconds
3. Medium: 0.25 seconds
4. Strong: 0.15 seconds
5. Very strong: 0.05 seconds
6. Custom, for setting a custom threshold in seconds

For example, the "Medium" level will identify all IBIs that are larger/smaller than 0.25 seconds compared to the local average. The correction is made by replacing the identified artefacts with interpolated values using a cubic spline interpolation (Daskalov and Christov, 1997). The cubic spline interpolation is a series of unique cubic polynomials fitted between each of the data points, with the stipulation that the curve obtained are continuous and appear smooth. These cubic splines can then be used to determine rates of change and cumulative change over an interval (Unser, 1999). Please note that the thresholds shown above are when 60 beats per minute (bpm) heart rate and are adjusted according to mean heart rate (that is lower thresholds for a higher heart rate). The correction level has been adjusted individually, because inter-individual difference in HRV is significant and therefore a fixed threshold did not work optimally for all participants. The optimal threshold

is the lowest correction level, which identifies all artefacts, but does not identify too many normal RR intervals as artefacts.

The sampling rate is 1Hz for the extracted HRV, and 250Hz for the EEG. In order to perform *Pearson Correlation Coefficient (PCC)*, which requires the same sampling rate for both signals, I have reduced the sampling rate of EEG to be the same as the one for the HRV.

### **7.2.1 Time Domain (TD) Analysis**

For TD analysis, the sampling rate was reduced by segmenting EEG signals using 1 second window and representing each window by its mean value (the mean amplitude value of each segment - 250 samples), unlike normal down sampling with which much of the data are thrown away. For each participant's EEG signal, this process was repeated for all 5 minute slots. The aim of the TD analysis is to identify whether the correlation changes with change in amplitude of EEG and HRV signals.

### **7.2.2 Frequency Domain (FD) Analysis**

The power spectrum for each frequency band of EEG - Delta (0.3-4 Hz), Theta (4-7.5 Hz), Alpha (7.5-13 Hz), Beta (13-30 Hz), and Gamma (30-50 Hz) were obtained by Power Spectrum Density (PSD)- the measure of signal's power content versus frequency (Stoica and Moses, 1997). Similarly, the power spectrum for each frequency band of HRV - VLF (0-0.04 Hz), LF (0.04-0.15 Hz), and HF (0.15-4Hz) were obtained by Power Spectrum Density (PSD) (Stoica and Moses, 1997). The reason for using PSD is to describe the distribution of power into each frequency band of EEG and HRV. The aim of the FD analysis is to identify whether the correlation changes with change in the frequency power of EEG and HRV signals. Furthermore, FD analysis is the focus of the researchers when it comes to the correlation between EEG and HRV signals, as shown in Chapter 2.

For FD analysis, the segmentation was done using two different approaches (Method 1 and Method 2) to perform PCC between EEG and HRV. The reason for using two different approaches is to check whether the correlation changes with the change in the way of segmenting the signals.

#### **7.2.2.1 Method 1**

This method is about segmenting EEG signals using 1 second window and representing each window by its mean value (the mean value of each segment - 250 samples), unlike normal down sampling with which much of the data are thrown away. For each participant's

EEG data, this process was repeated for all 5 minute slots. After windowing, spectral analysis was performed. From each frequency band of the EEG and the HRV, the mean of the amplitude value across the frequency range has been measured. From each 5 minute slot, a single value for each frequency range was obtained. This means from each five-minute slots, five values representing five frequency ranges. This process was repeated for all ten slots, and ten values obtained for each frequency range. Then, the correlation between these frequency values obtained from each slot (that is 10 frequency values for each frequency range of EEG and HRV) is performed. Figure 7.1 illustrates the experimental process using Method 1.

#### **7.2.2.2 Method 2**

This method is about segmenting EEG signals using 1 second window and perform spectral analysis on each 1 second window to extract frequency bands first. After windowing, from each 1 second window, the mean of the amplitude value within the frequency range has been measured for each frequency band of the EEG and the HRV. Once single amplitude value was gathered for each frequency band, there were 300 amplitude values obtained in total from 300 windows ( $1 \text{ Second} \times 5 \text{ Minute} = 300 \text{ Seconds}$ ). From each 5 minute slot, a single value for each frequency range was obtained. This means from each five-minute slots, five values representing five frequency ranges. This process was repeated for all ten slots, and ten values obtained for each frequency range. Then, the correlation between these frequency values obtained from each slot (that is 10 frequency values for each frequency range of EEG and HRV) is performed. Figure 7.2 illustrates the experimental process using Method 2.

In order to perform correlation based on wavelet transformed EEG and/or HRV signal, the *WT-Daubechies Wavelet* up to level 5 is performed on the signals before extracting frequency bands. However, loss of information from the signals has been found with level 4 and level 5. Therefore, the *WT-Daubechies Wavelet* up to level 3 is performed on the signals before extracting frequency bands for both methods Method 1 and Method 2, respectively. For the datasets I have, the low-pass filter worked very well because the low-pass filter contains the high scale and low frequencies of the signal. As shown in Figure 7.3, the low-pass filter works better than the high-pass filter on EEG signals. Therefore, we considered low-passed WT signals to perform the correlation.

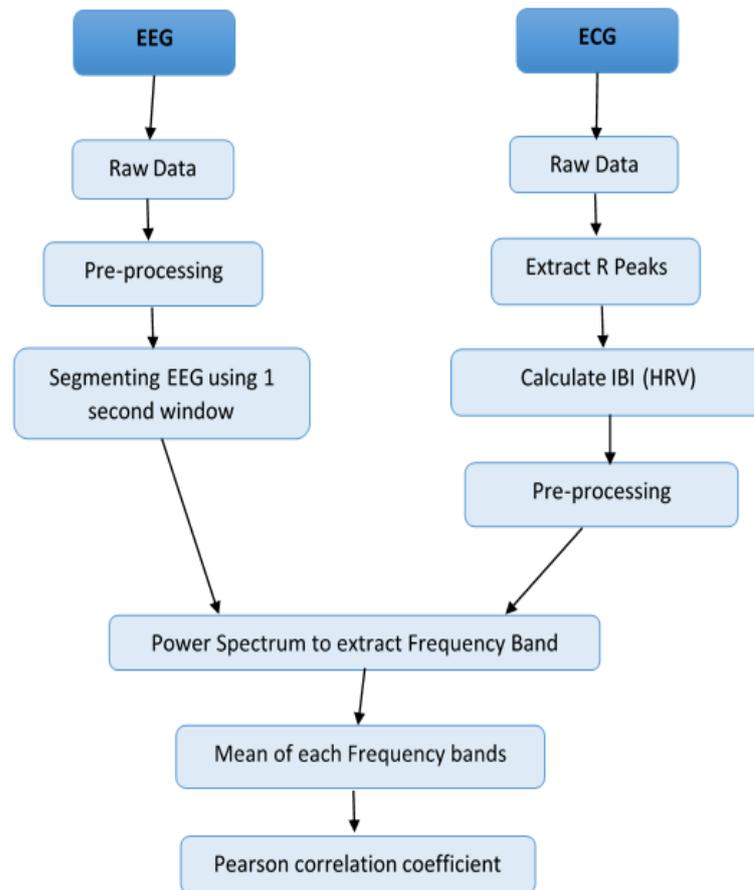


Figure 7.1: Method 1 - experiment steps of the correlation performance.

## 7.3 Experimental Results and Discussion

### 7.3.1 TD Analysis

For Dataset 1, I have performed PCC to investigate the correlation between each of the EEG electrodes with HRV, and plotted averaged result of correlation performance over all participants for each 5-minute slots (information about slot details is described in Chapter 4). The Figures 7.4 and 7.5 show the correlation performance of Dataset 1, where each slot is representing 5 minutes of EEG and HRV data. The x-axis (horizontal) denotes the number of EEG electrodes, and the y-axis (vertical) denotes the PCC values.

As shown in Figures 7.4 and 7.5, the correlation between EEG and HRV is not present in the TD for Dataset 1 and Dataset 2, respectively. The results indicate that PCC values are between -0.05 to 0.05 for both datasets, suggesting no correlation. However, in slot 6 and Slot 7 PCC values were obtained between -0.17 to 0.17 for Dataset 2. This increase in PCC values in two slots might be the effect of the acupuncture stimulation performed on

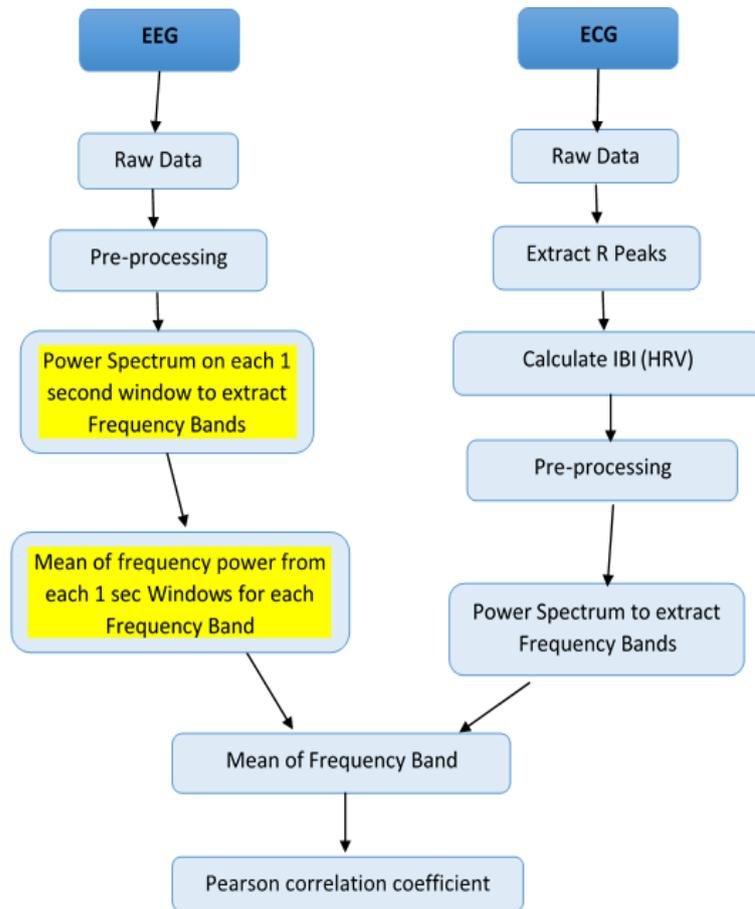


Figure 7.2: Method 2- experiment steps of the correlation performance.

participants (please see Chapter 4 for more detail on Dataset 2).

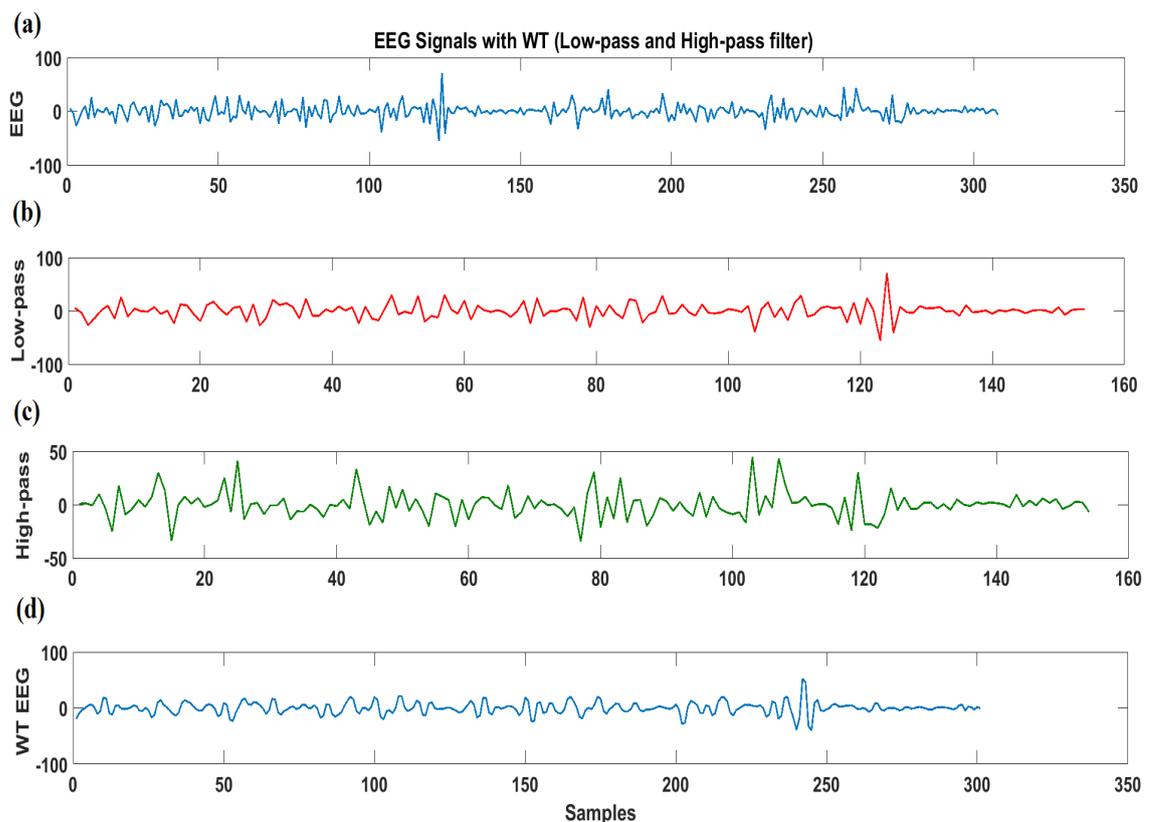


Figure 7.3: WT EEG signals showing: a) unfiltered EEG signal in blue colour, b) Low-pass filtered EEG shown in red colour (Approximation part, showing trend of the EEG signal), c) High-pass filtered EEG shown in green colour (Detailed part, showing fluctuation of the EEG signal), and d) WT EEG signal with low-passed filter.

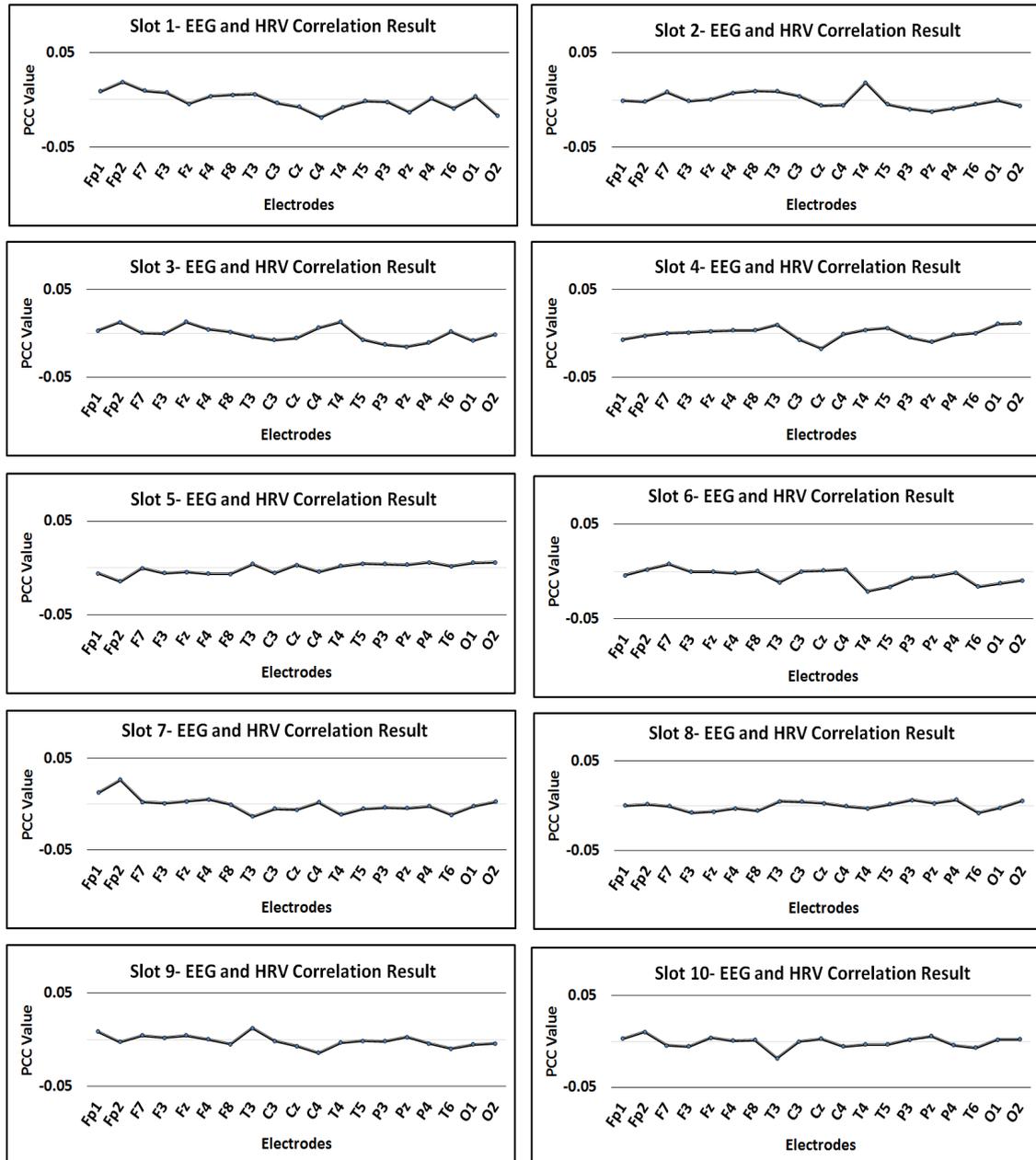


Figure 7.4: Dataset 1 - Time domain (TD) analysis of correlation performance between each of the EEG electrodes with HRV for each 5-minute slot.

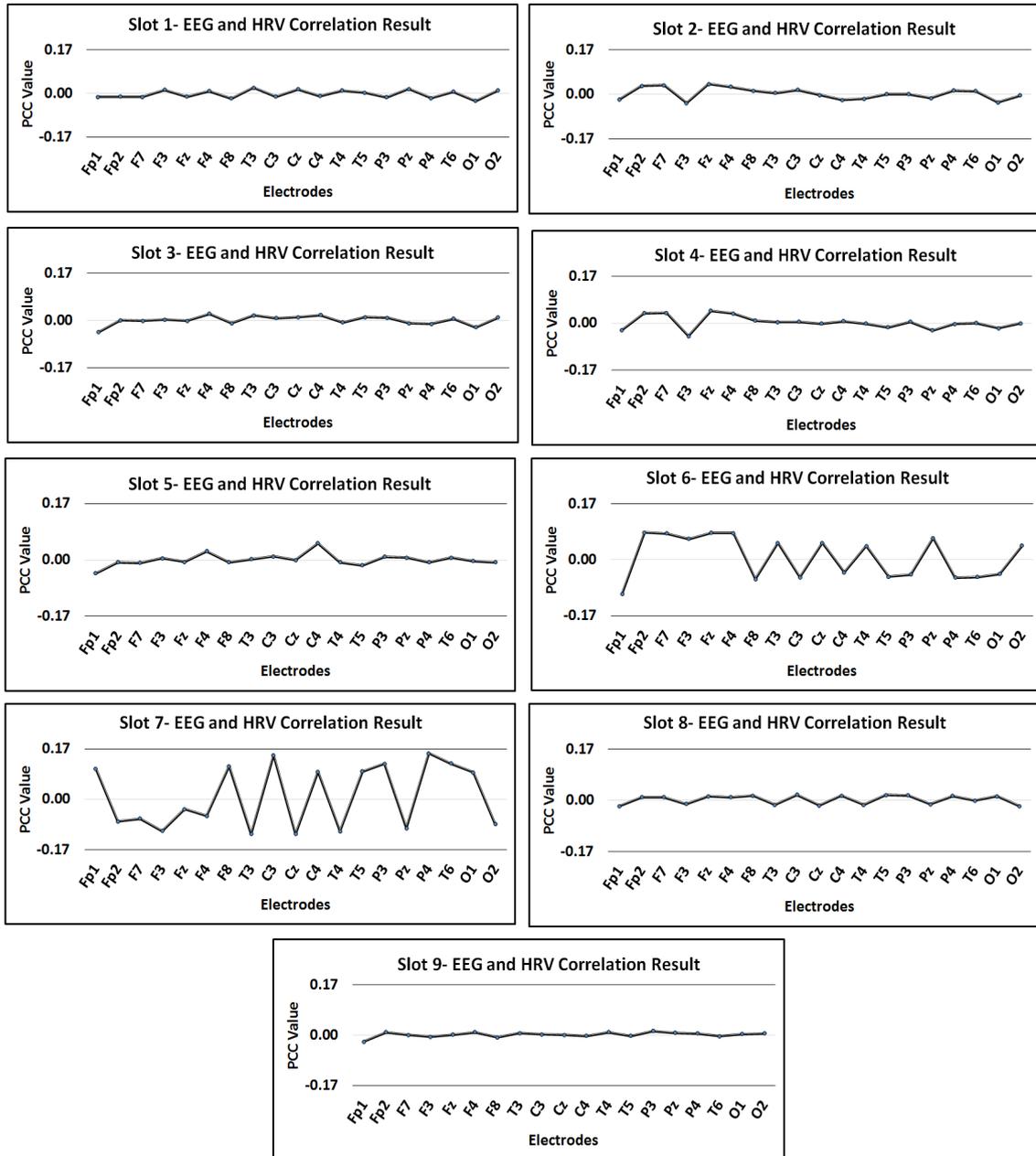


Figure 7.5: Dataset 2 - Time domain (TD) analysis of correlation performance between each of the EEG electrodes with HRV for each 5-minute slot.

### 7.3.2 FD Analysis

For each dataset, I used both Method 1 and Method 2 to investigate the correlation between each of the EEG frequencies (Delta, Theta, Alpha and Beta) with each frequency of the HRV (LF and HF) in three different experiments:

1. The correlation between pre-processed signals.
2. The correlation between pre-processed and WT signals of the EEG and HRV.
3. The correlation between pre-processed HRV with pre-processed and WT signals of EEG.

The Gamma frequency of EEG did not give any correlation effect. Therefore, it is not included in the results shown in Figures 7.6, 7.7 and Tables 7.2 - 7.7.

For both datasets, Experiment 2) correlation between both WT signals did not give better results than Experiments 1) and 3). This might be because HRV is usually less noisy. Moreover, when the WT has been performed on HRV, the signal became flatter and I have lost information. From all 3 experiments, I have found the most interesting results from Experiments 1) and 3).

For each dataset, there are many things to consider: different number of electrodes, different number of participants, different number of frequency combination (5 EEG Frequencies against two HRV frequencies) for the correlation analysis. Therefore, for each participant, correlation result is first obtained for each electrode for all frequency combination. Then for each frequency combination, I have calculated the average of the correlation value over all participants for each EEG electrode. Then each electrode has been ranked in the term of their performance, where, I have given the ranking based on the electrode correlation result. The average of electrode ranking for each frequency combination is then gathered and the five electrodes with the top five performances have been looked into closely. I have found some common electrodes in all the frequency combinations investigated. Figures 7.6 and 7.7 show the results for both Dataset 1 and 2, employing Method 1 and Method 2, respectively. Please refer Appendix F to see electrodes ranking before averaging.

As shown in Figures 7.6 and 7.7, for Dataset 1 (a) and (b), some electrodes (P3, Pz and O1) from the backside of the brain are giving better results than Dataset 2 (a) and (b), respectively. This is due to more randomness in the EEG signals from Dataset 1. Also, the location where TEAS have been performed might contribute to this result.

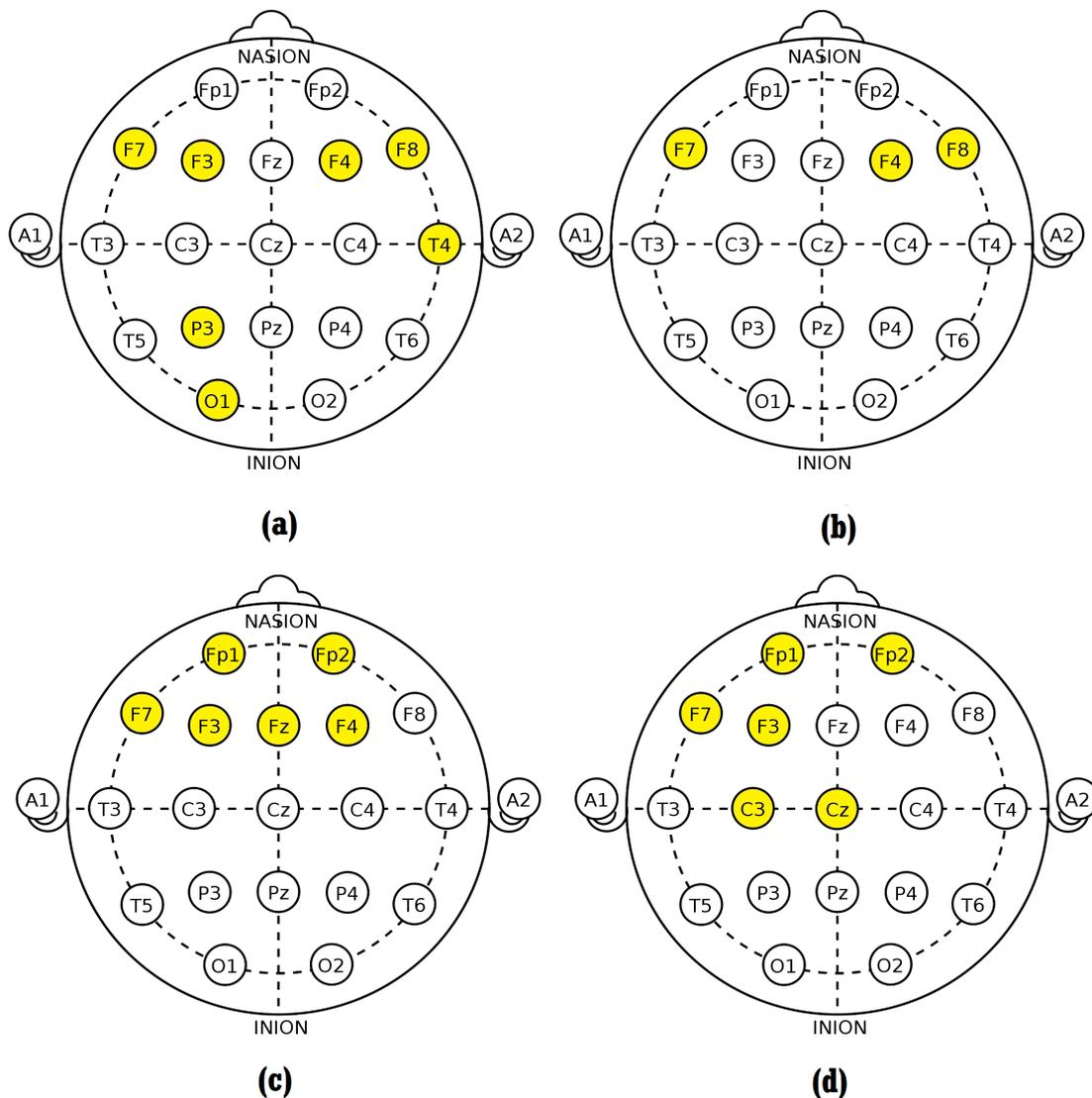


Figure 7.6: Method 1 - Best Electrodes Correlation Performance, highlighted in yellow colour: **(a)** Experiment 1 - Dataset 1 Correlation performance on pre-processed HRV and EEG, **(b)** Experiment 3 - Dataset 1 Correlation performance on pre-processed HRV and WT signals of EEG, **(c)** Experiment 1 - Dataset 2 Correlation performance on pre-processed HRV and EEG, **(d)** Experiment 3 - Dataset 2 Correlation performance on pre-processed HRV and WT signals of EEG.

Based on the results shown in Figures 7.6 and 7.7, we can see it that the frontal lobe of the brain is correlated with the heart. The frontal lobe involved in higher mental functions, such as concentration, creativity, speaking, muscle movement and in making plans and judgements, is a part of cerebral cortex (body's ultimate control and information processing) of the brain (McCraty et al., 2009). The usual Heart-Brain communication path is through the spinal cord. In order to have a relationship between frontal lobe of the brain

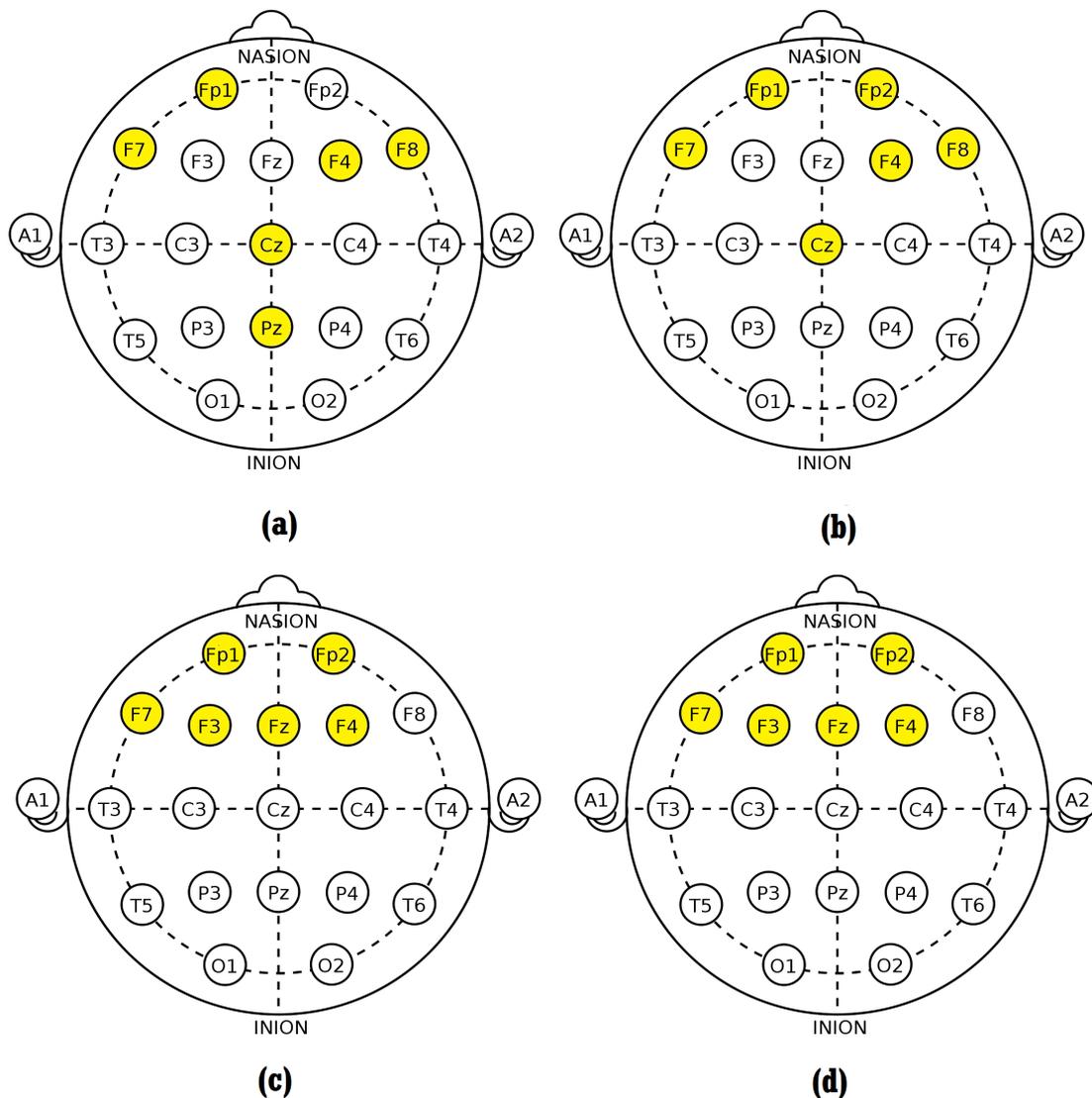


Figure 7.7: Method 2 - Best Electrodes Correlation Performance, highlighted in yellow colour: **(a)** Experiment 1 - Dataset 1 Correlation performance on pre-processed HRV and EEG, **(b)** Experiment 3 - Dataset 1 Correlation performance on pre-processed HRV and WT signals of EEG, **(c)** Experiment 1 - Dataset 2 Correlation performance on pre-processed HRV and EEG, **(d)** Experiment 3 - Dataset 2 Correlation performance on pre-processed HRV and WT signals of EEG.

and heart, I assume the communication might have done through "Medulla"(cardiovascular centre placed in medulla controls the heart beating) which is part of the brain stem. The signal has been then directed to the thalamus and then to the cerebral cortex (Lane et al., 2001), (ATKINSON and BRADLEY, 2004).

Tables 7.2 and 7.4 show the averaged correlation result of participants for each frequency comparison from Dataset 1 employing Method 1 and Method 2, respectively. Ta-

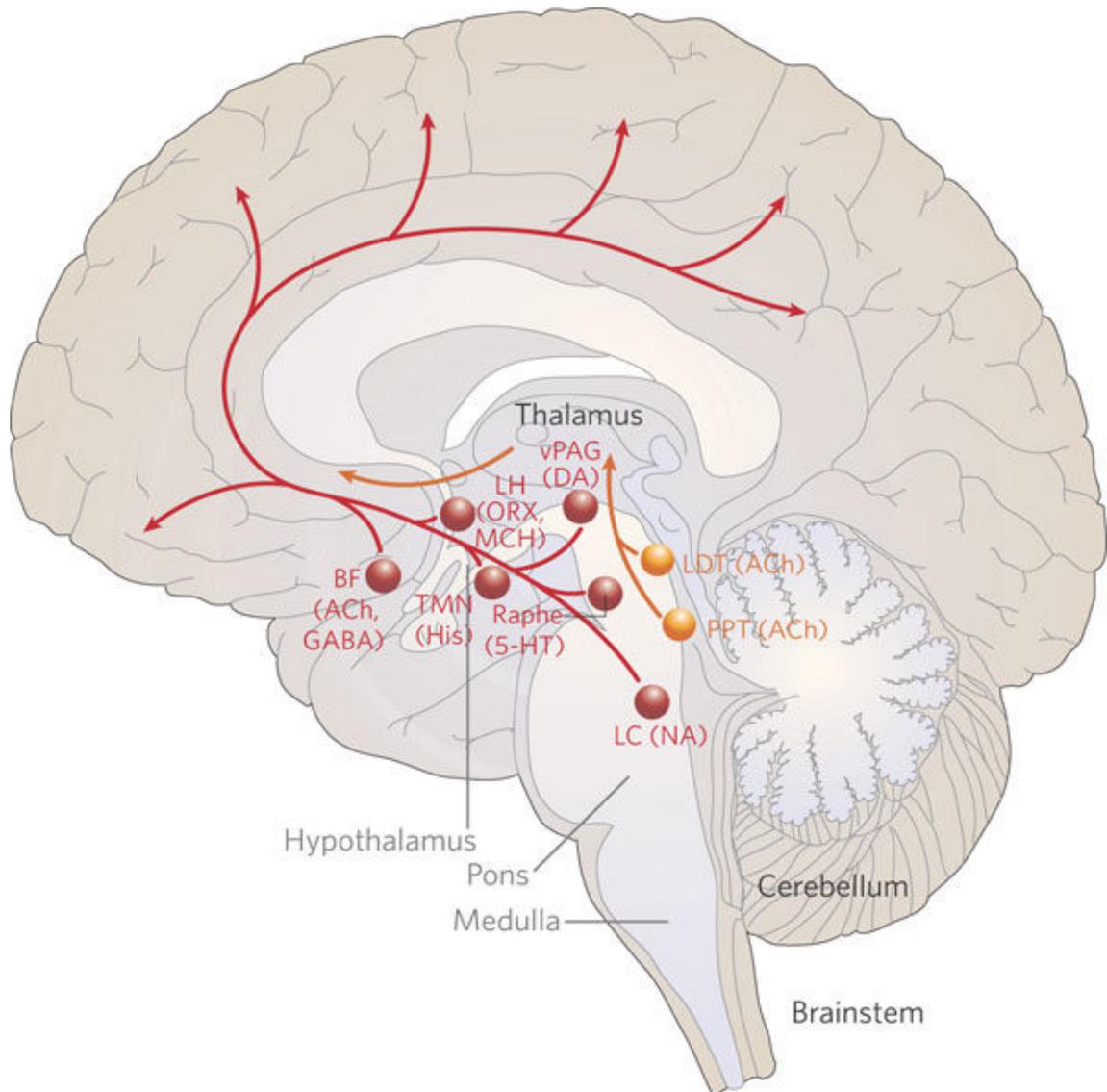


Figure 7.8: Key neuronal projections that maintain alertness, and possibly the path from the cardiovascular centre to the frontal lobe of the brain's communication. The figure is obtained from (Saper et al., 2005).

bles 7.3 and 7.5 show the averaged correlation result of participants for each frequency comparison from Dataset 1 employing Method 1 and Method 2, respectively. Where Level 0 indicates without WT, and Levels 1, 2 and 3 indicate with WT, shown as (a), (b), (c), and (d), respectively. The correlation results showed more negative values in Tables 7.4 - 7.5 (Method 2) than Tables 7.2 - 7.3 (Method 1). This is because they both are analysed using different approaches: Method 1 (as shown in Figure 7.1) used to average 1 second EEG signals, whereas, Method 2 (as shown in Figure 7.2) utilised averaging power spectrum results. Furthermore, the correlation value shown in Tables 7.2 - 7.5 is very low, where it

is possible that the biggest value shown in the tables could be the effect of  $\pm$  signs of the averaged peak values of frequency. Therefore, I considered absolute of the averaged peak values of frequency obtained and repeated the correlation performance between EEG and HRV frequency values. The Heat map with absolute correlation values are shown in Table 7.6 (employing Method 1) and Table 7.7 (Employing Method 2). Please note that I have shown only the averaged result for electrodes in Table 7.6 and Table 7.7. Where, Level 0 indicates without WT, and Level 1, 2 and 3 indicates with WT. The heat map of these results ("Red" is strongest and "Dark-Blue" means weakest) indicates the change in correlation performance with levels of WT. Please refer Appendix F to check the result before averaging across electrodes for Tables 7.6-7.7.

Tables 7.6 and 7.7 show the average absolute correlation result of participants for each frequency comparison from Dataset 1 (Table 7.4 (a) and Table 7.5 (a)) and Dataset 2 (Table 7.4 (b) and Table 7.5 (b)). Where, Level 0 means the correlation between pre-processed data, and Level 1 to 3 means, the correlation between pre-processed HRV with pre-processed and WT EEG. The heat map of these results ("Red" is strongest and "Dark-Blue" means weakest) indicates the correlation performance changes with the levels of WT. The difference in the correlation values shown in 7.6 and 7.7 is for Method 1 and Method 2, respectively, indicates different phases of analysis as discussed earlier. We found the signal became flat after level 2 and lost information when levels have been increased. Therefore, I have not considered the result of levels 3 in Figure 7.6 (b) and (d), and Figure 7.7 (b) and (d).

Results shown in Tables 7.2-7.7 are showing negative correlation at most of the EEG electrodes. However, positive correlation is found at some of the EEG correlation, as shown in Figure 7.6 and 7.7. However, few frequencies of EEG have shown some correlation, which are Delta, Theta, Alpha, and Beta. They have shown a correlation at both LF and HF of HRV. Each of these frequencies represents the activities of these signals. For example, Delta will be higher if the person is in deep sleep, Theta will appear if the person is daydreaming, Alpha will appear if the person is calm, relaxed or in creative visualisation, and Beta will show if the person is working or feeling more alert. For HRV, LF and HF represent the sympathetic and parasympathetic activities of the autonomic nervous system (ANS), respectively.

## 7.4 Conclusion

In this chapter, no correlation found between EEG and HRV when analysed in TD. In recent years, TD was not the focus of the correlation performance between EEG and HRV (See

EEG	HRV	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14	E15	E16	E17	E18	E19	Average
DELTA	LF	0.19	0.24	0.29	0.16	0.06	0.07	0.23	0.16	0.29	0.15	0.12	0.29	0.15	0.19	0.08	0.06	0.21	0.19	0.09	<b>0.17</b>
	HF	0.14	0.13	0.27	0.08	-0.07	-0.10	0.10	0.01	0.18	0.12	0.14	0.22	-0.06	0.01	0.14	0.05	-0.01	-0.04	0.02	<b>0.07</b>
THETA	LF	0.13	0.15	0.23	0.12	0.02	-0.08	0.12	0.00	0.05	-0.03	0.18	0.19	-0.08	-0.06	-0.10	0.02	0.11	-0.03	-0.10	<b>0.04</b>
	HF	0.16	0.11	0.20	0.08	-0.03	0.04	0.09	0.08	0.24	-0.01	0.12	0.19	0.01	0.04	0.07	0.01	0.02	-0.10	-0.04	<b>0.07</b>
ALPHA	LF	0.15	0.08	0.17	0.03	0.15	0.16	0.23	0.08	0.09	-0.05	0.16	0.15	0.05	0.17	0.00	0.11	0.10	0.03	0.08	<b>0.10</b>
	HF	0.12	0.09	0.16	-0.12	-0.04	-0.02	0.09	-0.05	-0.03	0.01	0.06	0.05	-0.04	-0.01	0.03	-0.05	-0.09	-0.04	-0.06	<b>0.00</b>
BETA	LF	0.06	-0.08	0.12	0.03	-0.02	-0.11	0.07	0.17	0.04	-0.17	-0.08	-0.05	-0.14	0.16	0.14	-0.05	-0.10	-0.11	-0.14	<b>-0.01</b>
	HF	0.03	0.00	-0.04	0.12	0.00	0.02	-0.08	0.01	-0.06	-0.20	0.12	-0.07	0.00	0.00	0.00	0.01	-0.04	-0.16	-0.02	<b>-0.02</b>

(a)

EEG	HRV	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14	E15	E16	E17	E18	E19	Average
DELTA	LF	0.18	0.31	0.25	0.20	0.04	0.04	0.25	0.11	0.31	0.17	0.10	0.30	0.13	0.18	0.08	0.06	0.23	0.19	0.10	<b>0.17</b>
	HF	0.13	0.15	0.26	0.11	-0.16	-0.08	0.13	-0.05	0.18	0.16	0.19	0.18	-0.01	0.04	0.15	0.09	-0.03	0.01	0.07	<b>0.08</b>
THETA	LF	0.12	0.23	0.23	0.17	-0.04	-0.07	0.09	0.00	0.02	0.02	0.14	0.12	-0.11	-0.08	-0.06	0.03	0.11	-0.06	-0.10	<b>0.04</b>
	HF	0.19	0.16	0.17	0.12	-0.05	0.08	0.09	0.02	0.20	0.05	0.08	0.12	0.07	0.08	0.09	0.02	0.01	-0.07	0.04	<b>0.08</b>
ALPHA	LF	0.09	0.11	0.19	0.11	0.09	0.18	0.23	0.04	0.07	-0.02	0.11	0.14	0.02	0.14	-0.01	0.10	0.10	0.03	0.05	<b>0.09</b>
	HF	0.14	0.10	0.15	-0.09	-0.04	0.02	0.07	-0.13	-0.03	-0.03	0.00	0.00	-0.02	0.04	0.03	-0.02	-0.08	0.03	0.01	<b>0.01</b>
BETA	LF	-0.03	-0.14	0.09	0.06	0.08	0.02	0.08	0.14	0.03	-0.06	-0.08	-0.11	-0.16	0.05	0.07	-0.05	-0.05	-0.06	-0.05	<b>-0.01</b>
	HF	0.04	0.08	0.00	0.15	0.12	0.13	-0.10	-0.01	-0.04	-0.12	0.11	-0.07	0.14	0.12	-0.02	0.02	0.10	0.15	0.14	<b>0.05</b>

(b)

EEG	HRV	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14	E15	E16	E17	E18	E19	Average
DELTA	LF	0.23	0.20	0.24	0.20	0.09	0.10	0.21	0.14	0.29	0.09	0.13	0.27	0.09	0.19	0.09	0.13	0.27	0.15	0.09	<b>0.17</b>
	HF	0.18	0.13	0.26	0.10	-0.10	-0.12	0.08	0.00	0.21	0.17	0.18	0.26	-0.01	0.01	0.12	0.08	0.04	-0.03	0.04	<b>0.08</b>
THETA	LF	0.13	0.09	0.20	0.19	-0.01	-0.04	0.00	-0.01	-0.01	-0.05	0.17	0.05	-0.12	-0.15	-0.06	0.07	0.04	-0.17	-0.16	<b>0.01</b>
	HF	0.15	0.08	0.12	0.13	-0.04	0.07	0.06	0.06	0.19	0.02	0.11	0.12	0.12	-0.01	0.06	0.02	0.10	-0.08	0.00	<b>0.07</b>
ALPHA	LF	0.15	-0.03	0.14	0.05	0.11	-0.01	-0.01	0.04	0.04	-0.13	0.07	-0.02	-0.05	-0.02	0.03	0.16	-0.04	-0.11	-0.07	<b>0.02</b>
	HF	0.09	0.12	0.09	0.01	-0.02	0.08	-0.02	0.00	-0.06	-0.13	-0.05	-0.06	0.01	-0.06	-0.04	0.02	-0.02	-0.04	-0.05	<b>-0.01</b>
BETA	LF	0.01	-0.10	0.17	0.02	0.10	0.04	-0.01	0.03	0.04	-0.01	-0.11	0.04	-0.01	0.02	0.03	0.09	-0.06	-0.05	0.02	<b>0.01</b>
	HF	0.00	0.08	0.08	-0.07	-0.03	0.08	0.04	-0.04	-0.04	-0.05	-0.13	-0.05	-0.03	-0.01	-0.13	0.03	-0.01	0.00	-0.05	<b>-0.02</b>

(c)

EEG	HRV	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14	E15	E16	E17	E18	E19	Average
DELTA	LF	0.27	0.18	0.22	0.22	0.10	0.20	0.20	0.15	0.27	0.07	0.13	0.27	0.16	0.19	0.16	0.16	0.29	0.13	0.12	<b>0.18</b>
	HF	0.22	0.06	0.25	0.14	-0.11	-0.08	0.05	-0.03	0.22	0.08	0.23	0.23	0.01	0.02	0.16	0.13	0.00	-0.03	0.03	<b>0.08</b>
THETA	LF	0.27	0.09	0.18	0.13	0.15	0.06	0.04	0.03	0.01	0.00	0.35	0.01	0.08	-0.01	-0.01	0.09	0.02	-0.05	-0.03	<b>0.07</b>
	HF	0.08	-0.02	0.08	-0.07	0.01	0.00	-0.02	0.04	0.18	-0.01	0.14	0.04	0.16	-0.03	0.00	-0.03	0.01	-0.02	0.05	<b>0.03</b>
ALPHA	LF	0.30	0.10	0.24	0.14	0.07	0.01	0.05	0.04	0.05	0.03	0.28	0.01	0.05	0.02	-0.06	0.09	0.00	-0.01	0.03	<b>0.08</b>
	HF	0.01	-0.07	0.10	-0.07	-0.07	-0.04	0.03	0.02	0.12	-0.05	0.08	0.03	0.15	-0.08	0.00	-0.05	-0.01	-0.05	0.06	<b>0.01</b>
BETA	LF	0.27	0.08	0.20	0.12	0.13	0.05	0.05	0.04	0.00	0.00	0.33	0.02	0.08	-0.02	0.00	0.10	0.01	-0.03	-0.02	<b>0.07</b>
	HF	0.08	-0.04	0.09	-0.07	0.00	0.00	-0.01	0.04	0.16	0.01	0.12	0.02	0.12	-0.05	0.01	0.00	0.00	-0.04	0.06	<b>0.03</b>

(d)

Table 7.2: Dataset 1- Method 1 - Averaged participants' correlation performance: (a) Level 0 (No WT), (b) Level 1 WT, (c) Level 2 WT, (d) Level 3 WT)

Chapter 2 for more details on the TD correlation analysis). I assume this could be because no useful information can be found when comparing the amplitude of these signals.

The main conclusion of this work is that electrical activity in the frontal lobe of the brain is correlated with the HRV for the given two datasets in the FD analysis. To the best of my knowledge this is a new result. This suggests that most probably the electrical signals could be transmitted through the cerebral cortex, Thalamus, and Medulla of the brain (Saper et al., 2005). Figure 7.8 shows the possible path of the key neuronal projections that maintain alertness. I believe this neuronal projection shows the possible path of why frontal lobe is showing correlation with heart, as shown in this chapter. However, I have

EEG	HRV	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14	E15	E16	E17	E18	E19	Average
DELTA	LF	-0.06	0.06	0.19	-0.17	-0.05	0.00	0.09	0.09	-0.13	-0.20	-0.07	0.03	-0.15	0.20	0.08	0.01	0.02	-0.08	0.15	<b>0.00</b>
	HF	0.03	0.02	0.27	0.20	-0.10	0.06	0.12	0.26	-0.09	-0.04	-0.12	0.04	0.00	0.14	0.09	0.04	0.22	-0.02	-0.11	<b>0.05</b>
THETA	LF	-0.04	-0.07	0.03	-0.04	-0.11	-0.06	0.00	0.06	-0.08	-0.20	-0.09	0.09	-0.20	0.15	-0.06	-0.01	0.01	-0.04	-0.08	<b>-0.04</b>
	HF	-0.04	0.00	-0.03	-0.01	0.04	-0.08	0.11	0.16	0.01	-0.19	-0.03	0.04	-0.03	0.08	-0.08	-0.14	0.05	0.08	-0.02	<b>0.00</b>
ALPHA	LF	-0.16	-0.29	0.06	0.07	0.00	-0.04	-0.04	-0.05	-0.21	-0.10	-0.07	0.14	-0.34	0.09	-0.01	0.09	-0.08	-0.08	0.02	<b>-0.05</b>
	HF	-0.01	-0.06	0.02	0.11	0.13	0.09	0.14	0.07	0.04	-0.05	0.06	0.03	-0.07	0.03	0.02	0.06	0.04	0.07	0.03	<b>0.04</b>
BETA	LF	-0.14	-0.28	0.09	0.06	-0.10	0.08	-0.08	-0.07	-0.21	-0.14	-0.04	0.15	-0.33	0.05	-0.01	-0.02	-0.17	0.03	-0.10	<b>-0.06</b>
	HF	0.05	-0.05	0.09	0.17	0.02	0.12	0.13	0.08	0.02	-0.10	0.14	0.05	-0.10	0.00	0.00	-0.01	-0.05	0.14	0.01	<b>0.04</b>

(a)

EEG	HRV	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14	E15	E16	E17	E18	E19	Average
DELTA	LF	-0.07	0.06	0.10	-0.16	-0.07	0.06	0.08	0.13	-0.09	-0.16	0.00	-0.01	-0.11	0.14	0.04	0.05	0.00	-0.04	0.06	<b>0.00</b>
	HF	-0.05	0.00	0.23	0.20	-0.15	0.21	0.13	0.36	-0.10	-0.06	0.00	-0.03	0.01	0.10	0.07	0.02	0.23	0.04	-0.25	<b>0.05</b>
THETA	LF	-0.09	-0.10	0.05	-0.03	-0.23	-0.06	0.02	0.04	-0.07	-0.19	-0.08	0.02	-0.19	0.16	-0.04	-0.02	-0.02	-0.08	-0.07	<b>-0.05</b>
	HF	-0.07	-0.01	-0.02	-0.01	-0.06	0.06	0.10	0.15	0.00	-0.18	-0.03	-0.03	-0.03	0.08	-0.07	-0.12	0.00	0.09	-0.01	<b>-0.01</b>
ALPHA	LF	-0.15	-0.31	0.06	0.07	-0.06	-0.07	-0.04	-0.06	-0.21	-0.09	-0.13	0.16	-0.32	0.09	-0.02	0.05	-0.09	-0.08	0.00	<b>-0.06</b>
	HF	-0.01	-0.08	0.03	0.11	0.07	0.08	0.13	0.10	0.04	-0.03	0.03	0.03	-0.07	0.02	0.00	0.01	0.02	0.06	0.03	<b>0.03</b>
BETA	LF	-0.09	-0.30	0.09	0.07	-0.21	-0.01	-0.08	-0.13	-0.23	-0.16	-0.08	0.22	-0.33	0.05	-0.02	0.01	-0.15	-0.02	-0.04	<b>-0.07</b>
	HF	0.10	-0.05	0.09	0.19	-0.07	0.10	0.11	0.06	0.00	-0.12	0.12	0.10	-0.09	0.00	-0.01	0.00	-0.01	0.13	0.05	<b>0.04</b>

(b)

EEG	HRV	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14	E15	E16	E17	E18	E19	Average
DELTA	LF	-0.16	0.06	0.05	-0.11	-0.04	0.11	0.11	0.14	-0.11	-0.08	-0.07	-0.02	-0.13	0.13	0.04	0.03	0.04	-0.02	0.08	<b>0.00</b>
	HF	-0.13	-0.07	0.02	0.12	-0.20	0.24	0.06	0.36	-0.13	0.04	0.04	-0.04	0.02	0.04	0.05	-0.03	0.15	0.01	-0.24	<b>0.02</b>
THETA	LF	-0.09	-0.11	0.06	-0.04	-0.25	-0.07	0.02	0.03	-0.07	-0.19	-0.08	0.04	-0.19	0.16	-0.04	0.00	-0.01	-0.09	-0.07	<b>-0.05</b>
	HF	-0.07	0.01	0.00	-0.01	-0.06	0.06	0.11	0.15	0.01	-0.19	-0.03	-0.02	-0.03	0.08	-0.07	-0.10	0.01	0.09	-0.01	<b>0.00</b>
ALPHA	LF	-0.11	-0.33	0.06	0.04	-0.09	-0.01	-0.03	-0.08	-0.21	-0.08	-0.11	0.15	-0.33	0.08	-0.05	0.09	-0.03	-0.02	0.08	<b>-0.05</b>
	HF	-0.02	-0.06	0.03	0.11	0.02	0.12	0.18	0.06	0.03	-0.02	0.04	0.03	-0.08	0.01	-0.02	0.09	0.06	0.11	0.10	<b>0.04</b>
BETA	LF	-0.13	-0.31	0.04	0.03	-0.14	-0.08	-0.14	-0.24	-0.07	-0.17	0.16	-0.30	0.08	-0.07	0.06	-0.03	-0.06	0.02	-0.08	<b>-0.08</b>
	HF	0.01	-0.05	-0.02	0.13	0.03	0.05	0.10	-0.02	0.00	-0.09	0.02	0.02	-0.05	0.01	-0.01	0.12	0.14	0.18	0.13	<b>0.04</b>

(c)

EEG	HRV	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14	E15	E16	E17	E18	E19	Average
DELTA	LF	-0.18	0.06	0.06	-0.10	-0.03	0.10	0.13	0.14	-0.11	-0.05	-0.08	-0.02	-0.14	0.12	0.04	0.02	0.08	-0.04	0.08	<b>0.00</b>
	HF	-0.13	-0.08	0.00	0.11	-0.19	0.23	0.06	0.34	-0.13	0.06	0.04	-0.05	0.01	0.01	0.03	-0.03	0.20	-0.01	-0.24	<b>0.01</b>
THETA	LF	-0.05	-0.13	0.10	-0.03	-0.13	-0.03	0.04	0.09	-0.10	-0.27	-0.06	0.08	-0.20	0.10	-0.05	-0.11	-0.01	0.05	-0.09	<b>-0.04</b>
	HF	0.01	0.04	0.07	0.04	0.01	-0.02	0.23	0.20	0.04	-0.34	-0.02	0.01	-0.03	0.04	-0.08	-0.16	-0.03	0.15	-0.06	<b>0.01</b>
ALPHA	LF	-0.04	-0.11	0.13	-0.01	-0.01	0.03	0.02	0.09	-0.10	-0.25	-0.07	0.08	-0.20	0.06	-0.07	-0.05	0.04	0.17	-0.06	<b>-0.02</b>
	HF	-0.01	0.02	0.08	0.04	0.04	0.02	0.21	0.13	0.03	-0.27	-0.06	0.02	-0.03	0.03	-0.10	-0.19	-0.08	0.16	-0.05	<b>0.00</b>
BETA	LF	-0.03	-0.11	0.11	-0.03	0.01	-0.01	0.03	0.01	-0.09	-0.22	-0.13	0.09	-0.20	0.11	-0.03	-0.03	0.04	0.07	-0.01	<b>-0.02</b>
	HF	0.02	0.05	0.10	0.06	0.10	-0.02	0.24	0.09	0.07	-0.21	-0.09	0.01	-0.04	0.07	-0.06	-0.10	-0.01	0.09	0.03	<b>0.02</b>

(d)

Table 7.3: Dataset 2- Method 1 - Averaged participants' correlation performance: (a) Level 0 (No WT), (b) Level 1 WT, (c) Level 2 WT, (d) Level 3 WT)

lack of expertise to find out possible reason for this, but I think this might be of interest to the people working in the medical area.

The second conclusion from this work is that, WT signals also give correlation from the frontal lobe of the brain when analysed in FD. To the best of my knowledge, the correlation between WT signals of EEG and ECG/HRV has not yet been investigated.

A more tentative conclusion of this work is that four frequencies in the EEG Delta, Theta, Alpha and Beta are correlated with both LF and HF of HRV. Whereas, most of previous studies, (Yang et al., 2002),(Ako et al., 2003),(Jurysta et al., 2003),(Abdullah et al., 2010) and (Chua et al., 2012), have shown a negative correlation between these frequency

EEG	HRV	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14	E15	E16	E17	E18	E19	Average
DELTA	LF	0.06	-0.09	-0.06	-0.05	-0.03	-0.04	-0.11	0.02	-0.05	0.06	0.03	0.04	0.05	-0.05	-0.08	-0.03	-0.03	0.07	0.05	-0.01
	HF	-0.08	0.07	0.05	0.03	0.06	0.05	0.01	0.07	0.02	0.08	0.04	0.03	0.12	0.04	-0.03	0.09	0.06	0.12	0.14	0.05
THETA	LF	0.00	-0.06	-0.07	-0.05	-0.07	-0.06	-0.05	-0.04	-0.06	-0.01	-0.10	-0.04	-0.08	0.00	-0.05	-0.12	-0.10	0.00	-0.14	-0.06
	HF	-0.15	-0.07	0.01	0.01	0.00	0.01	0.03	0.07	0.01	-0.06	-0.10	-0.02	-0.17	-0.12	-0.11	-0.12	-0.10	-0.08	-0.15	-0.06
ALPHA	LF	-0.06	0.05	0.00	-0.04	-0.06	-0.02	0.05	-0.05	-0.01	-0.02	0.09	0.06	-0.06	-0.02	-0.08	-0.01	0.04	-0.11	-0.09	-0.02
	HF	0.04	-0.11	0.03	0.03	-0.02	0.01	0.05	0.01	0.05	-0.09	0.04	0.02	-0.13	-0.12	-0.11	-0.10	-0.06	-0.16	-0.15	-0.04
BETA	LF	-0.08	0.06	-0.02	-0.03	-0.07	-0.05	0.04	-0.05	0.01	-0.02	-0.03	-0.03	-0.09	-0.03	-0.03	-0.01	0.05	-0.12	-0.08	-0.03
	HF	0.10	-0.07	0.02	0.04	-0.02	0.01	0.05	0.01	0.06	-0.02	-0.02	-0.02	-0.14	-0.06	0.02	-0.06	-0.03	-0.14	-0.13	-0.02

(a)

EEG	HRV	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14	E15	E16	E17	E18	E19	Average
DELTA	LF	0.05	-0.09	-0.06	-0.05	-0.03	-0.04	-0.11	0.02	-0.05	0.06	0.03	0.03	0.05	-0.06	-0.09	-0.03	-0.03	0.07	0.05	-0.01
	HF	-0.08	0.07	0.05	0.04	0.06	0.05	0.01	0.07	0.02	0.08	0.04	0.03	0.11	0.03	-0.03	0.09	0.06	0.12	0.14	0.05
THETA	LF	0.00	-0.06	-0.07	-0.05	-0.07	-0.06	-0.05	-0.04	-0.06	-0.01	-0.11	-0.05	-0.08	0.00	-0.05	-0.12	-0.09	-0.01	-0.14	-0.06
	HF	-0.16	-0.09	0.01	0.01	0.00	0.01	0.03	0.06	0.01	-0.07	-0.11	-0.03	-0.17	-0.12	-0.11	-0.12	-0.10	-0.09	-0.15	-0.06
ALPHA	LF	-0.06	0.05	0.00	-0.04	-0.07	-0.02	0.05	-0.05	-0.01	-0.03	0.07	0.05	-0.06	-0.02	-0.08	-0.01	0.04	-0.11	-0.09	-0.02
	HF	0.04	-0.11	0.02	0.03	-0.02	0.01	0.05	0.01	0.05	-0.09	0.03	0.01	-0.13	-0.12	-0.11	-0.10	-0.06	-0.16	-0.15	-0.04
BETA	LF	-0.08	0.06	-0.02	-0.03	-0.07	-0.05	0.04	-0.05	0.01	-0.02	-0.03	-0.03	-0.08	-0.04	-0.03	-0.02	0.05	-0.12	-0.08	-0.03
	HF	0.10	-0.07	0.02	0.04	-0.02	0.01	0.05	0.01	0.06	-0.02	-0.02	-0.03	-0.14	-0.06	0.02	-0.05	-0.03	-0.14	-0.12	-0.02

(b)

EEG	HRV	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14	E15	E16	E17	E18	E19	Average
DELTA	LF	0.04	-0.11	-0.06	-0.05	-0.03	-0.04	-0.10	0.01	-0.04	0.06	0.05	0.03	0.05	-0.07	-0.10	-0.04	-0.03	0.04	0.04	-0.02
	HF	-0.08	0.07	0.04	0.04	0.05	0.05	0.01	0.07	0.03	0.09	0.06	0.03	0.10	0.02	-0.03	0.10	0.07	0.10	0.13	0.05
THETA	LF	-0.05	-0.02	-0.07	-0.05	-0.07	-0.07	-0.03	-0.04	-0.05	-0.02	-0.13	-0.07	-0.09	0.00	-0.05	-0.11	-0.08	-0.05	-0.13	-0.06
	HF	-0.15	-0.12	0.00	0.01	-0.01	0.00	0.03	0.06	0.02	-0.07	-0.12	-0.04	-0.17	-0.12	-0.11	-0.12	-0.10	-0.12	-0.15	-0.07
ALPHA	LF	-0.07	0.06	0.00	-0.04	-0.07	-0.02	0.07	-0.06	0.00	-0.02	0.03	0.02	-0.06	-0.02	-0.09	-0.01	0.04	-0.12	-0.10	-0.02
	HF	0.05	-0.10	0.02	0.03	-0.02	0.00	0.05	0.00	0.06	-0.08	0.00	-0.01	-0.13	-0.11	-0.10	-0.10	-0.06	-0.15	-0.15	-0.04
BETA	LF	-0.06	0.06	0.00	-0.04	-0.07	-0.02	0.06	-0.04	0.00	-0.01	0.00	0.01	-0.05	-0.02	-0.06	0.01	0.04	-0.12	-0.09	-0.02
	HF	0.06	-0.09	0.01	0.03	-0.02	0.01	0.05	0.02	0.05	-0.06	-0.02	-0.02	-0.12	-0.12	-0.07	-0.09	-0.06	-0.15	-0.14	-0.04

(c)

EEG	HRV	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14	E15	E16	E17	E18	E19	Average
DELTA	LF	-0.02	-0.15	-0.05	-0.05	-0.04	-0.04	-0.06	-0.02	-0.03	0.04	0.11	0.05	0.01	-0.11	-0.12	-0.10	-0.02	-0.07	-0.01	-0.03
	HF	-0.11	-0.05	0.04	0.03	0.04	0.04	0.03	0.05	0.04	0.07	0.09	0.04	0.07	-0.05	-0.07	0.02	0.05	-0.01	0.08	0.02
THETA	LF	-0.10	0.01	-0.05	-0.05	-0.08	-0.07	0.02	-0.05	-0.05	-0.02	-0.10	-0.03	-0.09	-0.03	-0.05	-0.09	-0.04	-0.09	-0.13	-0.06
	HF	-0.08	-0.13	0.00	0.00	-0.02	-0.01	0.04	0.02	0.03	-0.07	-0.09	-0.02	-0.17	-0.12	-0.07	-0.12	-0.10	-0.14	-0.15	-0.06
ALPHA	LF	-0.09	-0.02	-0.06	-0.05	-0.08	-0.06	0.01	-0.06	-0.05	-0.04	-0.11	-0.06	-0.07	-0.03	-0.07	-0.10	-0.07	-0.07	-0.12	-0.06
	HF	-0.14	-0.14	0.00	-0.01	-0.03	-0.02	0.03	0.01	0.02	-0.07	-0.10	-0.02	-0.17	-0.13	-0.08	-0.13	-0.11	-0.14	-0.15	-0.07
BETA	LF	-0.11	0.01	-0.05	-0.05	-0.08	-0.07	0.02	-0.05	-0.05	-0.04	-0.10	-0.03	-0.10	-0.03	-0.06	-0.09	-0.04	-0.09	-0.13	-0.06
	HF	-0.09	-0.13	0.00	0.00	-0.02	-0.01	0.04	0.02	0.03	-0.07	-0.08	-0.02	-0.17	-0.12	-0.07	-0.12	-0.11	-0.14	-0.15	-0.06

(d)

Table 7.4: Dataset 1- Method 2 - Averaged participants' correlation performance: (a) Level 0 (No WT), (b) Level 1 WT, (c) Level 2 WT, (d) Level 3 WT)

bands due to the condition in which these signals have been analysed.

In summary, the number of EEG electrodes used by other people to investigate correlation was limited. My results cover a gap in the research concerning the correlation between the EEG and the HRV using 10%-20% electrode placement system. My work suggests a correlation between the frontal lobe of the EEG and the HRV, with and without WT signals. This could be because the frontal lobe is related to higher mental functions of the cerebral cortex and responsible for muscle movements of the body (Stuss and Benson, 1986).

EEG	HRV	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14	E15	E16	E17	E18	E19	Average
DELTA	LF	0.10	0.09	0.11	0.08	0.10	0.08	0.07	0.13	0.08	0.10	0.07	0.15	0.07	0.08	0.10	0.09	0.06	0.10	0.09	<b>0.09</b>
	HF	0.01	0.04	-0.01	-0.01	-0.03	0.01	0.09	-0.11	-0.02	0.00	-0.02	0.04	-0.02	-0.02	-0.02	-0.02	-0.03	0.02	0.01	<b>0.00</b>
THETA	LF	0.04	0.08	0.04	0.03	0.04	0.08	0.12	-0.01	0.03	0.01	0.02	-0.03	0.02	0.01	0.00	0.02	0.07	-0.01	0.03	<b>0.03</b>
	HF	-0.07	-0.09	-0.03	-0.06	-0.06	-0.10	-0.14	0.05	-0.05	-0.04	-0.06	-0.11	-0.08	-0.07	-0.04	-0.04	-0.06	-0.10	-0.12	<b>-0.07</b>
ALPHA	LF	-0.08	-0.06	-0.07	-0.07	-0.19	-0.05	-0.03	-0.12	-0.07	-0.06	-0.08	-0.14	-0.07	-0.06	-0.06	-0.06	-0.05	-0.08	-0.06	<b>-0.08</b>
	HF	0.04	0.03	0.05	0.04	-0.01	0.03	-0.01	0.10	0.04	0.04	0.02	-0.01	0.03	0.03	0.03	0.03	0.03	0.02	0.02	<b>0.03</b>
BETA	LF	-0.04	-0.05	-0.06	-0.06	-0.08	-0.06	-0.05	-0.08	-0.06	-0.06	-0.06	-0.08	-0.06	-0.06	-0.06	-0.06	-0.06	-0.06	-0.06	<b>-0.06</b>
	HF	0.02	-0.01	0.04	0.04	0.06	0.04	0.03	0.05	0.04	0.04	0.03	0.02	0.04	0.04	0.04	0.03	0.03	0.04	0.04	<b>0.04</b>

(a)

EEG	HRV	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14	E15	E16	E17	E18	E19	Average
DELTA	LF	0.10	0.09	0.11	0.08	0.10	0.08	0.07	0.13	0.08	0.10	0.07	0.15	0.07	0.08	0.10	0.09	0.06	0.10	0.09	<b>0.09</b>
	HF	0.01	0.04	-0.01	-0.01	-0.03	0.02	0.09	-0.11	-0.02	0.00	-0.02	0.04	-0.02	-0.01	-0.02	-0.02	-0.03	0.02	0.01	<b>0.00</b>
THETA	LF	0.05	0.08	0.04	0.03	0.04	0.08	0.12	-0.01	0.03	0.01	0.02	-0.03	0.02	0.01	0.00	0.02	0.07	-0.01	0.03	<b>0.03</b>
	HF	-0.07	-0.10	-0.03	-0.06	-0.06	-0.10	-0.14	0.05	-0.05	-0.04	-0.06	-0.11	-0.08	-0.07	-0.04	-0.04	-0.06	-0.10	-0.12	<b>-0.07</b>
ALPHA	LF	-0.07	-0.06	-0.07	-0.07	-0.18	-0.05	-0.03	-0.11	-0.07	-0.06	-0.07	-0.13	-0.07	-0.06	-0.06	-0.06	-0.05	-0.08	-0.06	<b>-0.08</b>
	HF	0.04	0.03	0.05	0.04	0.00	0.03	0.00	0.10	0.04	0.04	0.02	-0.01	0.03	0.03	0.03	0.03	0.03	0.02	0.02	<b>0.03</b>
BETA	LF	-0.04	-0.05	-0.06	-0.06	-0.07	-0.06	-0.05	-0.08	-0.06	-0.06	-0.06	-0.09	-0.06	-0.06	-0.06	-0.06	-0.06	-0.06	-0.06	<b>-0.06</b>
	HF	0.03	0.00	0.04	0.04	0.06	0.04	0.03	0.05	0.04	0.04	0.03	0.02	0.04	0.04	0.04	0.03	0.04	0.04	0.04	<b>0.04</b>

(b)

EEG	HRV	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14	E15	E16	E17	E18	E19	Average
DELTA	LF	0.13	0.10	0.11	0.09	0.18	0.09	0.07	0.13	0.08	0.11	0.07	0.15	0.07	0.09	0.10	0.09	0.06	0.10	0.09	<b>0.10</b>
	HF	0.02	0.03	0.00	0.01	0.16	0.05	0.10	-0.11	-0.01	0.01	-0.01	0.05	-0.02	-0.01	-0.01	-0.02	-0.03	0.03	0.02	<b>0.01</b>
THETA	LF	0.05	0.08	0.03	0.02	-0.04	0.09	0.12	-0.02	0.03	0.01	0.01	-0.05	0.01	0.00	-0.01	0.01	0.08	-0.02	0.02	<b>0.02</b>
	HF	-0.09	-0.12	-0.03	-0.06	-0.03	-0.11	-0.14	0.06	-0.05	-0.04	-0.06	-0.12	-0.08	-0.08	-0.03	-0.04	-0.06	-0.11	-0.13	<b>-0.07</b>
ALPHA	LF	-0.06	-0.06	-0.06	-0.06	-0.09	-0.06	-0.05	-0.09	-0.06	-0.06	-0.06	-0.10	-0.06	-0.06	-0.06	-0.06	-0.06	-0.07	-0.06	<b>-0.06</b>
	HF	0.04	0.03	0.04	0.03	0.03	0.03	0.02	0.06	0.04	0.04	0.03	0.02	0.03	0.04	0.03	0.03	0.03	0.03	0.03	<b>0.03</b>
BETA	LF	-0.06	-0.06	-0.06	-0.06	-0.07	-0.06	-0.05	-0.08	-0.06	-0.06	-0.06	-0.08	-0.06	-0.06	-0.06	-0.06	-0.06	-0.06	-0.06	<b>-0.06</b>
	HF	0.04	0.03	0.04	0.03	0.03	0.03	0.03	0.05	0.04	0.04	0.03	0.02	0.03	0.03	0.03	0.03	0.03	0.03	0.03	<b>0.03</b>

(c)

EEG	HRV	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14	E15	E16	E17	E18	E19	Average
DELTA	LF	0.22	0.20	0.25	0.15	0.14	0.16	0.10	0.09	0.13	0.19	0.05	0.24	0.09	0.18	0.21	0.20	0.11	0.22	0.16	<b>0.16</b>
	HF	0.14	0.27	0.16	0.23	0.18	0.34	0.31	-0.08	0.18	0.21	0.14	0.26	0.13	0.18	0.16	0.09	0.05	0.30	0.30	<b>0.19</b>
THETA	LF	0.02	0.07	-0.09	-0.03	-0.05	0.09	0.11	-0.07	0.00	-0.07	-0.09	-0.13	-0.13	-0.11	-0.12	-0.16	0.00	-0.15	-0.08	<b>-0.05</b>
	HF	-0.09	-0.14	0.05	-0.07	-0.02	-0.14	-0.15	0.08	-0.06	-0.05	-0.09	-0.14	-0.08	-0.12	-0.01	-0.04	-0.05	-0.09	-0.16	<b>-0.07</b>
ALPHA	LF	0.00	0.05	-0.11	-0.04	-0.03	0.09	0.11	-0.05	-0.01	-0.06	-0.10	-0.15	-0.09	-0.11	-0.13	-0.13	-0.03	-0.14	-0.09	<b>-0.05</b>
	HF	-0.09	-0.14	0.04	-0.06	0.01	-0.11	-0.14	0.13	-0.04	-0.02	-0.07	-0.14	-0.01	-0.10	-0.01	-0.04	-0.01	-0.07	-0.12	<b>-0.05</b>
BETA	LF	0.07	0.08	-0.06	0.00	-0.01	0.11	0.12	-0.06	0.03	-0.04	-0.08	-0.12	-0.10	-0.06	-0.09	-0.08	0.01	-0.14	-0.08	<b>-0.03</b>
	HF	-0.08	-0.13	0.08	-0.05	0.00	-0.12	-0.14	0.08	-0.01	-0.03	-0.09	-0.12	-0.05	-0.10	-0.01	-0.06	-0.06	-0.08	-0.15	<b>-0.06</b>

(d)

Table 7.5: Dataset 2- Method 2 - Averaged participants' correlation performance: (a) Level 0 (No WT), (b) Level 1 WT, (c) Level 2 WT, (d) Level 3 WT

**Dataset 1-HRV & Wavelet Transformed EEG Correlation**

EEG	HRV	Level 0	Level1	Level2	Level3
DELTA	LF	0.29	0.30	0.29	0.29
	HF	0.30	0.31	0.30	0.30
THETA	LF	0.28	0.29	0.27	0.28
	HF	0.27	0.29	0.29	0.29
ALPHA	LF	0.25	0.24	0.27	0.31
	HF	0.26	0.25	0.27	0.32
BETA	LF	0.24	0.23	0.27	0.29
	HF	0.27	0.24	0.26	0.29

**(a)**

**Dataset 2-HRV & Wavelet Transformed EEG Correlation**

EEG	HRV	Level 0	Level1	Level2	Level3
DELTA	LF	0.28	0.25	0.27	0.27
	HF	0.26	0.29	0.30	0.30
THETA	LF	0.30	0.30	0.30	0.31
	HF	0.24	0.22	0.22	0.24
ALPHA	LF	0.28	0.28	0.28	0.30
	HF	0.21	0.20	0.22	0.23
BETA	LF	0.30	0.29	0.27	0.30
	HF	0.26	0.25	0.23	0.24

**(b)**

Table 7.6: Method 1 - Heat Map Results of Averaged participants' correlation performance: **(a)** Dataset 1, **(b)** Dataset 2. Colour coding is from Red to Dark Blue, Red=Strongest, Dark-Blue=Weakest)

**Dataset 1-HRV & Wavelet Transformed EEG Correlation**

EEG	HRV	Level 0	Level1	Level2	Level3
DELTA	LF	0.27	0.27	0.27	0.26
	HF	0.24	0.24	0.24	0.25
THETA	LF	0.30	0.30	0.31	0.29
	HF	0.27	0.27	0.26	0.25
ALPHA	LF	0.29	0.29	0.29	0.29
	HF	0.26	0.25	0.25	0.25
BETA	LF	0.29	0.28	0.28	0.29
	HF	0.25	0.25	0.25	0.25

**(a)**

**Dataset 2-HRV & Wavelet Transformed EEG Correlation**

EEG	HRV	Level 0	Level1	Level2	Level3
DELTA	LF	0.23	0.23	0.24	0.31
	HF	0.29	0.29	0.29	0.35
THETA	LF	0.19	0.19	0.20	0.32
	HF	0.32	0.32	0.32	0.34
ALPHA	LF	0.22	0.22	0.21	0.32
	HF	0.29	0.29	0.29	0.33
BETA	LF	0.21	0.21	0.21	0.30
	HF	0.29	0.29	0.29	0.35

**(b)**

Table 7.7: Method 2 - Heat Map Results of Averaged participants' correlation performance: (a) Dataset 1, (b) Dataset 2. Colour coding is from Red to Dark Blue, Red=Strongest, Dark-Blue=Weakest)

## Chapter 8

# Efficient Methods for Calculating Sample Entropy in Time Series Data Analysis

*Results shown in Figures 8.1, 8.3, and Table 8.1 are already published in the conference proceeding of BICA (9th International Conference on Biological Inspired Cognitive Architecture, Prague, Czech Republic, August 2018).*

### 8.1 Introduction

In recent years, different algorithms attempting to improve SE have been proposed. Quadratic Sample Entropy (QSE) was introduced to reduce the influence of arbitrary constants of sequence comparison and tolerance on SE, as well as to reduce the skewing of results when either the top or the bottom of the conditional probabilities was very small or very large (Lake, 2011). Another attempt to improve SE was with the introduction of Fuzzy Entropy (FuzzyEn) (Chen et al., 2007), using the concept of fuzzy sets to determine a fuzzy measurement of the similarity of two vectors based on their shapes. Multi Scale Entropy (MSE) established by (Takahashi et al., 2010), was a useful extension of SE to multiple time scales, in recognition of the likelihood that the dynamical complexity of biological signals may operate across a range of temporal scales. Although new methods for calculating SE have been proposed, improving the efficiency (computational time) of SE calculation methods has not been considered so far.

In this chapter, the type of nonlinear complex measures of variability exploited is SE. The aim of the research is not to propose another new method derived from SE, but to efficient method improving the computational time for the SE calculation. I will compare the computational time for SE calculation using the new and original SE methods. This

Chapter shows an analysis of calculating correlation between Electroencephalogram (EEG) and Heart Rate Variability (HRV) based on their SE values in time domain.

### 8.1.1 Dataset Information

The dataset consists of EEG and ECG recordings from 15 participants. This data was obtained over 5 minute time slots in a relaxed state with eyes opened. The dataset used in this chapter is Dataset 1 as described in Chapter 3.

## 8.2 Sample Entropy and Proposed Implementation

SE is considered as an effective method for investigating different types of time series data. A lower SE value indicates a high frequency of similarity in time series (Richman and Moorman, 2000). The description of how SE works can be found in Chapter 5.

### 8.2.1 Efficient and Parsimonious way for Sample Entropy Calculation

SE measures the probabilities of matches for a time series data using point-wise approach. Instead of finding differences between each point  $m$  with *Approximate Entropy*, the point-wise approach of SE calculates the probability for each point separately. This can be time-consuming when a long sequence of points needs to be compared. The computation time for SE can be reduced without losing much information from the signals by using the three methods proposed in this section. Thus, the calculation time for SE could be shortened and computational expense would be cheaper. Figure 8.1 illustrates how these three methods work.

**SE-Method 1** is about shortening the time series signal without the loss of too much information to the point-wise approach. For example, instead considering the original data length of the 5-minute signal ( $250Hz \times 300Sec = 75000$  data points), it could be shortened by dividing the original data length by 1.1 (ignoring last 25 seconds data) ( $75000/1.1 = 68181$ , equivalent to 275 seconds data points). Binary chop (Knuth, 1998) is performed ignoring the number of neighbours for comparison, in order to find out, at which point the most accurate result with the shortest length for SE could be obtained. Due to much less combination results, consequently, less comparison, the computational time is much cheaper than the original method.

**SE-Method 2** is about SE calculation on a moving window. For example, using a 2 seconds moving window, SE is calculated for a window size of 500 points ( $2 \times (250Hz) =$

500 data points). Using this method, I calculate SE on individual windows and find out which window size gives the SE values that are most similar to the original SE value.

**SE-Method 3** is to calculate the mean of input data for a given window first before performing SE. This data window could be as long as a minute or as short as a second. For example, if the mean of each 1 second data (250 points) is gathered, then it will give us a reduced length of  $n = 300$  data points on which to perform SE calculation, and not  $n = 75000$ . By this way the SE computational time should be reduced dramatically.

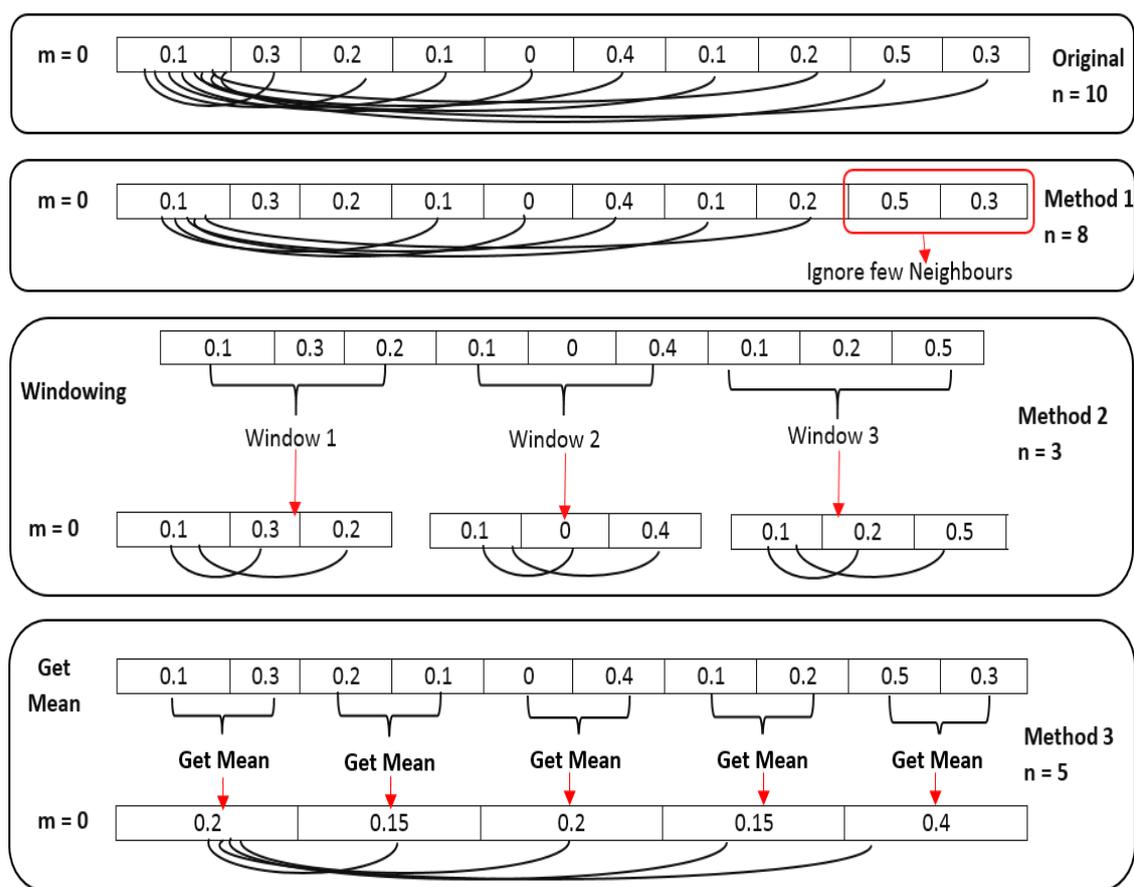


Figure 8.1: An example of how the SE can be calculated efficiently.

### 8.3 Experimental Results

The EEG signals were pre-processed to remove artefacts caused by electrical activity in muscles, including the eye, jaw and other muscle movements using Independent Component Analysis (ICA) as mentioned by (Hyvärinen and Oja, 2000). It is relatively straightforward to remove these artefacts using ICA.

### 8.3.1 Experiments using three proposed SE Calculation Methods

For each of the five minutes of EEG data, the following three experiments have been undertaken, and the results are shown in Table 8.1 and Figure 8.2. For the purpose of comparison, SE values of the original SE performance are also shown. All code is run on a personal computer: Windows 7 Enterprise, Intel (R) Core (TM) i7-3770T, 64-bit Operating System, 16 GB RAM.

**Experiment 1** is the implementation of SE-Method 1, by restricting the number of neighbours for comparisons on SE calculation. It is found that ignoring the last 25 seconds of data ( $300\text{Seconds} - 275\text{Seconds} = 25\text{Seconds}$ ) still achieves similar results as if they are included, but with the improving computational time of 13 seconds. The ignorance of 25 second data was considered because due to binary chop process, difference chops have been applied, and the SE was applied to each chopped data. During which, no change in SE found between 275 - 300 seconds data. Therefore, 25 seconds data were ignored in SE calculation.

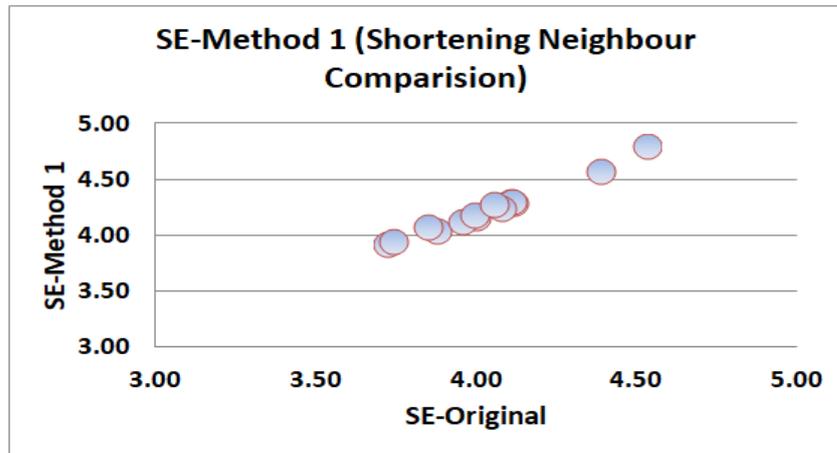
**Experiment 2** is about experimenting SE-Method 2, I considered 10 different window sizes (i.e. 2 Sec, 10 Sec, 20 Sec, 30 Sec, 40 Sec, 50 Sec, 60 Sec, 70 Sec, 80 Sec and 90 Sec windows) on which the SE calculation was performed, in order to find out which window size gives the SE values that are most similar to the original one. As shown in the Table 8.1 the smaller the window size, the shorter the calculation time. Moreover, it is found that the smallest window size gives most similar results to the original SE calculation.

**Experiment 3** demonstrates SE-Method 3, calculating the mean of each window (e.g. 1 Sec = mean of 250 points). The experiment is done with 8 different window sizes (i.e. 0.06 Sec, 0.12 Sec, 0.25 Sec, 0.55 Sec, 1 Sec, 2 Sec, 3 Sec and 4 Sec) on which to calculate the mean, as shown in Table 8.1. The SE is then performed on the mean values of the signal. As expected, the bigger the window size, the shorter the calculation time. Moreover, the best match with the original SE calculation results is the mean of each 1 second window.

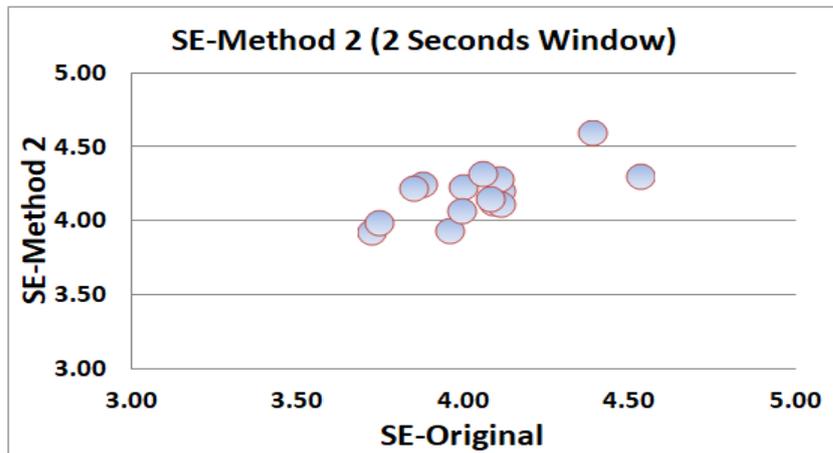
Figure 8.2 shows the results of Experiments 1-3 for all 15 participants, where the results for individual participants is the average for all electrodes. For each of these methods except SE-Method 1, different window sizes on which to calculate SE are utilised, as shown in Table 8.1. Figure 8.2 shows the results of Experiment 1 (SE-Method 1), Experiment 2 (SE-Method 2 for 2 seconds moving window), and Experiment 3 (SE-Method 3 for the mean of each 1-second window). The SE result for all other window sizes are available in Appendix C.

Experiments 1-3 demonstrate strong positive correlation between the results obtained using the original method and each of the new three SE approaches. The correlation values are 0.99, 0.68, and 0.96 for SE-Method 1, SE Method 2 (2 seconds moving window) and

(a)



(b)



(c)

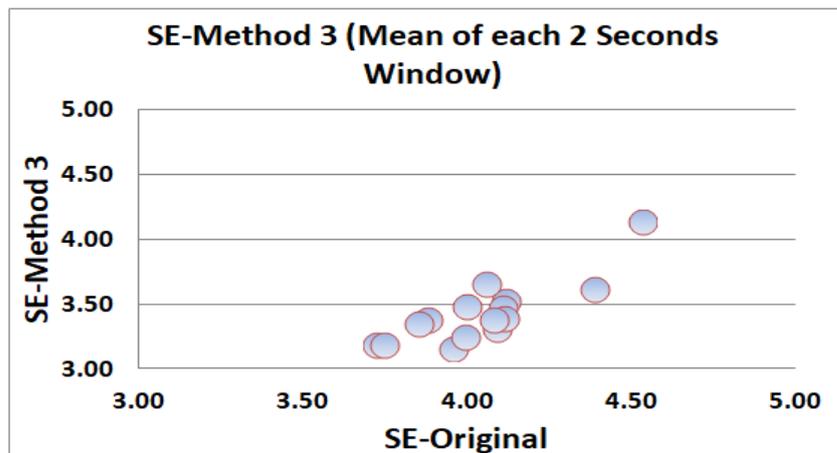


Figure 8.2: Comparison of original SE calculation and proposed SE methods, showing average results for all electrodes are shown for the individual participants : (a) Method 1- Shortened Neighbours comparison, (b) Method 2- Using Moving Window of 2 seconds of data, and (c) Method 3- Calculating SE from Mean values of each 2 seconds of data. Red= Original SE performance, Blue= Proposed Method's SE Performance.

Table 8.1: Computation time for the SE calculation using the Original approach and the three efficient methods.

Experiments	Details	Computation Time
SE-Original	Original Performance	75 Sec
SE-Method 1	Shortening the neighbour comparison	62 Sec
SE-Method 2	2 Seconds Moving Window	0.002 Sec
	10 Seconds Moving Window	0.08 Sec
	20 Seconds Moving Window	0.38 Sec
	30 Seconds Moving Window	0.72 Sec
	40 Seconds Moving Window	1.30 Sec
	50 Seconds Moving Window	1.94 Sec
	60 Seconds Moving Window	3 Sec
	70 Seconds Moving Window	4 Sec
	80 Seconds Moving Window	6 Sec
90 Seconds Moving Window	9 Sec	
SE-Method 3	Mean of Each 0.06 Seconds Window	0.38 Sec
	Mean of Each 12 Seconds Window	0.11 Sec
	Mean of Each 25 Seconds Window	0.02 Sec
	Mean of Each 50 Seconds Window	0.007 Sec
	Mean of Each 1 Seconds Window	0.003 Sec
	Mean of Each 2 Seconds Window	0.008 Sec
	Mean of Each 3 Seconds Window	0.01 Sec
	Mean of Each 4 Seconds Window	0.02 Sec

SE-Method 3 (mean of each 2 second window), respectively, along with the probability of 0. Whilst SE-Method 1 and SE-Method 2 is not following the trend of SE values and do not improve the computational time for SE calculation significantly, SE-Method 3 clearly works best in terms of trend and computational time, providing the most predictive value for SE performance to those provided by original SE.

### 8.3.2 Experiment 4

The aim of experiment 4 is to compare the correlation performance of the new methods and original methods for SE calculation through SE values of EEG and HRV data using Pearson Correlation Coefficient (PCC). The previous three experiments show that SE-Method 3 is the best one in terms of improving SE calculation time without losing much information. Hence, only SE-Method 3 is considered in this experiment. PCC works on the same length of the signals, and SE-Method 3 could give us the same length of samples for EEG and HRV. In the previous chapter, the TD correlation between EEG and HRV data has been performed, and results suggested no correlation. In this chapter, the TD correlation analysis has been conducted through SE values. In order to demonstrate a correlation between EEG

and HRV in TD, for each of the five minutes of EEG and HRV data, the following steps have been undertaken for both original and the new approach (SE-Method 3) of SE calculation.

1. For each electrode's EEG, divide data into 10 equal width bins to perform SE calculation. Similarly, for HRV data, divide data into 10 equal width bins to perform SE calculation. By this way, 10 SE values of EEG and 10 SE values of HRV will be gathered. This process has been repeated for each participant.
2. For each Participant, compute PCC on the 10 SE values of the EEG and 10 SE values of the HRV obtained in step 1.

Pearson's correlation is used to perform the correlation coefficients. It measures how closely two different observables are related to each other. Pearson's correlation coefficient  $R$  ranges between 1 (when the matching entities are exactly the same) and  $-1$  (when the matching entities are inverses of each other). A value of zero indicates no relationship existing between the entities.

Once the correlation values are gathered for each electrode, the best performance electrodes have been ranked, where the ranking has been given based on electrode correlation values, the bigger the value, the higher the rank. The top three best performance electrodes' results have been looked closely. Some common best electrode rankings are found for all the participants investigated. Figure 8.3 shows the results of Experiment 4, showing common best electrode ranking for all participants, highlighted in yellow colour.

It is found that the electrical activity in the frontal lobe of the brain appears to be correlated with the HRV in time domain. Moreover, the new approach (SE-Method 3) of SE calculation is giving more focused results than the original SE calculation.

## 8.4 Discussion and Conclusion

The main conclusion of this work is that parsimonious results for SE can be achieved using the proposed new methods of preprocessing the data prior to SE calculation. SE-Method 3 works best because it improves the SE performance, giving good predictive values without changing the trends visible in the SE, calculated using the original standard approach. SE-Method 1 can provide SE values very close to those obtained using the original SE approach, but it does not improve computational time much. Similarly, SE-Method 2 is not robust because it does not improve the trend nor computational time significantly.

The second conclusion from this work is that there is a strong positive correlation ( $R=0.96$ , Probability = 0) between results obtained using the original and the new (SE-Method 3) SE approaches. Also, we found low positive correlations between SE values of

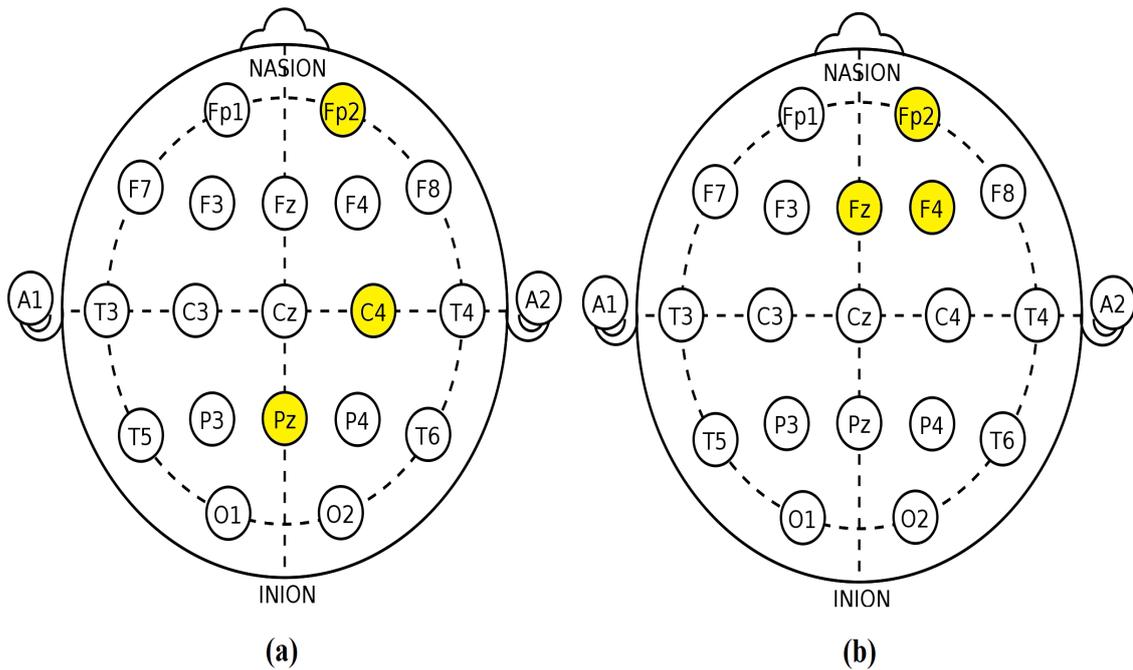


Figure 8.3: Electrode Ranking based on correlation performance between SE values of EEG and HRV, showing best three performing electrodes across participants, highlighted in yellow colour: (a) Ranking based on the original approach for SE calculation, and (b) Ranking based on the new approach (SE-Method 3) for SE calculation.

EEG and HRV in the time domain. The results shown in the previous chapter (Chapter 7) suggested that the electrical activity in the frontal lobe of the brain is correlated with the HRV at frequency domain. The work in this Chapter shows that the electrical activity in the frontal lobe of the brain appears to be correlated with the HRV in the time domain as well.

In summary, SE-Method 1 and SE-Method 2 does not improve the trend or the computational time for SE calculation much. SE-Method 3 does not give values similar to those provided by the original SE approach, but it provides the most predictive value for SE performance. Although the result is not exactly similar as the original SE performance, the trend is the same. Therefore, we demonstrated that the most efficient way for SE calculation is SE-Method 3.

# Chapter 9

## Heart Rate Variability Time Series Analysis using Embedding Dimension

*Results shown in Figures 9.1 and 9.2 are already published in the conference proceeding of BICA (9th International Conference on Biological Inspired Cognitive Architecture, Prague, Czech Republic, August 2018, Elsevier's Procedia Computer Science).*

### 9.1 Introduction

Heart Rate Variability (HRV) is a measurement sequence with one or more visible variable of an underlying dynamic system, whose state changes with time. In practice, it is difficult to know what variables determine the actual dynamic system. Embedding Dimension (ED) plays an important role in time series analysis (Chun-Hua and Xin-Bao, 2004). The aim of this chapter is to find the nature of underlying dynamical systems for HRV time series using the ED. For analysing predicting variables responsible for HRV time series, False Nearest Neighbour (FNN) method of estimating ED is adapted. The description and illustration of ED along with FNN method are available in Chapter 5. In this chapter, four datasets are used to analyse the HRV signals of participants with and without medical conditions. The HRV time series taken from participants are over a fixed period.

#### 9.1.1 Dataset Information

This Chapter utilised four datasets including two (Dataset 2, and Dataset 3) without a medical condition and two (Dataset 9, and Dataset 10) with a medical condition. The detail of these datasets is available in Chapter 3.

## 9.2 Experiments and Results

### 9.2.1 HRV Analysis using ED

The purpose of this experiment is:

- To find out underlying variables in the HRV time series.
- To find out if there is any difference between individual participant's HRV based on their ED results?
- To find out if there is any difference between the ED result of individual 5-minute slots of each participant?

For each participant, the percentage of FNN obtained as an increase in the ED. I use the MATLAB code to obtain ED of HRV (Mirwais, 2012). Once the percentage of FNN for the ED for each 5-minute slot is obtained, the results are plotted for each participant, as shown in Figures 9.1, 9.2, 9.3, and 9.4. A comparison on ED result of different participants and each 5-minute slots has analysed in detail.

For Dataset 2 and Dataset 3, the ED results of 4 participants, two from each dataset (Participant 1 and Participant 2 from Dataset 2, and Participant 3 and Participant 4 from Dataset 3) are shown in Figure.9.1 and Figure 9.2, respectively. For all the other participants from these datasets, results are similar to what as shown in Figures 9.1 and 9.2. The results of other participants of Dataset 2 and Dataset 3 are available in Appendix D (See Figure D.1 and Figure D.2, respectively).

For Dataset 9, the ED results of two participants (Participant 2, and Participant 4) are as shown in Figure 9.3 (a) and (b), respectively. For all the other participants from Dataset 9, results are similar to what is shown in Figure 9.3 (a) and (b). The results of other participants are available in Appendix C (See Figure C.3).

For Dataset 10, the ED results of two participants (Participant 2, and Participant 4) are as shown in Figure 9.4 (a) and (b), respectively. For all the other participants from Dataset 10, results are similar to what is shown in Figure 9.4 (a) and (b). The results of other participants are available in Appendix C (See Figure C.4). Note, for each participant in Figure 9.4 (a) and (b), comparison of ED performance is among HRV time series measured with different electrode (Please refer to Chapter 4 for more detail about the four electrodes), and not between the slots as shown in Figures 9.1, 9.2, and 9.3 respectively. The reason for not showing the slot's information is because the datasets only contain a 5-minute recording of each participant, and therefore, lack of slots information is available with this dataset.

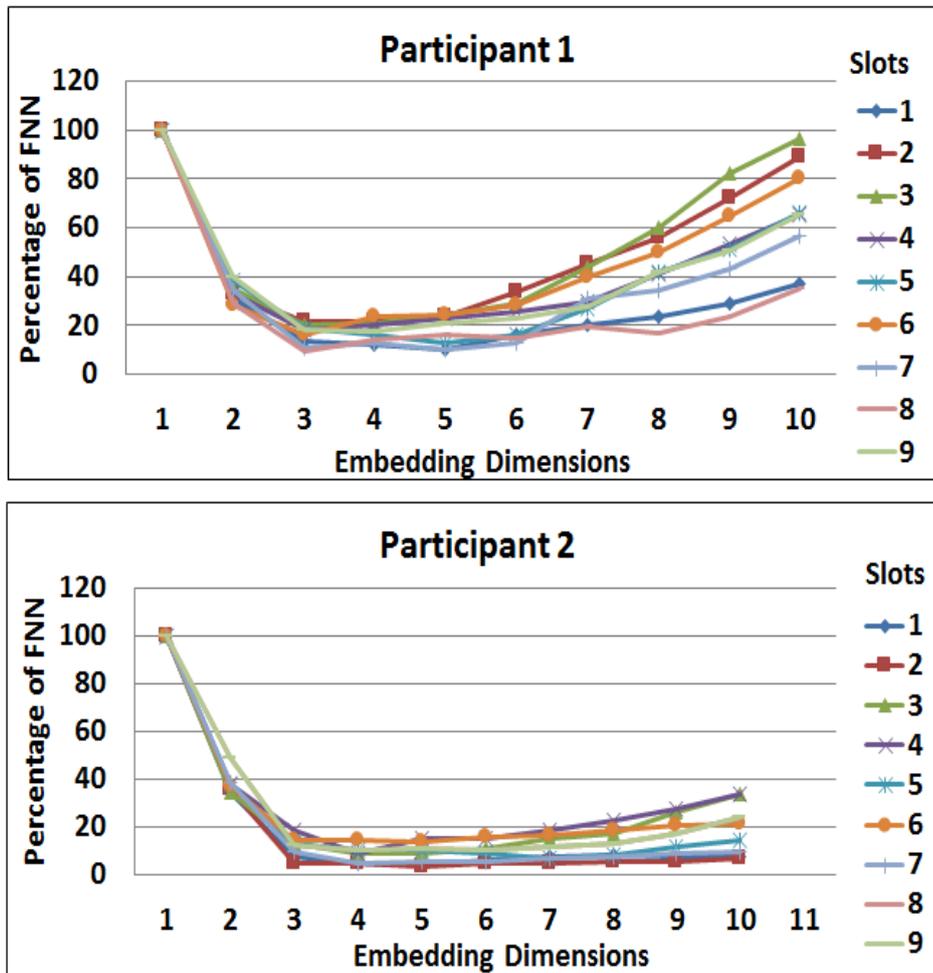


Figure 9.1: Dataset 2- Embedding Dimension Result: Two participants (Participant 1 and 2) ED result from Dataset 2.

In Figures 9.1, 9.2, 9.3, and 9.4., the X axes (Horizontally) represent ED from 1 to 10, and Y axes (Vertically) represent the percentage of FNN for the ED. In Figures 9.1, 9.2, and 9.3 the 8-9 different colours (Curves) in the graph correspond to slots. In Figure 9.4, the five different colours (Curves) in the graph correspond to different Electrodes. There are three important findings from these results:

1. The first notable finding for all participants/Datasets, is that the optimal ED is about 4. It is also notable that this optimal ED is independent of the stimulus locations. As shown in Figures 9.1, and 9.3, the different colours represent different slots, containing different stimulus locations. Therefore, it suggests the independence of stimulation. These results indicate that an ED of 4 is most appropriate for HRV data for all slots and all participants.

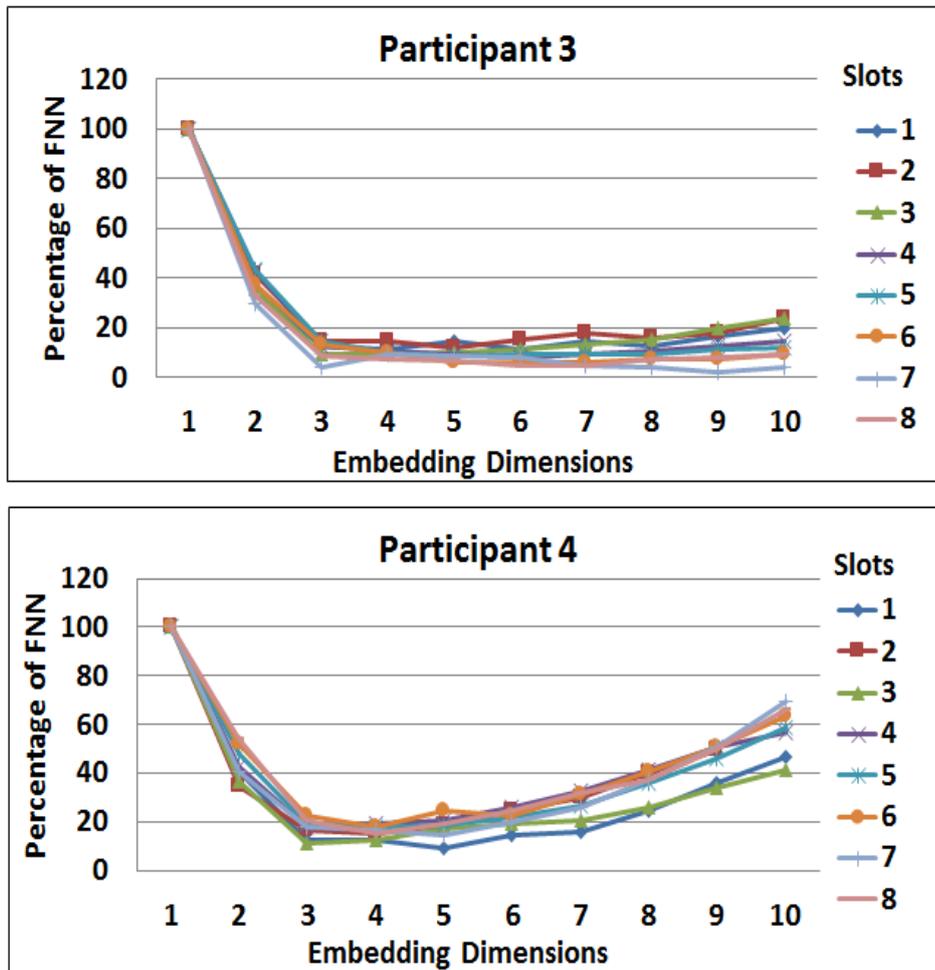
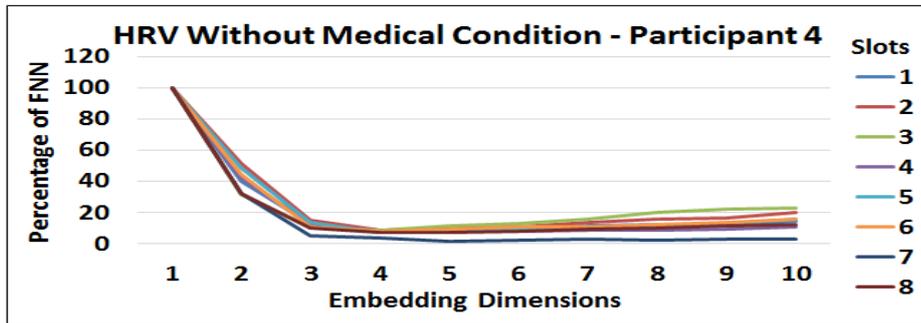
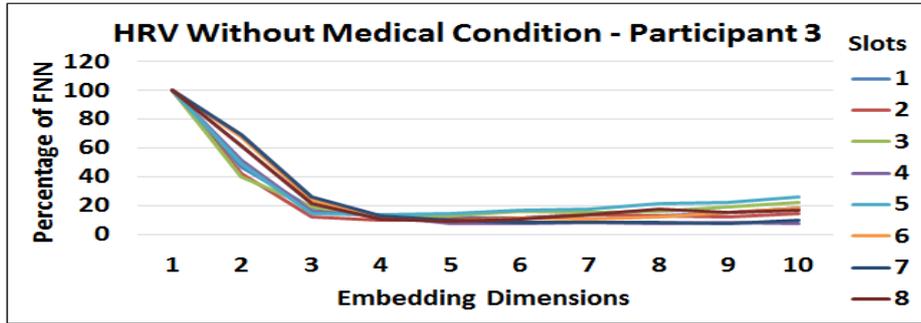
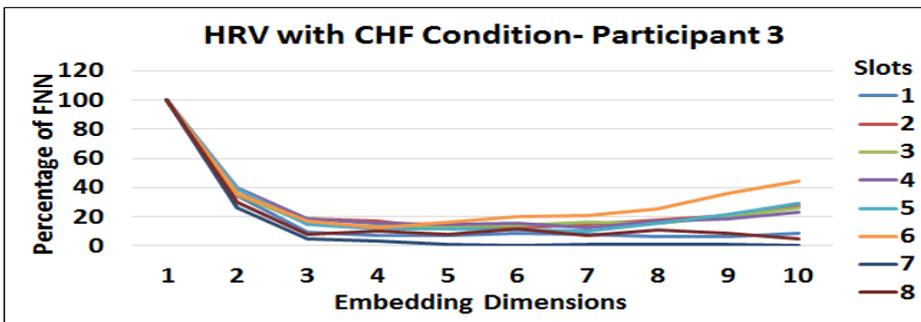
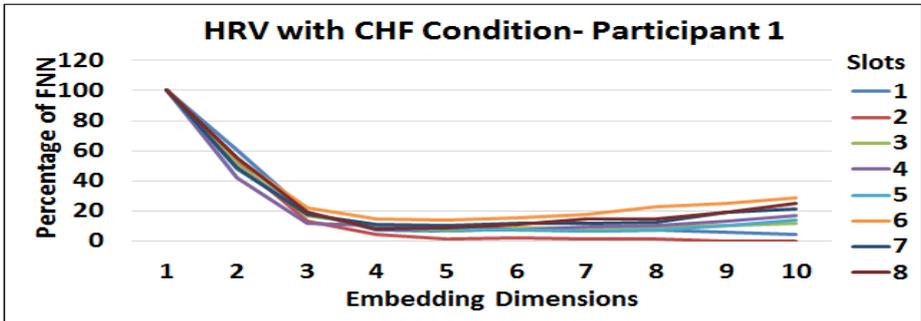


Figure 9.2: Dataset 3 - Embedding Dimension Result: Two participant's (Participant 3 and 4) ED result from Dataset 3.

- In Figures 9.1 and 9.2, the bottom figures (Participant 2 and Participant 4) show increasing numbers of FNN while ED rises above its optimal value of 4, whereas the top figures (Participant 1 and Participant 3) do not display this. An increase in the number of false nearest neighbours with increasing ED usually is suggestive of noise in the data (Abarbanel et al., 1993).
- In Figures 9.3 and 9.4, nevertheless of the medical condition of the data, the results are similar to datasets without a medical condition. However, less noise has been found in the participants with a medical condition such as for Fetal ECG (Figure 9.4). These results suggest that an ED of 4 is most appropriate for HRV data with the medical condition as well.



(a)



(b)

Figure 9.3: Dataset 9 - Embedding Dimension Result: (a) Two Participants (Participant 3 and 4) ED result for HRV without medical condition, (b) Two Participants (Participant 1 and 3) ED result for HRV with a medical condition Congestive Heart Failure (CHF).

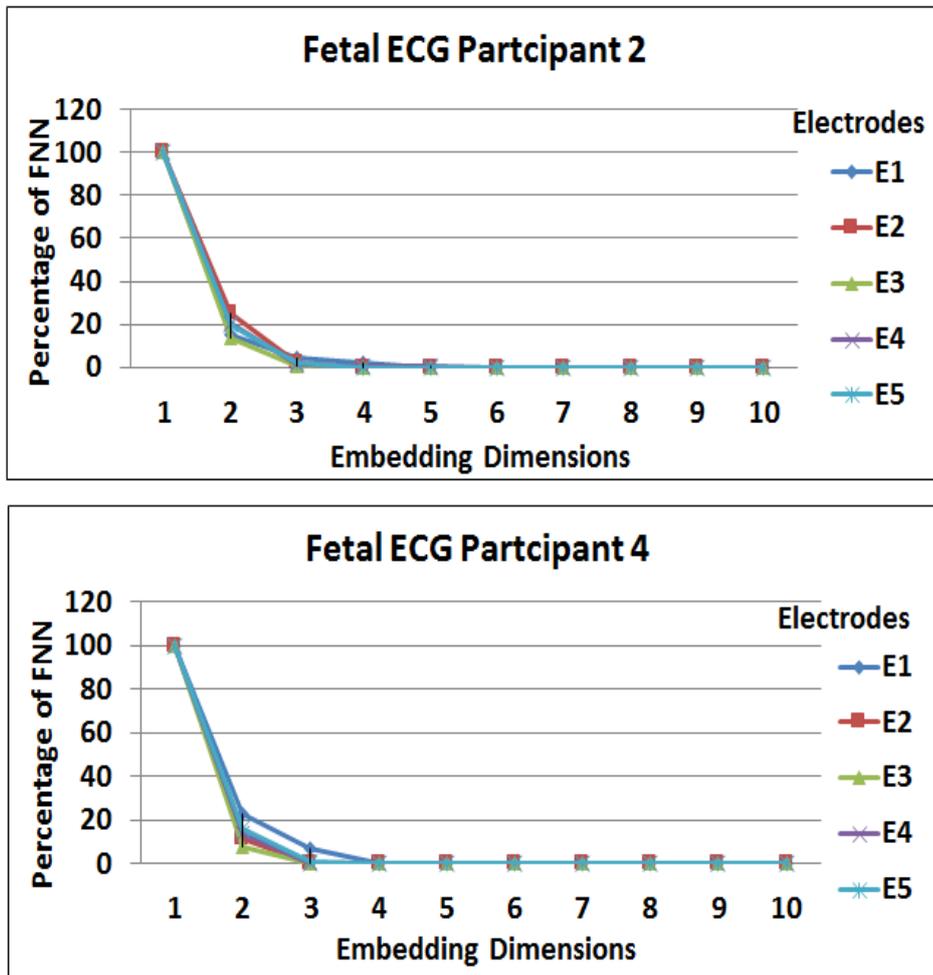


Figure 9.4: Embedding Dimension Result: Two participants (Participant 2 and participant 4 (5 measured Electrodes, as shown in different colour)) from Dataset 10.

### 9.2.2 HRV Prediction using Linear Regression

The standard Linear Regression is utilised to produce a linear predictor for the embedded data. It was suggested by Kil (Kil et al., 1997) that the size of a window, in which the time series embedded, plays an important role to project a time series. Therefore, to predict the window size for the HRV series to achieve the best result, Linear Regression is used in this chapter. To perform Linear Regression, I have considered a 5 minute HRV time series data. The embedding dimension, the size of the input layer, is increased from 2-units to 6 units. For each window, data are split into a training set of 250 vectors and a test set of 106 vectors; moving window size by one vector point. The size of the window is the size of the previous window excluding the first data point and adding the last data point to make the similar length windows. For example, for the window 2, the size of the window will be as follows:

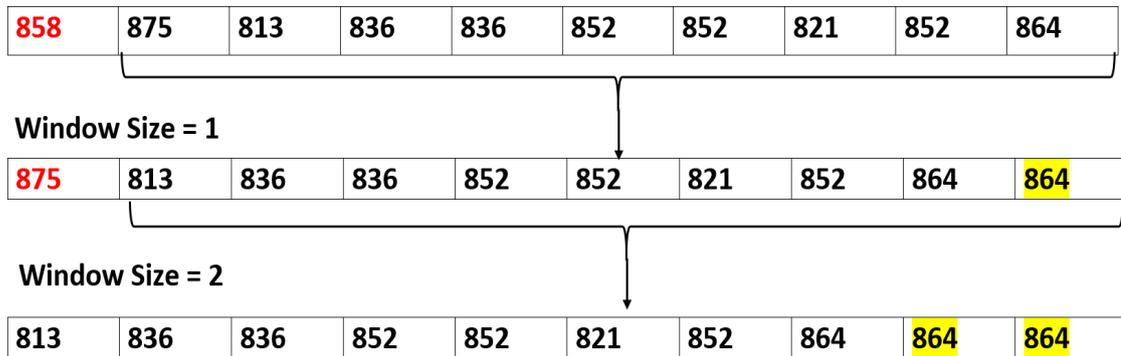


Figure 9.5: Window Size selection for the HRV prediction, where red coloured data points are not considered, and the yellow highlighted data points are repeated in the window size of 1 and 2.

After results from five different window sizes are obtained, the relative errors are calculated using the following equation:

$$RelativeErrors = \sqrt{\frac{\sum_t (observation_t - prediction_t)^2}{\sum_t (observation_t - prediction_{t+1})^2}} \quad (9.1)$$

The results for five different window size are as shown in Figure 9.6. It is clear that the best regressor has four inputs and changing this number either way harms the projection of HRV time series. Figure 9.6 shows that error rises as the window increases, and this cannot, here, be due to contamination in the higher dimensions. It is probably due to over-fitting on the training set.

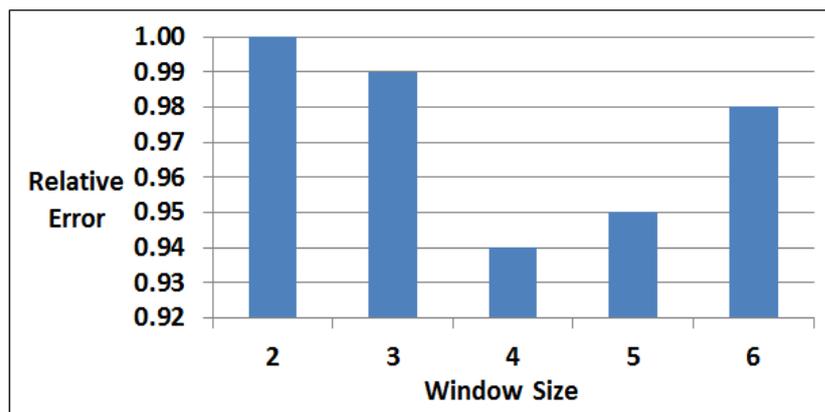


Figure 9.6: Relative error of the predictor for varying window size for HRV Time series.

The linear prediction of the HRV suggests that the window size of 4 will be enough to fit the HRV time series data. Also, this reflects the ED result of 4 as the minimum ED for

the HRV Data with/without any medical condition shown in Figures 9.1, 9.2, 9.3, and 9.4, respectively.

### 9.3 Discussion and Conclusion

Results shown in this chapter, prove that ED can produce evidence of dynamic variables which contribute to the HRV time series. Moreover, the embedding of the HRV time series into a four-dimensional space produced the smallest number of FNN. Based on Takens (Takens, 1981), if the original dynamic system had a dimension of  $N$ , then an embedding of size  $2N$  will be fully regained the original system. The results shown in this chapter estimated ED of 4 ( $2 \times N = 2 \times 2 = 4$ ), which suggests that the underlying dynamic system of HRV has  $N = 2$  features. This result is impressive because HRV is driven by two underlying variables, the sympathetic and parasympathetic neural pathways. HRV is a marker of sympathetic and parasympathetic influences on the modulation of heart rate (Zhong et al., 2004), and this is reflected in the ED result. The effect of the sympathetic pathway is to increase heart rate and blood pressure (Fight or Flight response), whereas the parasympathetic path acts to decrease heart rate and blood pressure (Rest and Digest response). Therefore, the main finding here is that, in all circumstances, an Embedding of the HRV time series in a four-dimensional space generates the smallest number of false nearest neighbours. This finding strongly suggests that the Autonomic Nervous System that drives the heart is a two-dimensional dynamic system.

From the participants questionnaires from Dataset 2 and Dataset 3, a variety of subjective responses to the acupuncture stimulation were found. However, this did not appear to have much effect on the HRV time series, which robustly kept its two-dimensional dynamic system.

In some circumstances, the percentage of FNN increase as the ED became more massive than the optimal value. This increase suggests noise in the data that may have come from the ECG measuring equipment.

It was suggested by Kil (Kil et al., 1997) that the performance of a time series prediction is affected by the window size, in which the time series embedded. So the best predictor would be the one that used correct ED. The experiments, reported in this Chapter, using a Linear Regression to predict the HRV series, confirmed that a window size of four gave the best prediction result.

# Chapter 10

## Conclusion and Future Work

This Chapter is the concluding chapter of the thesis. It contains a summary of each preceding chapters, contributions to the knowledge and future work in Section 10.1, 10.2 and 10.3, respectively. It is worth mentioning that the work shown in this thesis is an interdisciplinary study between Computer Science and bio-informatics. Hence, researching on both sides has been a challenging experience for me.

### 10.1 Summary of Chapters

This section summarises the main points in the thesis.

- **Chapter 1** explains the necessity of time series data analysis, specifically on EEG and ECG. In this chapter, state-of-the-art research in this area is briefly introduced, and how those work motivated me to extend the research is depicted.
- **Chapter 2** opens the required background to understand time series data: EEG, ECG (HRV). This chapter is divided into a few parts to echo individual components of the interdisciplinary area. The first part is the general description of dynamic, linear, nonlinear and chaos systems. In the second and third parts, information about EEG and ECG (HRV) signals is provided, respectively. Furthermore, these parts also show the existing research work on analysing EEG and ECG (HRV) signals. The fourth (final) part describes the research work done on investigating correlation between EEG and ECG (HRV) signals.
- **Chapter 3** presents the datasets used in this thesis. This chapter is divided into two parts with the first part providing the description of the dataset (Dataset 1) recorded by myself and the second part explaining all the other datasets (Datasets 2-10), either given to me or obtained from the internet. The datasets from the second part contain

EEG and ECG (HRV) data of participants with and without medical conditions. The medical conditions include Autism, Epilepsy, Epileptic Seizure, Congestive Heart Failure and Fetal ECG, respectively.

- **Chapter 4** talks about what data preprocessing is and its importance in analysing time series data. Commonly used data preprocessing techniques for the time series data is listed in this chapter. Further to this, the well-known data preprocessing techniques utilised for the EEG and ECG (HRV) data, such as Independent Component Analysis (ICA), Fast Fourier Transform (FFT) and Wavelet Transform (WT), along with an example for each, are described. The utilisation of these techniques in my thesis is as follow: FFT and WT are utilised only in Chapter 7, whereas ICA in Chapters 6, 7 and 8. WT and ICA are employed on the raw data in order to remove artefacts, whereas FFT is used to extract useful component of the data.
- **Chapter 5** gives a description of data analysis and its importance in finding information from time series data. It enumerates commonly used data analysis methods for the time series data. Further to this, well-known data analysis methods utilised for the EEG and ECG (HRV) data, such as Approximate Entropy, Sample Entropy (SE), Pearson's Correlation Coefficient (PCC), Cross-Correlation (CC), and Embedding Dimension (ED), along with an example for each, are described. The utilisation of these methods in my thesis is as follow: SE is used in Chapter 8, PCC in Chapter 7, CC in Chapter 6, and ED in Chapter 9. SE is adopted for measuring time series complexity, ED for unveiling underlying variables of the time series, and PCC as well as CC for finding correlation between two time series.
- **Chapter 6** is the first result chapter of the thesis, in which correlation between EEG signals measured through electrodes is investigated using Cross-Correlation method. This chapter shows correlation performance between EEG signals of participants with and without medical conditions. The main findings of this chapter are:
  1. The correlation decreases linearly when the distance between the electrode increases. This means that electrical activity correlates with the physical distance between electrodes.
  2. The presence of linear dependency is only found in the participants without medical conditions. For the participants with medical conditions, such as Epilepsy and Epileptic Seizures, linear dependency does not exist.

3. The correlation found is independent of brain hemispheres for all six datasets utilised in this chapter. This result suggests that most probably the electrical signals are transmitted through the white matter of the brain. This means that in practice it might not matter which side of the medial plane you place the electrodes.
- **Chapter 7** is the second result chapter of the thesis, in which Pearson's Correlation Coefficient between EEG and HRV signals in Time Domain (TD) and Frequency Domain (FD) are shown. In FD, both with and without WT processed data are analysed, whilst TD is only for the data without WT due to the time limitation of this research. In FD, to investigate correlation between EEG frequencies and HRV frequencies, three experiments are conducted: 1) Correlation of preprocessed signals HRV and EEG, 2) Correlation of preprocessed as well as WT on both signals HRV and EEG, and 3) Correlation of preprocessed as well as WT only on EEG and HRV. For the reason of different sample rates of EEG and HRV, two different approaches (named as Method 1 and Method 2) are utilised to segment EEG signals and to calculate Pearson's Correlation Coefficient for each of the EEG frequencies with each of the HRV frequencies in FD.
    1. In TD, no correlation has been found between EEG and HRV.
    2. In FD, for both with and without WT processed EEG, electrical activities in the frontal lobe of the brain correlates with the HRV without WT process. This result suggests that the electrical signals might be transmitted through the cerebral cortex, Thalamus, and Medulla of the brain (Saper et al., 2005).
    3. In FD, a more tentative conclusion of this work is that EEG frequencies (Delta, Theta, Alpha and Beta) are positively correlated with HRV frequencies (Low Frequency and High Frequency). Whereas, most of previous studies, (Yang et al., 2002), (Ako et al., 2003), (Jurysta et al., 2003), (Abdullah et al., 2010) and (Chua et al., 2012), shown a negative correlation between these frequencies. The reason for this contradictory results could be the condition in which these signals have been analysed.
  - **Chapter 8** is the third result chapter of the thesis. The aim of this chapter is to design a new calculation method for time series complexity to improve computational time. In order to find an efficient method for SE calculation, three different calculation methods (named as SE-Method 1, SE-Method 2, and SE-Method 3) are experimented

on preprocessed EEG and HRV data. Out of the three methods, the one having the best performance (SE-Method 3) is chosen to compare against the original SE method in the context of calculating a correlation between EEG and HRV in the Time Domain (TD). The main findings are:

1. Parsimonious results for SE calculation is achieved using a proposed new method. SE-Method 3 gives the best SE values among the three proposed methods. In addition, it improves the computational time without changing the original prediction trends visible in the SE.
  2. The work in this chapter shows that the electrical activities in the frontal lobe of the brain appears to be related to the HRV in the TD. This finding is in consistency with the results shown in Chapter 7 which suggests that the electrical activities in the frontal lobe of the brain correlates with the HRV in FD.
- **Chapter 9** is the last result chapter of the thesis, which unveils the underlying variables determining HRV using Embedding Dimension (ED). False Nearest Neighbours (FNN) method is exploited for analysing variables responsible for the HRV time series. In this chapter, four datasets are used to analyse the HRV signals of participants with and without medical conditions. The main findings are:
    1. For Both participants with and without medical condition, the results present that the HRV has an ED of four, which suggests that the underlying dynamic system has two variables. This is because, if the original dynamic system has a dimension of  $N$ , then an embedding size of  $2N$  is fully regained from the original system (Takens, 1981). This finding is interesting because it was found that HRV is a marker of sympathetic and parasympathetic influences on the modulation of heart rate (Zhong et al., 2004).
    2. The results show an increase in the percentage of FNN when the ED increases, which might be the effect of noise in the data (Abarbanel et al., 1993).

## 10.2 Contribution to knowledge

My contributions are:

- **Discovered that the correlation between EEG signals measured through electrodes is linearly dependent on the straight-line (Euclidean) distance between**

**them** for participants without medical conditions, but not for participants with medical conditions.

Over the past few years, correlation between EEGs was investigated in frequency domain using limited number of electrodes, and limited number of combinations of electrode pairs. Their focus was primarily on electrode combinations within either the left or right brain hemisphere. To my knowledge, no research has investigated the correlation between EEG signals and distance between electrodes. Furthermore, no one has compared the correlation performance for participants with and without medical conditions. In my research, I have filled up these gaps by using a full range of electrodes and all possible combinations of electrode pairs analysed in Time Domain (TD).

- **Demonstrated that EEG at the front area of the brain has a stronger correlation with HRV than the other area** in frequency domain.

The recent research for correlation between EEG and HRV has focused on Fourier analysis of the frequencies presented in these signals, to analyse their functionalities under certain conditions. Some research has investigated correlation between EEG and HRV limited to certain brain areas and demonstrated the existence of correlation between EEG and HRV. But no research has indicated whether or not the correlation changes with brain area in either frequency domain or time domain. Although Wavelet Transformations (WT) have been performed on time series data including EEG and HRV signals to extract certain features respectively by other research, so far correlation between WT signals of EEG and HRV has not been analysed. My research covers these gaps by conducting a thorough investigation of all electrodes on the human scalp.

- **Designed a new calculation method of sample entropy which could improve computational time significantly in the context of calculating a correlation between EEG and HRV.**

Recent research has proposed new calculation methods for Sample Entropy, aiming to improve the accuracy. To my knowledge, no one has attempted to reduce the computational time of SE calculation. My application required a fast calculation method because of the huge data length. Therefore, I designed three methods for SE calculation: Method 1 (SE-Method 1) is about shortening the time series signals without loss of too much information, Method 2 (SE-Method 2) is about SE calculation on the moving window, and Method 3 (SE-Method 3) is about calculating means for a

given window first before calculating SE. Method 1 and Method 2 do not improve computational time. Method 3 (SE-Method 3) improves the computational time for SE calculation significantly whilst has similar SE value pattern with time as that of the original SE.

- **Two variables determining ECG (HRV) were unveiled .**

Embedding dimension (ED) is utilised to study complex systems that appear ubiquitous in nature, but limited to certain dynamic systems (e.g. analysing variables affecting stock values). No literature has investigated the nature of the underlying dynamic system of HRV. My research highlights this matter by analysing actual variables determining HRV using ED technique. ED is a well-known technique for dynamic system analysis. In this thesis, I have not proposed a new idea on finding ED, but presented experimental results of using it on analysing HRV. The results strongly suggest that the autonomic nervous system driving the heart is a two-dimensional dynamic system. This finding is interesting because HRV is a marker of sympathetic and parasympathetic influences on the modulation of heart rate (Zhong et al., 2004), and this is reflected in the ED results.

## 10.3 Future Work

The future work to perform after this research includes:

- Utilise more datasets to support the research questions in this thesis:
  1. Extend correlation analysis between EEG and HRV signals on other datasets recorded under the same condition. In this thesis, only Dataset 1 and Dataset 2 contains EEG and HRV signals under the same condition.
  2. Apply the new, improved method of SE calculation on other datasets to see the effect of it, compared with the original SE calculation method.
- Consider more nonlinear methods to assess the variability of EEG and ECG/HRV:
  1. Cross-Correlation (CC) method is used to investigate correlation between EEG signals in TD. The reason for using CC is to consider different time lags. It is worth extending correlation performance using other method, such as Pearson's Correlation Coefficient.

2. Pearson's Correlation Coefficient was utilised to analyse correlation between EEG and HRV in frequency domain. It would be interesting to extend the correlation analysis using other methods such as, Mutual Information, Correlation Dimensions, Autocorrelation, and Spearman's Correlation.
3. Use other entropy methods, such as Shannon Entropy and SE-based Multi scale Entropy, to assess variability of EEG and HRV, and compare the results of these methods with the new SE calculation method proposed in this thesis.

# Appendix A

## Dataset 1- Recording Photos

After completing my training with Dr Tony Steffert, I started recording EEG and ECG data. In order to make the recording smoothly, I followed the steps below:

1. Position Electroencephalography (EEG) and Electrocardiogram (ECG) equipment on a table: (1) Gel jar (a jar of 10/20 NuPrep), stick, alcohol wipes, cotton wool ball, two sponges, Harness belt for holding EEG cap tight. (2) ECG cables (Red, Black, and White (ground) electrodes), Respiration belt, Blood Volume Pulse (BVP) cable, Temperature cable.
2. Equipment checks: NeXus10 batteries (checked/replaced after every 3 sessions), NeXus10 / BioTrace (plug all sensors leads into NeXus10 except for Respiration belt) function, Mitsar/WinEEG function.
3. Position harness (standard size ), Respiration belt (tight, c level of the umbilicus); strain gauge just to one side of midline (note where the cable comes off for NeXus10).
4. Position cushion on a lap, and clean participants' wrists with an alcohol wipe (P6 acupoint, or medial to tendons; avoid tendons); put swab back in its sleeve for later use.
5. Position temp sensor and BVP plethysmograph on the nondominant middle finger; the temperature to socket F; make participants' hands positioned palms down.
6. Place NeXus on Mitsar storage box beneath participant chair and start NeXus and BioTrace to check to record before capping up.
7. For EEG Capping: (1) Measure head circumference and check the size on cap label, (2) Position 2 rings around Fp1/Fp2 electrodes. Approach capping up from behind/right. Ensure nasion to level of Fpz and inion to level Oz is 10% of nasion

- toinion distance, using a tape measure, (3) Then gel ear clip cup and clip onto the earlobe (with a cotton wool ball between clip and neck), pull cap ring down and attach to harness. Remember to cross attachments to harness, (4) Attach D plug to Mitsar socket, and ear leads to A1 and A2 (keeping L and R parallel, not crossed over). 9. Tie leads together just behind the cap, keeping them loose. (5) Check cap symmetrical (sideways) and correctly positioned back/front.
8. Explain black vs. white or yellow (more important is that all should be the same colour rather than all white) using impedance screen. Note that Fp1 and Fp2 are the ‘most tickly’ (use that word, to emphasise not invasive). Once a participant is looking at the traces on screen, need to balance saying too much about eye-blink and to relax jaw with them relaxing, can talk about smiley muscles and frowny muscles.
  9. Start with Ground (behind Fz). Need to ensure ground contact has a good connection if there is no white at all on impedance screen. If there is some white, the ground is okay. Ears should be white (this is must, as reference to ears).
  10. Use ‘transport tape’ (non-tacky, 3M) to hold specs to be outside of the cap, and check for any cap slippage.
  11. Ensure participant is comfortable – supply back cushion if required.
  12. EEG monitoring; (a) Start EEG monitoring immediately, while organising BioTrace, then start BioTrace recording a little earlier than WinEEG recording. Stop (WinEEG) and Pause (BioTrace) simultaneously. (b) Use ‘Stop’ button between slots; if need to re-check impedance screen (Stop, check, and then restart with a new recording/new label, so best not to have to do this). If a participant needs to sneeze or move, use the ‘Pause’ button, and then push Green to continue. (c) At the end of the recording, ‘Save As’ with the label copied and pasted into three fields in the ‘Participant Card’ already (participant ID, Date of Birth and gender, but NOT an identifiable name).
  13. BioTrace : (1) use the ‘Pause’ button between slots. When the session ends, ‘Stop’ and save. (2) Export IBI into the export folder. Ensure codes for slots identical for BioTrace and WinEEG, copying and pasting as necessary between slots (WinEEG) and sessions (BioTrace).
  14. When Finished: (1) Switch off NeXus10 gently (just holding down Off button), (2) Unplug EEG leads from Mitsar first, then undo harness; take ear clips off; remove cap with tissue in hand (so that participant doesn’t get to rub a hand over the head

before you do); check how much gel there is or not. (3) Use wet wipe on scalp/ears, then offer it to a participant to complete cleansing. (4) Remove all NeXus10 sensors carefully and position safely. (5) soak ECG electrodes thoroughly.

15. Cap maintenance: (1) Gently wrap cap, then at the home immerse cap (NOT D plug!) in water, wash the sponges out; keep participant's stick to go through the holes while washing, (2) Drip dry over bath, taking care with where leads are joined to electrodes, (3) Take care when cleaning ear clip electrodes – only tin, so don't want to wear surface away. (Don't leave too long for the gel to dry out), (4) Use alcohol wipes on a syringe, wash and rotate the needle (OR: arrange drying facility in the storage room). Cap should have dried by the time the next, but one person comes through. Use a hairdryer (not too close to the cap!) if necessary.

In this section, the process of EEG and ECG recording is shown through some photos taken during the recording process. Participants shown in the pictures have given their permission to share their pictures in my research.



Figure A.1: Location of the Recording at the University of Hertfordshire Premises.



Figure A.2: Device used to control Transcutaneous Electroacupuncture Stimulation (TEAS) Stimulation Frequency.

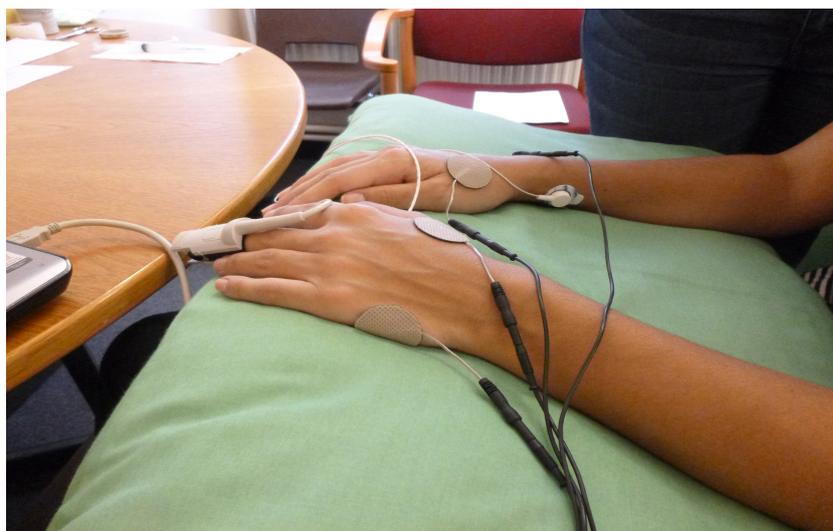


Figure A.3: ECG electrodes attached on both wrists, TEA electrodes attached on both hands, and sensor attached on the indexed finger.

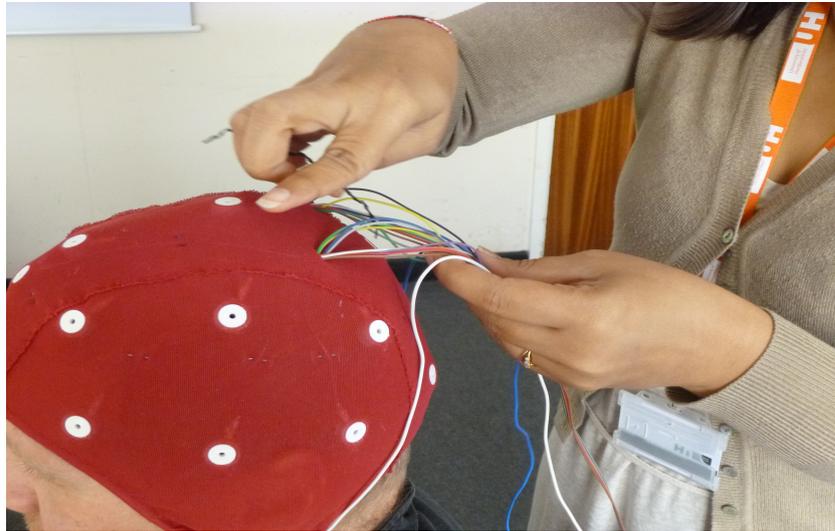


Figure A.4: Tie leads together just behind cap, keeping them loose.



Figure A.5: Explaining one of the participants the next step of getting a good connection between the surface of the scalp and EEG electrodes. Explaining (black = no connection vs white = very good connection or yellow = good connection colour) what colours represented on the screen in front of them.



2014.07.25

Figure A.6: Adjustment between the gel and the hair is needed to make the connection between the electrodes and the surface of the scalp. Therefore, rubbing the gel, and checking the connection on the laptop screen (This process can be seen by participants also).

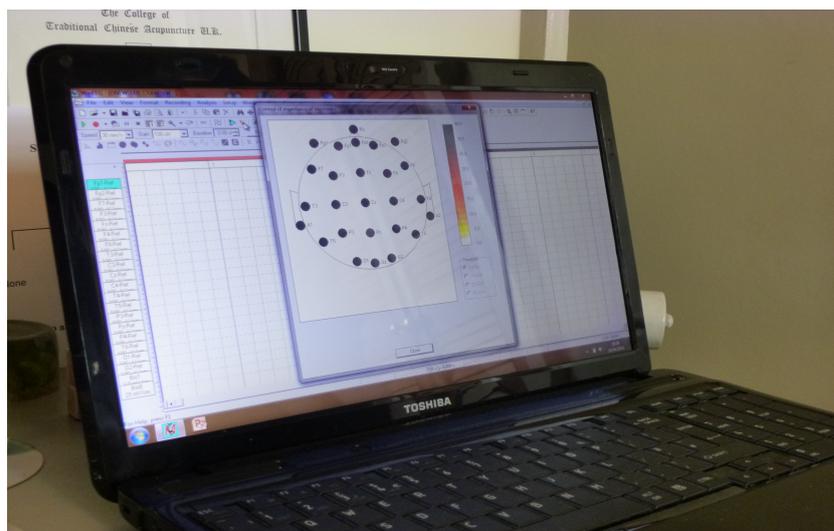


Figure A.7: Impedance screen showing all 21 electrodes before them, being connected on the surface of the scalp (Black colour = No connection).

# Appendix B

## Cross-correlation performance of Electrodes Pair against Time lags and Distance

### B.1 Cross-correlation performance of Electrodes Pair against Time Lags

In all electrode pairs from Dataset 1, and Dataset 4-8, the result is similar to what is shown in Chapter 6 (Figure 6.2). Therefore, in this section, I have chosen only two datasets (Dataset 1 and Dataset 4) randomly, to show Cross-correlation performance for one electrode from each lobe of the brain from. For example, the chosen electrodes from Dataset 1 are: Fp2 (from Front polar-right side of the brain), T3 (from Temporal-left side of the brain), T4 (from Temporal-Right lobe of the brain), O1 (from Occipital-left side of the brain) and O2 (from Occipital-right side of the brain). The chosen electrodes from Dataset 4 are: F8 (from Frontal-right side of the brain), P3 (from Parietal-left side of the brain), and P4 (from Parietal-Right lobe of the brain).

#### B.1.1 Dataset 1

The result in the figures B.1-B.5 suggests that all signals measured from the particular lobe of the brain are positively correlated with the signals measured within the same lobe of the brain, but negatively correlated with signals measured from another lobe of the brain. For example, Figure B.1 shows the information for all electrode pairs for the electrode Fp2, where all signals from the frontal lobe of the brain are positively correlated with Fp2, and all signals from the back lobe are negatively correlated.

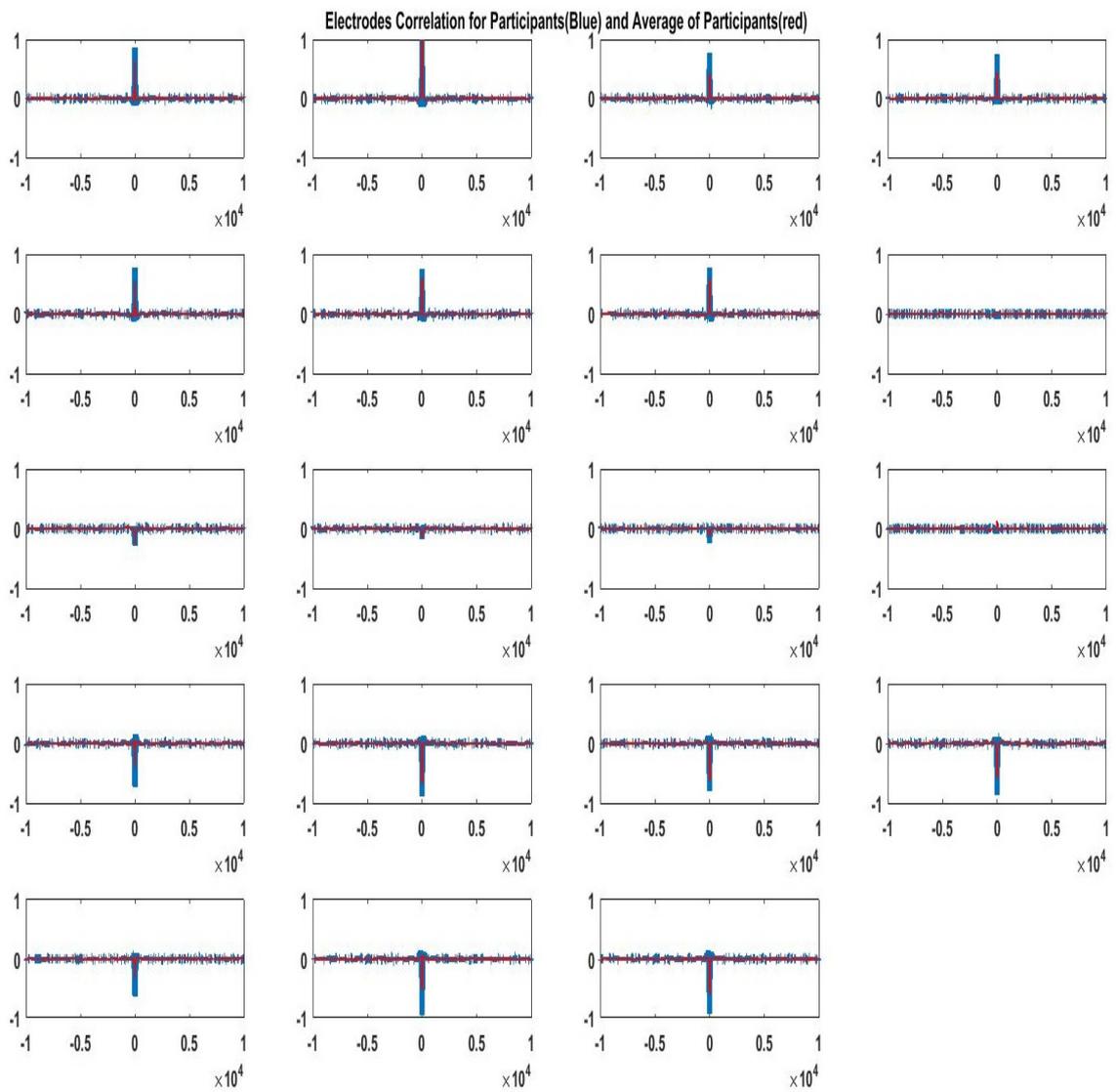


Figure B.1: Cross-correlation at all possible Lags for all electrode pairs for Electrode Fp2. Blue color shows the result for one of the participant, and Orange color shows the average of 15 participants result from Dataset 1

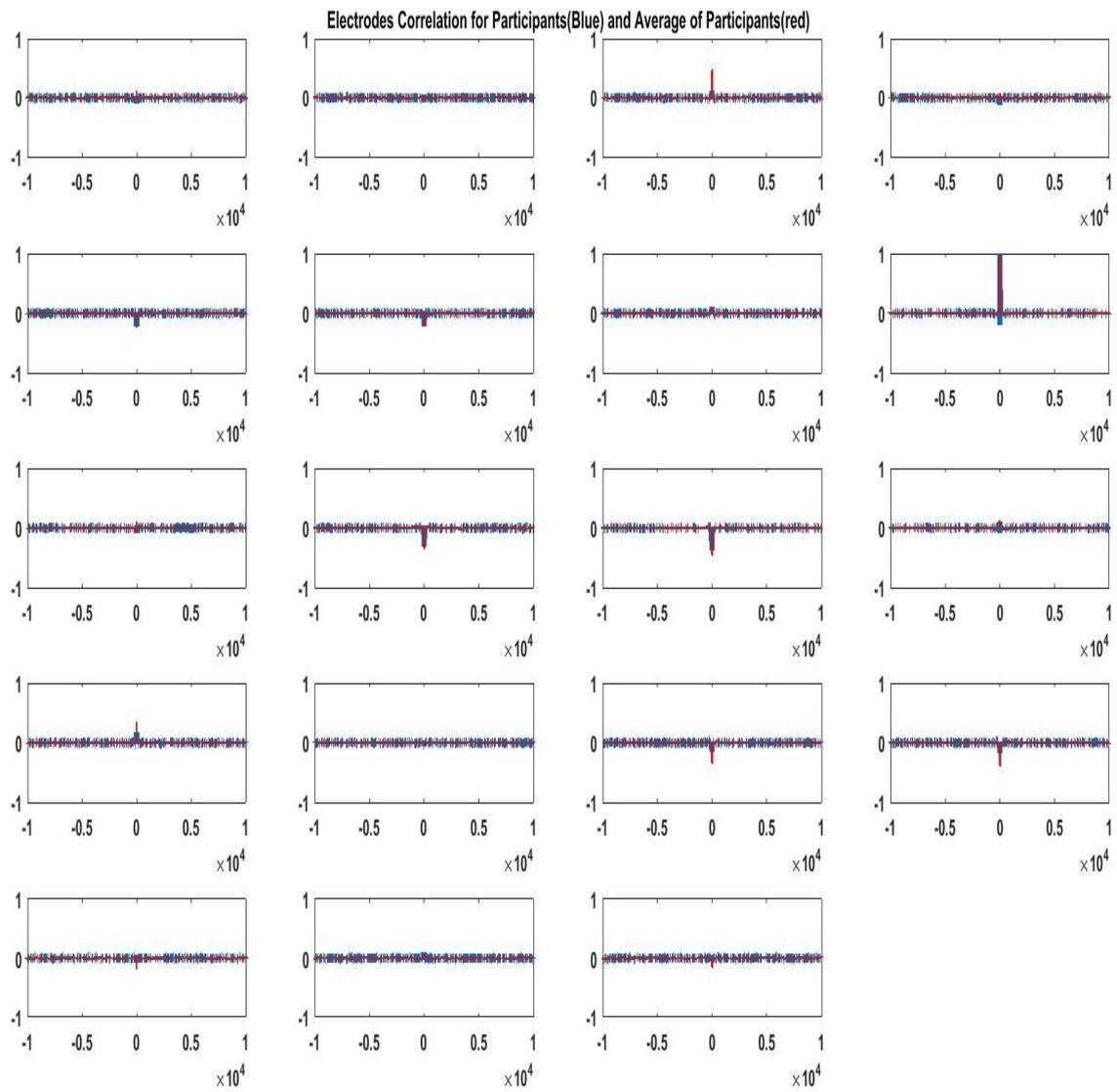


Figure B.2: Cross-Correlation at all possible Lags for all electrode pairs for Electrode T3. Blue color shows the result for one of the participant, and Orange color shows the average of 15 participants result from Dataset 1

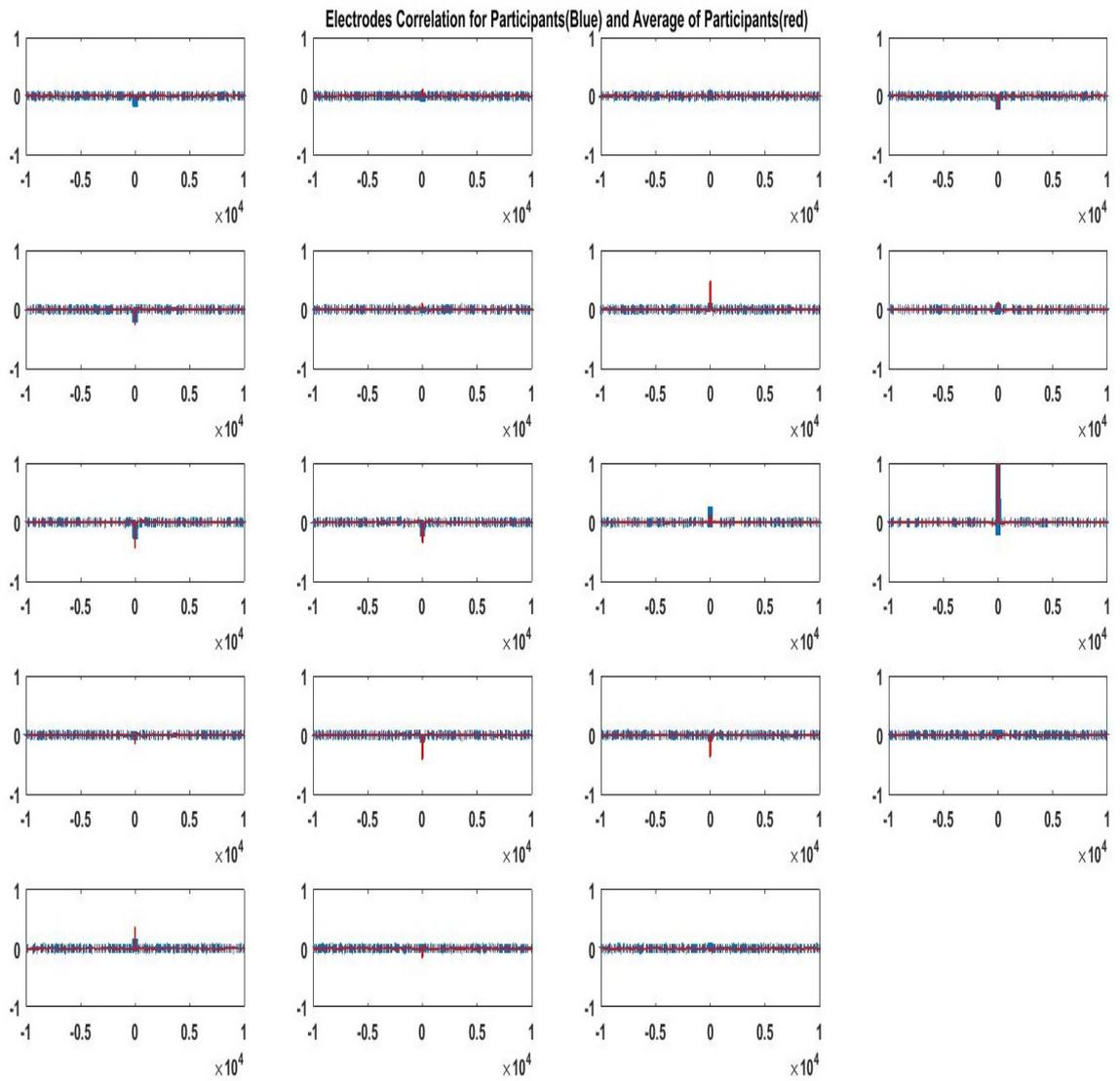


Figure B.3: Cross-Correlation at all possible Lags for all electrode pairs for Electrode T4. Blue color shows the result for one of the participant, and Orange color shows the average of 15 participants result from Dataset 1

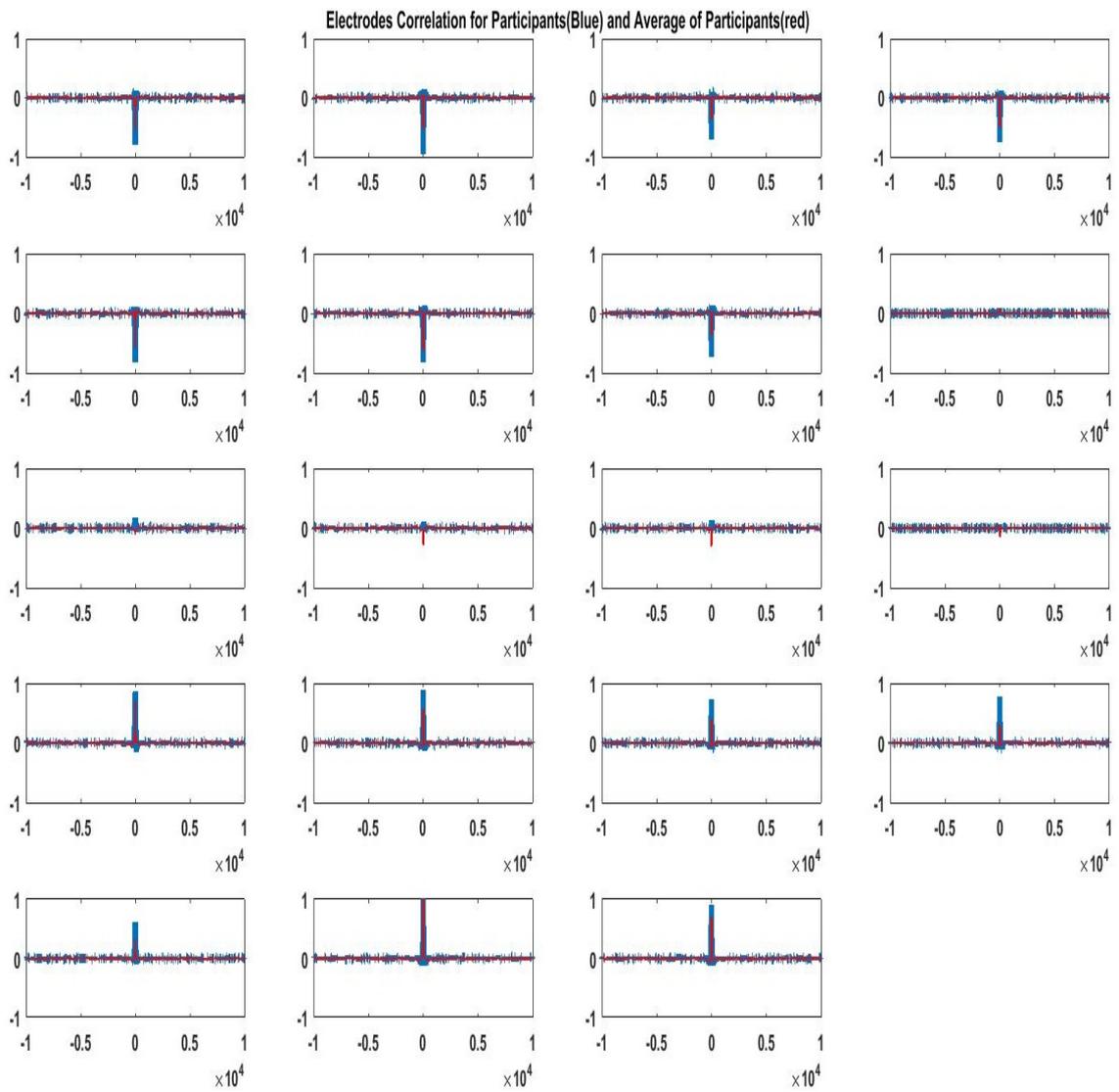


Figure B.4: Cross-Correlation at all possible Lags for all electrode pairs for Electrode O1. Blue color shows the result for one of the participant, and Orange color shows the average of 15 participants result from Dataset 1

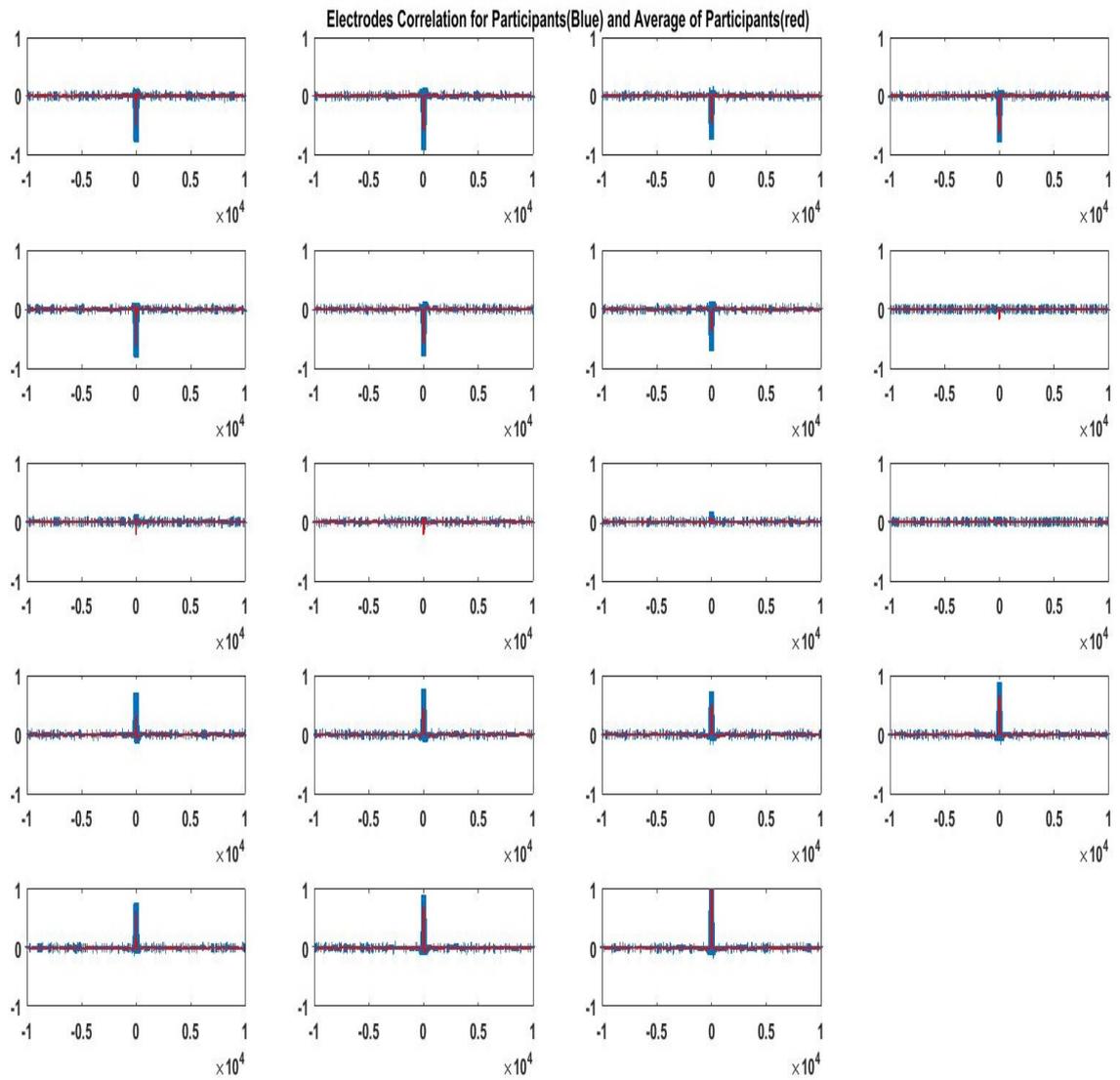


Figure B.5: Cross-Correlation at all possible Lags for all electrode pairs for Electrode O2. Blue color shows the result for one of the participant, and Orange color shows the average of 15 participants result from Dataset 1

### **B.1.2 Dataset 4**

The result in the figures B.6-B.8 suggests that all signals measured from the particular lobe of the brain are positively correlated with the signals measured within the same lobe of the brain, but negatively correlated with signals measured from another lobe of the brain. For example, Figure B.6 shows the information for all electrodes pairs for the electrode F8, where all signals from frontal lobe (Top five plots) of the brain are showing a strong positive correlation with F8, and all signals from the back lobe (Bottom five plots) are not showing such as strong connections.

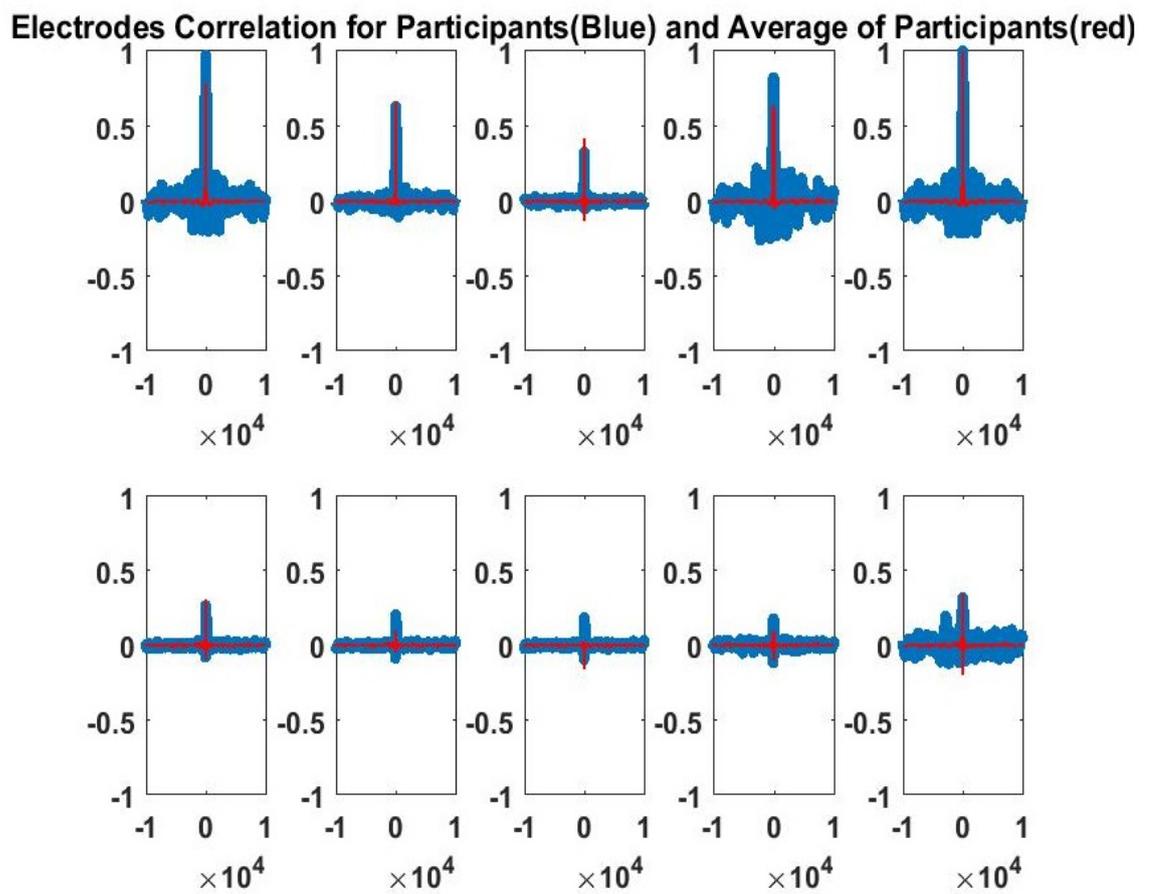


Figure B.6: Cross-Correlation at all possible Lags for all electrode pairs for Electrode F8. Blue color shows the result for one of the participant, and Orange color shows the average of 15 participants result from Dataset 4

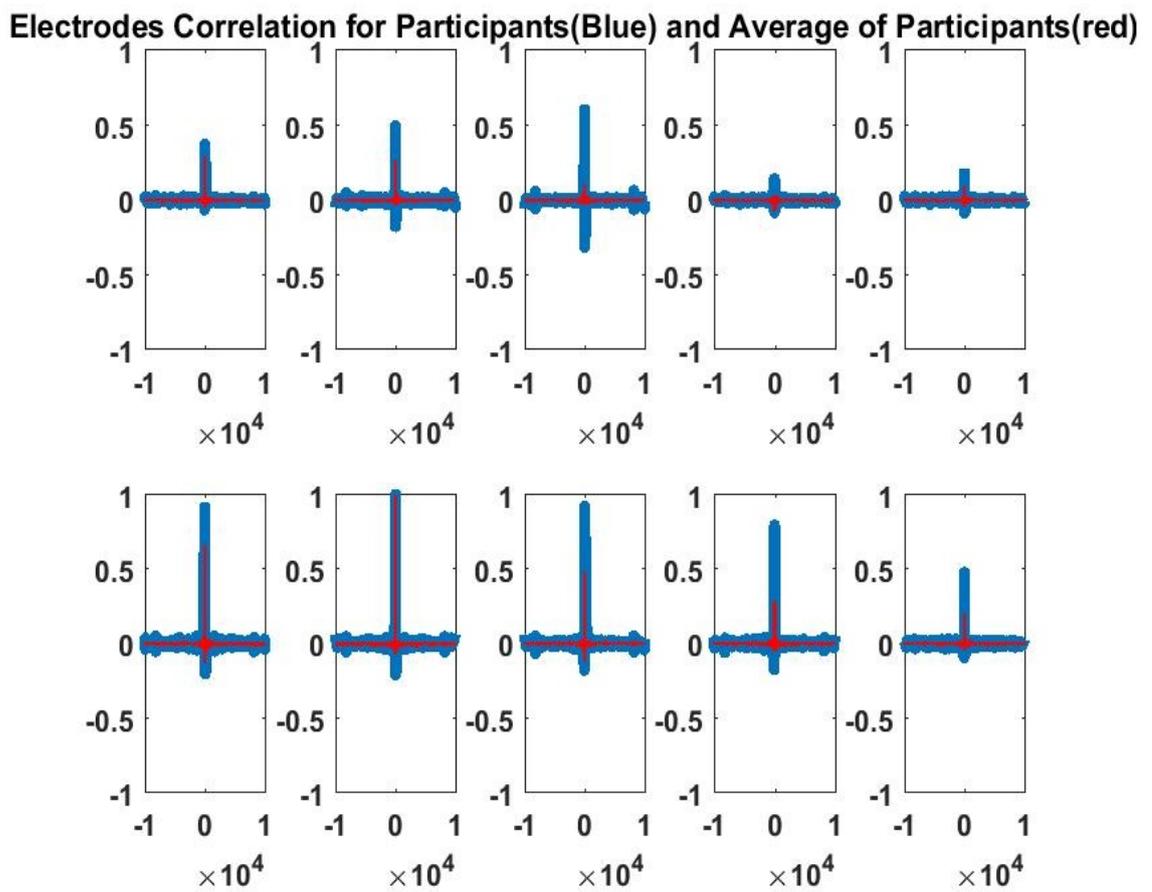


Figure B.7: Cross-Correlation at all possible Lags for all electrode pairs for Electrode P3. Blue color shows the result for one of the participant, and Orange color shows the average of 15 participants result from Dataset 4

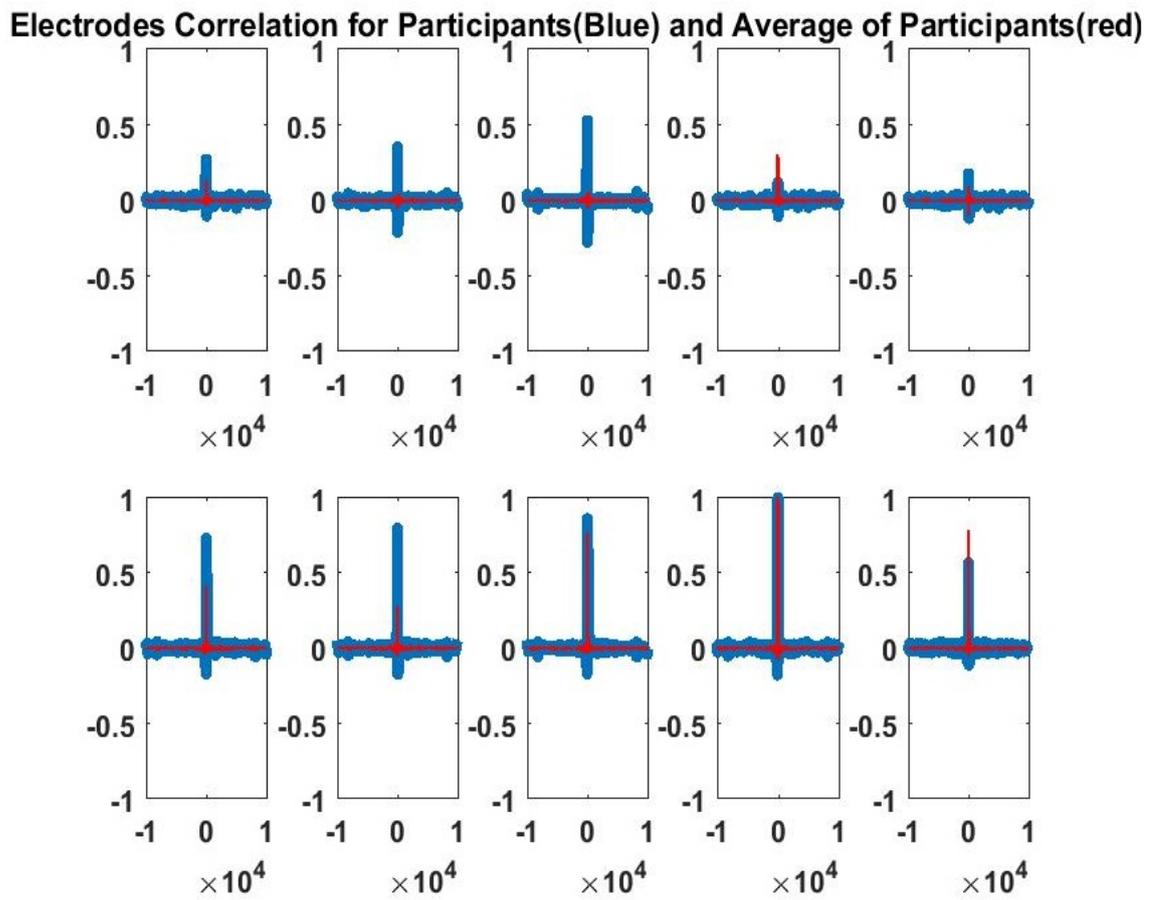


Figure B.8: Cross-Correlation at all possible Lags for all electrode pairs for Electrode P4. Blue color shows the result for one of the participant, and Orange color shows the average of 15 participants result from Dataset 4

## **B.2 Cross-Correlation performance of Electrodes Pair against Distance between them**

Cross-Correlation performance of Electrodes Pair against Distance between two electrodes shown in the following figures is a straight line distance between two electrodes, not the distance as measured over the surface of the scalp. In Chapter 6, the results for Electrode Fp1 and F7 are shown (Figure 6.5-6.10) from respective datasets. In this section result of all other electrodes Cross-Correlation performance against distance is shown.

### **B.2.1 Dataset 1**

For this dataset, Cross-Correlation performance between electrodes at varying distance for electrode Fp1 is shown in Chapter 6 (Figure 6.5). Therefore, in this section Cross-Correlation performance for remaining 18 electrodes are shown.

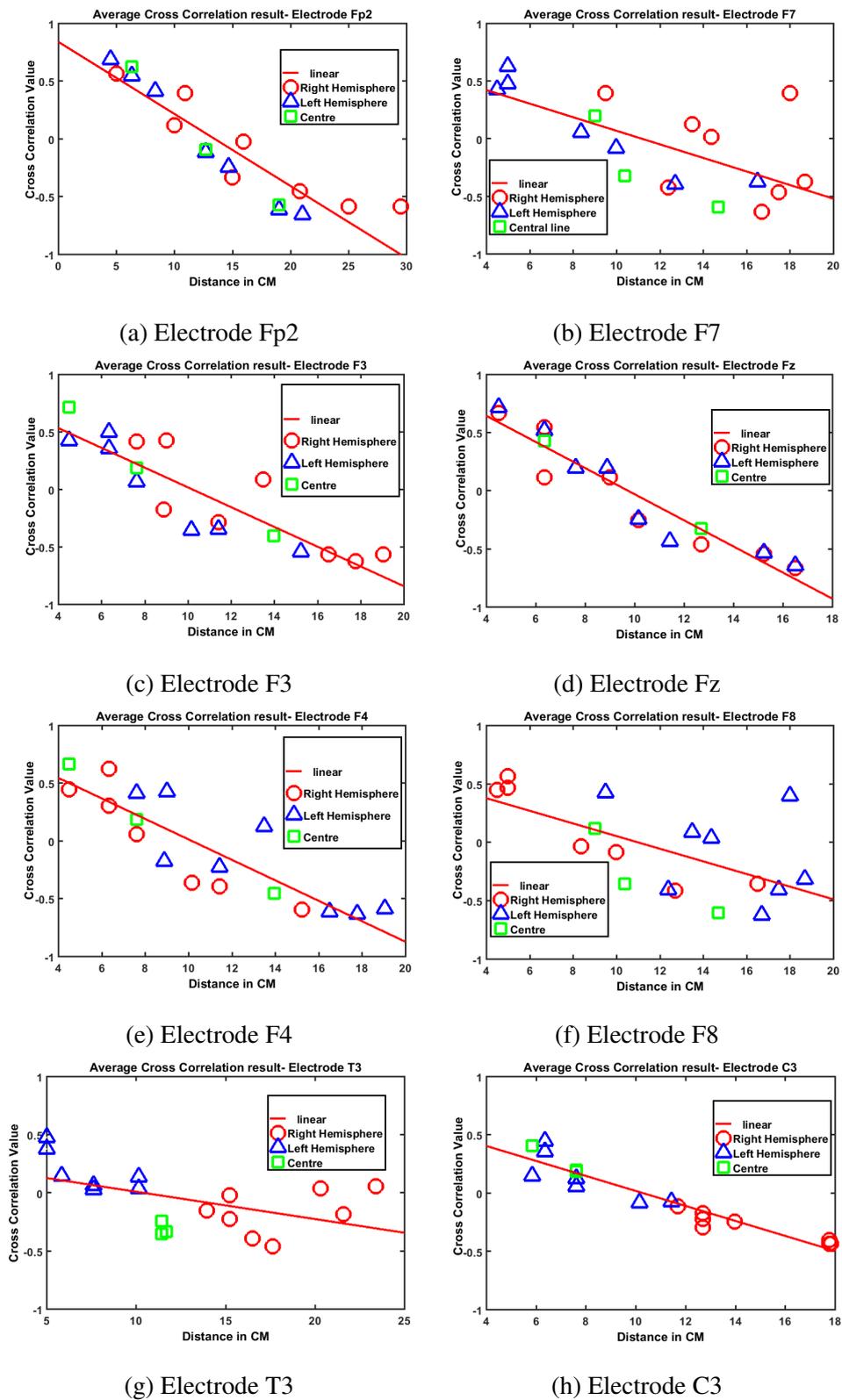


Figure B.9: Cross-Correlation between electrodes at varying distance on Dataset 1 for: (a) Electrode Fp2, (b) Electrode F7, (c) Electrode F3, (d) Electrode Fz, (e) Electrode F4, (f) Electrode F8, (g) Electrode T3, and (h) Electrode C3

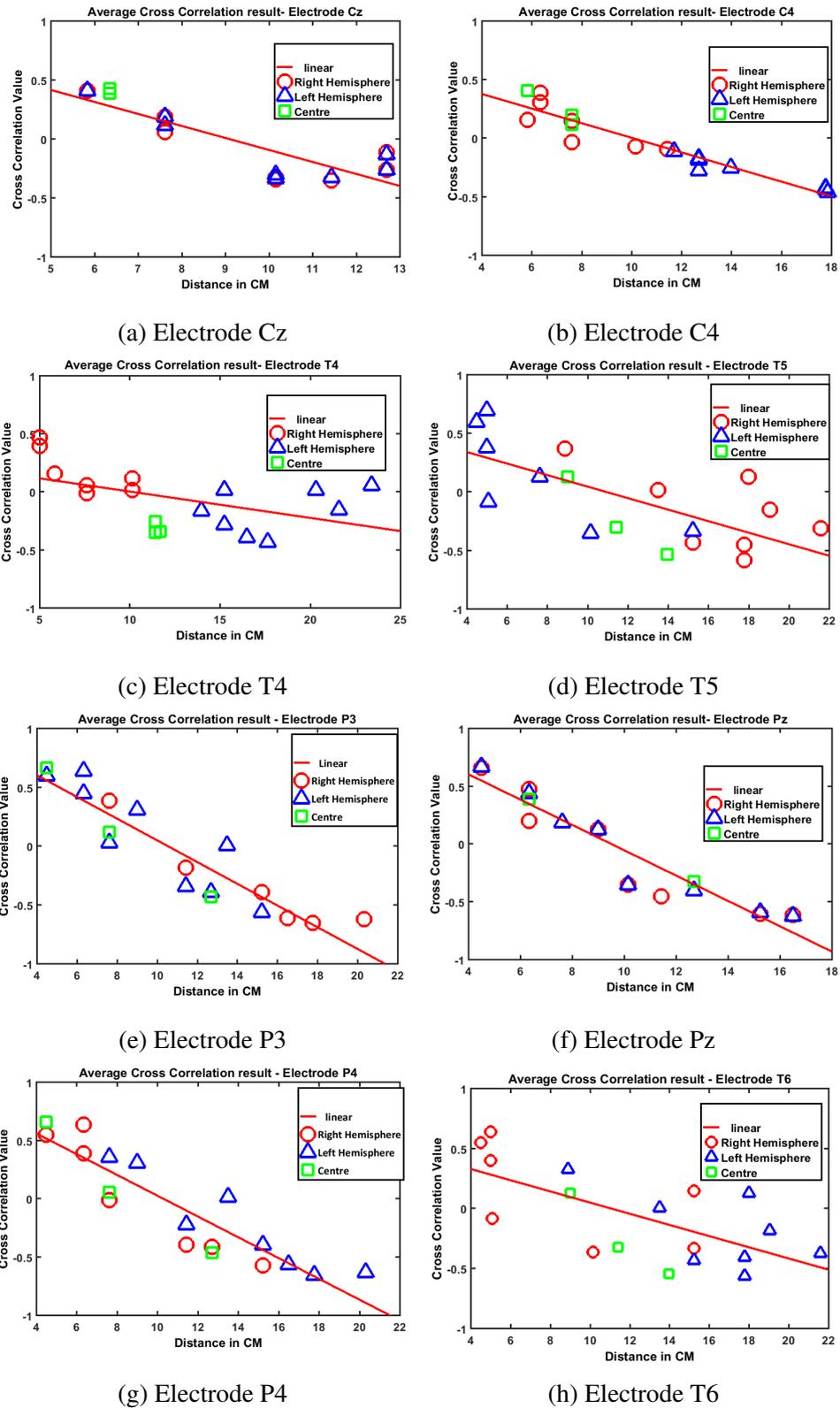
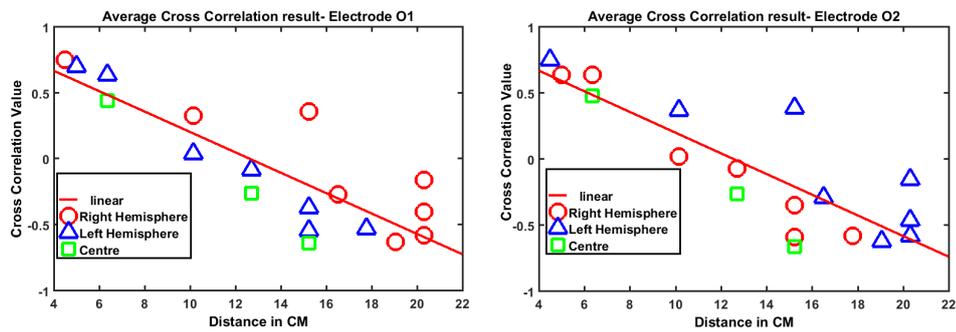


Figure B.10: Cross-Correlation between electrodes at varying distance on Dataset 1 for: (a) Electrode Cz, (b) Electrode C4, (c) Electrode T4, (d) Electrode T5, (e) Electrode P3, (f) Electrode Pz, (g) Electrode P4, and (h) Electrode T6



(a) Electrode O1

(b) Electrode O2

Figure B.11: Cross-Correlation between electrodes at varying distance on Dataset 1 for: (a) Electrode O1, and (b) Electrode O2.

## B.2.2 Dataset 4

For this dataset, Cross-Correlation performance between electrodes at varying distance for electrode F7 is shown in Chapter 6 (Figure 6.6). Therefore, in this section Cross-Correlation performance for the remaining 9 electrodes are shown.

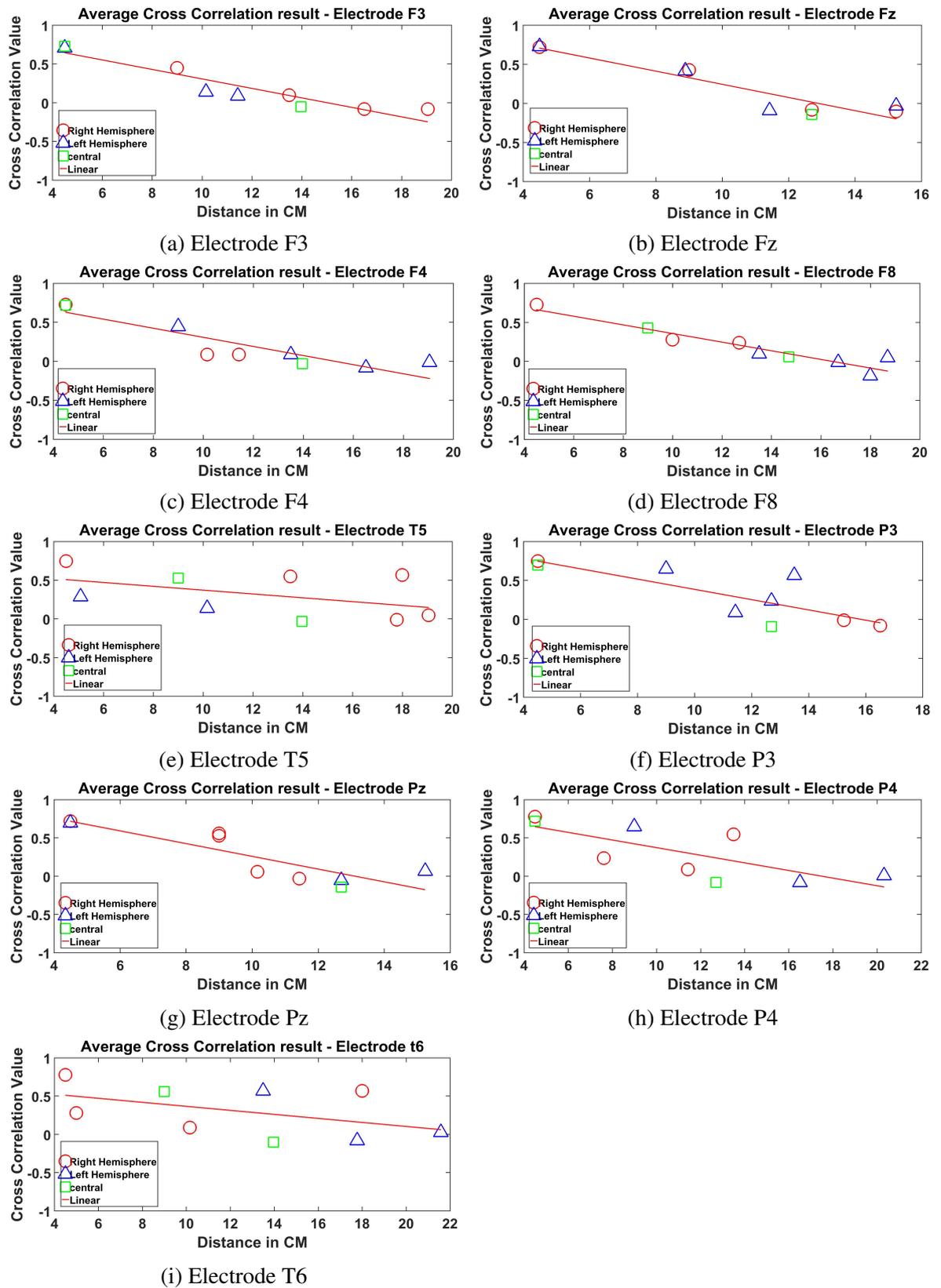


Figure B.12: Cross-Correlation between electrodes at varying distance on Dataset 4 for: (a) Electrode F3, (b) Electrode Fz, (c) Electrode F4, (d) Electrode F8, (e) Electrode T5, (f) Electrode P3, (g) Electrode Pz, (h) Electrode P4, and (i) Electrode T6

### **B.2.3 Dataset 5**

For this dataset, Cross-Correlation performance between electrodes at a varying distance for electrode Fp1 is shown in Chapter 6 (Figure 6.7). Therefore, in this section Cross-Correlation performance for the remaining 14 electrodes are shown.

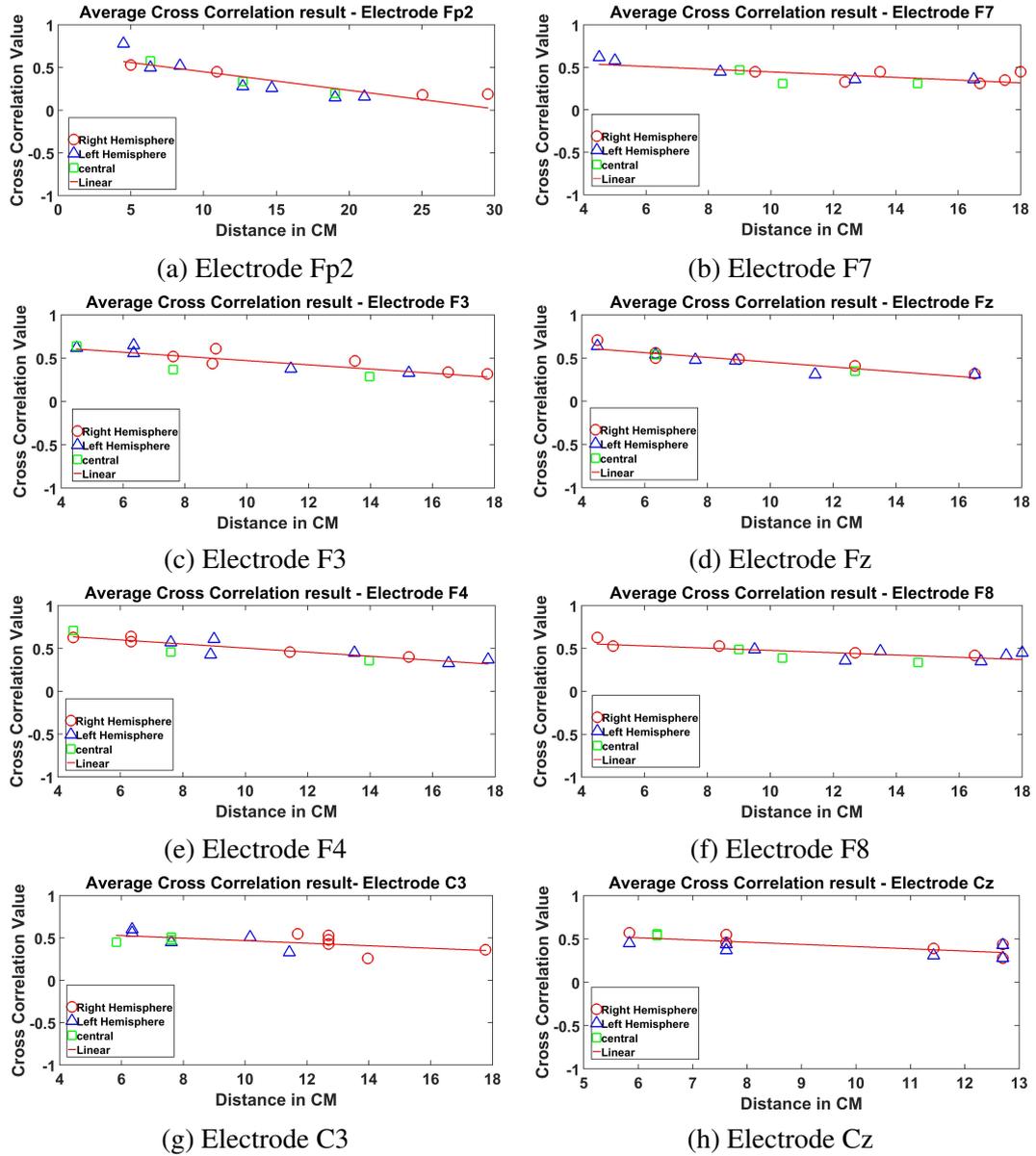


Figure B.13: Cross-Correlation between electrodes at varying distance on Dataset 5 for: (a) Electrode Fp2, (b) Electrode F7, (c) Electrode F3, (d) Electrode Fz, (e) Electrode F4, (f) Electrode F8, (g) Electrode C3, and (h) Electrode Cz.

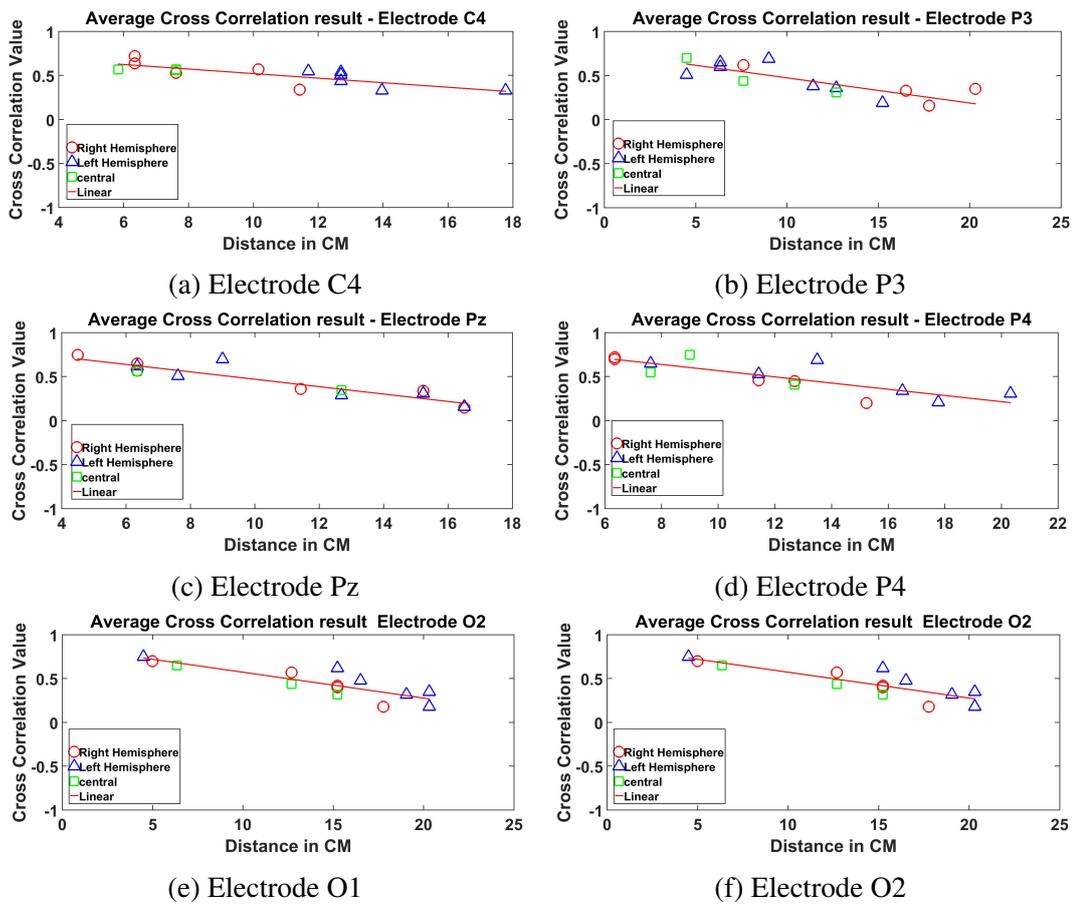


Figure B.14: Cross-Correlation between electrodes at varying distance on Dataset 5 for: (a) Electrode C4, (b) Electrode P3, (c) Electrode Pz, (d) Electrode P4, (e) Electrode O1, and (f) Electrode O2.

## **B.2.4 Dataset 6**

For this dataset, Cross-Correlation performance between electrodes at a varying distance for electrode Fp1 is shown in Chapter 6 (Figure 6.8). Therefore, in this section Cross-Correlation performance for the remaining 18 electrodes are shown.

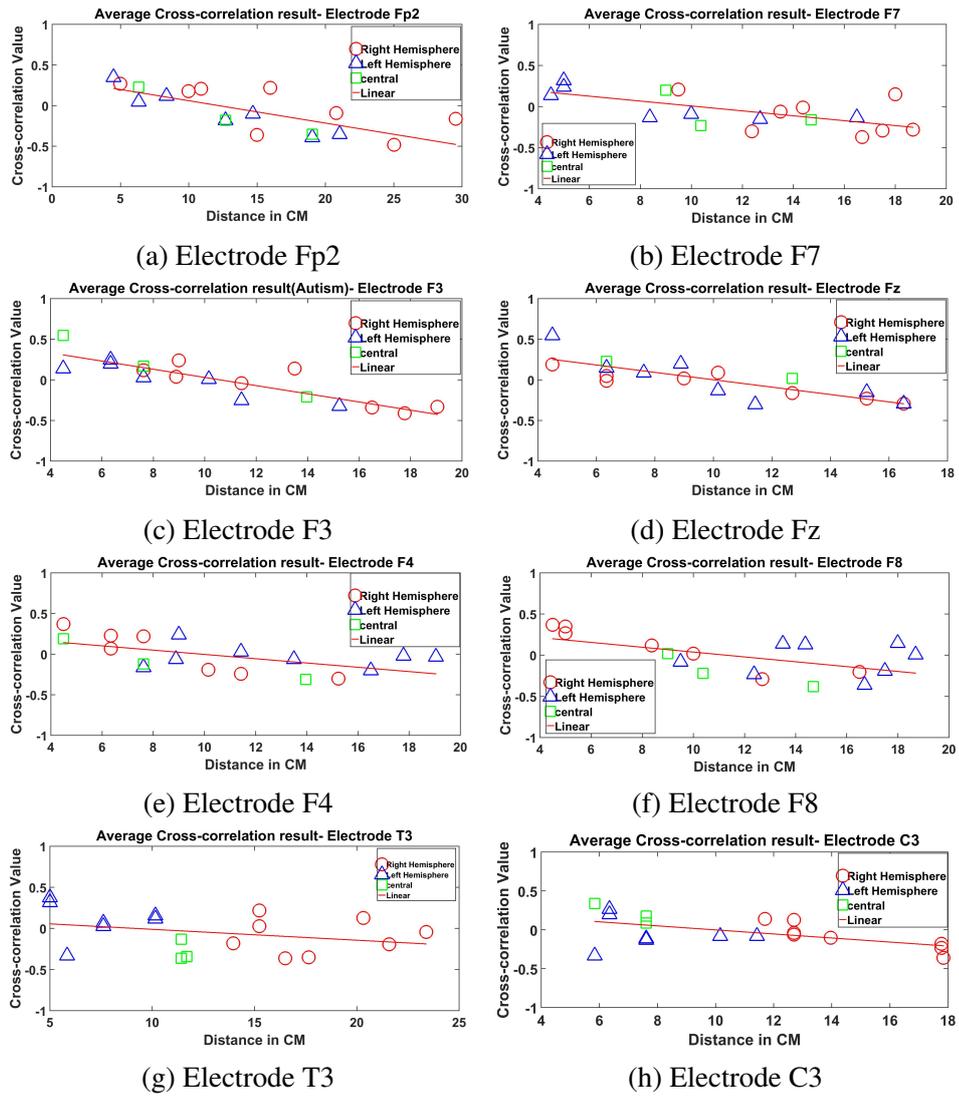


Figure B.15: Cross-Correlation between electrodes at varying distance on Dataset 6 for: (a) Electrode Fp2, (b) Electrode F7, (c) Electrode F3, (d) Electrode Fz, (e) Electrode F4, (f) Electrode F8, (g) Electrode T3, and (h) Electrode C3

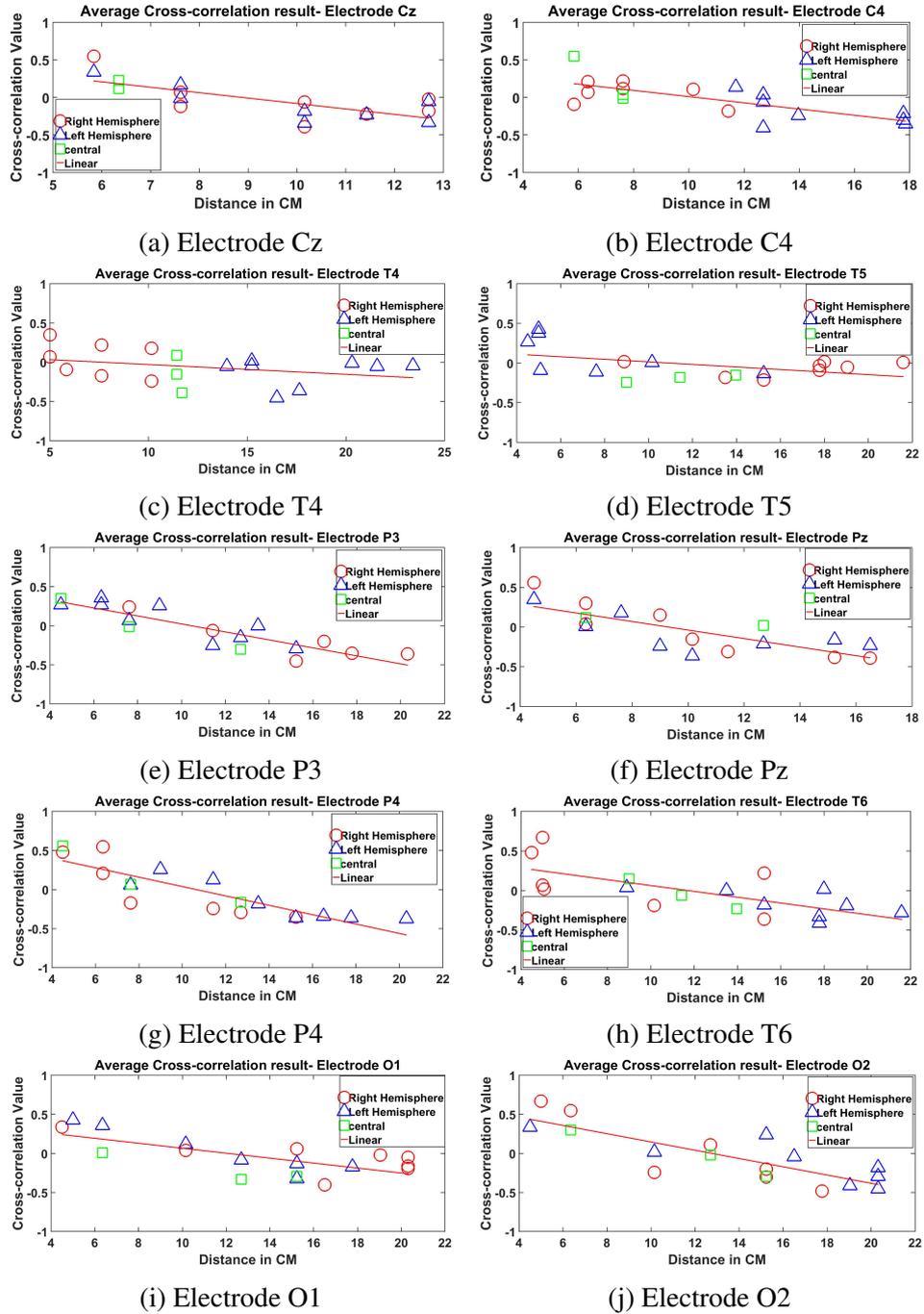


Figure B.16: Cross-Correlation between electrodes at varying distance on Dataset 6 for: (a) Electrode Cz, (b) Electrode C4, (c) Electrode T4, (d) Electrode T5, (e) Electrode P3, (f) Electrode Pz, (g) Electrode P4, (h) Electrode T6, (i) Electrode O1, and (j) Electrode O.

### **B.2.5 Dataset 7**

For this dataset, Cross-Correlation performance between electrodes at a varying distance for electrode Fp1 is shown in Chapter 6 (Figure 6.9). Therefore, in this section Cross-Correlation performance for the remaining 18 electrodes are shown.

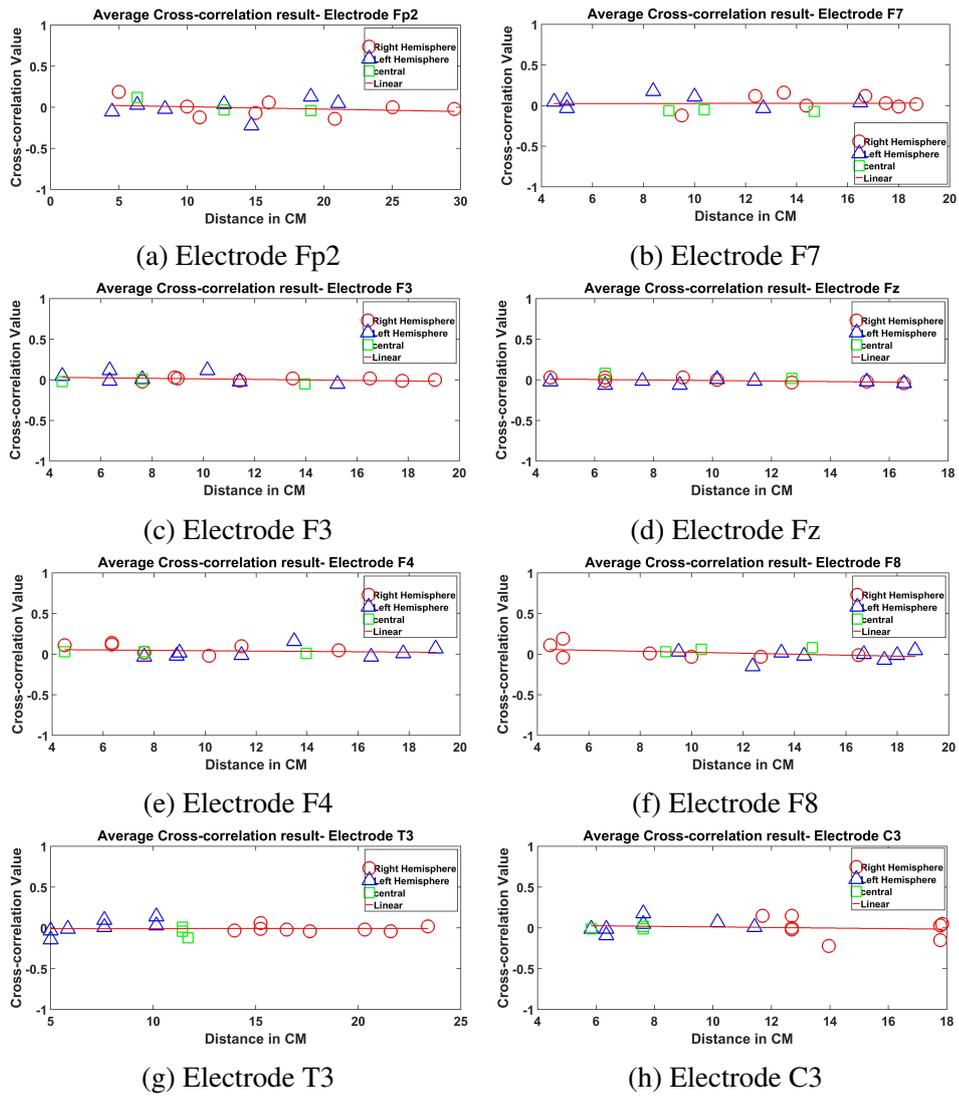


Figure B.17: Cross-Correlation between electrodes at varying distance on Dataset 7 for: (a) Electrode Fp2, (b) Electrode F7, (c) Electrode F3, (d) Electrode Fz, (e) Electrode F4, (f) Electrode F8, (g) Electrode T3, and (h) Electrode C3

## **B.2.6 Dataset 8**

For this dataset, Cross-Correlation performance between electrodes at a varying distance for electrode Fp1 is shown in Chapter 6 (Figure 6.10). Therefore, in this section Cross-Correlation performance for the remaining 18 electrodes are shown.

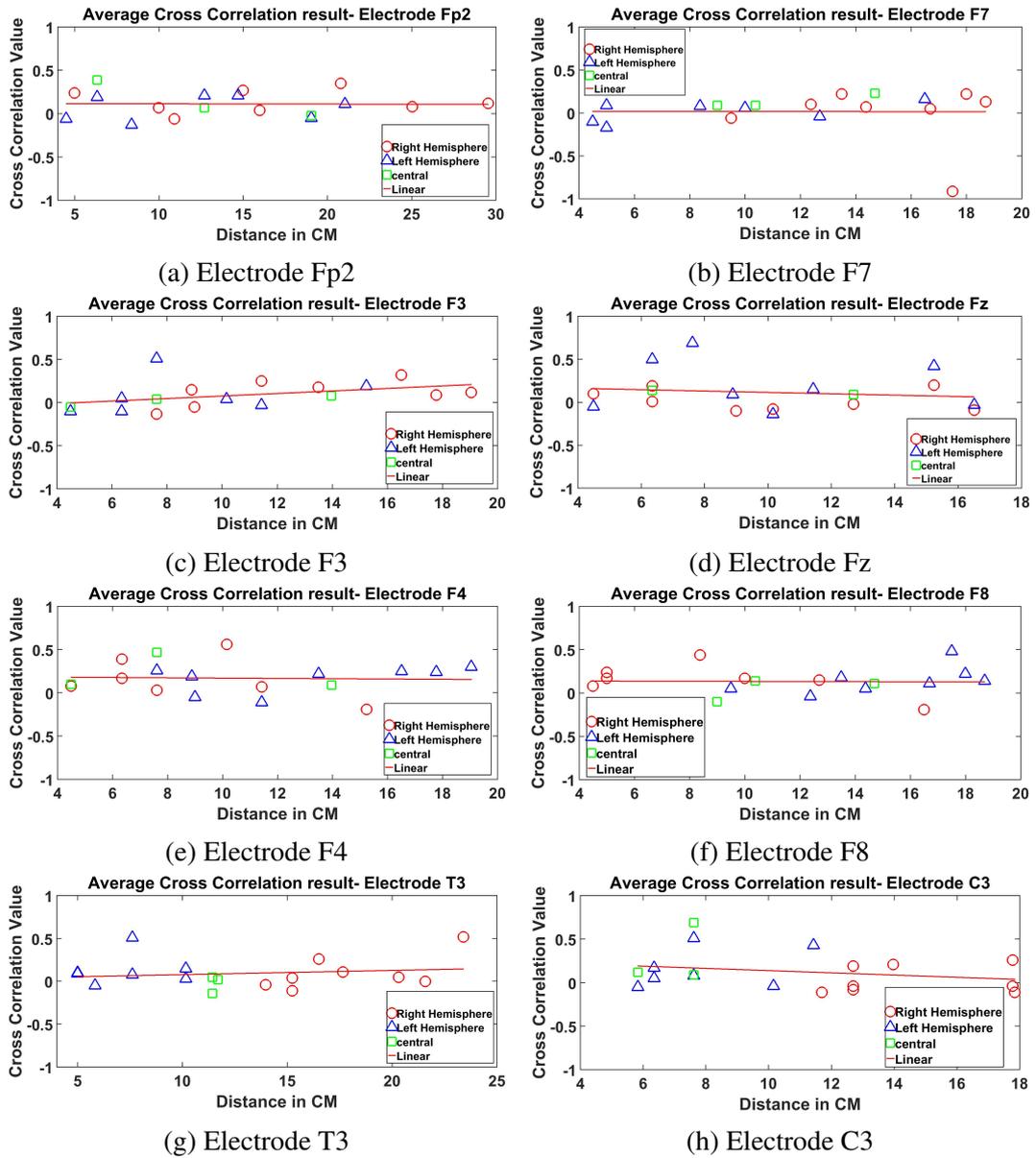


Figure B.18: Cross-Correlation between electrodes at varying distance on Dataset 8 for: (a) Electrode Fp2, (b) Electrode F7, (c) Electrode F3, (d) Electrode Fz, (e) Electrode F4, (f) Electrode F8, (g) Electrode T3, and (h) Electrode C3

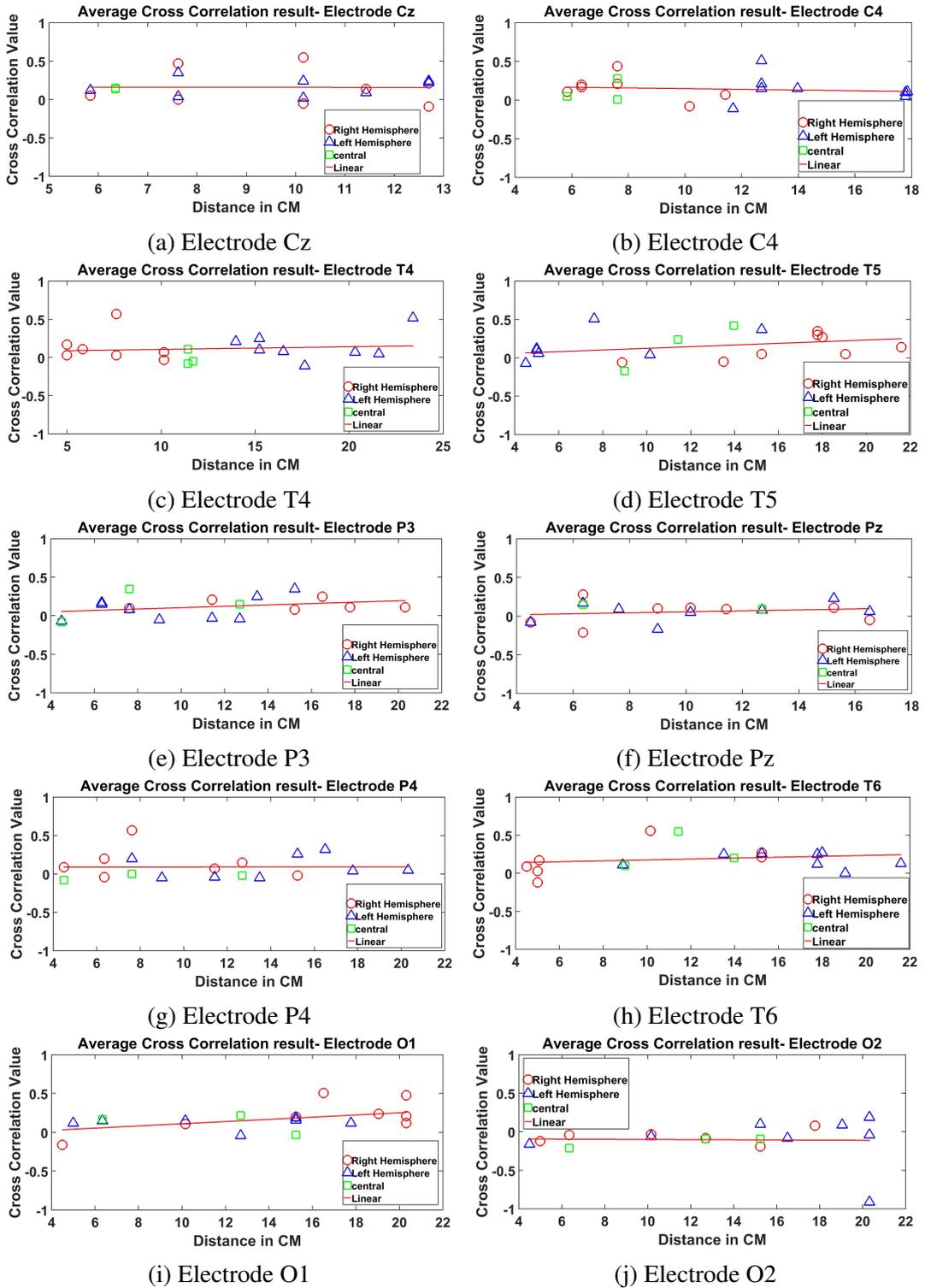


Figure B.19: Cross-Correlation between electrodes at varying distance on Dataset 8 for: (a) Electrode Cz, (b) Electrode C4, (c) Electrode T4, (d) Electrode T5, (e) Electrode P3, (f) Electrode Pz, (g) Electrode P4, (h) Electrode T6, (i) Electrode O1, and (j) Electrode O.

# Appendix C

## Additional SE performance of SE-Method 2 and SE-Method 3

In this section, SE performance with SE-Method 2 and SE-Method 3 is shown. In order to show the SE results, I have selected two 5-minute slots from Dataset 1. In Chapter 8, I have shown SE performance only on one 5-minute slot (Slot 1) for each SE-Methods. In this section, I am showing the SE performance of the remaining window sizes of SE-Method 2 and SE-Method 3 for Slot 1, which are not shown in Chapter 8. In addition, I am also showing SE performance for SE-Method 1, SE-Method 2, and SE-Method 3 for 5-minute data from Slot 5. The results of all other slots are similar to the one shown here.

### C.1 Slot 1 Result

In Chapter 8, the result of the SE-Method 1, SE-Method 2 (2 Seconds Moving Window), and SE-Method 3 (Mean of Each 2 Seconds Window) is shown. In this section, SE performance for the remaining windows of SE-Method 2 and SE-Method 3 are shown.

#### C.1.1 SE-Method 2

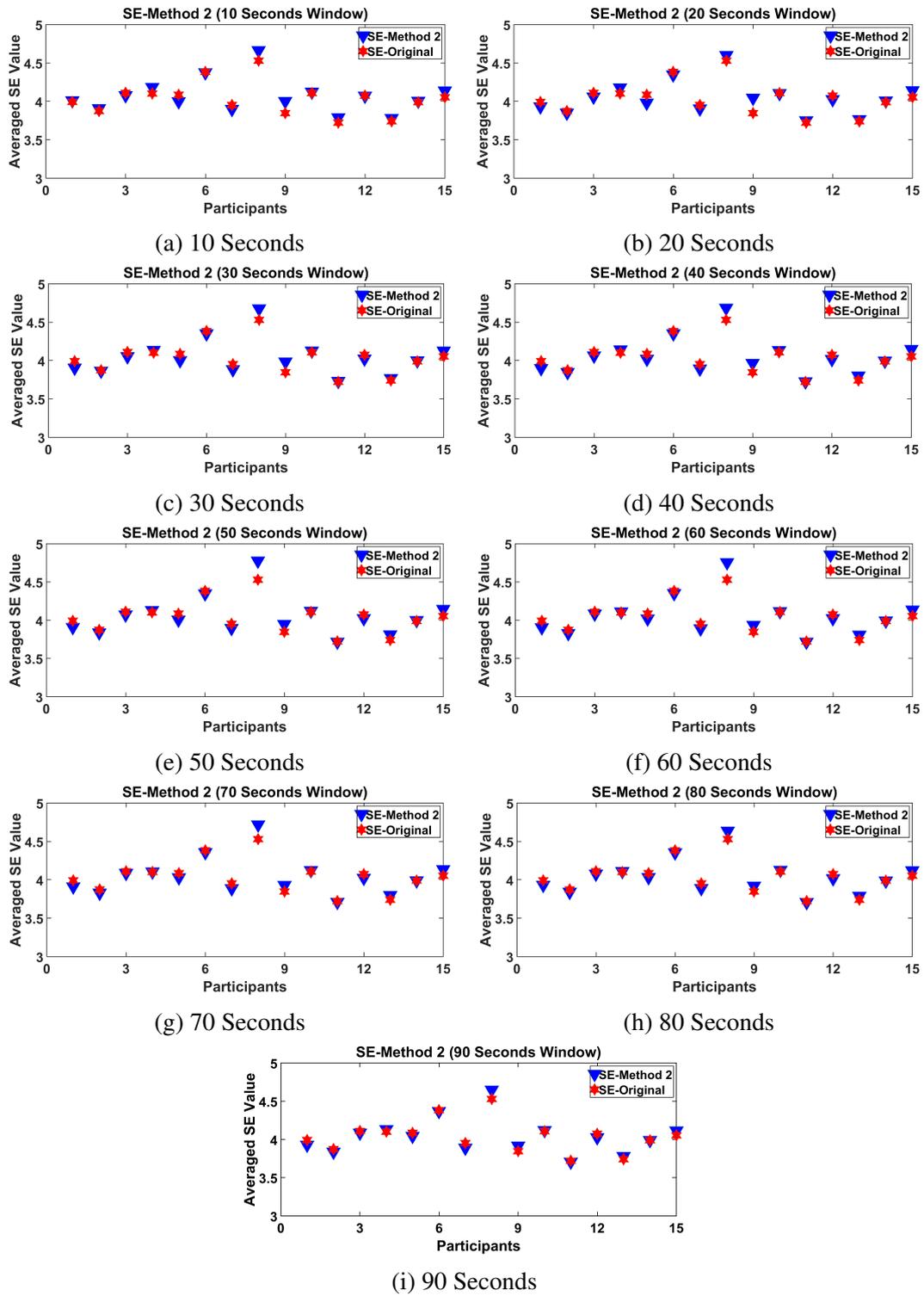


Figure C.1: SE-Method 2 - Comparison of original SE calculation and proposed SE methods, showing average results for all electrodes are shown for the individual participants: SE-Method 2 on moving windows for: (a) 10 Seconds, (b) 20 Seconds, (c) 30 Seconds, (d) 40 Seconds, (e) 50 Seconds, (f) 60 Seconds, (g) 70 Seconds, (h) 80 Seconds, and (i) 90 Seconds.

## C.1.2 SE-Method 3

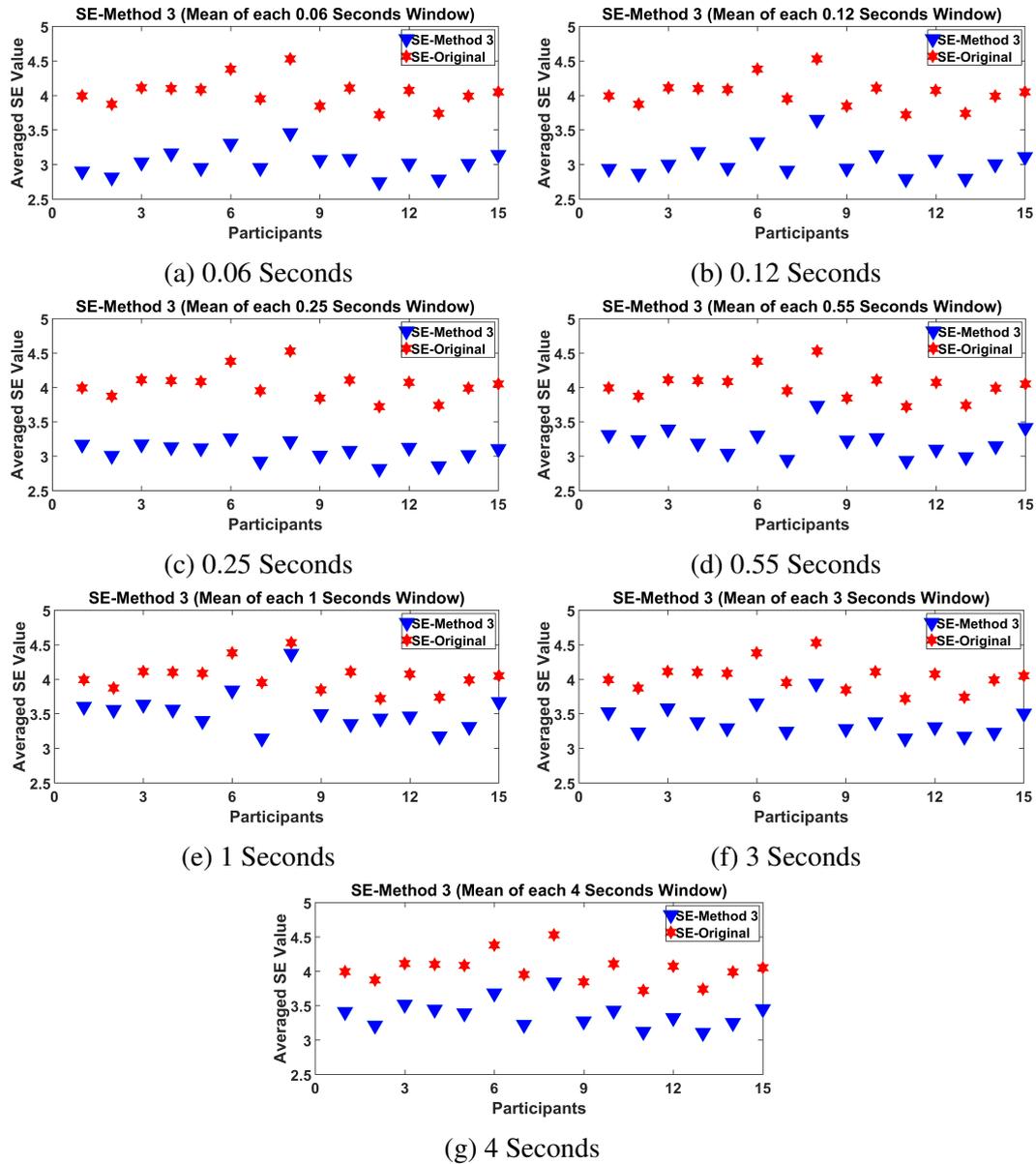


Figure C.2: Slot 1 - Comparison of original SE calculation and proposed SE methods, showing average results for all electrodes are shown for the individual participants: SE-Method 3 on moving windows for: (a) 0.06 Seconds, (b) 0.12 Seconds, (c) 0.25 Seconds, (d) 0.55 Seconds, (e) 1 Seconds, (f) 3 Seconds, and (g) 4 Seconds.

## C.2 Slot 5 Result

In Chapter 8, the results shown are for only Slot 1 data. In this section, SE performance for Slot 5 EEG data is shown for SE-Method 1, SE-Method 2 and SE-Method 3. The Slot 5 result is chosen randomly, and the SE results for all other slots are similar to what is shown here.

### C.2.1 SE-Method 1

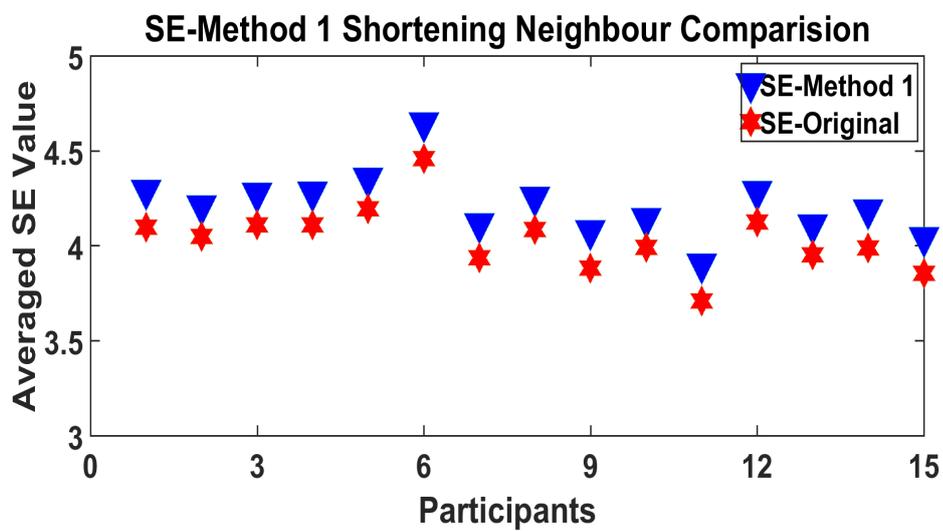


Figure C.3: SE-Method 1 - Comparison of original SE calculation and proposed SE methods, showing average results for all electrodes are shown for the individual participants: SE-Method 1-Shortened Neighbours comparison.

### C.2.2 SE-Method 2

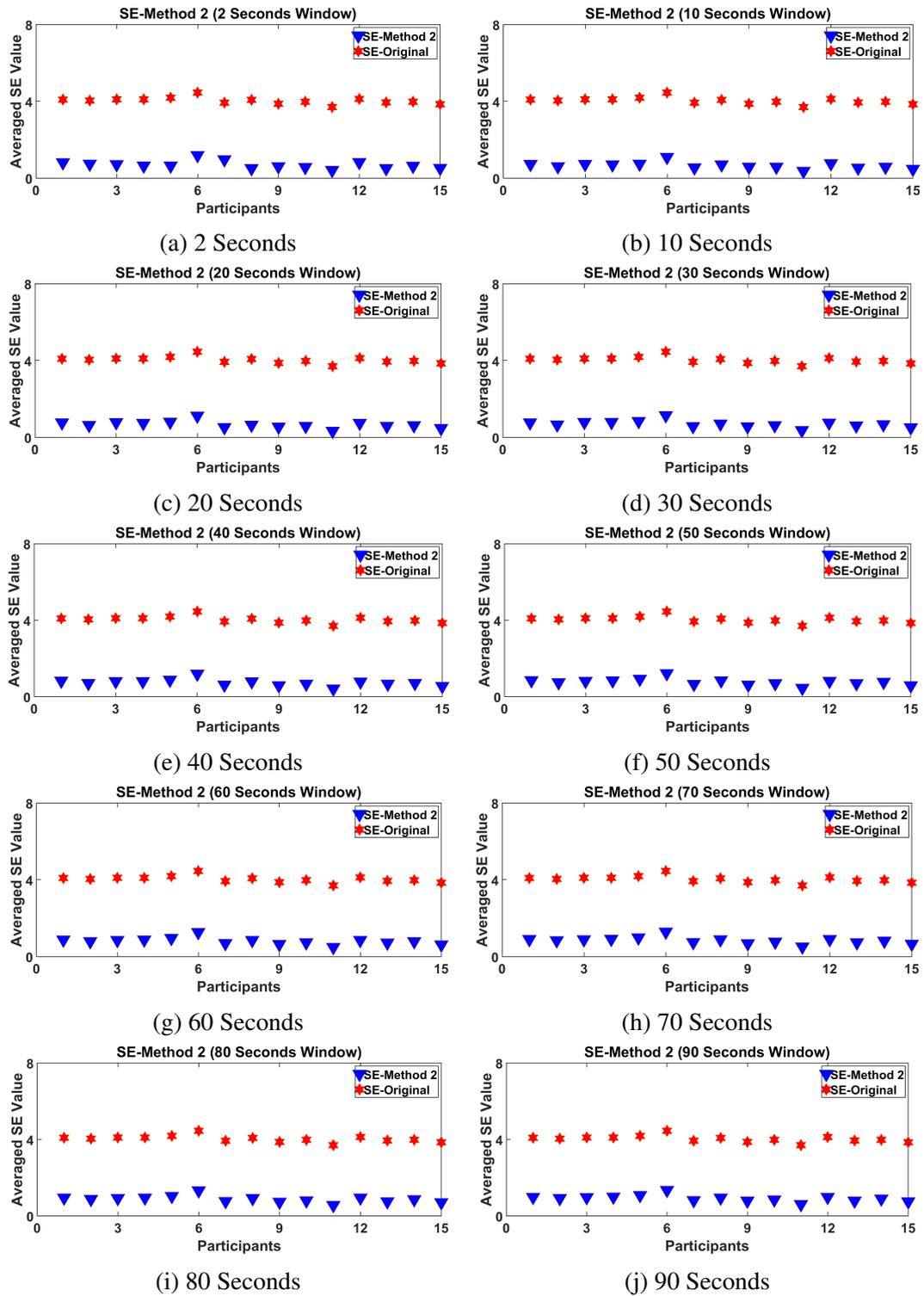


Figure C.4: SE-Method 2 - Comparison of original SE calculation and proposed SE methods, showing average results for all electrodes are shown for the individual participants: SE-Method 2 on moving windows for: (a) 2 Seconds, (b) 10 Seconds, (c) 20 Seconds, (d) 30 Seconds, (e) 40 Seconds, (f) 50 Seconds, (g) 60 Seconds, (h) 70 Seconds, (i) 80 Seconds, and (j) 90 Seconds.

## C.2.3 SE-Method 3

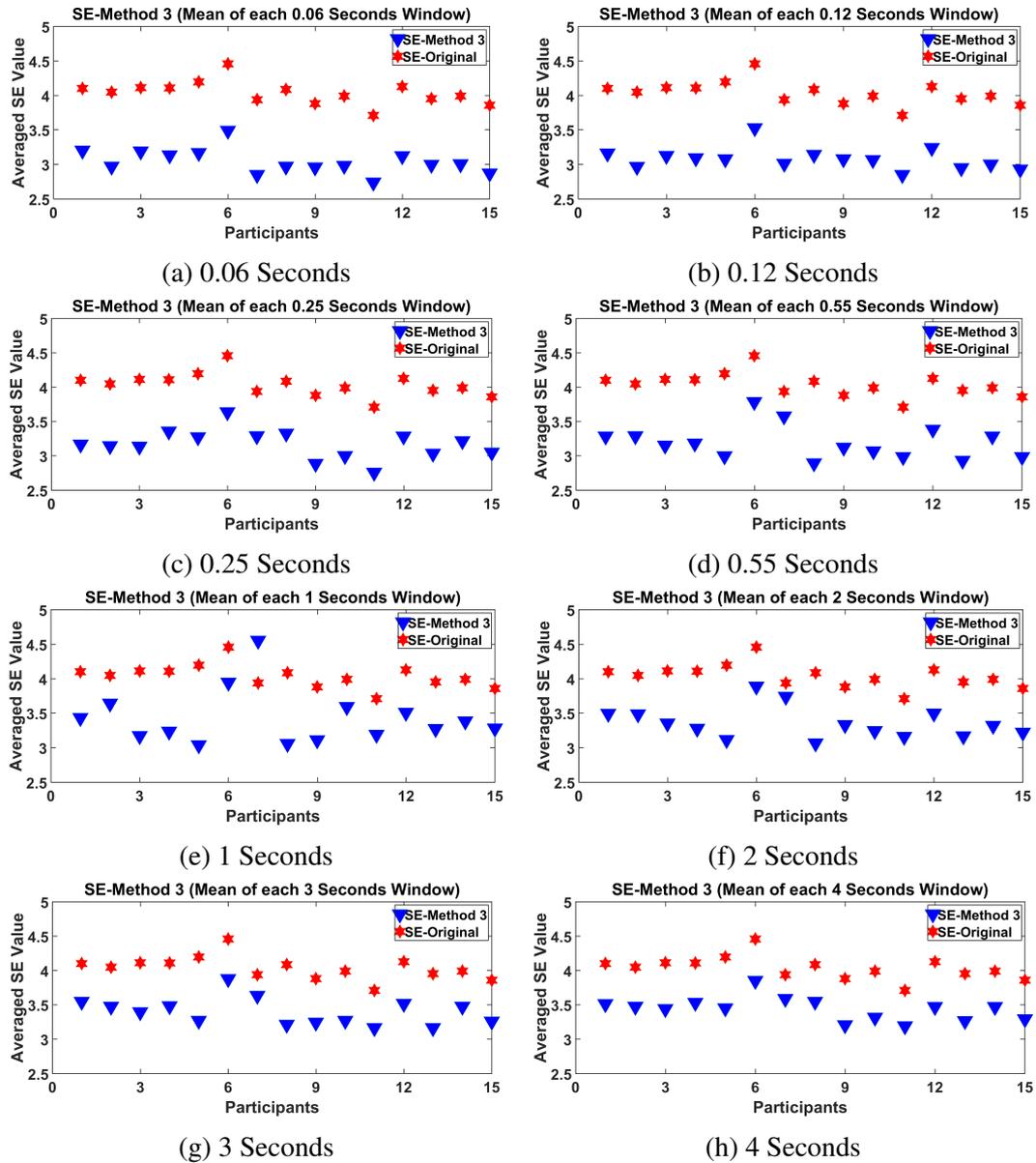


Figure C.5: Slot 5 - Comparison of original SE calculation and proposed SE methods, showing average results for all electrodes are shown for the individual participants: SE-Method 3 on moving windows for: (a) 0.06 Seconds, (b) 0.12 Seconds, (c) 0.25 Seconds, (d) 0.55 Seconds, (e) 1 Seconds, (f) 2 Seconds (g) 3 Seconds, and (h) 4 Seconds.

## **Appendix D**

### **Additional ED Results for Chapter 9**

The ED results of all other participants from Datasets 4, 5, 9 and 10 are shown below.

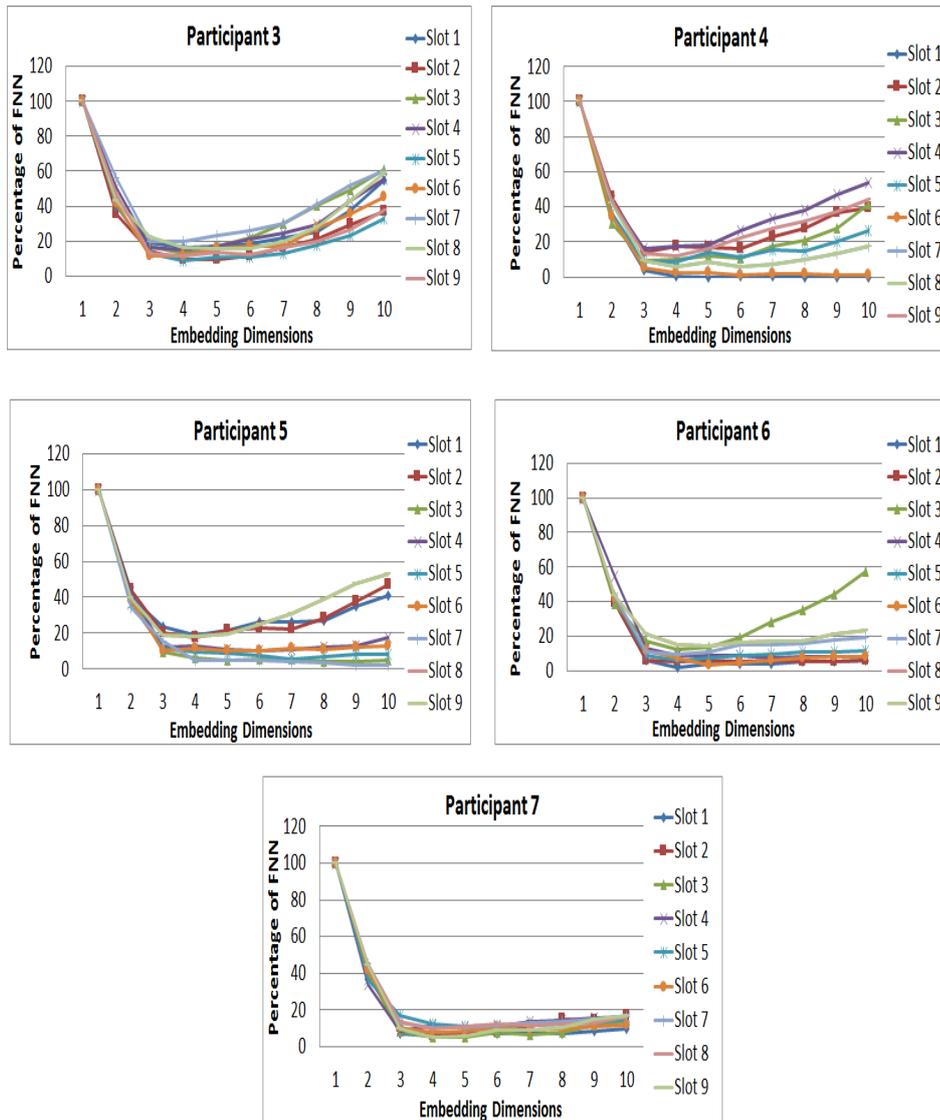


Figure D.1: Dataset 4- Embedding Dimension Result: Five participant's (Participants 3, 4, 5, 6 and 7) ED result from Dataset 4.

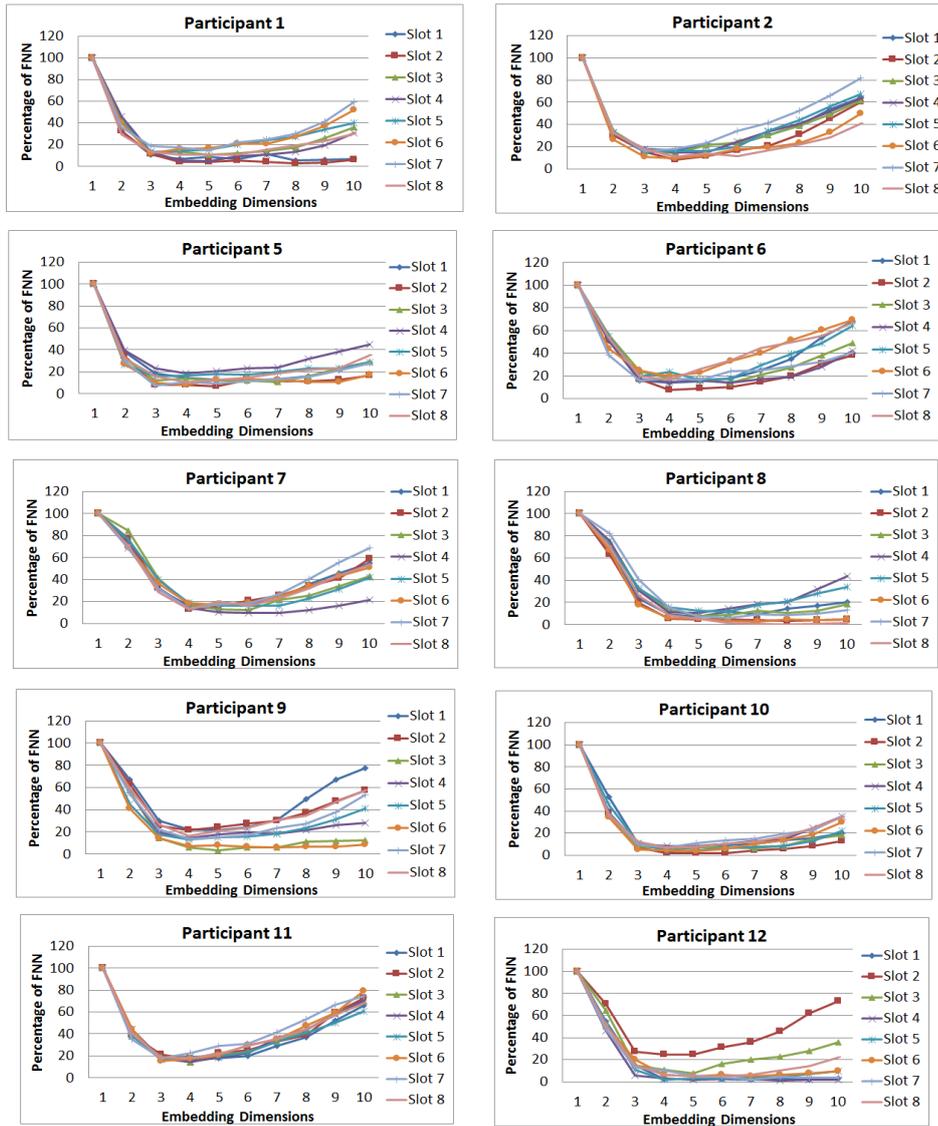


Figure D.2: Dataset 5- Embedding Dimension Result: Ten participant's (Participants 1, 2, 5, 6, 7, 8, 9, 10, 11 and 12) ED result from Dataset 4.

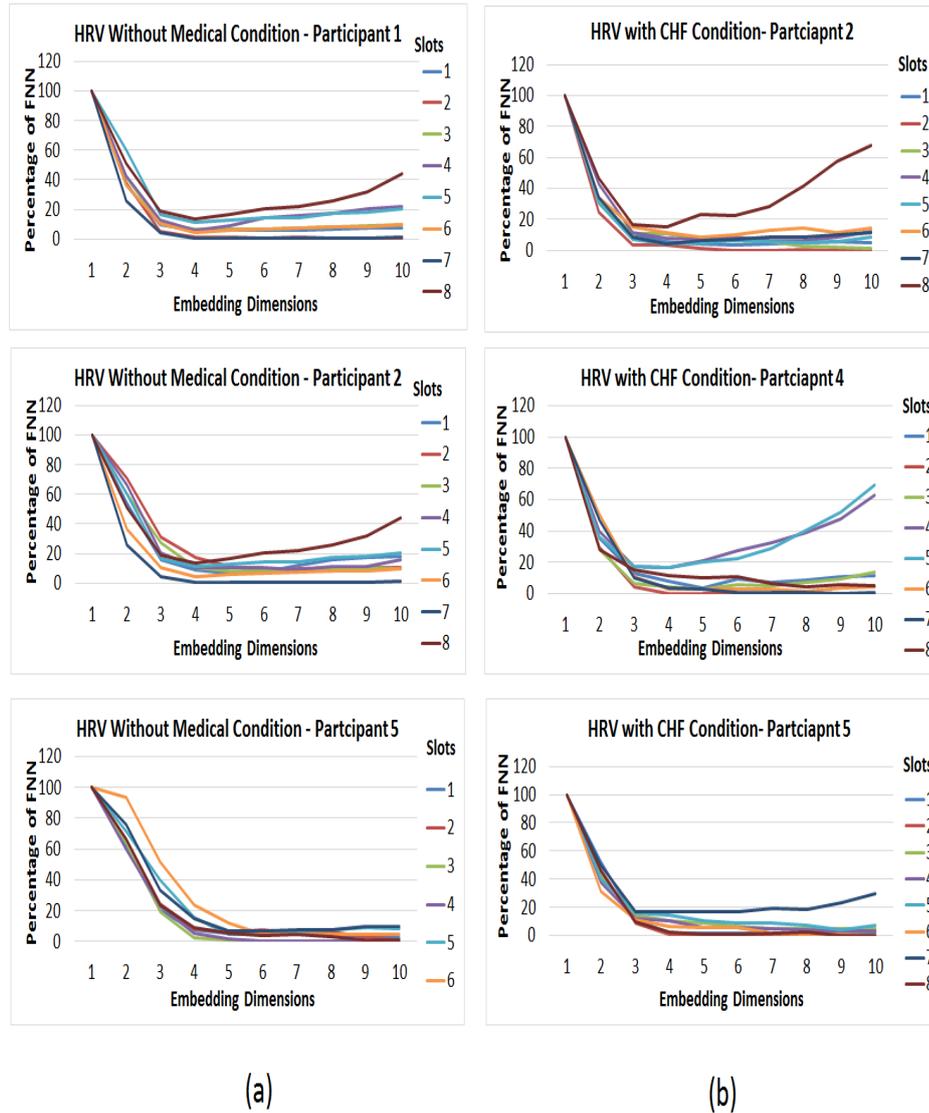


Figure D.3: Dataset 9 - Embedding Dimension Result: (a) Three Participant's (Participant 1, 2 and 5) ED result for HRV without medical condition, (b) Three Participant's (Participant 2, 4 and 5) ED result for HRV with medical condition Congestive Heart Failure (CHF).

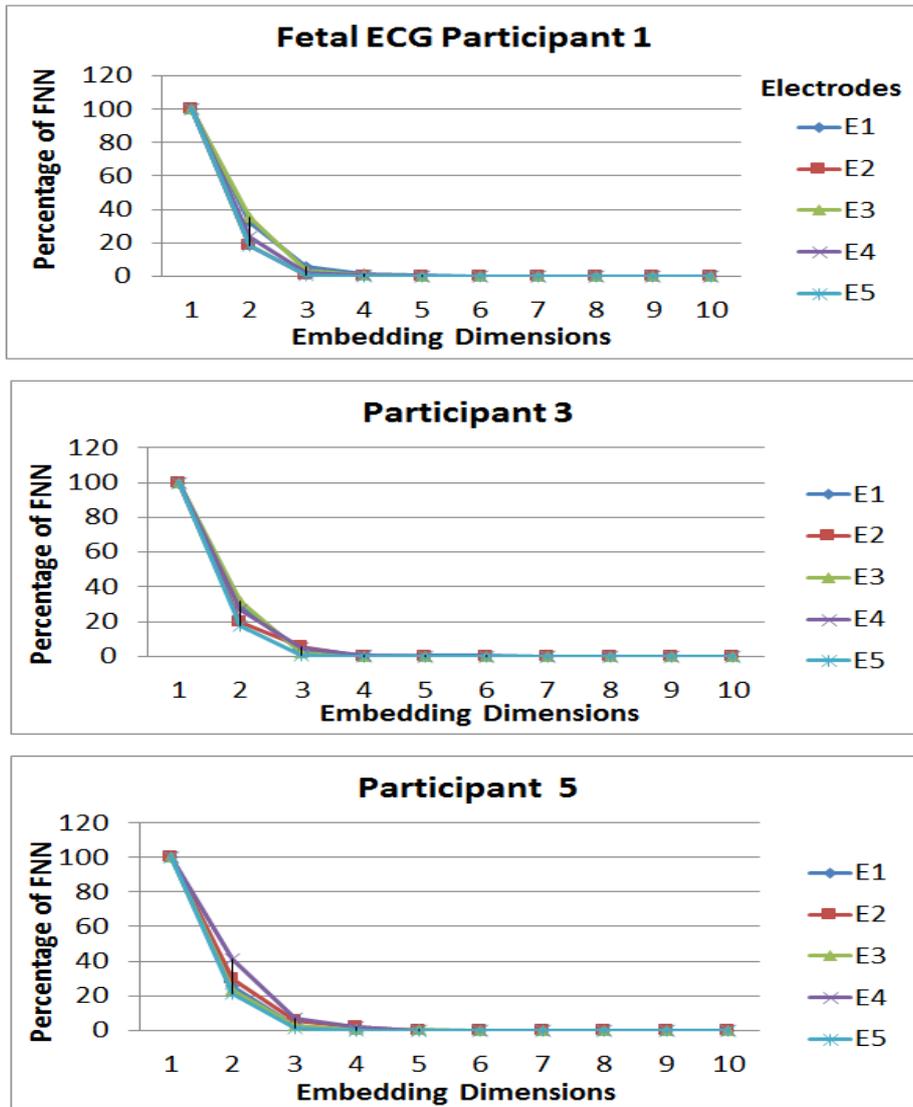


Figure D.4: Dataset 10- Embedding Dimension Result: Three participant's (Participants 1, 3 and 5 ) ED result from Dataset 10 showing 5 measured Electrodes, as shown in different colour).

# **Appendix E**

## **Publications**

8th Annual International Conference on Biologically Inspired Cognitive Architectures, BICA 2017

## The Correlation between EEG Signals as Measured in Different Positions on Scalp Varying with Distance

Ronakben Bhavsar<sup>1</sup>, Yi Sun<sup>1</sup>, Na Helian<sup>1</sup>, Neil Davey<sup>1\*</sup>, David Mayor<sup>1</sup> and Tony Steffert<sup>2</sup>

<sup>1</sup>The University of Hertfordshire, Hatfield, U.K.

<sup>2</sup>Open University, London, U.K.

*r.bhavsar2, y.2.sun, n.helian, n.davey@herts.ac.uk, dfmayor@ntlworld.com  
tony@qeeg.co.uk*

### Abstract

Biomedical signals such as electroencephalogram (EEG) are the time varying signal, and different position of electrodes give different time varying signals. There might be a correlation between these signals. It is likely that the correlation is related to the actual position of electrodes. In this paper, we show that correlation is related to the physical distance between electrodes as measured. This finding is independent of participants and brain hemisphere. Our results indicate that the EEG signal is not transmitted via neurons but through white matter in a brain.

© 2018 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license

(<http://creativecommons.org/licenses/by-nc-nd/3.0/>).

Peer-review under responsibility of the scientific committee of the 8th Annual International Conference on Biologically Inspired Cognitive Architectures

*Keywords:* EEG, Biomedical signal processing, Time Series Data Analysis, Cross-Correlation.

## 1 Introduction

Electroencephalogram (EEG) signals provide a measure of brain nerve cell electro-physiological activity that is accessible on the surface of the scalp (EEG indices of G-induced loss of consciousness (G-LOC), 1988), thus provide information about different types of brain activity. The electrical activity in the brain is recorded via measurement electrodes attached to the surface of the scalp. The EEG signals detected will vary, depending on the location of the electrodes on the scalp. Identifying changes in EEG signals has improved our understanding of the relationship of these signals to people's moods, and behavior (Han, 2012).

Research (Niedermeyer, 2005) suggests that various characteristics of EEG signals are representative of distinct states of brain activity. These distinct states can be quantified using linear or non-linear measures. Previous research has demonstrated a correlation between EEG signals (or brain activity) from different part of the brain (Na, 2002), (Bob, 2010), (Jeong, 2015). A high correlation between the signals from different electrodes indicates similar brain activity, and a low correlation indicates that the brain activity at the different measurement sites is relatively independent.

Researchers (Na, 2002), (Li, 2013) have demonstrated that brain activities within the same (local) region might be similar, but that they might be different among non-identical regions (globally). One question that we address here is whether the activities of the two brain hemispheres are similar.

White matter, which modulates the distribution of action potentials, is brain tissue that is composed of bundles of axons. It acts to coordinate communication between different brain regions (Fields, 2008). One issue we address here is how electrical activity can be communicated across the surface of the brain. We believe that white matter makes a significant contribution to this communication. Our research focuses on evaluating the correlation of EEG signals between different brain regions. The aim of this

study is to determine the relationship between EEG signals and electrode location on the scalp, and to check whether this relationship differs in the two brain hemispheres.

## 2 Related Work

A series of data points in time order, or time series, provides the view of a signal as it evolves over time, i.e. in the Time domain (TD). TD analysis is used to analyse the signal in its actual state, which is the earliest and direct way of analysing EEG signals - it is utilised to analyse changes in EEG signals, such as power (or amplitude) over time. In addition, the frequencies present in the signal are open to investigation (for example, by using the Fast Fourier Transform (FFT)). Such an analysis is said to take place in the Frequency domain (FD). FD analysis is used to identify frequencies present in the signals. Furthermore, it can be utilized to establish the relationship between EEG frequency and its corresponding power (amplitude), and so the energy distributions in EEG signals.

In recent research, the correlation between EEG signals has been analysed in FD using various methods, such as Mutual information, Coherence analysis, Wavelet coherence, Correlation coefficient, Auto-correlation and Cross-correlation. Mutual information has been utilized to examine information transmission between different cortical areas in both patients with schizophrenia and Alzheimer's disease (Na, 2002). This research found lower mutual information between EEG signals of patients with these conditions when compared to normal controls. Coherence analysis has been applied to study brain interactions between EEG signals (Nolte, 2004), indicating significant correlation in EEG Beta ( $\beta$ ) frequency range between the left and right motor areas of the human brain. Wavelet coherence has been applied to distinguish the EEG signals of normal controls and patients with conditions, such as Parkinson's related dementia and Alzheimer's disease (Jeong, 2015). Correlation coefficient has been utilized to discover changes in EEG signals and autonomic nervous activity, and the association of these with personality traits (Takahashi, 2005), with an increase in EEG theta ( $\theta$ ) power and EEG alpha ( $\alpha$ ) power predominantly in the frontal area. Cross-correlation has been utilized to study the degree of association between activities in symmetrical (left and right) parts of the brain (Li, 2013), and results indicated that there is a stronger correlation in the delta ( $\delta$ ) frequency range on the right side of a brain than the left. To our knowledge, limited research has been conducted to analyse EEG signals in the TD. Therefore it is important to perform a comparative analysis and an interpretation of EEG signals in the TD, not just the FD.

Cross-correlation can be performed to analyse the time delay between two related processes. In the present context, it offers a valuable and sensitive method for investigating EEG signals that are recorded at the same time from different electrodes that is independent of their amplitudes. To analyse EEG signals in the time domain, Cross-correlation stands out as the most appropriate correlation method, because of its ability to assess signal similarity at all possible time delays. Cross-correlation has been successfully applied in analysing EEG signals in the FD (Li, 2013), as well as TD (Bob, 2010). This method can be used to determine the relationship between activity in global and local areas, and also among the different local areas of the human brain.

It has been found that the numbers of electrodes and combinations of electrode pairs used to analyse EEG signals is different. Usually, the combination of electrode pairs depends on the total number of electrodes. For example, if there are 19 electrodes then the number of unique potential electrode pair is 171. According to recent research on EEG signal analysis, electrodes from the central part of the brain deserve the best consideration, possibly because minimum noise is found in the recorded signals (Klein, 2006). This was one of the reasons we explored papers in which EEG signals were analysed using limited numbers of electrodes and combinations of electrode pairs. For example, Na et al. (Na, 2002), examined 16 electrodes with 38 pairs of electrodes within the right hemisphere and within the left hemisphere. Their results showed less complex EEG activity in the left temporal regions. Bob et al. (Bob, 2010), inspected 8 electrodes and 16 electrodes pairs to examine the relation between EEG activity

at the Dissociative Experiences Scale (DES) in paranoid schizophrenia patients. Their results explored a significant correlation of DES in 9 EEG electrode pairs. Similar electrode pair effects have been found by Cuevas et al. (Cuevas, 2011), who studied 8 electrodes and 16 electrode pairs in their investigation of patterns of EEG signals in developing children's brains. Their results suggested an age-related increase in EEG power for 9 electrode pairs. Li et al. (Li, 2013), examined 16 electrodes and 4 electrode pairs, and proposed more significant changes in the EEG signals of electrodes from the right-side of the brain when compared to those of the left-side.

This brief review of research on the correlation of EEG signals indicates that investigations have been focused on the FD. Furthermore, limited information was found on the correlation of EEG signals in the TD. Additionally, the numbers of paired electrodes investigated, the number of datasets used, and use of Cross-correlation for analysing EEG signals are limited. Instead, researchers focus has been primarily on electrode combinations within the left and right brain hemispheres. The summary of research work reported in this paper to be replicated can be found at <https://ronak2.wixsite.com/mysite/research-blog>.

To our knowledge, very limited work has been done on the correlation of EEG signals using multiple electrode. This paper investigates the correlation of EEG signals in the TD using Cross-correlation. Three datasets have been used and are named as Data-set 1, Data-set 2 and Data-set 3. Each dataset involves a different number of electrodes. Therefore, the number of unique electrode pairs to perform Cross-correlation is different. From Data-sets 1, 2 and 3 we have 171 pairs, 45 pairs, and 105 pairs, respectively.

### 3 Data Collection

Three different datasets were obtained with each of them containing different numbers of participants and electrodes. All of these datasets follow the 10-20 electrode placement system shown in Fig 1. The 10-20 system is the recognized method to describe the location of electrodes (Klem, 1999). The values of 10% and 20% shown in Fig. 1 refer to the distances between adjacent electrodes: either 10% or 20% of the total front-to-back or right-to-left distance over the skull - front-to-back distance is based on the measurement from Nasion (point between forehead and nose) to Inion (lowest point of the skull from the back of the head indicated by a prominent bump), and right-to-left distance is based on the measurement between the left and right preauricular ear points.

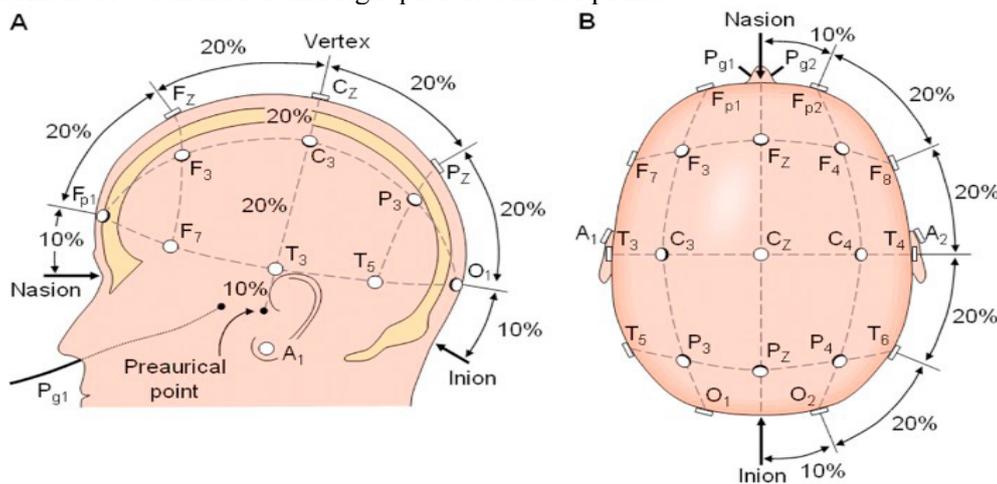


Figure 1: The international 10-20 system seen from **A** (left side of the head) and **B** (above the head). The letter F, T, C, P, O, A, Fp and Pg stands for frontal, temporal, central, parietal, occipital, earlobes, frontal polar, and nasopharyngeal, respectively (Klem, 1999).

Data-set 1 consists scalp EEG recordings of 16 participants obtained over 5 minutes in a relaxed state with eyes opened. 19 electrodes (Fp1, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, and O2) were used, following the 10-20 system. The sampling rate used was 250Hz, and the reference was to linked ear electrodes.

Data-set 2 consists of scalp EEG recordings of 20 participants, while they watched a short documentary movie. 10 electrodes (F7, F3, Fz, F4, F8, T5, P3, Pz, P4, and T6) were used following the 10-20 system. The sampling rate used was 500Hz, and the reference was to linked ear electrodes.

Data-set 3 presents a multi-modal dataset for the analysis of human affective states (Koelstra, 2012). 32 participants EEG signals were recorded while the participants watched 40 one-minute long excerpts of music videos. Out of 32 electrodes recorded, 15 electrodes (Fp1, Fp2, F7, F3, Fz, F4, F8, C3, Cz, C4, P3, Pz, P4, O1, and O2) following the 10-20 system were used for this study. The sampling rate used was 512Hz, and the reference was to linked ear electrodes.

## 4 Methodology

As described in the Related Work section, Cross-correlation measures how closely two different observables are related to each other at the same or different time, taking time lag into consideration. Normalized Cross-correlation (Lewis J. , 1995) is used in this work to find the similarity between two time series signals. The normalized Cross-correlation for time sequence  $x_t$  and  $y_t$  of signals  $x$  and  $y$ , respectively, is defined as follows:

$$R_{xy}(\mathcal{J}) = \frac{\frac{1}{N} \sum_{t=1}^{N-\mathcal{J}} [(x_t - \mu_x) (y_{(t+\mathcal{J})} - \mu_y)]}{\sigma_x \sigma_y} \quad (1)$$

$\mathcal{J}$  is the time lag,  $N$  is the length of signals  $x$  and  $y$ ,  $\mu_x$  is the mean of  $x_t$  and  $\mu_y$  is the mean of  $y_t$ .  $\sigma_x$  is the standard deviation of  $x_t$  and  $\sigma_y$  is the standard deviation of  $y_t$ . The values of the normalized Cross-correlation range between 1 (when the matching entities are exactly the same) and  $-1$  (when the matching entities are inverses of each other). A value of zero indicates no relationship existing between the entities. Note that the Cross-correlation can be evaluated for any length of  $x_t$  and  $y_t$ , and they are not required to be the same (Lewis J. , 1995).

## 5 Experiments & Results

The EEG signals were processed to remove artefacts, such as eye blinks, eye movements, jaw movements and muscle movements, by using Independent Component Analysis (ICA). Cross-correlation has been calculated on the processed EEG signals for the 171 electrode pairs of Data-set 1, the 45 electrode pairs of Data-set 2 and the 105 electrode pairs of Data-set 3.

In order to obtain the distance in centimeters (cm) between electrodes, a measuring tape was used. For most of our participants the head circumference range was 54-58cm, for which a medium-sized ‘electro-cap’ is appropriate. According to (Mitsar, 1996), the circumference of the medium-sized EEG cap is ideal for 64% of adults, whether male or female. Therefore, we utilized a medium-sized EEG cap made of an elastic material which stretches according to the participants head circumference, and measured distances using a straight line on the cap - not a curved line over the skull. Note that the distance between electrodes as shown in Fig. 3 is straight line distance between two electrodes, not the distance as measured over the surface of the scalp.

The maximum absolute correlation was found at lag 0. Fig. 2 shows the information for electrode pairs Fp1-Fp2, as an example. The other two datasets show similar results. In Fig. 2, the x-axis

(horizontally) denotes the time lag; a lag of 1 corresponds to 4 milliseconds - positive time lags (0 to 1000) and negative time lags (0 to -1000) indicate when one of the signals shifted to the right and left side of the reference signal Fp1, respectively. The y axis (vertically) denotes the cross-correlation value. The blue color line is for an individual participant, and the red is for the average of all participants' correlation performance for each electrode pair.

Electrode F7 from Data-set 2 have been chosen randomly to show the Cross-correlation performance. The other two datasets show similar results. The results show the averages for all participants. Fig. 3 show that there is an inverse relationship between Cross-correlation and distance. The linear regression has been plotted to fit the data with a probability of  $p < 0.001$ . This indicates that the Cross-correlation value decreases while the distance from F7 increases, irrespective of brain hemisphere.

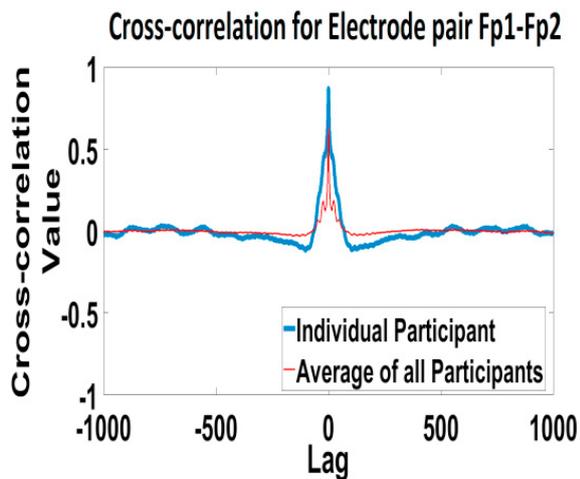


Figure 2: Positive Cross-correlation at Lag 0

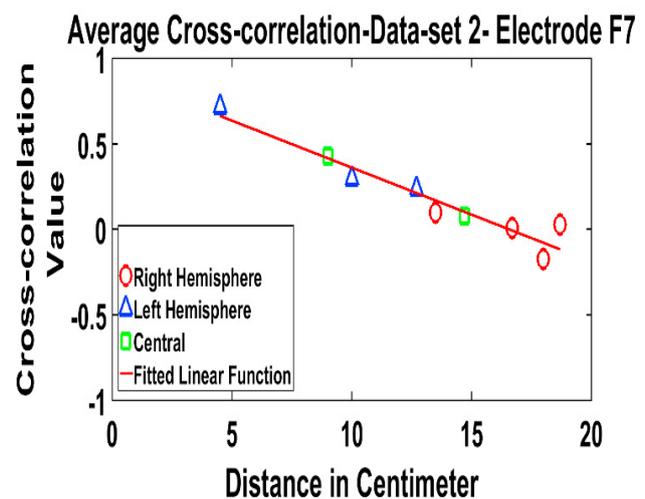


Figure 3: Cross-correlation between electrodes at varying distance on Data-set 2

The results of all other electrodes of all three datasets are similar to the ones shown in Fig. 3. The differences between each Cross-correlation value and the corresponding value on the fitted linear function line were calculated. The results show that the difference between them is very small (about 0.03 98%), which suggests that a close linear dependency does exist.

## 6 Discussion & Conclusion

One of the main conclusions of this work is that electrical activity correlates linearly with distance within the brain, i.e. when distance increases the correlation decreases. To our knowledge, previous research has not described this linear relationship in TD. Our results cover a gap in the research concerning the correlation of EEG signals in the TD using Cross-correlation and possible combinations of electrodes pairs; and also the linear dependence of Cross-correlation with electrodes location. It is important to consider physical separation as measured directly through the skull, and not over the surface of skull when you position electrodes on the skull.

The second conclusion from this work is that the correlation is independent of brain hemisphere. This suggests that most probably the electrical signals are transmitted through the white matter of the brain. We assume signal transmission is through white matter because of the commissural tracts within the white matter which connect the two hemispheres of the brain. This means in practice it does not matter which side of the medial plane you place the electrodes.

Our work suggests that this white matter in the brain is significant in the transmission of electrical activity. White matter is composed of bundles of axons which connect various grey matter areas (the locations of nerve cell bodies) of the brain to each other and carry nerve impulses between neurons. White matter might actively affect how the brain learns and functions, and modulates the distribution of action potentials, acting as a relay and coordinating communication between different brain regions (Fields, 2008). In summary, regardless of the anatomical substrates involved, our main finding is that the correlation between electrical activities in different parts of the brain is linearly related with the electrode distance between them. At the moment we are extending this work to find the correlation between EEG and Electrocardiogram (ECG) signals.

## References

- Bob, P. a. (2010). Dissociative symptoms and interregional EEG cross-correlations in paranoid schizophrenia. *Psychiatry research*, 177(1), 37--40.
- Cuevas, K. a. (2011). EEG and ECG from 5 to 10months of age: Developmental changes in baseline activation and cognitive processing during a working memory task. *International Journal of Psychophysiology*, 80(2), 119--128.
- Fields, R. D. (2008). White matter matters. *Scientific American*, 298(3), 54--61.
- Han, M. a. (2012). Extraction of the EEG signal feature based on echo state networks. *Journal of biomedical engineering*, 29, 206--211.
- Jeong, D.-H. a.-D.-U.-A. (2015). Wavelet Energy and Wavelet Coherence as EEG Biomarkers for the Diagnosis of Parkinson's Disease-Related Dementia and Alzheimer's Disease. *Entropy*, 18(1), 8.
- Klein, S. B. (2006). *Biological psychology*. Macmillan.
- Klem, G. H. (1999). The ten-twenty electrode system of the International Federation. *Electroencephalogr Clin Neurophysiol*, 52(3), 3--6.
- Koelstra, S. a.-S. (2012). Deap: A database for emotion analysis; using physiological signals. *IEEE Transactions on Affective Computing*, 3(1), 18--31.
- Lewis, J. (1995). *Fast normalized cross-correlation* (Vol. 10).
- Lewis, N. G. (1988). *EEG indices of G-induced loss of consciousness (G-LOC)*.
- Li, Y. a. (2013). EEG Physiological Signals Correlation under Condition of+ Gz Accelerations. *Journal of Multimedia*, 8(1), 64--71.
- Mitsar. (1996). (Mitsar Co., Ltd.) Retrieved April 21, 2017, from Mitsar Brain Diagnosticas Solutions: <http://www.mitsar-medical.com/eeg-accessories/#aecap>
- Na, S. H.-H.-J. (2002). EEG in schizophrenic patients: mutual information analysis. *Clinical Neurophysiology*, 113(12), 1954--1960.
- Niedermeyer, E. a. (2005). *Electroencephalography: basic principles, clinical applications, and related fields*. Lippincott Williams & Wilkins.
- Nolte, G. a. (2004). Identifying true brain interaction from EEG data using the imaginary part of coherency. *Clinical neurophysiology*, 115(10), 2292--2307.
- Takahashi, T. a. (2005). Changes in EEG and autonomic nervous activity during meditation and their association with personality traits. *International Journal of Psychophysiology*, 55(2), 199--207.

# An Investigation of How Wavelet Transform Can Affect the Correlation Performance of Biomedical Signals

## *The Correlation of EEG and HRV Frequency Bands in the Frontal Lobe of the Brain*

Ronakben Bhavsar, Neil Daveya and Yi Sun and Na Helian  
*The School of Computer Science, University of Hertfordshire, Hatfield, U.K.*

**Keywords:** EEG, HRV, Biomedical Signal Processing, Time series Data Analysis, Pearson Correlation, Wavelet Transform, Independent Component Analysis, Feature Extraction, Fast Fourier Transform.

**Abstract:** Recently, the correlation between biomedical signals, such as electroencephalograms (EEG) and electrocardiograms (ECG) time series signals, has been analysed using the Pearson Correlation method. Although Wavelet Transformations (WT) have been performed on time series data including EEG and ECG signals, so far the correlation between WT signals has not been analysed. This research shows the correlation between the EEG and HRV, with and without WT signals. Our results suggest electrical activity in the frontal lobe of the brain is best correlated with the HRV. We assume this is because the frontal lobe is related to higher mental functions of the cerebral cortex and responsible for muscle movements of the body. Our results indicate a positive correlation between Delta, Alpha and Beta frequencies of EEG at both low frequency (LF) and high frequency (HF) of HRV. This finding is independent of both participants and brain hemisphere.

## 1 INTRODUCTION

Biomedical signals are a record of electrical activity within human body, and they may indicate the state of health of human. Among many biomedical signals, Electroencephalograph (EEG) and Electrocardiograph (ECG) signals are considered in this work. EEG signals provide a measure of brain nerve cell electro-physiological activity, that is accessible on the surface of the scalp (Lewis et al., 1988), thus provide information about different types of brain activity. Identifying changes in EEG signals has improved our understanding of the relationship of these signals to people's moods, and behaviour (Han et al., 2012), (Ebersole and Pedley, 2003). ECG signals contains a plethora of information on the normal and pathological physiology of the heart and its health. Furthermore, ECG signals provide vital information with regards to the function and rhythm of the heart. The heart rate variability (HRV) has been extracted from the ECG signals. HRV describes the variation in time between consecutive heart beats, which is commonly referred to as the RR (R wave to R wave) or NN (Normal beat to normal beat) intervals.

In recent years, the correlation between the EEG

and the ECG have been conducted to analyse their functionality under certain conditions and to check whether this functionality is related to each other. Research (Kim et al., 2013), (Chua et al., 2012), (Abdullah et al., 2009), (Sakai et al., 2007), (Berg et al., 2005), (Edlinger and Guger, 2006), suggests that the correlation between spectral bands of EEG and HRV has been conducted to assess the interaction between them, and achieved remarkable correlation.

The recent research on correlation between these two signals as mentioned earlier has focused on the Fourier analysis of the frequencies presents in these signals. Whilst, the wavelet transform (WT), acts on frequency and time of the recorded signals. Therefore, WT has widely utilized for analysing biomedical or time series signals. The WT of the signal can be thought of as an extension of the classic Fourier transform (FT) - it works on multi-scale basis, instead of working on a single scale (Time or Frequency) as FT, and gives detailed and clear information of the signals. Therefore, WT of the signals is an important method not only to analyse EEG and ECG/HRV signals individually, but also to analyse the correlation between them. According to recent research (Thomas and Moni, 2016), (Chandra et al., 2017), (Mirsadeghi

et al., 2016), (Mporas et al., 2015), (Valderrama et al., 2012), (Nasehi and Pourghassem, 2011), (Cvetkovic et al., 2008), WT has been used to analyse either EEG or ECG signal, but the correlation between these transformed signals has not yet been conducted. In this paper we are not only focusing on the correlation between without wavelet transform signals but also between wavelet transformed signals.

## 2 RELATED WORK

A series of data points in time order, or time series, provide the view of a signal as it evolves over time, in the Time domain (TD). TD analysis is used to analyse the signal in its actual state - it is utilised to analyse changes in biomedical signals, such as the power (or amplitude) over time. In addition, the frequencies present in the signal are open to investigation (for example, by using the Fast Fourier Transform (FFT)). Such an analysis is said to take place in the Frequency domain (FD). The FD analysis is used to identify frequencies present in the signals. Furthermore, it can be utilized to establish the relationship between frequencies and its corresponding power (amplitude), and so the energy distributions in signals.

In recent research, the correlation between EEG and ECG/HRV signals have been analysed in the FD, as shown in Table 1, which indicates that the Pearson correlation is the best method for the FD analysis. In addition, different numbers of EEG electrodes have been used to analyse the relationship with the ECG/HRV. To the best of our knowledge, very limited work has been done on the correlation between EEG and ECG/HRV signals using 19 EEG electrodes. Moreover, no one has analysed these signals under the same condition (i.e. with TEAS acupuncture applied) that utilised in this research. This paper investigates the correlation between EEG and ECG/HRV signals in FD using Pearson correlation considering all 19 EEG electrodes under the same condition.

Based on the research as shown in Table 2 on WT, it is straightforward that the DWT based methods are well known for EEG and ECG feature extraction and analysis. Furthermore. Among the DWT based methods mentioned, *db wavelet* method has been considered by the researchers. It is obvious from the research on WT that key features of EEG and ECG signal can improve the analysis performance. Therefore, it is important to analyse not just either EEG or ECG as shown in Table 2, but also the correlation between EEG and ECG. To our knowledge, we have not yet found information on the correlation between wavelet transformed signals. In this work, we describes such

an analysis.

## 3 DATASET INFORMATION

Two different datasets were obtained with each of them containing different numbers of participants, stimulation location, and total time length as shown in Table 3. All of these datasets follow the 10-20 electrode placement system shown in Figure 1. The 10-20 system is the recognized method to describe the location of electrodes (Klem et al., 1999). The values of 10 and 20 percentage shown in Figure 1 refer to the distances between adjacent electrodes: either 10 or 20 percentage of the total front-to-back or right-to-left distance over the skull - front-to-back distance is based on the measurement from the Nasion (point between forehead and nose) to the Inion (lowest point of the skull from the back of the head indicated by a prominent bump), and right-to-left distance is based on the measurement between the left and right preauricular ear points.

Dataset 1 and 2 consist of EEG and ECG recordings from 16 and 7 participants, respectively. These data were obtained over ten 5 minutes slots with eyes open using Transcutaneous Electro Acupuncture (TEAS) method, including resting state data in the first and the last slot. The EEG and ECG recording were made simultaneously. 19 electrodes (Fp1, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, and O2) for EEG recording were used, following the 10-20 system. The sampling rate used for EEG was 250Hz, and the reference was to linked ear electrodes. For ECG data, two electrodes were placed on both side of the wrist (having one electrode as ground) to record the electrical activity of the heart over time, and the sampling rate used was 256Hz.

Table 3: Information about the Datasets.

Label	Dataset 1	Dataset 2
Number of Participants	16	7
EEG-Electrodes	19	19
EEG-Sampling Rate	250Hz	250Hz
ECG-Electrode	1	1
Stimulation Location	1	4
ECG-Sampling Rate	256Hz	256Hz
Total Time Length	50 minutes	45 minutes
Slot Time Length	5 minute	5 minute

The difference between these datasets, other than the participants, is the body location where TEAS stimulation has been performed. For Dataset 1, only one body location (Dominant Hand), and for Dataset 2, four different body location (Left Hand, Below Left Knee, Right Hand, and Below Right Knee) has been used to perform TEAS stimulation.

Table 1: Summary of Correlation Research on Biomedical Signals since 2003 to 2017.

RefDetail	TD	FD	Pearson Correlation Method	Other Correlation Method	EEG Electrodes Investigated
(Miyashita et al., 2003)	-	✓	✓	-	4
(Yang et al., 2002)	-	✓	✓	-	2
(Ako et al., 2003)	-	✓	✓	-	1
(Jurysta et al., 2003)	-	✓	-	Coherency Analysis	3
(Takahashi et al., 2005)	-	✓	✓	-	6
(Edlinger and Guger, 2006)	-	✓	✓	-	2
(Berg et al., 2005)	-	✓	✓	-	2
(Sakai et al., 2007)	-	✓	✓	-	19
(Abdullah et al., 2010)	-	✓	-	Cross-correlation	1
(Chua et al., 2012)	-	✓	-	✓	4
(Kim et al., 2013)	-	✓	-	Coherency Analysis	19
(Prinsloo et al., 2013)	✓	-	✓	-	3
(Liou et al., 2014)	-	✓	✓	-	19
(Triggiani et al., 2016)	-	✓	✓	-	19

Table 2: Summary of Research on Well known Wavelet Transformation Methods for Biomedical Signals since 2012 to 2017.

RefDetail	EEG	ECG/HRV	TD	FD	Feature Extraction Method
(Kutlu and Kuntalp, 2012)	-	✓	✓	-	DWT-Daub Wavelet
(Thomas et al., 2015)	-	✓	✓	-	DWT-Daub Wavelet
(Sudarshan et al., 2017)	-	✓	✓	-	DWT-Daub Wavelet
(Acharya et al., 2017)	-	✓	-	✓	DWT-Daub Wavelet
(Dolatabadi et al., 2017)	-	✓	✓	✓	Principal Component Analysis (PCA)
(Kumari et al., 2014)	✓	-	✓	✓	DWT-Daub Wavelet
(Mumtaz et al., 2017)	✓	-	✓	✓	DWT-Daub Wavelet
(Kevric and Subasi, 2017)	✓	-	-	✓	DWT-Daub Wavelet
(Faust et al., 2015)	✓	-	✓	-	DWT-Daub Wavelet

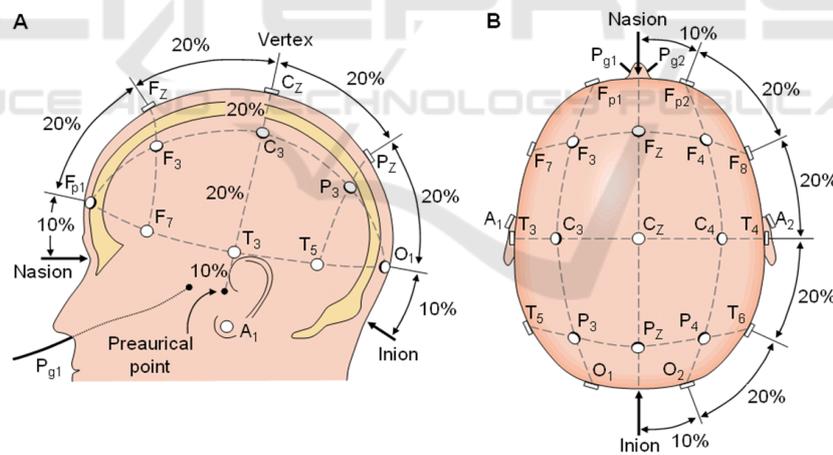


Figure 1: The international 10-20 system seen from A (left side of the head) and B (above the head). The letter F, T, C, P, O, A, Fp and Pg stands for frontal, temporal, central, parietal, occipital, earlobes, frontal polar, and nasopharyngeal, respectively. The figure is obtained from (Klem et al., 1999).

## 4 METHODS

### 4.1 Pearson Correlation

The Pearson's correlation coefficient measures how closely two different observables are related to each other. Correlation co-efficient range between 1 (when the matching entities are exactly the same) and  $-1$

(when the matching entities are inverses of each other). A value of zero indicates no relationship existing between the entities.

### 4.2 Wavelet Transform

The Wavelet Transform (WT) is designed to direct the problem of signals with nonstationarity. It in-

cludes representation of time function in terms of simple blocks, termed wavelets. These blocks are derived from a signal generating function called the *mother wavelet* by translation and dilation operations. Dilation, also known as scaling, compresses or stretches the mother wavelet and translation shifts it along the time axis (Daubechies, 1990), (Akay, 1997), (Unser and Aldroubi, 1996). The WT can be categorized into continuous and discrete. Continuous wavelet transform (CWT), implies that the scaling and translation parameters change continuously, and thus, represent considerable effort and vast amount of data calculation for every possible scale. Therefore, we used discrete wavelet transform (DWT). The WT of the signal can be thought of as an extension of the classic Fourier transform (FT) - it works on multi-scale basis, instead of working on a single scale (Time or Frequency) as FT. This is achieved by decomposition of the signal over dilated (scale) and translated (time) version of wavelet. An input signal is decomposed by using low pass filter and high pass filter followed by down sampling in each stage. The output of the first stage high pass filter gives the detail coefficient (D1), whereas the low pass filter gives the approximation coefficient (A1).

The prototype wavelet used in this study is Daubechies wavelet of order 4 (db4) based on our research on biomedical/time series signal analysis, as mentioned in Table 2.

## 5 EXPERIMENTAL SET-UP

The experimental steps are shown in Figure 2. The EEG signals were pre-processed to remove artefacts caused by the electrical activity in muscles including eye, jaw and muscle movements using Independent Component Analysis (ICA). It was straightforward to remove these using ICA (Hyvärinen and Oja, 2000). The power spectrum for each frequency band of EEG - Delta (0.3-4 Hz), Theta (4-7.5 Hz), Alpha (7.5-13 Hz), Beta (13-30 Hz), and Gamma (30-50 Hz) were then obtained by Power Spectrum Density (PSD) (Stoica and Moses, 1997).

To extract HRV from ECG signals, we used method designed by Lin et al. (Lin et al., 2010). The results of the automatic analysis were reviewed and any errors in R-wave detection and QRS labelling were then removed manually. R-R interval data obtained from the edited time sequence of R-wave and QRS labelling were then transferred to a personal computer. In order to remove artefact from extracted HRV signal, each R-R interval has been compared against a local average interval. If an R-R interval

differs from the local average more than a specified threshold (Threshold in seconds) value, then that R-R interval is defined as an artefact and is replaced with an interpolated value using a cubic spline interpolation. The power spectrum for each frequency band of HRV - Very Low Frequency (VLF) ranges 0-0.04 Hz, Low Frequency (LF) ranges 0.04-0.15 Hz, and High Frequency 0.15-4 Hz were then obtained by PSD (Power Spectrum Density).

The sampling rate is 1Hz for the extracted HRV, and 250Hz for the EEG. In order to perform correlation between these different sampling rate signals, it was required to change the sampling rate for either the EEG or HRV signals. Therefore, we decided to segmenting EEG signals using 1 second window and represent each window by its means value (the mean value from each 250 samples), unlike normal down sampling, where much of the data is thrown away. For each participant's EEG data, this process has been repeated for all 5 minutes slots. After windowing, the spectral analysis was performed. From each frequency bands of the EEG and the HRV, the mean of the amplitude value within the frequency range has been measured, single value for each of these frequency band, and for each 5 minute is obtained. Then, the correlation between these frequency values is performed.

In order to perform correlation based on wavelet transformed EEG and/or HRV signal, the WT-Daubechies (db) Wavelet up to level 5 is performed on the signals before extracting frequency bands as mentioned in Figure 2. For the datasets we have, the low pass filter worked very well. Therefore, we considered low passed WT signals to perform the correlation.

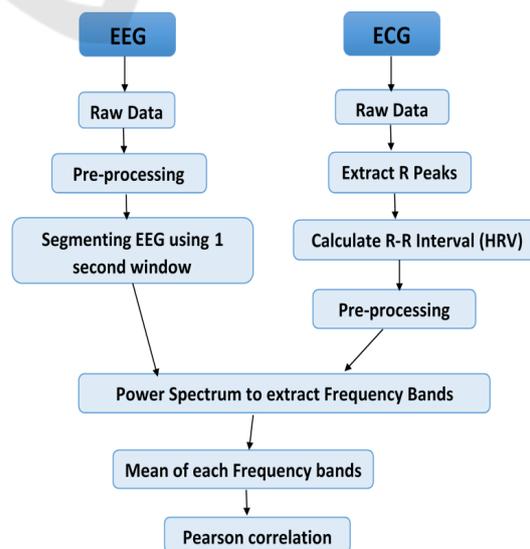


Figure 2: Experiment steps for the correlation performance.

## 6 EXPERIMENTAL RESULTS AND DISCUSSION

For each dataset, we investigated the correlation between each of the EEG frequencies (Delta, Theta, Alpha and Beta) with each frequencies of the HRV frequencies (LF and HF) in three different experiments: 1). The Correlation between Pre-processed Signals, 2). The Correlation between Pre-processed and WT signals of the EEG and HRV, and finally 3). The Correlation between Pre-processed HRV with Pre-processed and WT signals of the EEG. The Gamma frequency of EEG did not give us the correlation effect. Therefore, it is not considered in the result shown in Figure 3 and Table 4.

For both datasets, the experiment 2). correlation between both WT signals did not give better results, because HRV is tend to be less noisy. Therefore, when the WT has been performed on HRV, information has been lost and the signal became more flat. The most interesting result has been found from experiments 1). and 3).

For each frequency combination correlation, the average of participants for each EEG electrode has been calculated. Then the best performance electrode has been ranked- where, the ranking has been given based on electrode correlation result. The average of electrode ranking for each frequency combination is then gathered and five best performance electrodes result has been looked closely. We have found some common electrodes in all of the frequency combination we have investigated. Figure 3 shows the result of this investigation for Dataset 1 and 2.

As shown in Figure 3, for dataset 2, some electrodes from the back side of the brain are giving stronger result than dataset 1. This is due to more randomness in the EEG signals from dataset 2. Also, the location where TEAS has been performed might contributed to this result.

Based on results shown in Figure 3, it can be seen that the frontal lobe of the brain is correlated with the heart. The frontal lobe involved in higher mental functions, such as concentration, creativity, speaking, muscle movement and in making plans and judgements, is a part of cerebral cortex (body's ultimate control and information processing) of the brain (McCraty et al., 2009). The usual Heart-Brain communication path is through spinal cord. In order to have relationship between frontal lobe of the brain and heart, we assume the communication might have done through 'Medulla'(cardiovascular center placed in medullacontrols the heart beating) which is part of brain stem. The signal has been then directed to the Thalamus and then to the cerebral cortex (Lane et al.,

2001), (ATKINSON and BRADLEY, 2004).

Table 4 shows the average correlation result of participants for each frequency comparison from dataset 1 and 2. Where, Level 0 means correlation between pre-processed data, and Level 1 to 5 means, correlation between pre-processed HRV with pre-processed and WT EEG. The heat map of these result ("Red" is strongest and "Dark-Blue" means weakest) as shown in Table 4, indicates the correlation performance changes with the levels of WT. We found the signal became flat after level 2, and lost information when levels has been increased. Therefore, we have not considered result of levels 3, 4 and 5 in Figure 3 (b) and (d).

Results shown in Table 4 are indicative and not statistically significant, according to these, three frequencies of EEG have shown some correlation, such as Delta, Alpha, and Beta, have shown correlation at both LF and HF of HRV. Each of these frequencies represent the activities of these signals. For example, Delta will be higher if the person is in deep sleep, Alpha will appear if the person is calmed, relaxed or in creative visualisation, and Beta will show if the person is working or feeling more alert. For HRV, LF and HF represent the sympathetic and parasympathetic activities of autonomic nervous system (ANS), respectively.

## 7 CONCLUSIONS

The main conclusion of this work is that electrical activity in the frontal lobe of the brain is correlated with the HRV for the given two datasets. To the best of our knowledge this is a new result. This suggests that most probably the electrical signals could be transmitted through the cerebral cortex, Thalamus, and Medulla of the brain (Saper et al., 2005). The possible path of the key neuronal projections that maintain alertness is shown in Figure 4.

The second conclusion from this work is that, WT signals also give correlation from the frontal lobe of the brain. To the best of our knowledge, the correlation between WT signals of EEG and ECG/HRV has not yet been investigated.

A more tentative conclusion of this work is that three frequencies of the EEG Delta, Alpha and Beta are correlated with the LF and HF of HRV, for dataset 1 and dataset 2, respectively. Whereas, most of previous studies, (Yang et al., 2002),(Ako et al., 2003),(Jurysta et al., 2003),(Abdullah et al., 2010) and (Chua et al., 2012), have shown negative correlation between these frequency bands due to the condition in which these signals have been analysed.

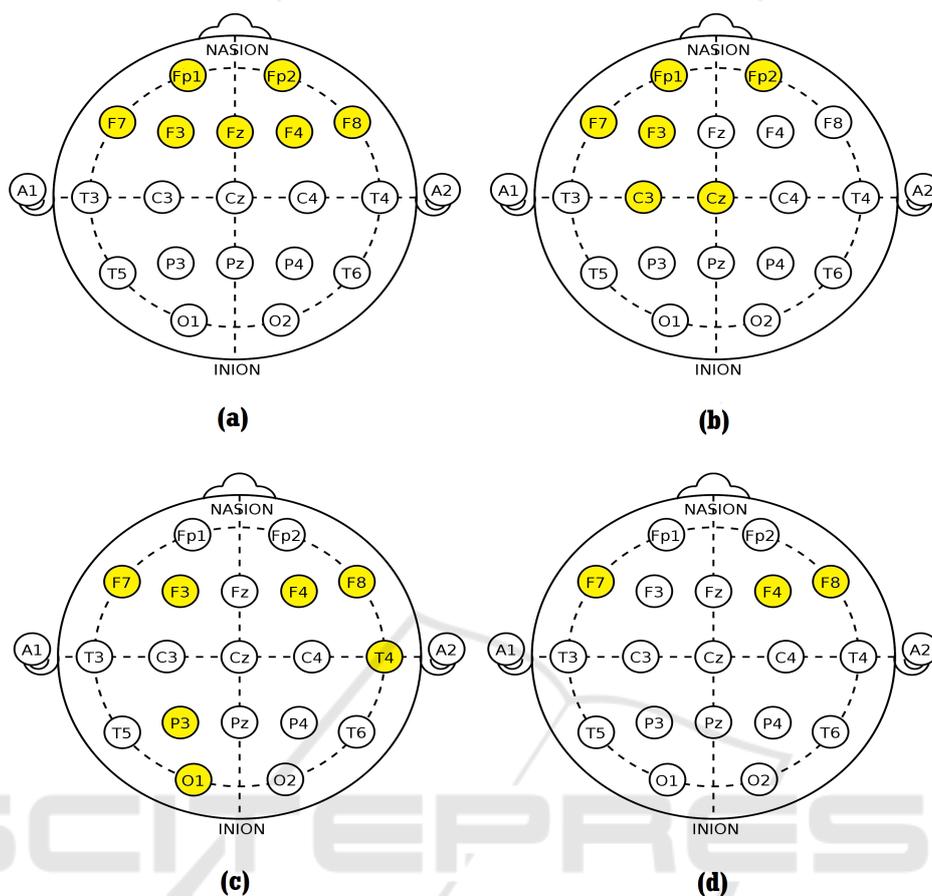


Figure 3: Best Electrodes Correlation Performance, highlighted in yellow colour: (a) Dataset 1 Correlation performance on pre-processed HRV and EEG, (b) Dataset 1 Correlation performance on pre-processed HRV and WT signals of EEG, (c) Dataset 2 Correlation performance on pre-processed HRV and EEG, (d) Dataset 2 Correlation performance on pre-processed HRV and WT signals of EEG.

Dataset 1-HRV & Wavelet Transformed EEG Correlation								Dataset 2-HRV & Wavelet Transformed EEG Correlation							
EEG	HRV	Level 0	Level1	Level2	Level3	Level4	Level5	EEG	HRV	Level 0	Level1	Level2	Level3	Level4	Level5
DELTA	LF	0.17	0.17	0.17	0.18	0.18	0.01	DELTA	LF	0.00	0.00	0.00	0.00	0.01	0.01
	HF	0.07	0.08	0.08	0.08	0.11	0.08		HF	0.05	0.05	0.02	0.01	0.00	0.06
THETA	LF	0.04	0.04	0.01	0.07	0.17	0.03	THETA	LF	-0.04	-0.05	-0.05	-0.04	0.02	0.08
	HF	0.07	0.08	0.07	0.03	0.15	0.03		HF	0.00	-0.01	0.00	0.01	0.10	0.13
ALPHA	LF	0.10	0.09	0.02	0.08	0.15	0.01	ALPHA	LF	-0.05	-0.06	-0.05	-0.02	0.03	0.11
	HF	0.00	0.01	-0.01	0.01	0.10	0.00		HF	0.04	0.03	0.04	0.00	0.16	0.13
BETA	LF	-0.01	-0.01	0.01	0.07	0.03	-0.05	BETA	LF	-0.06	-0.07	-0.08	-0.02	0.03	0.02
	HF	-0.02	0.05	-0.02	0.03	0.07	0.01		HF	0.04	0.04	0.04	0.02	0.16	0.11

Table 4: Heat Map Results of Averaged participants correlation performance: Dataset 1 (Left), and Dataset 2 (on Right). Colour coding from Red to Dark Blue, Red=Strongest, Dark-Blue=Weakest).

In summary, the number of EEG electrodes used by other people to investigate correlation was limited. Our results cover a gap in the research concerning the correlation between the EEG and the HRV using all EEG electrodes. Our work suggests a correlation be-

tween the frontal lobe of the EEG and the HRV, with and without WT signals. We assume this is because the frontal lobe is related with higher mental functions of cerebral cortex and responsible for muscle movements of the body (Stuss and Benson, 1986).

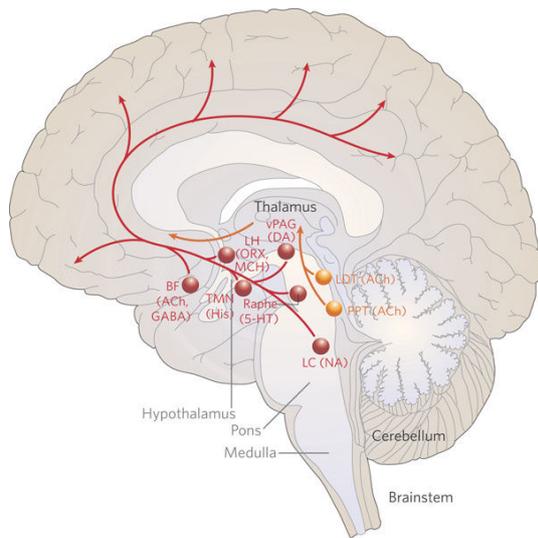


Figure 4: Key neuronal projections that maintain alertness, and possibly the path from cardiovascular center to the frontal lobe of the brain's communication. The figure is obtained from (Saper et al., 2005).

## REFERENCES

- Abdullah, H., Holland, G., Cosic, I., and Cvetkovic, D. (2009). Correlation of sleep eeg frequency bands and heart rate variability. In *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE*, pages 5014–5017. IEEE.
- Abdullah, H., Maddage, N. C., Cosic, I., and Cvetkovic, D. (2010). Cross-correlation of eeg frequency bands and heart rate variability for sleep apnoea classification. *Medical & biological engineering & computing*, 48(12):1261–1269.
- Acharya, U. R., Fujita, H., Adam, M., Lih, O. S., Sudarshan, V. K., Hong, T. J., Koh, J. E., Hagiwara, Y., Chua, C. K., Poo, C. K., et al. (2017). Automated characterization and classification of coronary artery disease and myocardial infarction by decomposition of eeg signals: a comparative study. *Information Sciences*, 377:17–29.
- Akay, M. (1997). Wavelet applications in medicine. *IEEE spectrum*, 34(5):50–56.
- Ako, M., Kawara, T., Uchida, S., Miyazaki, S., Nishihara, K., Mukai, J., Hirao, K., Ako, J., and Okubo, Y. (2003). Correlation between electroencephalography and heart rate variability during sleep. *Psychiatry and clinical neurosciences*, 57(1):59–65.
- ATKINSON, M. and BRADLEY, R. T. (2004). Electrophysiological evidence of intuition: Part 2. a system-wide process? *THE JOURNAL OF ALTERNATIVE AND COMPLEMENTARY MEDICINE*, 10(2):325–336.
- Berg, J., Neely, G., Wiklund, U., and Landström, U. (2005). Heart rate variability during sedentary work and sleep in normal and sleep-deprived states. *Clinical physiology and functional imaging*, 25(1):51–57.
- Chandra, S., Sharma, G., Sharma, M., Jha, D., and Mittal, A. P. (2017). Workload regulation by sudarshan kriya: an eeg and ecg perspective. *Brain informatics*, 4(1):13.
- Chua, E. C.-P., Tan, W.-Q., Yeo, S.-C., Lau, P., Lee, I., Mien, I. H., Puvanendran, K., and Gooley, J. J. (2012). Heart rate variability can be used to estimate sleepiness-related decrements in psychomotor vigilance during total sleep deprivation. *Sleep*, 35(3):325–334.
- Cvetkovic, D., Übeyli, E. D., and Cosic, I. (2008). Wavelet transform feature extraction from human ppg, ecg, and eeg signal responses to elf pemf exposures: A pilot study. *Digital signal processing*, 18(5):861–874.
- Daubechies, I. (1990). The wavelet transform, time-frequency localization and signal analysis. *IEEE transactions on information theory*, 36(5):961–1005.
- Dolatabadi, A. D., Khadem, S. E. Z., and Asl, B. M. (2017). Automated diagnosis of coronary artery disease (cad) patients using optimized svm. *Computer methods and programs in biomedicine*, 138:117–126.
- Ebersole, J. S. and Pedley, T. A. (2003). *Current practice of clinical electroencephalography*. Lippincott Williams & Wilkins.
- Edlinger, G. and Guger, C. (2006). Correlation changes of eeg and ecg after fast cable car ascents. In *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the*, pages 5540–5543. IEEE.
- Faust, O., Acharya, U. R., Adeli, H., and Adeli, A. (2015). Wavelet-based eeg processing for computer-aided seizure detection and epilepsy diagnosis. *Seizure*, 26:56–64.
- Han, M., Sun, L., and Hong, X. (2012). Extraction of the eeg signal feature based on echo state networks. *Sheng wu yi xue gong cheng xue za zhi= Journal of biomedical engineering= Shengwu yixue gongchengxue zazhi*, 29(2):206–211.
- Hyvärinen, A. and Oja, E. (2000). Independent component analysis: algorithms and applications. *Neural networks*, 13(4):411–430.
- Jurysta, F., Van De Borne, P., Migeotte, P.-F., Dumont, M., Lanquart, J.-P., Degaute, J.-P., and Linkowski, P. (2003). A study of the dynamic interactions between sleep eeg and heart rate variability in healthy young men. *Clinical neurophysiology*, 114(11):2146–2155.
- Kevric, J. and Subasi, A. (2017). Comparison of signal decomposition methods in classification of eeg signals for motor-imagery bci system. *Biomedical Signal Processing and Control*, 31:398–406.
- Kim, D.-K., Lee, K.-M., Kim, J., Whang, M.-C., and Kang, S. W. (2013). Dynamic correlations between heart and brain rhythm during autogenic meditation. *Frontiers in human neuroscience*, 7.
- Klem, G. H., Lüders, H. O., Jasper, H., Elger, C., et al. (1999). The ten-twenty electrode system of the international federation. *Electroencephalogr Clin Neurophysiol*, 52(3):3–6.

- Kumari, P., Kumar, S., and Vaish, A. (2014). Feature extraction using empirical mode decomposition for biometric system. In *Signal Propagation and Computer Technology (ICSPCT), 2014 International Conference on*, pages 283–287. IEEE.
- Kutlu, Y. and Kuntalp, D. (2012). Feature extraction for eeg heartbeats using higher order statistics of wpd coefficients. *Computer methods and programs in biomedicine*, 105(3):257–267.
- Lane, R., Reiman, E., Ahern, G., and Thayer, J. (2001). 21. activity in medial prefrontal cortex correlates with vagal component of heart rate variability during emotion. *Brain and Cognition*, 47(1-2):97–100.
- Lewis, N. G., McGovern, J. B., Miller, J. C., Eddy, D. R., and Forster, E. M. (1988). Eeg indices of g-induced loss of consciousness (g-loc). Technical report, SCHOOL OF AEROSPACE MEDICINE BROOKS AFB TX.
- Lin, C.-W., Wang, J.-S., and Chung, P.-C. (2010). Mining physiological conditions from heart rate variability analysis. *IEEE Computational Intelligence Magazine*, 5(1):50–58.
- Liou, L.-M., Ruge, D., Kuo, M.-C., Tsai, J.-C., Lin, C.-W., Wu, M.-N., Hsu, C.-Y., and Lai, C.-L. (2014). Functional connectivity between parietal cortex and the cardiac autonomic system in uremics. *The Kaohsiung journal of medical sciences*, 30(3):125–132.
- McCraty, R., Atkinson, M., Tomasino, D., and Bradley, R. T. (2009). The coherent heart brain interactions, psychophysiological coherence, and the emergence of system-wide order. *Integral Review: A Transdisciplinary & Transcultural Journal for New Thought, Research, & Praxis*, 5(2).
- Mirsadeghi, M., Behnam, H., Shalhaf, R., and Moghadam, H. J. (2016). Characterizing awake and anesthetized states using a dimensionality reduction method. *Journal of Medical Systems*, 40(1):1.
- Miyashita, T., Ogawa, K., Itoh, H., Arai, Y., Ashidagawa, M., Uchiyama, M., Koide, Y., Andoh, T., and Yamada, Y. (2003). Spectral analyses of electroencephalography and heart rate variability during sleep in normal subjects. *Autonomic Neuroscience*, 103(1):114–120.
- Mporas, I., Tsirka, V., Zacharaki, E. I., Koutroumanidis, M., Richardson, M., and Megalooikonomou, V. (2015). Seizure detection using eeg and ecg signals for computer-based monitoring, analysis and management of epileptic patients. *Expert Systems with Applications*, 42(6):3227–3233.
- Mumtaz, W., Xia, L., Yasin, M. A. M., Ali, S. S. A., and Malik, A. S. (2017). A wavelet-based technique to predict treatment outcome for major depressive disorder. *PloS one*, 12(2):e0171409.
- Nasehi, S. and Pourghassem, H. (2011). Real-time seizure detection based on eeg and ecg fused features using gabor functions. In *Intelligent Computation and Bio-Medical Instrumentation (ICBMI), 2011 International Conference on*, pages 204–207. IEEE.
- Prinsloo, G. E., Rauch, H. L., Karpul, D., and Derman, W. E. (2013). The effect of a single session of short duration heart rate variability biofeedback on eeg: a pilot study. *Applied psychophysiology and biofeedback*, 38(1):45–56.
- Sakai, S., Hori, E., Umeno, K., Kitabayashi, N., Ono, T., and Nishijo, H. (2007). Specific acupuncture sensation correlates with eegs and autonomic changes in human subjects. *Autonomic Neuroscience*, 133(2):158–169.
- Saper, C. B., Scammell, T. E., and Lu, J. (2005). Hypothalamic regulation of sleep and circadian rhythms. *Nature*, 437(7063):1257.
- Stoica, P. and Moses, R. L. (1997). *Introduction to spectral analysis*, volume 1. Prentice hall Upper Saddle River, NJ.
- Stuss, D. T. and Benson, D. F. (1986). *The frontal lobes*. Raven Pr.
- Sudarshan, V. K., Acharya, U. R., Oh, S. L., Adam, M., Tan, J. H., Chua, C. K., Chua, K. P., and San Tan, R. (2017). Automated diagnosis of congestive heart failure using dual tree complex wavelet transform and statistical features extracted from 2s of ecg signals. *Computers in Biology and Medicine*, 83:48–58.
- Takahashi, T., Murata, T., Hamada, T., Omori, M., Kosaka, H., Kikuchi, M., Yoshida, H., and Wada, Y. (2005). Changes in eeg and autonomic nervous activity during meditation and their association with personality traits. *International Journal of Psychophysiology*, 55(2):199–207.
- Thomas, M., Das, M. K., and Ari, S. (2015). Automatic eeg arrhythmia classification using dual tree complex wavelet based features. *AEU-International Journal of Electronics and Communications*, 69(4):715–721.
- Thomas, P. and Moni, R. (2016). Methods for improving the classification accuracy of biomedical signals based on spectral features. *Technology*, 7(1):105–116.
- Triggiani, A. I., Valenzano, A., Del Percio, C., Marzano, N., Soricelli, A., Petito, A., Bellomo, A., Başar, E., Mundi, C., Cibelli, G., et al. (2016). Resting state rolandic mu rhythms are related to activity of sympathetic component of autonomic nervous system in healthy humans. *International Journal of Psychophysiology*, 103:79–87.
- Unser, M. and Aldroubi, A. (1996). A review of wavelets in biomedical applications. *Proceedings of the IEEE*, 84(4):626–638.
- Valderrama, M., Alvarado, C., Nikolopoulos, S., Martinerie, J., Adam, C., Navarro, V., and Le Van Quyen, M. (2012). Identifying an increased risk of epileptic seizures using a multi-feature eeg–ecg classification. *Biomedical Signal Processing and Control*, 7(3):237–244.
- Yang, C. C., Lai, C.-W., Lai, H. Y., and Kuo, T. B. (2002). Relationship between electroencephalogram slow-wave magnitude and heart rate variability during sleep in humans. *Neuroscience letters*, 329(2):213–216.



Postproceedings of the 9th Annual International Conference on Biologically Inspired Cognitive Architectures, BICA 2018 (Ninth Annual Meeting of the BICA Society)

## Efficient Methods for Calculating Sample Entropy in Time Series Data Analysis

Ronakben Bhavsar<sup>a,\*</sup>, Na Helian<sup>a</sup>, Yi Sun<sup>a</sup>, Neil Davey<sup>a</sup>, Tony Steffert<sup>b</sup>, David Mayor<sup>a</sup>

<sup>a</sup>University of Hertfordshire, Hatfield, AL10 9AB, UK

<sup>b</sup>The Open University, Milton Keynes, MK7 6AA, UK

### Abstract

Recently, different algorithms have been suggested to improve Sample Entropy (SE) performance. Although new methods for calculating SE have been proposed, so far improving the efficiency (computational time) of SE calculation methods has not been considered. This research shows such an analysis of calculating a correlation between Electroencephalogram (EEG) and Heart Rate Variability (HRV) based on their SE values. Our results indicate that the parsimonious outcome of SE calculation can be achieved by exploiting a new method of SE implementation. In addition, it is found that the electrical activity in the frontal lobe of the brain appears to be correlated with the HRV in a time domain.

© 2019 The Authors. Published by Elsevier B.V.

This is an open access article under the CC BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0/>)

Peer-review under responsibility of the scientific committee of the 9th Annual International Conference on Biologically Inspired Cognitive Architectures.

**Keywords:** Sample Entropy; EEG; HRV; Time Series Data Analysis; Pearson Correlation.

### 1. Introduction and Related Work

Nonlinear dynamic analyses have been widely used to study the complex behaviours and different structures of biological systems [2]. Nonlinear dynamic analysis proves to be a powerful approach for the assessment of different physiological time series as it can determine the hidden patterns related to underlying mechanisms [3] [13]. The chaotic behaviour of cardiac system and brain waves indicate nonlinearity [1]. With the given nature of nonlinearity, Electroencephalogram (EEG) and Heart Rate Variability (HRV) turn out to be appropriate for nonlinear time series analysis [1]. The different types of nonlinear complex measures of variability are Lyapunov exponent, Correlation Dimension D2, Approximate Entropy (AE), Sample Entropy (SE), Multiscale entropy (MSE), Poincare plots, Detrended Fluctuation Analysis (DFA) and many more.

\* Corresponding author. Tel.: +44-757-436-9688.

E-mail address: [r.bhavsar2@herts.ac.uk](mailto:r.bhavsar2@herts.ac.uk)

SE has been used widely to investigate various biological conditions in human body like arrhythmia studied through ECG (Electrocardiogram) [3], Alzheimer's patients' EEG background activity [1], analyzing human postural sway data [12] and studying HRV in the case of obstructive sleep apnoea syndrome [2]. SE is also used to detect the termination of a particular medical condition like seizures [15] and to test the effect of a therapy like ketogenic diet used for controlling intractable seizures [14]. These studies have concluded that SE is robust quantifier of complexity, which offers an accurate nonlinear metric for quantification [3]. It gives a good dynamical signature and is a helpful tool that provides insights into various biological time series [1],[12]. Therefore, SE is considered as an effective method for investigating different types of time series data.

In recent years, different algorithms attempting to improve SE have been proposed. Quadratic Sample Entropy (QSE) was introduced to reduce the influence of arbitrary constants of sequence comparison and tolerance on SE, as well as to reduce the skewing of results when either the top or the bottom of the conditional probabilities was very small or very large [9]. Another attempt to improve SE was with the introduction of Fuzzy Entropy (FuzzyEn) [5], using the concept of fuzzy sets in order to determine a fuzzy measurement of similarity of two vectors based on their shapes. Multi Scale Entropy (MSE) established by [14], was a useful extension of SE to multiple time scales, in recognition of the likelihood that dynamical complexity of biological signals may operate across a range of temporal scales.

In this work, the type of nonlinear complex measures of variability exploited is Sample Entropy (SE). The aim of the research is not to propose another new method derived from SE, but efficient method improving the computational time for the SE calculation. The computational time for SE calculation using the new and original SE methods will be compared on calculating the correlation between SE values of EEG and HRV in time domain.

## 2. Dataset Information

Our datasets consist of EEG and ECG recordings from 15 participants. This data was obtained over 5 minute time slots in a relaxed state with eyes opened. 19 electrodes (Fp1, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, and O2) for EEG recording were used, following the standard 10%-20% system [7], as shown in Fig.1. The values of 10% and 20% shown in Fig.1. refer to the distances between adjacent electrodes: either 10% or 20% percentage of the total front-to-back or right-to-left distance over the skull - front-to-back distance is based on the measurement from the Nasion (point between forehead and nose) to Inion (lowest point of the skull from the back of the head indicated by a prominent bump), and right-to-left distance is based on the measurement between the left and right pre-auricular points.

The sampling rate used for EEG was 250Hz, and the reference was linked to ear electrodes. For ECG data, one electrode was positioned on the volar surface of each forearm (with an additional electrode as ground on the dominant side) to record the electrical activity of heart over time, and the sampling rate was 256Hz.

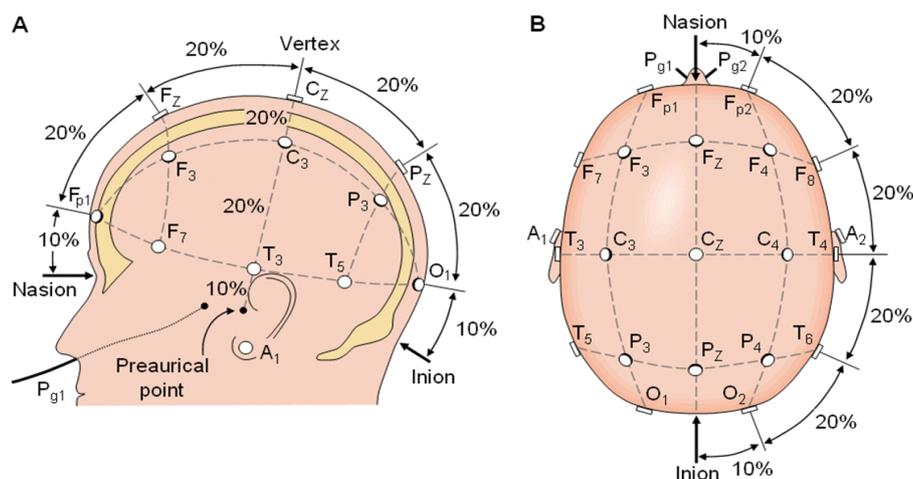


Figure 1. The international 10%-20% system seen from A (left side of the head) and B (above the head). The letter F, T, C, P, O, A, Fp and Pg stands for frontal, temporal, central, parietal, occipital, earlobes, frontal polar, and nasopharyngeal, respectively. The figure is obtained from [7].

### 3. Sample Entropy and Proposed Implementation

SE is considered as an effective method for investigating different types of time series data. A lower SE value indicates a high frequency of similarity in time series [13].

For a time series of length  $n$ ,  $SampleEntropy(m, r, n)$  can be defined as the negative logarithm of conditional probability that two sequences are similar for  $m$  point [10] within a tolerance value  $r$ , excluding any self-matches [13]. The equation can be represented according to [13], as:

$$SE(m, r, n) = -\ln\left(\frac{A}{B}\right), \tag{1}$$

where,  $m$  is the length of sequences to be compared,  $r$  is the tolerance value for accepting matches,  $n$  is the length of original data, and A and B are defined as follow:

$$A = \left\{ \frac{[(n-m-1)(n-m)]}{2} \right\} A^m(r), \text{ and } B = \left\{ \frac{[(n-m-1)(n-m)]}{2} \right\} B^m(r)$$

where,  $A^m(r)$  is the probability that the two sequences match for  $m + 1$  points, and  $B^m(r)$  is the probability that the two sequences match for  $m$  points. Each SE value indicates relative consistency with respect to any value of  $(m, r)$ . That is, if a record has a lower SE value than another record for a part of fixed  $m$  and  $r$  values, it will be lower for any part of fixed  $m$  and  $r$  values [10]. SE is independent of the data length and shows an elimination of self-matching. In order to approximate the conditional probabilities of matches, SE uses a point-wise approach [13].

#### 3.1. An Example of SE Calculation

In this section, we explain how the SE is calculated in practise by giving an example with a simple time series. Let the input time series be  $x(n) = \{0.1, 0.1, 0.2, 0.5, 0.22\}$ , with  $m = 2, r = 0.2, n = 5$ .

The value of  $m$  specifies the length of sequences to be considered for SE. Thus, the default value of  $m$  is 2 (i.e. the maximum length of sequence considered is 2). The value of  $r$  represents the tolerance value below that a match is deemed. The input point sequence for  $A^m(r)$  is  $n$  points, while  $B^m(r)$  considers  $n - 1$  points of the input sequences. That means, for  $A^m(r)$  and  $B^m(r)$  the input point sequence is  $\{0.1, 0.1, 0.2, 0.5, 0.22\}$  and  $\{0.1, 0.1, 0.2, 0.5\}$ , respectively.

As shown in Table 1, to calculate the probability for A and B, the number of matches obtained for respective sequences of  $A^m(r)$  and  $B^m(r)$  is counted as “1”. For a particular length of sequence  $m$ , the point matches are calculated by calculating the absolute difference between the points in the sequences. The difference should be below the tolerance value  $r$  (in this case 0.2). The calculation of similar segments can be summarized as:  $|x(i) - x(j)| < r$ . Considering the following sequence  $(x_k(i), x_k(j)) = [(0.1, 0.1), (0.2, 0.5)]$ , where  $i$  and  $j$  are the point sequence, and  $k$  is the index for these point sequences. To test the match,  $(|0.1 - 0.2|, |0.1 - 0.5|) = (0.1, 0.4)$  is calculated. It can be observed that  $x_1(i)$  and  $x_1(j)$  (i.e. 0.1 and 0.2) satisfy the condition but  $x_2(i)$  and  $x_2(j)$  (i.e. 0.1 and 0.5) do not satisfy the condition because the absolute difference is greater than the tolerance value  $r$ . Since the point sequence is not a complete match under the tolerance value  $r$ , this sequence is not considered as a match.

Table 1 represents the point sequence match at a given length of sequence for  $m(0$  to 2) for the tolerance value  $r(0.2)$ . In third and fourth columns, “1” represents a match and “0” represents no match at tolerance value  $r$ .

According to equation (1), SE value can be calculated as follows:

$$SE(0, 0.2, 5) = p(0) = -\ln\left(\frac{A[0]}{((n*n-1)/2)}\right) = -\ln(6/10) = 0.5108$$

$$SE(1, 0.2, 5) = p(1) = -\ln\left(\frac{A[1]}{B[0]}\right) = -\ln(1/3) = 1.0986$$

$$SE(2, 0.2, 5) = p(2) = -\ln\left(\frac{A[2]}{B[1]}\right) = -\ln(0/0) = \text{inf}$$

From the SE values obtained from the above examples, it can be seen that a low SE value is obtained at  $m = 0$ , which increases with the increase of  $m$ . This indicates that for a longer point sequence, the similarity decreases for this time series.

Table 1. Point sequences at  $m = 0; 1; \text{ and } 2$  along with the count of match obtained for  $A^m(r)$  and  $B^m(r)$ . Here, X represent that the point sequence was not considered for  $B(m)$ . Columns  $A(m)$  and  $B(m)$  indicates count for the total number of matches obtained.

Length of sequence for $m$	Point Sequences	Point Matches at $r = 0.2$			
		$A^m(r)$	$B^m(r)$	$A(m)$	$B(m)$
$m=0$	[0.1, 0.1]	1	1	6	3
	[0.1, 0.2]	1	1		
	[0.1, 0.5]	0	0		
	[0.1, 0.22]	1	X		
	[0.1, 0.2]	1	1		
	[0.1, 0.5]	0	0		
	[0.1, 0.22]	1	X		
	[0.2, 0.5]	0	0		
	[0.2, 0.22]	1	X		
	[0.5, 0.22]	0	X		
$m=1$	[(0.1, 0.1),(0.1, 0.2)]	(1, 1)	(1, 1)	1	1
	[(0.1, 0.1),(0.2, 0.5)]	(1, 0)	(1, 0)		
	[(0.1, 0.1),(0.5, 0.22)]	(1, 0)	X		
	[(0.1, 0.2),(0.2, 0.5)]	(1, 0)	(1, 0)		
	[(0.1, 0.2),(0.5, 0.22)]	(1, 0)	X		
	[(0.2, 0.5),(0.5, 0.22)]	(0, 0)	(0, 0)		
$m=2$	[(0.1, 0.1, 0.2),(0.1, 0.2, 0.5)]	(1, 1, 0)	(1, 1, 0)	0	0
	[(0.1, 0.1, 0.2),(0.2, 0.5, 0.22)]	(1, 0, 1)	X		
	[(0.1, 0.2, 0.5),(0.2, 0.5, 0.22)]	(1, 0, 0)	(1, 0, 0)		

### 3.2. Efficient and Parsimonious way for Sample Entropy Calculation

SE measures the probabilities of matches for a time series data using point-wise approach. This can be time consuming when long sequence of points need to be compare, and can be done more efficiently. Computation time for SE can be reduced without losing much information from the signals by using the three methods proposed in this section. Thus, calculation time for SE could be shortened and the computational expense would be more cheaper. Fig.2. illustrates on how these three methods work.

**SE-Method 1** is about shortening the time series signal without loss of too much information for the point-wise approach. For example, instead of considering the original data length( $n$ ) of the 5-minute signal ( $=75000$  data points), it could be shortened by dividing 1.1 on the original data length ( $75000/1.1=68181$  equivalent to 28 seconds data points). Binary chop [8] is performed in order to find out at which point the most accurate result for SE could be obtained.

**SE-Method 2** is about SE calculation on a moving window, calculating SE on individual windows to find out which window gives the SE values that are most similar to the original SE value. For example, using a 2 seconds moving window, SE is calculated for a window size of 500 points ( $2*(250\text{Hz}) = 500$  data points).

**SE-Method 3** is to calculate the mean for a given window first before performing SE. This data window could be as long as a minute or as short as a second. For example, if the mean of each 1 sec data (250 points) is gathered, then it will give us a reduced length of  $n = 300$  data points on which to perform SE calculation, and not  $n = 75000$ . This way the SE computational time should be reduced dramatically.

## 4. Experimental Results

The EEG signals were pre-processed to remove artefacts caused by electrical activity in muscles, including of the eye, jaw and other muscle movements using Independent Component Analysis (ICA) as mentioned in [6]. It is relatively straightforward to remove these artefacts using ICA.

To extract HRV from ECG signals, the method designed by Lin et al. [11] was adopted. The results of the automatic analysis were reviewed and then any errors in ECG R-wave detection and QRS complex (combination of three graph-

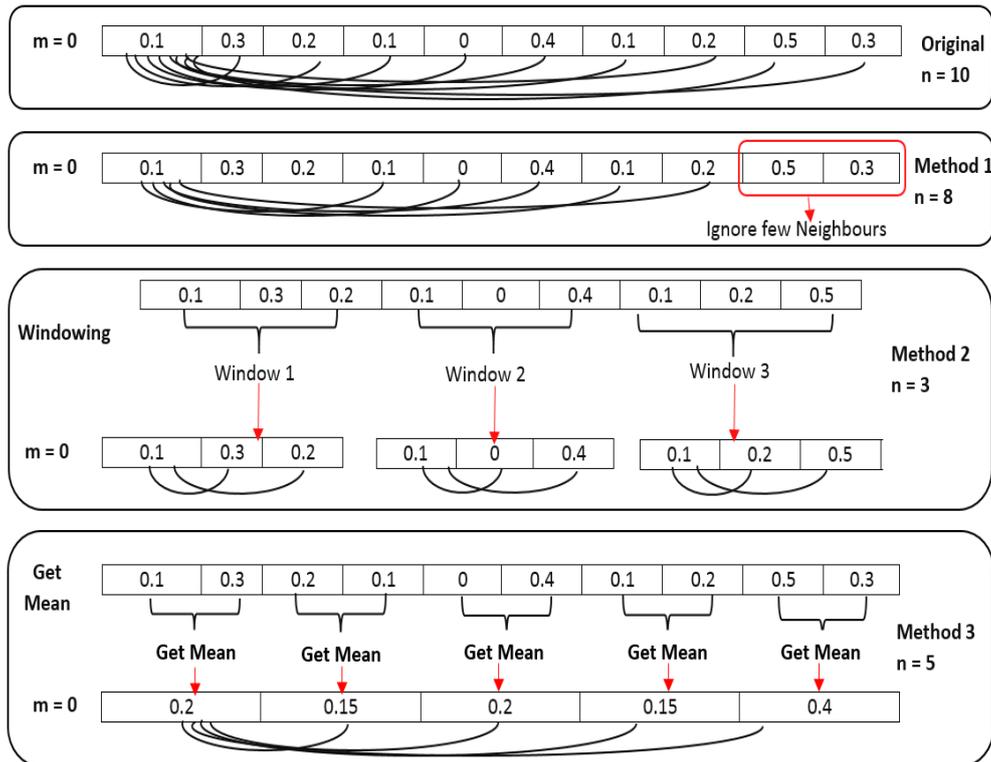


Figure 2. An example of how the SE can be calculated efficiently.

ical deflections: Q, R and S waves) labelling were removed manually. R-R interval data obtained from the edited time sequence of R-wave and QRS labelling was then transferred to a personal computer. In order to remove artefacts from extracted HRV signal, each R-R interval was compared against a local average interval. If a R-R interval differs from the local average more than a specified threshold (in seconds) value, then that R-R interval is defined as an artefact and is replaced with an interpolated value using a cubic spline interpolation.

#### 4.1. Experiments using three proposed SE Calculation Methods

For each of the five minutes of EEG data; the following three experiments have been undertaken, results are shown in Table 2. For the purpose of comparison, SE values of the original SE performance is also shown. All code is run on a personal computer: Windows 7 Enterprise, Intel(R) Core (TM) i7-3770T, 64-bit Operating System, 16 GB RAM.

**Experiment 1** is the implementation of SE-Method 1, by restricting number of neighbours for comparisons on SE calculation. It is found that ignoring the last 25 to 30 seconds of data still achieves as accurate results as if they are included, but with improving computational time by 10 seconds.

**Experiment 2** is about experimenting SE-Method 2, 10 different window size are considered (i.e. 2 Sec, 10 Sec, 20 Sec, 30 Sec, 40 Sec, 50 Sec, 60 Sec, 70 Sec, 80 Sec and 90 Sec windows) on which to perform the SE calculation, to find out which window gives the SE values that are most similar to the original SE value, as shown in the Table 2. It is found that the smaller the window size, the shorter the calculation time. In addition, the most similar results to the original SE calculation results is the smallest window size.

**Experiment 3** is demonstrating SE-Method 3, calculating the mean of each window of a second of data (i.e 1 Sec= mean of 250 points). The experiment is done with 8 different window sizes (i.e. 0.06 Sec, 0.12 Sec, 0.25 Sec, 0.55 Sec, 1 Sec, 2 Sec, 3 Sec and 4 Sec) on which to calculate the mean, as shown in Table 2. The SE is then performed on the mean values of the signal. It is found that bigger the window size, the shorter the calculation time. Moreover, the best match to the original SE calculation results is at the mean of each 1 seconds window.

Experiment 1-3 demonstrated a strong positive correlation between the results obtained using the original and each of the new three SE approaches with the correlation values of 0.99, 0.68, and 0.96 for SE-Method 1, SE Method 2 and

Table 2. Computation time for the SE calculation using the Original approach and our three Experimental methods.

Experiments	Details	Computation Time
SE-Original	Original Performance	75 Sec
SE-Method 1	Shortening the neighbour comparison	62 Sec
SE-Method 2	2 Seconds Moving Window	0.002 Sec
	10 Seconds Moving Window	0.08 Sec
	20 Seconds Moving Window	0.38 Sec
	30 Seconds Moving Window	0.72 Sec
	40 Seconds Moving Window	1.30 Sec
	50 Seconds Moving Window	1.94 Sec
	60 Seconds Moving Window	3 Sec
	70 Seconds Moving Window	4 Sec
	80 Seconds Moving Window	6 Sec
90 Seconds Moving Window	9 Sec	
SE-Method 3	Mean of Each 0.06 Seconds Window	0.38 Sec
	Mean of Each 12 Seconds Window	0.11 Sec
	Mean of Each 25 Seconds Window	0.02 Sec
	Mean of Each 50 Seconds Window	0.007 Sec
	Mean of Each 1 Seconds Window	0.003 Sec
	Mean of Each 2 Seconds Window	0.008 Sec
	Mean of Each 3 Seconds Window	0.01 Sec
	Mean of Each 4 Seconds Window	0.02 Sec

SE-Method 3, respectively, along with the probability of 0. Whilst SE-Method 1 and SE-Method 2 do not improve the trend and the computational time for SE calculation, SE-Method 3 clearly works best because it provides the most predictive value, and trend for SE performance to those provided by original SE performance with improving computational time.

#### 4.2. Experiment 4

The aim of experiment 4 is to compare the performance of new methods and original methods for SE calculation. The previous three experiments shows that SE-Method 3 is the best one in terms of improving SE calculation time without losing much information. Hence, Only SE-Method 3 is considered for this experiment because correlation coefficient works on similar length of the signals, and SE-Method 3 gives us the same length of samples for EEG as of HRV. In order to demonstrate correlation between EEG and HRV, for each of the five minutes of EEG and HRV data; the following steps have been undertaken for both original and the new approach (SE-Method 3) of SE calculation.

1. For each electrode's data, divide data into 10 equal bins to perform SE calculation. This process has been repeated for each participant.
2. Compute correlation coefficients on the 10 SE values of the EEG and 10 SE values of the HRV obtained in step 1.

Pearson's correlation is used to perform the correlation coefficients. It measures how closely two different observables are related to each other. Pearson's correlation co-efficient  $R$  ranges between 1 (when the matching entities are exactly the same) and  $-1$  (when the matching entities are inverses of each other). A value of zero indicates no relationship existing between the entities.

Once the electrode's SE correlation performance are gathered, the best performance electrodes have been ranked-where, the ranking has been given based on electrode correlation values, the bigger the value, the higher the rank. The three best performance electrodes' results have been looked closely. Some common electrode rankings are found for all the participants investigated, highlighted in yellow colour, as shown in the Fig.3.

It is found that the electrical activity in the frontal lobe of the brain appears to be correlated with the HRV in time domain. Moreover, the new approach (SE-Method 3) of SE calculation is giving more focused result than the original SE calculation.

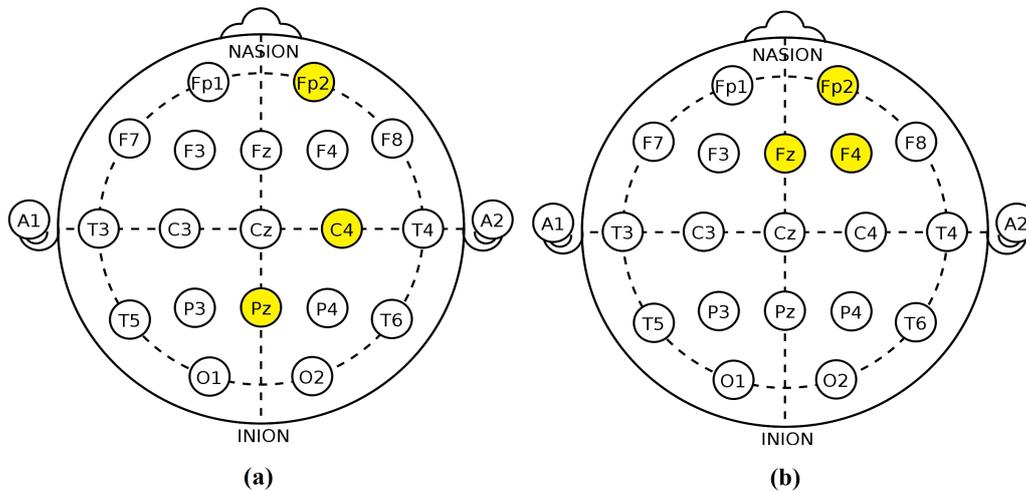


Figure 3. Electrode Ranking based on correlation performance between SE values of EEG and HRV, showing best three performing electrodes across participants, highlighted in yellow colour: (a) Ranking based on the original approach for SE calculation, and (b) Ranking based on the new approach (SE-Method 3) for SE calculation.

## 5. Discussion and Conclusion

The main conclusion of this work is that parsimonious results for SE can be achieved using the proposed new methods of pre-processing the data prior to SE calculation. SE-Method 3 clearly works best for improving the performance because it gives good predictive values without changing the trends visible in SE calculated using the original standard approach. SE-Method 1 provides SE values very close to those obtained using the original SE approach, but it does not improve computational time much. Similarly, SE-Method 2 is not robust because neither the trends nor computational time are improved significantly.

The second conclusion from this work is that there is a strong positive correlation ( $R=0.96$ , Probability = 0) between results obtained using the original and the new (SE-Method 3) SE approaches. Also, we found low positive correlations between SE values of EEG and HRV in the time domain. The results shown in our previous work suggested that the electrical activity in the frontal lobe of the brain is correlated with the HRV. It shows that the electrical activity in the frontal lobe of the brain appears to be correlated with the HRV in time domain, which is in consistent with our previous finding on frequency domain in paper [4].

In summary, SE-Method 1 and SE-Method 2 do not improve the trend or the computational time for SE calculation. SE-Method 3 does not give values similar to those provided by the original SE approach, but it does provide the most predictive value for SE performance. Although the result is not exactly similar as the original SE performance, the trend is. Therefore, we believe the most efficient way for SE calculation is SE-Method 3.

## References

- [1] D Abásolo, R Hornero, P Espino, D Alvarez, and J Poza. Entropy analysis of the eeg background activity in alzheimer's disease patients. *Physiological measurement*, 27(3):241, 2006.
- [2] Haitham M Al-Angari and Alan V Sahakian. Use of sample entropy approach to study heart rate variability in obstructive sleep apnea syndrome. *Biomedical Engineering, IEEE Transactions on*, 54(10):1900–1904, 2007.
- [3] Raul Alcaraz and Jose Joaquin Rieta. A review on sample entropy applications for the non-invasive analysis of atrial fibrillation electrocardiograms. *Biomedical Signal Processing and Control*, 5(1):1–14, 2010.
- [4] Ronakben Bhavsar, Neil Davey, Yi Sun, and Na Helian. An investigation of how wavelet transform can affect the correlation performance of biomedical signals. 2018.

- [5] Weiting Chen, Zhizhong Wang, Hongbo Xie, and Wangxin Yu. Characterization of surface emg signal based on fuzzy entropy. *IEEE Transactions on neural systems and rehabilitation engineering*, 15(2):266–272, 2007.
- [6] Aapo Hyvärinen and Erkki Oja. Independent component analysis: algorithms and applications. *Neural networks*, 13(4):411–430, 2000.
- [7] George H Klem, Hans Otto Lüders, HH Jasper, C Elger, et al. The ten-twenty electrode system of the international federation. *Electroencephalogr Clin Neurophysiol*, 52(3):3–6, 1999.
- [8] Donald E Knuth. Sorting and searching, 2nd edn. the art of computer programming, vol. 3, 1998.
- [9] DE Lake. Improved entropy rate estimation in physiological data. In *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, pages 1463–1466. IEEE, 2011.
- [10] Douglas E Lake, Joshua S Richman, M Pamela Griffin, and J Randall Moorman. Sample entropy analysis of neonatal heart rate variability. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, 283(3):R789–R797, 2002.
- [11] Che-Wei Lin, Jeen-Shang Wang, and Pau-Choo Chung. Mining physiological conditions from heart rate variability analysis. *IEEE Computational Intelligence Magazine*, 5(1):50–58, 2010.
- [12] Sofiane Ramdani, Benoît Seigle, Julien Lagarde, Frédéric Bouchara, and Pierre Louis Bernard. On the use of sample entropy to analyze human postural sway data. *Medical engineering & physics*, 31(8):1023–1031, 2009.
- [13] Joshua S Richman and J Randall Moorman. Physiological time-series analysis using approximate entropy and sample entropy. *American Journal of Physiology-Heart and Circulatory Physiology*, 278(6):H2039–H2049, 2000.
- [14] Tetsuya Takahashi, Raymond Y Cho, Tomoyuki Mizuno, Mitsuru Kikuchi, Tetsuhito Murata, Koichi Takahashi, and Yuji Wada. Antipsychotics reverse abnormal eeg complexity in drug-naive schizophrenia: a multiscale entropy analysis. *Neuroimage*, 51(1):173–182, 2010.
- [15] Cheol Seung Yoo, Dong Chung Jung, Yong Min Ahn, Yong Sik Kim, Su-Gyeong Kim, Hyeri Yoon, Young Jin Lim, and Sang Hoon Yi. Automatic detection of seizure termination during electroconvulsive therapy using sample entropy of the electroencephalogram. *Psychiatry research*, 195(1):76–82, 2012.



Postproceedings of the 9th Annual International Conference on Biologically Inspired Cognitive Architectures, BICA 2018 (Ninth Annual Meeting of the BICA Society)

# Time Series Analysis using Embedding Dimension on Heart Rate Variability

Ronakben Bhavsar<sup>a,\*</sup>, Neil Davey<sup>a</sup>, Na Helian<sup>a</sup>, Yi Sun<sup>a</sup>, Tony Steffert<sup>b</sup>, David Mayor<sup>a</sup>

<sup>a</sup>University of Hertfordshire, Hatfield, AL10 9AB, UK

<sup>b</sup>The Open University, Milton Keynes, MK7 6AA, UK

## Abstract

Heart Rate Variability (HRV) is the measurement sequence with one or more visible variables of an underlying dynamic system, whose state changes with time. In practice, it is difficult to know what variables determine the actual dynamic system. In this research, Embedding Dimension (ED) is used to find out the nature of the underlying dynamical system. False Nearest Neighbour (FNN) method of estimating ED has been adapted for analysing and predicting variables responsible for HRV time series. It shows that the ED can provide the evidence of dynamic variables which contribute to the HRV time series. Also, the embedding of the HRV time series into a four-dimensional space produced the smallest number of FNN. This result strongly suggests that the Autonomic Nervous System that drives the heart is a two features dynamic system: sympathetic and parasympathetic nervous system.

© 2019 The Authors. Published by Elsevier B.V.

This is an open access article under the CC BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0/>)

Peer-review under responsibility of the scientific committee of the 9th Annual International Conference on Biologically Inspired Cognitive Architectures.

**Keywords:** Time series analysis; HRV; Embedding Dimension; False Nearest Neighbours; Parasympathetic; Sympathetic; Linear Regression.

## 1. Introduction and Related Work

Heart Rate Variability (HRV) can be measured using Electrocardiography (ECG). ECG records the electrical activities of the heart, where each beat of the heart is initiated by an electric signal from the heart muscle (Vague). HRV is the estimation of neurocardiac function that reflects heart-brain interactions and autonomic nervous system dynamics [12]. The measurement of HRV is a valuable investigative tool in clinical cardiology as it gives a fundamental method to evaluate the physiological state of the heart directly. Many neurological and psychological investigations have used HRV to assess the effects of stress, emotion, and work on the autonomic nervous system [11]. The heart rate and rhythm are mainly under the control of the Autonomic Nervous System (ANS) and is the part of the Peripheral

\* Corresponding author. Tel.: +44-757-436-9688.

E-mail address: [r.bhavsar2@herts.ac.uk](mailto:r.bhavsar2@herts.ac.uk)

Nervous System (PNS) that acts as a control system that functions mostly below the level of consciousness to control physical functions. ANS contains two primary role on components: Sympathetic and Parasympathetic Nervous system. Both the sympathetic and parasympathetic nervous systems innervate the heart. The parasympathetic nervous system functions in regulating heart rate through the vagus nerve, with increased vagal activity producing a slowing of heart rate. The sympathetic nervous system has an excitatory influence on heart rate and contractility, and it serves as the final common pathway for controlling the smooth muscle tone of the blood vessels [15].

Time series, such as HRV, is the measurements sequence of one or more visible variables of an underlying dynamic system, whose state changes with time. These time series will be the results of the interaction of many underlying variables. For example, a stock market is affected by many interacting factors, such as economic data, exchange rates and so on. In practice, it is difficult to know what variables determine the actual dynamic system. It is shown by [16], if only one scalar value can be measured from an active system, then by windowing a sufficient number of consecutive values, the nature of the original multivariable dynamic system can be recaptured. In fact, [16] also mentioned if the original dynamic system had a dimension of  $N$ , then an embedding of size  $2N$  will be fully regained the original system. The size of this window is called the Embedding Dimensions (ED) [1]. ED estimation has successfully used in neural network approaches for time series prediction [7]. They concluded that optimal performance could be achieved using the correct ED. Furthermore, ED has been adopted by [18], for recurrence plot generation from the reconstructed phase space to represent many real application scenarios when not all variables to describe a system were available. The number of independent variables sufficient for modeling the hair cell response has been estimated utilizing ED approach [6]. Moreover, ED has been considered in a multilayer perceptron neural network to measure hyperchaotic Rssler system state variables [2].

The ED plays a vital role in nonlinear time series analysis[5], as discussed earlier. With its extensive use of finding the nature of an underlying dynamical system, ED is used in this work for HRV time series analysis. The False Nearest Neighbours (FNN) method of estimating ED has been adapted for analysing and predicting variables responsible for the HRV time series. HRV time series taken from participants over a fixed period.

## 2. Dataset Information

The HRV time series used in this work are of the data where participants undergoing acupuncture in the experimental settings. Three different methods of acupuncture have been used: Electro-Acupuncture(EA), Transcutaneous Electro-Acupuncture (TEA) and Manual Acupuncture (MA). EA is a method of inserting needles at specific points on the body. The needles have then been connected to a device that generates continuous electric pulses. These devices are used to adjust the frequency and intensity of the delivered impulse [3]. TEA is a safe, standardized acupuncture technique in which there is no needle insertion. It involves applying cutaneous electrical stimulation by placing electrodes at classical Chinese acupoints [14]. The electrodes (patches) are attached to the participant's skin when the unit is switched on; mild electrical current travels through the electrodes wires into the body. MA is an acupuncture method similar to EA, in which needles have inserted at specific points on the body. These needles are then twisted by or otherwise manipulated by the acupuncturist instead of passing electric pulses through the needle [17].

### 2.1. Dataset 1

This dataset consists of HRV data of 7 participants, derived using TEA method of acupuncture. HRV monitoring was carried out in nine 5-minute slots: three baseline slots and six acupuncture stimulation slots. The stimulation parameters (e.g., body location) are kept constant within each intervention but varies between interventions. Each participant visited twice, during which the TEA stimulation of either 2.5Hz or 10Hz is applied (randomised order used) at six different body locations (Slot 3 to 8) with eyes closed. The baseline measurements are slots 1, 2, and 9.

1. No stimulation with Eyes Closed.
2. No stimulation with Eyes Open.
3. TEA stimulation on Left Hand and below Left Knee.
4. TEA stimulation on Right Hand and below Right Knee.
5. TEA stimulation on 3. and 4. together.

6. TEA stimulation on Upper Body (Left and Right Hands).
7. TEA stimulation on Lower Body (below Left and Right Knees).
8. TEA stimulation on 6. and 7. Together.
9. No stimulation with Eyes Open.

## 2.2. Dataset 2

This dataset consists of HRV data of 12 participants, derived using both EA and MA method of acupuncture. All participants attended for four visits, during which stimulation performed at a different location (in randomized order): Right (Below Right Knee and Right Hand), Left (below Left Knee and Left Hand), Upper Body (Right and Left Hands) and Lower Body (below Left and Right Knees). HRV monitoring was carried out in eight 5-minute sequential slots with stimulation at a single location: EA stimulation of 2.5Hz, 10Hz, 20Hz and 80Hz is applied (Slot 3 to 6), MA stimulation applied in two slots (Slot 2 and Slot 7), and baseline measurements are slots 1 and slot 8.

1. No stimulation with Eyes Open 1
2. MA Stimulation 1
3. EA Stimulation at 2.5Hz
4. EA Stimulation at 10Hz
5. EA Stimulation at 20Hz
6. EA Stimulation at 80Hz
7. MA Stimulation 2
8. No stimulation with Eyes Open 2

## 3. Embedding Dimension

ED is used to find out the nature of an underlying dynamical system. The method FNN is used to determine how many dimensions are sufficient to embed a particular time series [8]. The FNN is designed to determine how many features are enough to integrate a specific time series [8]. The basic idea behind FNN is that points in a state space should be close to each other because their dynamical state is similar, not because they have been projected close to each other as an artefact of constructing the embedding with a dimension which is too low. In an embedding of dimension  $D$ , each point is established as a vector.

The FNN algorithm can be summarized as follows:

1. Find the nearest neighbour for each point in an embedding of dimension  $D$ ;
2. Find the percentage of those nearest neighbours which do not remain the nearest neighbour within embedding of dimension  $D+1$ , such points turns as false nearest neighbours;
3. Increase the embedding dimension until the number of false nearest neighbour is sufficiently small.

### 3.1. An Example of Embedding Dimension Calculation

In order to find the correct embedding dimension,  $n$ , an incremental search, from  $n = 1$ , is performed. A set of time lagged vectors  $x_n$ , for a given  $n$ , is formed. The nearest neighbour relation within the set of  $x_n$ 's is then computed. When the correct value of  $n$  has been reached, the addition of an extra dimension to the embedding should not cause these nearest neighbours to spring apart. Any pair whose additional separation is of a high relative size is deemed FNN. Specifically, if  $x_n$  has nearest neighbour  $\tilde{x}_n$ , then the relative additional separation when the embedding dimension is incremented is given by [1]:

$$FNN = \left| \frac{d(x_n, \tilde{x}_n) - d(x_{n+1}, \tilde{x}_{n+1})}{d(x_n, \tilde{x}_n)} \right|, \quad (1)$$

When this value exceeds an absolute value, then  $x_n$  and  $\tilde{x}_n$  are denoted as FNN. Where,  $x$  is the time series,  $n$  is the index for  $x$ , and  $d$  is the euclidean distance. In order to calculate nearness of neighbours, Euclidean Distance is used [10]. For example a time series is a sequence of values  $x_n(t)$ , where  $x$  is the time series,  $n$  is the index for  $x$ , and  $t$  represent time. Theoretically,  $x$  may be a value which varies continuously with  $t$ . An ED of 2 forms vectors  $(x_0, x_1), (x_1, x_2)$  and so on. An ED of 3 forms the vectors  $(x_0, x_1, x_2)$  and so on. Since this is numeric vectors, the distance apart of any pair of these vectors could be calculated. So for each vector in a given embedding, the nearest neighbour can be found. However, some of these nearest neighbour may be false neighbour, in a sense that they are not nearest neighbour in the embedding with one extra dimension [7]. A geometric explanation of the concept that is at the core of the FNN technique is as shown in Fig.1.

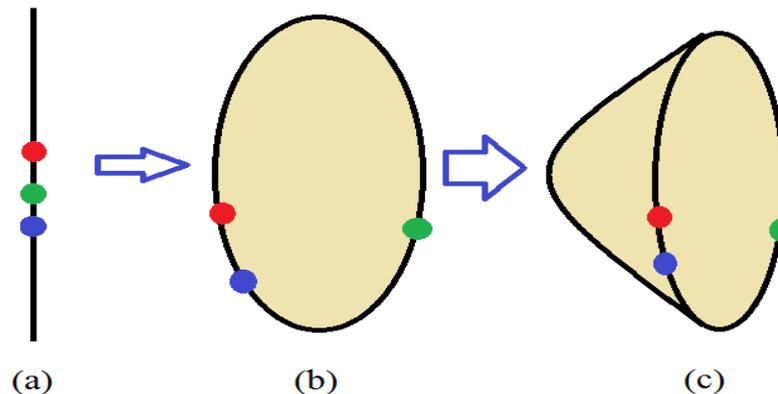


Figure 1. Geometric explanation of the FNN Algorithm [8]: (a) In one-dimensional, Red and Green Point are nearest neighbour, (b) In two-dimensional Red and Green point are not nearest neighbour (i.e. false nearest neighbour), but Red and Blue are nearest neighbour, and (c) In three-dimensional Red and Blue are still nearest neighbour, so they are real nearest neighbour.

The line at the bottom represents a dimensional state space  $X_1$  (Red Point),  $X_2$  (Green Point), and  $X_3$  (Blue Point) and the nearest neighbour of the  $X_1$  (Red Point) is the  $X_2$  (Green Point). Next, the time series embedded into two-dimensional state space, represented by the oval in the middle of the picture. The  $X_1$  (Red Point) and  $X_2$  (Green point) are no longer near to each other. So, the  $X_2$  (Green point) is labelled as a false nearest neighbour because it was only near to the  $X_1$  (Red Point) due to the projection of the time series onto the line.

Next, the nearest neighbour for each point in the two-dimensional state space found. Now the nearest neighbour to the  $X_1$  (Red Point) is the  $X_3$  (Blue Point). The time series is now embedded into a three-dimensional state space as represented in the rotated parabola at the top of the picture. The  $X_1$  (Red Point) and the  $X_3$  (Blue Point) are still near to each other, and so the  $X_3$  (Blue Point) is not a false nearest neighbour. This process continues until either there are no further malicious nearest neighbour or the data set becomes so sparse in a high dimensional space that no points can be considered to be near neighbours, to begin with. The resulting percentage of FNN for each ED is then plotted against the corresponding ED to create FNN plot.

### 3.2. An Example of Lorenz Attractor

The well known Lorenz Attractor as shown in Fig.2. has three underlying cross-coupled variables. However, the attractor itself is almost two dimensional. The minimum dimension of the attractor for a Lorenz data set as shown in Fig.3. is 4 or 5, which suggests that the actual underlying dimension system has features of around 2 [7].

## 4. Experiments and Results

### 4.1. HRV Analysis using ED

For analysing the HRV time series using the ED, for each participant in each 5-minute slot, the number of FNN has been calculated as the ED has increased. MATLAB code is used to gather ED performance of HRV [13]. Once, the

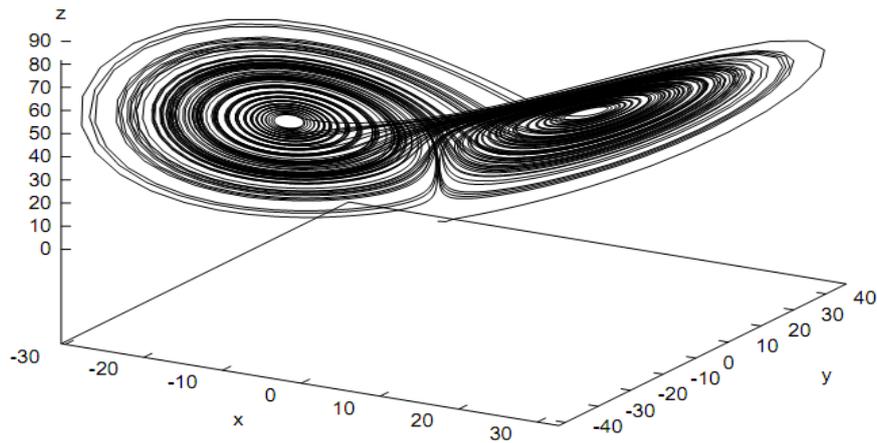


Figure 2. A visualisation of the Lorenz attractor in 3-dimensional phase space  $x(t), y(t), z(t)$  [7].

<i>Embedding Dimension</i>	<i>Percentage of False Nearest Neighbours</i>
	<i>Clean Data</i>
2	77%
3	3.3%
4	0.3%
5	0.3%

Figure 3. The percentage of false nearest neighbours in the Lorenz data set [7].

percentage of FNN for the ED is gathered for each 5-minute slot, the result is plotted for each participant, as shown in Fig.4. A comparison is made on the changes of various stimulus locations for each participant (i.e., between each slot’s result), and among the ED result of different participants. The first notable result for all participants is that the optimal ED is about 4. It is also noteworthy that this optimal ED is independent of stimulus location.

The ED results of 4 participants, two from each dataset (Participant 1 and Participant 2 from Dataset 1, and Participant 3 and Participant 4 from Dataset 2) as shown in Fig.4. For all the other participants from both datasets, results are similar to as shown in Fig.4. It is important to note that, some health-related problems found for few participants. For example, Participant 1 from dataset 1 has Thyroid insufficiency and menstrual irregularity, and participant 4 from dataset 2 have Asthma.

In the Fig.4., the X axes (Horizontally) represent ED from 1 to 10, and Y axes (Vertically) represent the percentage of FNN for the ED 1-10. The 8-9 different colors (Curves) in the graph correspond to slots containing baseline and acupuncture stimulation locations for an individual participant. There are two important findings from these results:

1. The first notable result for all participants is that the optimal ED is about 4. It is also notable that this optimal ED is independent of the specific stimulus location. This result suggests that an ED of 4 or 5 is most appropriate for HRV data for all slots and all participants.
2. In Fig.4., right-hand side figures (Participant 2 and Participant 4) shows increasing numbers of false nearest neighbours as ED increases above its optimal value of 4, whereas left-hand side figures (Participant 1 and Participant 3) does not display this. An increase in the number of false nearest neighbours with increasing ED is normally suggestive of noise in the data [1].

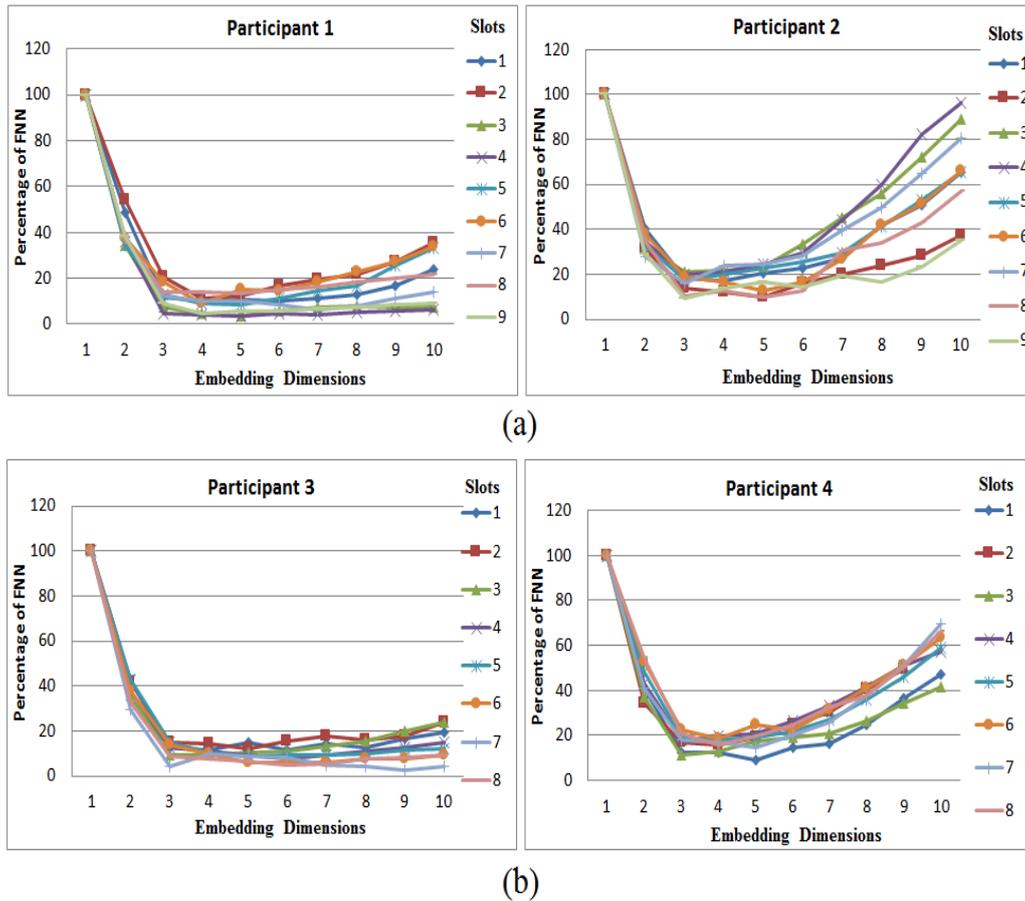


Figure 4. Embedding Dimension Result: (a) Two participant's (Participant 1 and 2) ED result from Dataset 1, and (b) Two participant's (Participant 3 and 4) ED result from Dataset 2.

#### 4.2. HRV Prediction using Linear Regression

The standard Linear Regression is utilized to produce a linear predictor for our embedded data. It was suggested by [9], that the performance of a time series predictor is affected by the size of a window, in which the time series embedded. Therefore, Linear Regression is used to predict the window size for HRV series to achieve the best result. For each window, data is split into a training set of 250 vectors and a test set of 106 vectors. The size of the embedding varies between 2 and 6. Errors are calculated as relative error [4], and results for five different window size has shown in Fig.5. It is clear that the best regressor has four inputs, and changing this number either way harms the performance.

The linear prediction of the HRV suggests that window size of 4 will be enough to fit the HRV time series data. Also, this reflects the ED result 4 is the minimum ED for the HRV Data as shown in Fig.4.

### 5. Discussion and Conclusion

Our result indicating that the HRV has an estimated ED of 4 suggests that the underlying dynamic system has 2 features; based on [16], if the original dynamic system had a dimension of  $N$ , then an embedding of size  $2N$  will be fully regained the original system. This result is interesting because HRV is driven by two underlying systems, the sympathetic and parasympathetic neural pathways; HRV is a marker of sympathetic and parasympathetic influences on the modulation of heart rate [19], and this reflects in the ED result. The effect of the sympathetic pathway is to increase heart rate and blood pressure (Fight or Flight response), whereas the parasympathetic path acts to decrease heart rate and blood pressure (Rest and Digest response). Therefore, the main finding here is that in all circumstances an Embedding of the HRV time series into a four-dimensional space produced the smallest number of false nearest neighbours.

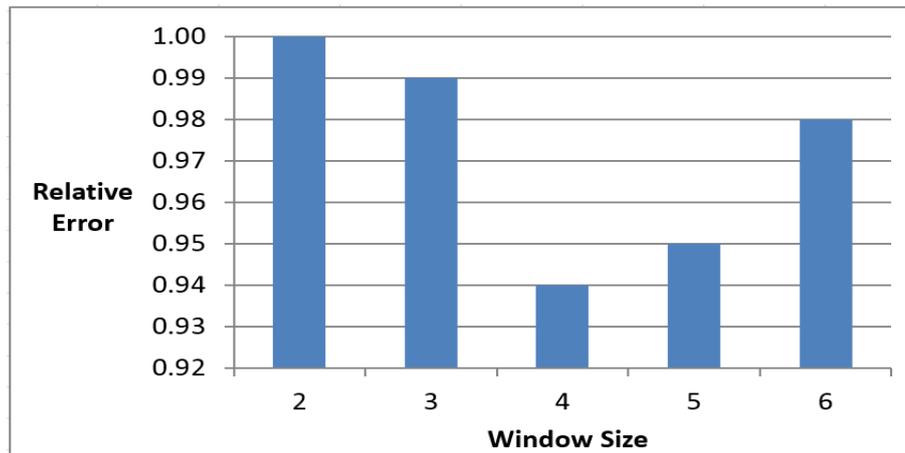


Figure 5. Relative error of the predictor for varying window size for HRV Time series.

This finding strongly suggests that the Autonomic Nervous System that drives the heart is a two-dimensional dynamic system.

From the participant's questionnaire, a variety of subjective responses to the acupuncture stimulation found. However, this did not appear to have much effect on the HRV time series, which robustly kept its two-dimensional dynamic system.

In some circumstances, the number of FNN increase as the ED became more massive than the optimal value. This increase is suggestive of noise in the data that may have come from the ECG measuring equipment.

It was suggested by [9], that the performance of a time series predictor is affected by the window size, in which the time series embedded. So that the best predictor would be the one that used correct ED. Our experiments, reported here, using a Linear Regression to predict the HRV series confirmed this as a window size of four gave the best result.

## References

- [1] Henry DI Abarbanel, Reggie Brown, John J Sidorowich, and Lev Sh Tsimring. The analysis of observed chaotic data in physical systems. *Reviews of modern physics*, 65(4):1331, 1993.
- [2] Massimo Camplani and Barbara Cannas. The role of the embedding dimension and time delay in time series forecasting. *IFAC Proceedings Volumes*, 42(7):316–320, 2009.
- [3] Lynn Casimiro, Les Barnsley, Lucie Brosseau, Sarah Milne, Vivian Welch, Peter Tugwell, and George A Wells. Acupuncture and electroacupuncture for the treatment of rheumatoid arthritis. *The Cochrane Library*, 2005.
- [4] Chris Chatfield and Andreas S Weigend. Time series prediction: Forecasting the future and understanding the past: Neil a. gershenfeld and andreas s. weigend, 1994, the future of time series, in: As weigend and na gershenfeld, eds., (addison-wesley, reading, ma), 1-70., 1994.
- [5] Bian Chun-Hua and Ning Xin-Bao. Determining the minimum embedding dimension of nonlinear time series based on prediction method. *Chinese Physics*, 13(5):633, 2004.
- [6] Justin Faber and Dolores Bozovic. Chaotic dynamics of inner ear hair cells. *Scientific reports*, 8(1):3366, 2018.
- [7] Ray J Frank, Neil Davey, and Stephen P Hunt. Time series prediction and neural networks. *Journal of Intelligent and Robotic Systems*, 31(1-3):91–103, 2001.
- [8] Matthew B Kennel, Reggie Brown, and Henry DI Abarbanel. Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Physical review A*, 45(6):3403, 1992.
- [9] Rhee M Kil, Seon Hee Park, and Seunghwan Kim. Optimum window size for time series prediction. In *Engineering in Medicine and Biology Society, 1997. Proceedings of the 19th Annual International Conference of the IEEE*, volume 4, pages 1421–1424. IEEE, 1997.
- [10] Joseph B Kruskal. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 29(1):1–27, 1964.
- [11] Marek Malik and A John Camm. Heart rate variability. *Clinical cardiology*, 13(8):570–576, 1990.
- [12] Rollin McCraty, Mike Atkinson, Dana Tomasino, and William P Stuppy. Analysis of twenty-four hour heart rate variability in patients with panic disorder. *Biological psychology*, 56(2):131–150, 2001.
- [13] Mirwais. Finds minimum Embedding Dimension with false nearest neighbours method. <https://uk.mathworks.com/matlabcentral/fileexchange/37239-minimum-embedding-dimension>, 2012. [Online; accessed 18-May-2018].
- [14] MML Ng, Mason CP Leung, and DMY Poon. The effects of electro-acupuncture and transcutaneous electrical nerve stimulation on patients with painful osteoarthritic knees: a randomized controlled trial with follow-up evaluation. *The Journal of Alternative & Complementary Medicine*, 9(5):641–649, 2003.

- [15] Brian F Robinson, Stephen E Epstein, G David Beiser, and Eugene Braunwald. Control of heart rate by the autonomic nervous system: studies in man on the interrelation between baroreceptor mechanisms and exercise. *Circulation Research*, 19(2):400–411, 1966.
- [16] Floris Takens. Detecting strange attractors in turbulence. In *Dynamical systems and turbulence, Warwick 1980*, pages 366–381. Springer, 1981.
- [17] Elizabeth A Tough, Adrian R White, T Michael Cummings, Suzanne H Richards, and John L Campbell. Acupuncture and dry needling in the management of myofascial trigger point pain: a systematic review and meta-analysis of randomised controlled trials. *European Journal of Pain*, 13(1):3–10, 2009.
- [18] Dadiyorto Wendi, Norbert Marwan, and Bruno Merz. In search of determinism-sensitive region to avoid artefacts in recurrence plots. *International Journal of Bifurcation and Chaos*, 28(01):1850007, 2018.
- [19] Yuru Zhong, Hengliang Wang, Ki Hwan Ju, Kung-Ming Jan, and Ki H Chon. Nonlinear analysis of the separate contributions of autonomic nervous systems to heart rate variability using principal dynamic modes. *IEEE transactions on biomedical engineering*, 51(2):255–262, 2004.

# The Correlation between EEG Signals Varying with Distance for Datasets With and Without Medical Condition

Ronakben Bhavsar<sup>1\*</sup>, Yi Sun<sup>1</sup>, Na Helian<sup>1</sup>, Neil Davey<sup>1</sup>, David Mayor<sup>1</sup> and Tony Steffert<sup>2</sup>

<sup>1</sup>University of Hertfordshire

<sup>2</sup>The Open University

[\\*r.bhavsar2@herts.ac.uk](mailto:r.bhavsar2@herts.ac.uk)

Electroencephalogram (EEG) are time varying signal, and give different signals at the different position of electrodes. There might be a correlation between a pair of these signals; more likely related to the actual positions of electrodes. In this paper, we show that the correlation is related to the physical distance between electrodes as measured on the scalp for datasets without medical condition, but might not for datasets with medical conditions. This finding is independent of participants and brain hemisphere. Our results indicate that the EEG signal is not transmitted via neurons but through white matter in a brain.

Keywords: EEG; Independent Component Analysis (ICA); Cross-correlation; Time Series Data; Biomedical Data.

## Introduction

An Electroencephalogram (EEG) is a time varying signal, and the electrodes at different positions give different time varying signals. Our previous work indicated that there was a correlation between these signals [1]. However, that research only focused on datasets without any medical conditions. In this work, we analyse datasets not only without medical conditions, but also with medical conditions, such as Epilepsy, Autism, and Seizure.

## Dataset Information

This research utilised six datasets including; 3 (Dataset 1, Dataset 2, and Dataset 3) without medical condition [2], and 3 (Dataset 4, Dataset 5, and Dataset 6) with medical condition [3] [4], as shown in Table 1.

Labels	Dataset 1	Dataset 2	Dataset 3	Dataset 4	Dataset 5	Dataset 6
Medical Condition	None	None	None	Epilepsy	Autism	Seizure
Participants	16	20	32	5	13	12
Electrodes	19	10	15	19	19	19
Paired Electrodes	171 Pairs	45 Pairs	105 Pairs	171 Pairs	171 Pairs	171 Pairs

Table 1. Datasets Information

## Experiments and Results

The EEG signals were processed to remove artefacts, such as eye blinks, eye movements, jaw movements and muscle movements, by using Independent Component Analysis (ICA). In order to obtain distance in centimetres (cm) between electrodes, a measuring tape was used to measure distance using a straight line distance between two electrodes on a cap - not the distance as measured over the surface (curved line) of the scalp. Cross-correlation has been calculated on the processed EEG signals. Figure 1 shows the average correlation results of participants for electrode Fp1 on analysing all six datasets, where electrode Fp1 has been chosen randomly across all 19 electrodes for all datasets to show the Cross-correlation performance - other electrodes have similar results. Figure 1 (A), (B), (C), and (E) demonstrates that there is an inverse linear relationship between Cross-correlation value and distance. Whereas Figure 1 (D), and (F) do not indicate any linear dependency.

## Discussion and Conclusion

One of the main conclusions of this work is that electrical activity correlates linearly with straight line physical distance (that is when the distance increases the correlation decreases) for participants without medical conditions. However, participants with medical conditions such as Epilepsy and Seizure, the linear dependency might not exist. We lack of the expertise to provide possible reasons for this, but think this might be of interest to the people working in medical area. The second conclusion from this work is that the correlation is independent of brain hemisphere for all datasets. This suggests that most probably the electrical signals are transmitted through the white matter of the brain [3]. We assume that signal transmission is through white matter because of the commissural tracts within the white matter which connect the two hemispheres of the brain. This means, in practice it does not matter which side of the median plane you place the electrodes.

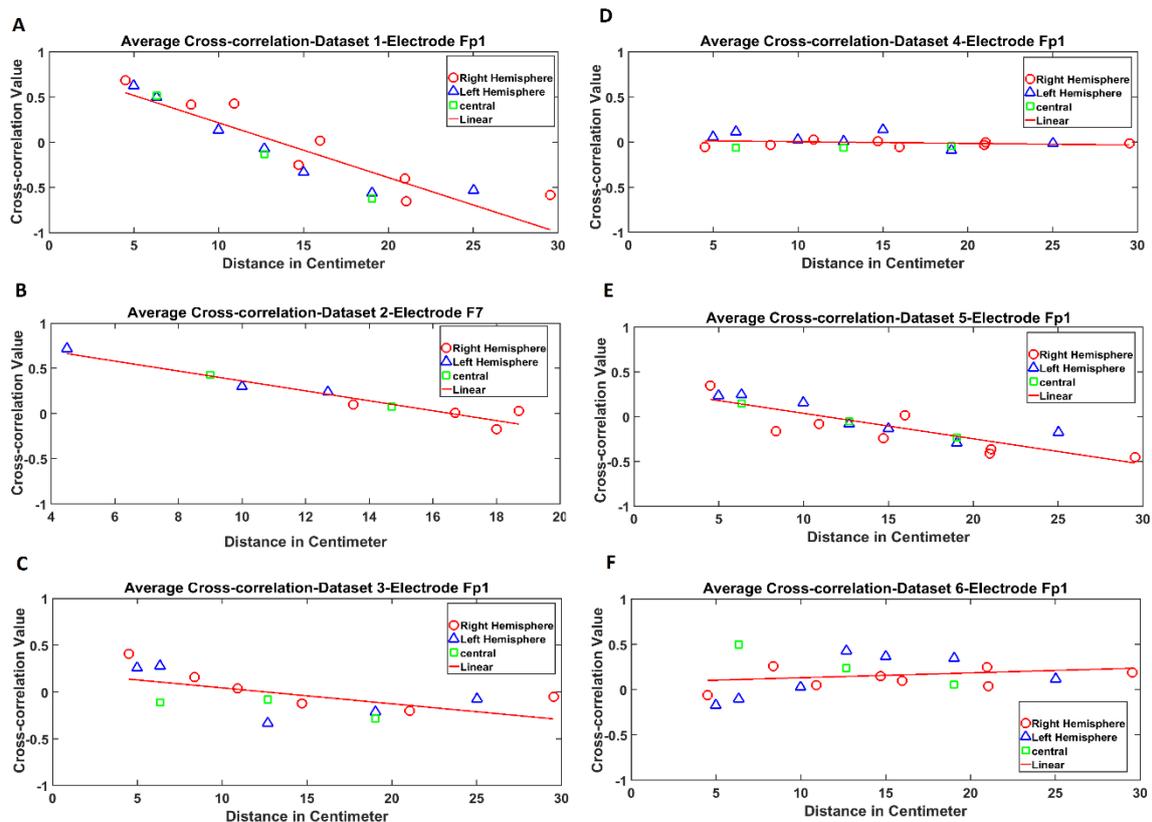


Figure 1. Cross-correlation between electrodes at varying distance on all datasets. Where, (A) Dataset 1, (B) Dataset 2, (C) Dataset 3, are without any medical condition, and (D) Dataset 6, (E) Dataset 7, and (F) Dataset 8, are with medical condition Autism, Epilepsy, and Epileptic Seizures, respectively.

## References

- [1] Bhavsar, R., Sun, Y., Helian, N., Davey, N., Mayor, D. and Steffert, T., "The Correlation between EEG Signals as Measured in Different Positions on Scalp Varying with Distance.," *Procedia Computer Science.*, vol. 123, pp. 92-97, 2018.
- [2] Koelstra, "Deap: A database for emotion analysis; using physiological signals," in *IEEE Transactions on Affective Computing*, 2012.
- [3] Fields, R.D., 2008., "White matter matters.," *Scientific American*, vol. 293, no. 3, pp. 54-61, 2008.
- [4] D. D. a. E. K. Taniskidou, "UCI machine learning repository," 2017. [Online]. Available: <http://epileptologie-bonn.de/cms/upload/workgroup/lehnertz/eegdata.html>. [Accessed 2018].
- [5] A. L. Goldberger, "Component of a new research resource for complex physiologic signals.," 2000. [Online]. Available: <https://physionet.org/pn6/chbmit/>. [Accessed 2018].

# The Effects of Electro-Acupuncture Related methods on the EEG Signals

Ronakben Bhavsar<sup>1</sup>, Na Helian<sup>1</sup>, Yi Sun<sup>1</sup>, Neil Davey<sup>1</sup>, David Mayor<sup>2</sup>, Tony Steffert<sup>3</sup>

r.p.bhavsar@herts.ac.uk

<sup>1</sup>Science and Technology Research Institute, University of Hertfordshire, UK

<sup>2</sup>Hon Research Fellow, Division of Physiotherapy, University of Hertfordshire, UK

<sup>3</sup>Independent EEG Consultant and Researcher in Physiological Signification, Computing Department, Open University



## 1. Motivation

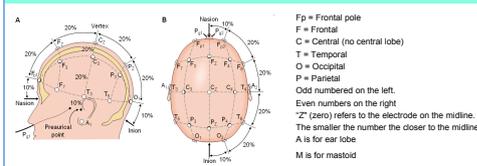
In this research the Electroencephalography (EEG) signals of patients have been recorded when acupuncture is performed.

Here we try to detect variation in the EEG signals following a Transcutaneous Electrical Acupoint System (TEAS) at different stimulation locations.

Here we use Sample Entropy to measure the complexity of the EEG signals.

## 2. Experimental Techniques

### EEG



EEG is a method of recording brain activity (NHS, 2012). The electrical signals that travel through the active brain cells are recorded by placing small electrodes onto the scalp. The signals are measured using only 19 electrodes because 1 electrode is considered as ground. The readings obtained in EEG can be used to investigate some brain conditions. A EEG signal is a symptomatic indicator which helps in determining brain activities under a physiological condition.

"This EEG cap is stretchable and can fit any size of head"

### TEAS for Acupuncture



TEAS is a safe, standardized acupuncture technique in which there is no needle insertion. It involves applying cutaneous electrical stimulation by placing electrodes at classical Chinese acupoints.

<http://emedicine.medscape.com/article/325107-overview>

### Sample Entropy as non-linear measure

Sample Entropy measures the predictability of a time series. If the Sample Entropy is 0, then the time series is completely predictable, as is the case for example if the time series is constant. On the other hand, if the time series is completely unpredictable the sample entropy is 1 (for a binary series).

EEG signals represent the complex dynamic behaviours of a biological system. In order to understand the complex behaviours Sample Entropy is used.

The Sample Entropy analysis proves to be a powerful approach for assessment of different physiological time series as it can determine the hidden patterns related to the underlying mechanism. With the extensive use of Sample Entropy in the field of medical science, this non-linear measure proves to be appropriate to investigate the EEG series with certain parameters such as bodily location.

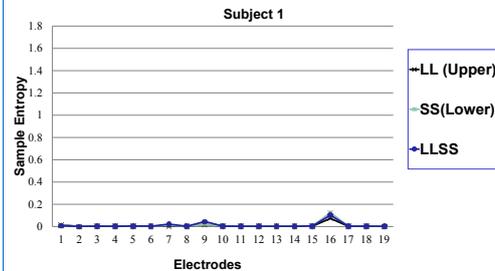
## 3. Description of the Datasets

The EEG data is taken from signals provided by 7 patients on whom acupuncture was performed. During this visit EEG monitoring was carried out in six 5 minute 'slots' with stimulation at six different locations:

1. First stimulation – Left (L14, a point on the left hand, and ST36, a point on the left leg)
2. Second stimulation – Right (L14, a point on the right hand, and ST36, a point on the right leg)
3. Third stimulation- Bilateral (both 'Left' and 'Right')
4. Fourth stimulation- LL or Upper (left L14 and right L14)
5. Fifth stimulation-SS or Lower (left ST36 and right ST36)
6. Sixth stimulation-LLSS (both upper and Lower)

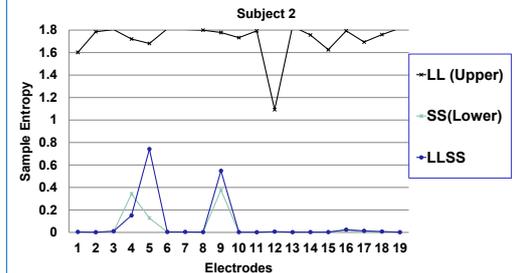
## 4. Results

The figure shows the Sample Entropy for a typical patient over 19 electrodes. The first thing to note is that the Sample Entropy is low showing that the EEG time series is fairly predictable. It can also be seen that there is not much difference in location or stimulus site.



The Location Effect of electrical acupoints on the EEG signals of Subject 1. (No response towards location). The lines are a guide for the eye and do not represent a graph

Subject 2 however showed a completely different response. When stimulated on the Upper body the EEG became extremely disordered. The Sample Entropy is between the range of 1.6 to 1.8. This suggests a strong response to the TEAS stimulation at this location. Interestingly electrode 5 and electrode 9 gives higher Sample Entropy than other electrodes.



The location Effect of electrical acupoints on the EEG signals of Subject 2. (responded towards locations)

Sample Entropy ↓	Six Locations →					
	Left	Right	Left & Right	LL (Upper)	SS (Lower)	LLSS
Mean SE	0.02	0.02	0.04	0.2	0.03	0.03
Min	0.0007	0.008	0.0008	0.1	0.0008	0.0008
Max	0.2	0.3	0.4	0.5	0.2	0.3
SD	0.05	0.07	0.1	0.09	0.08	0.1

Table 1: The average Sample Entropy result for all patients on six locations.

## 5. Conclusion & Discussion

- Subject 1 didn't give response to any of the location
- Subject 2 responded highly when upper body was stimulated. (Higher Sample Entropy value obtained)
- Result in Table 1 suggest the average Sample Entropy value.

The higher sample entropy value the higher disorder in EEG recording of patient. This result can help to identify their mental state (such as anxious, actively thinking, relax, sleeping) and suggest that different patient reacted differently on different location when they were under acupuncture treatment.

## 6. References

1. Jasper, H.H. The ten-twenty electrode system of the International Federation of societies for electroencephalography and clinical neurophysiology. Electroencephalogram. Clin. Neurophysiology., 1958, 10: 370-375.
2. Richman, JS; Moorman, JR (2000). "Physiological time-series analysis using approximate entropy and sample entropy". American journal of physiology. Heart and circulatory physiology 278 (6): H2039-49. PMID 10843903

# **Appendix F**

## **Publications**

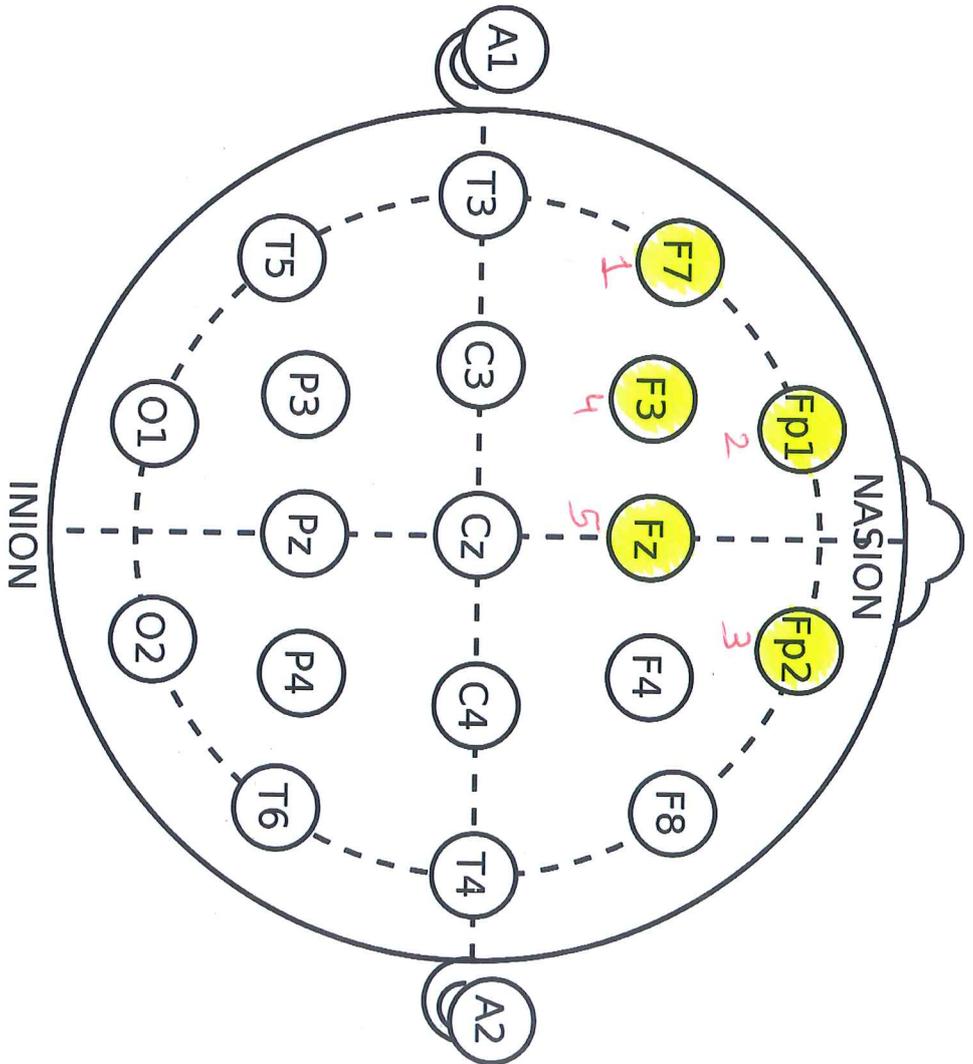
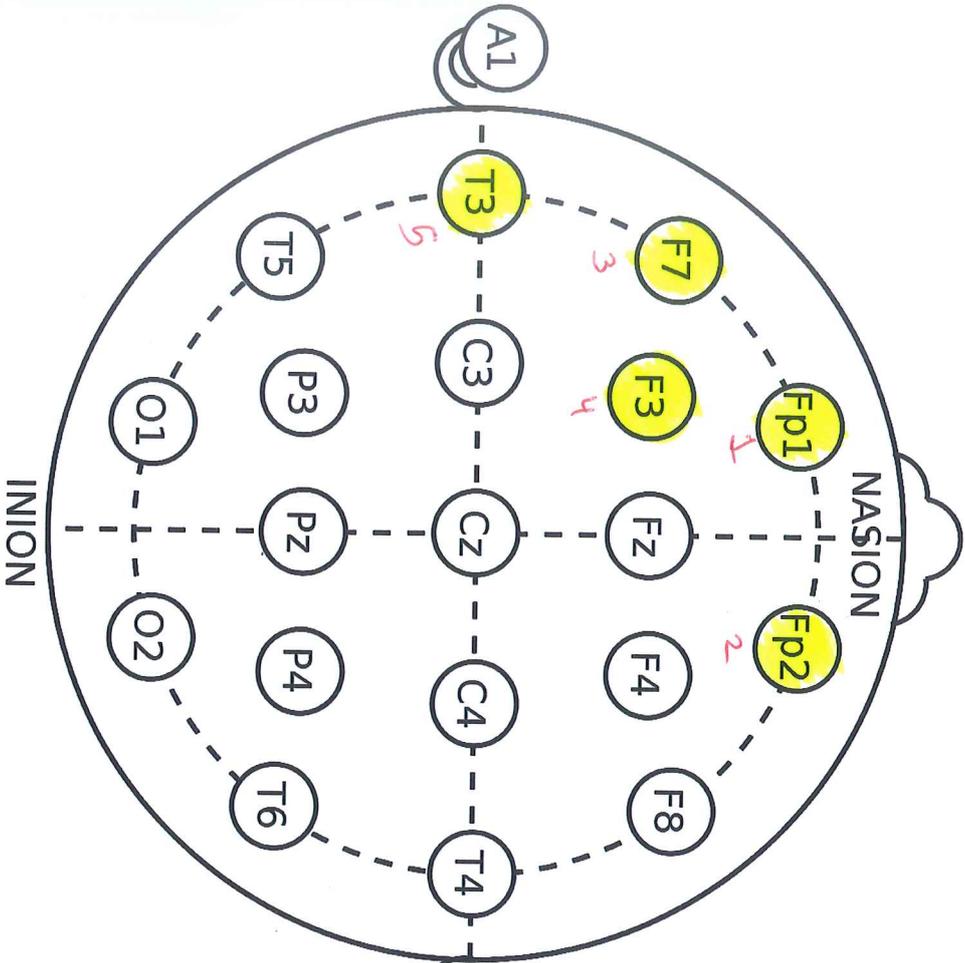
Method 1-Experiment 1

Delta (EEG)

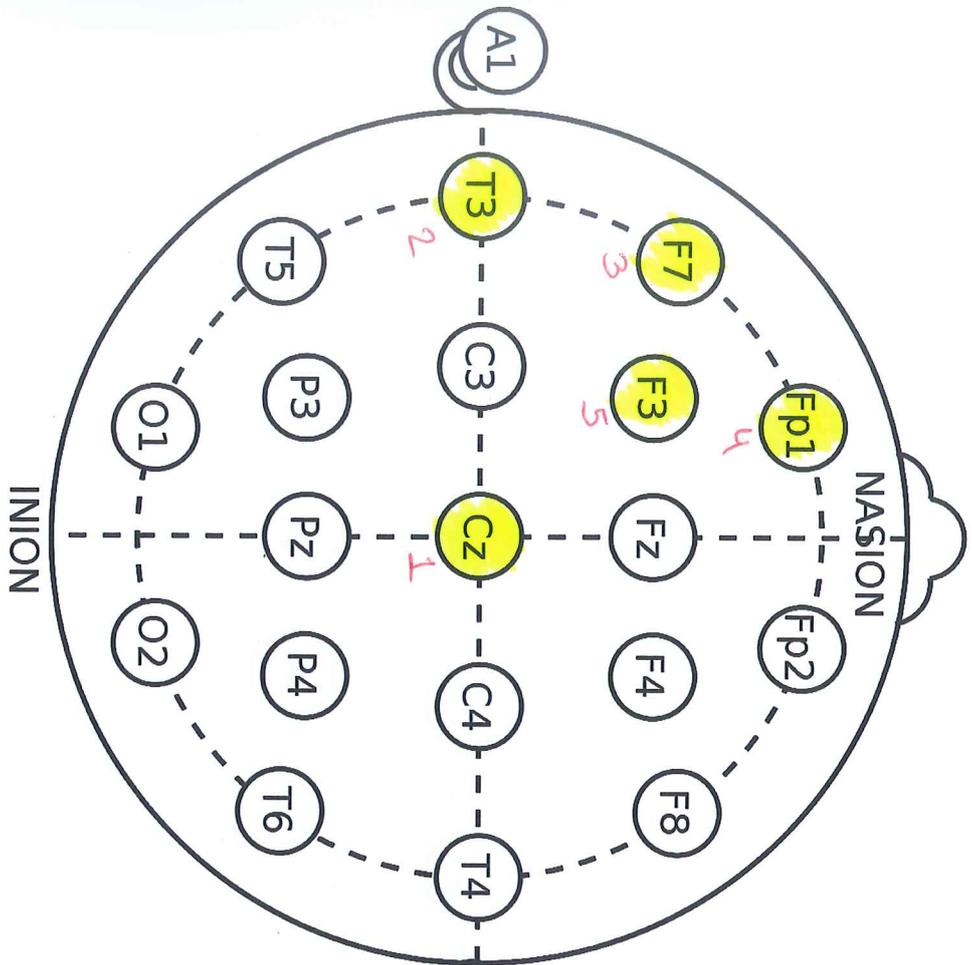
Each 5 min Data

LF

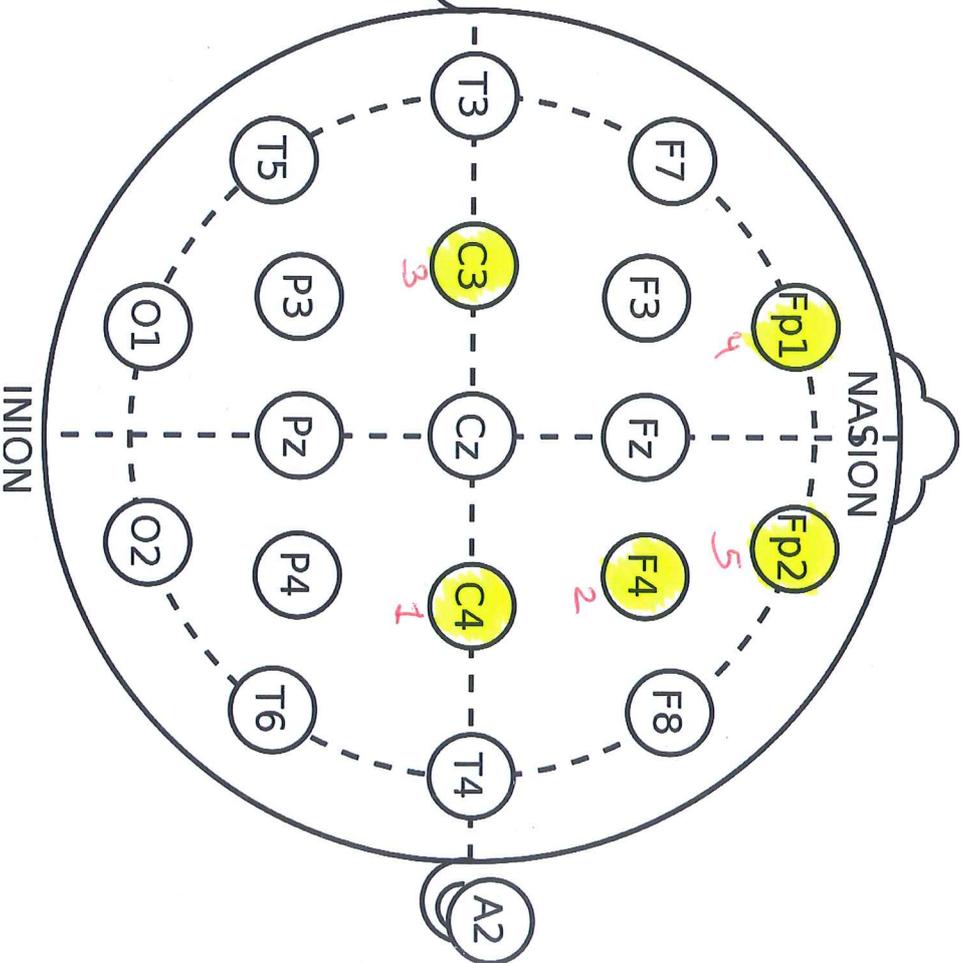
HF



(Average Result)



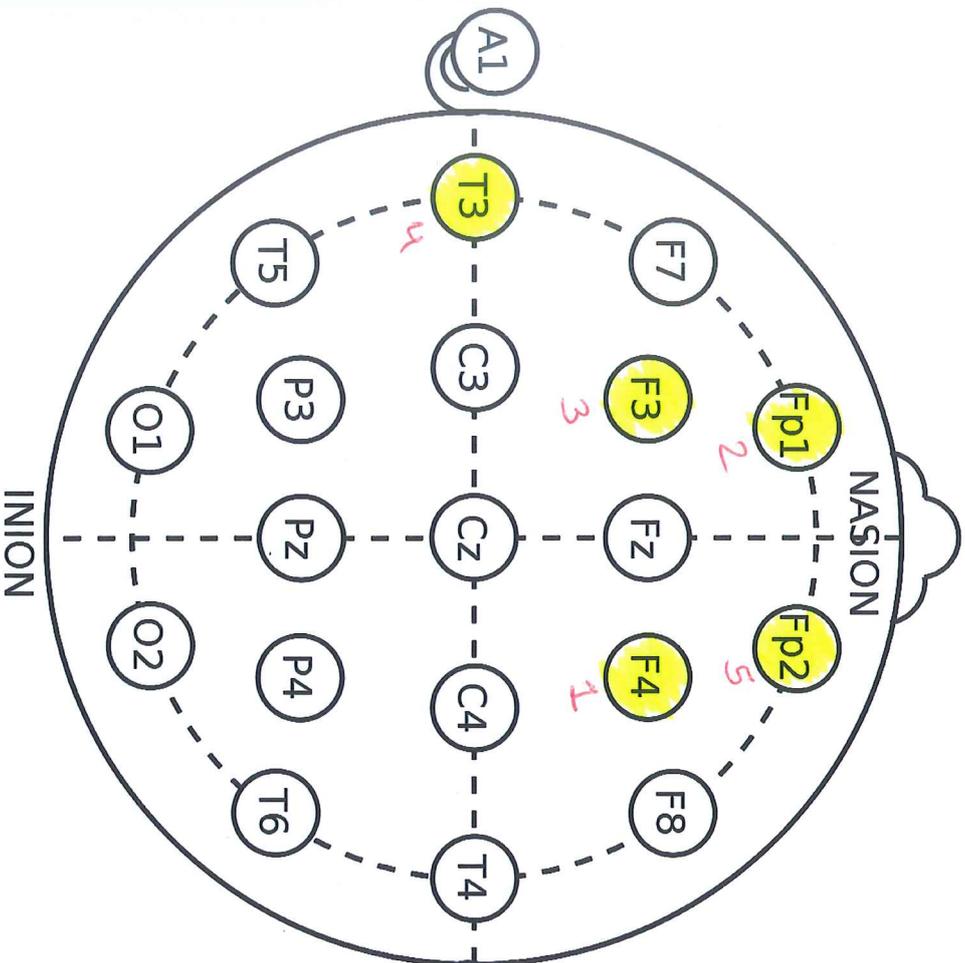
LF



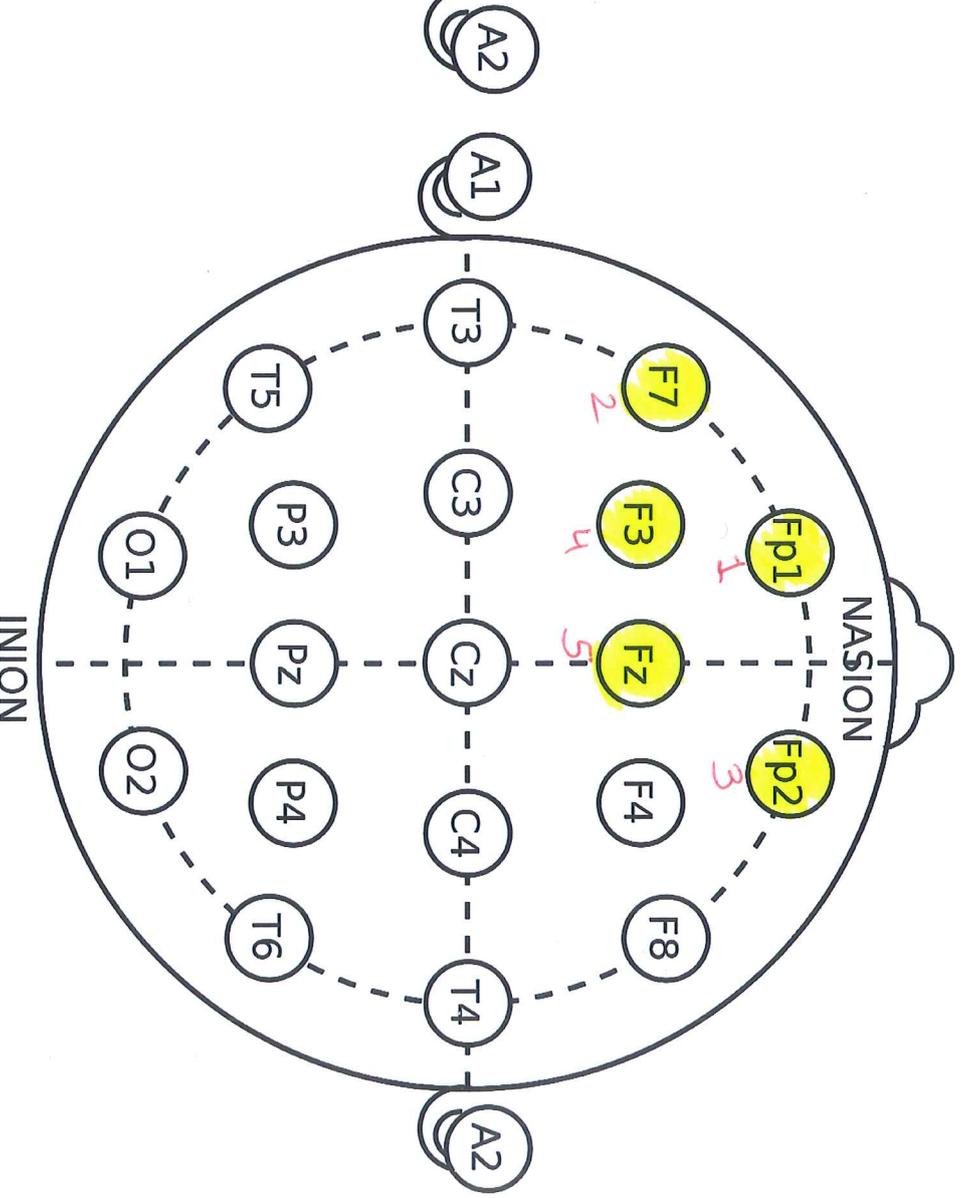
HF

Delta (EEG)

Each 2.5 min Data



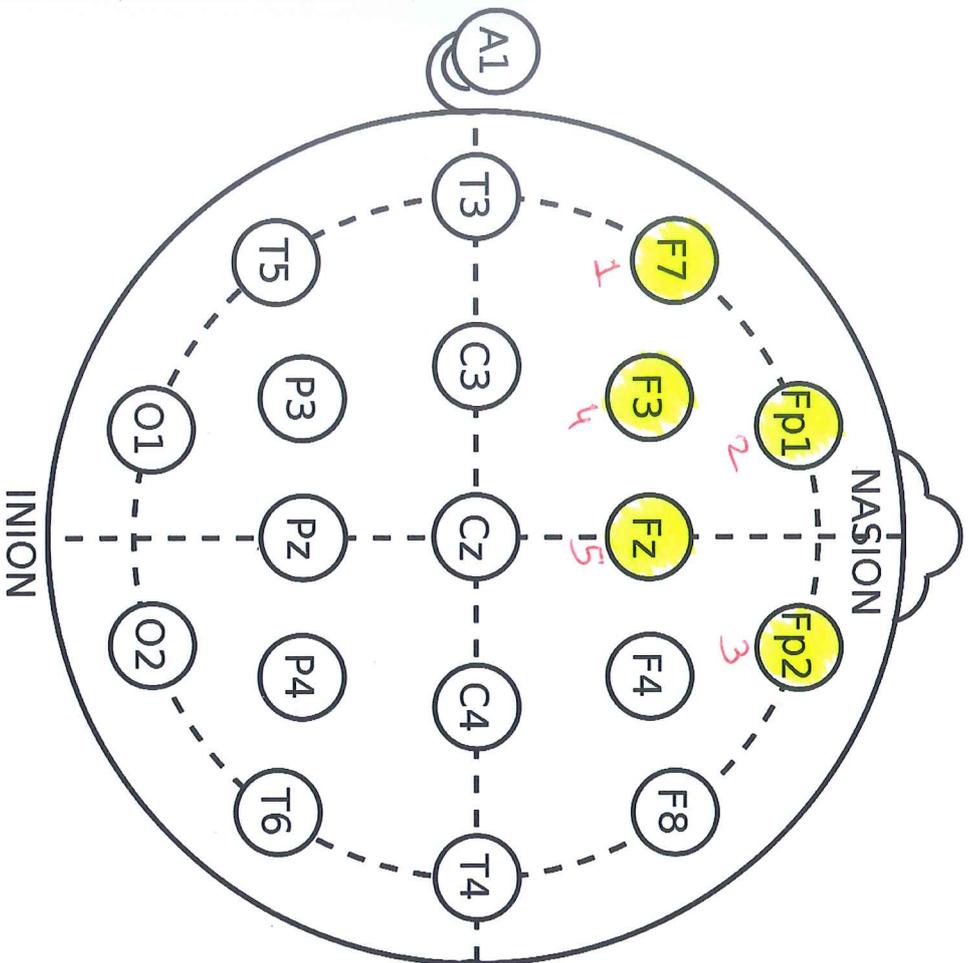
LF



HF

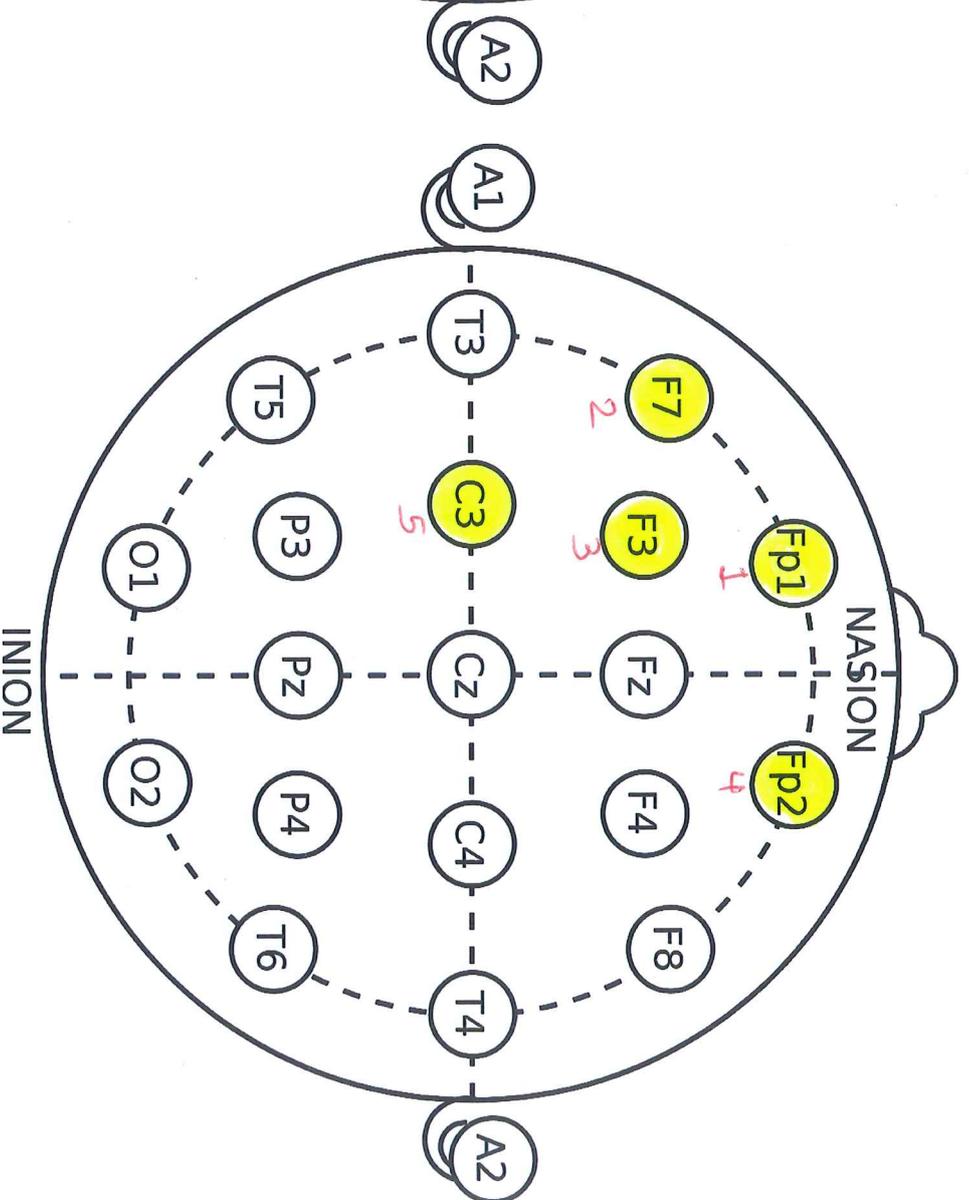
Delta (EEG)

Each 1 min Data



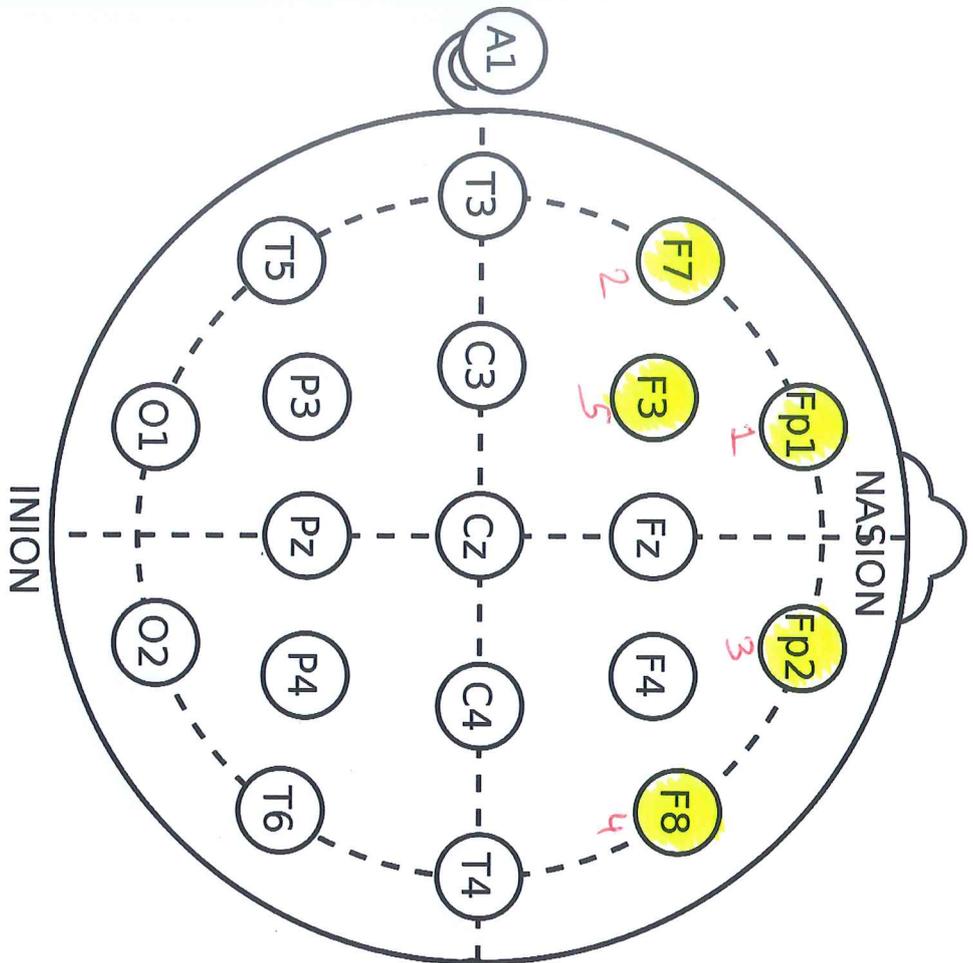
LF

Theta  
(EEG)



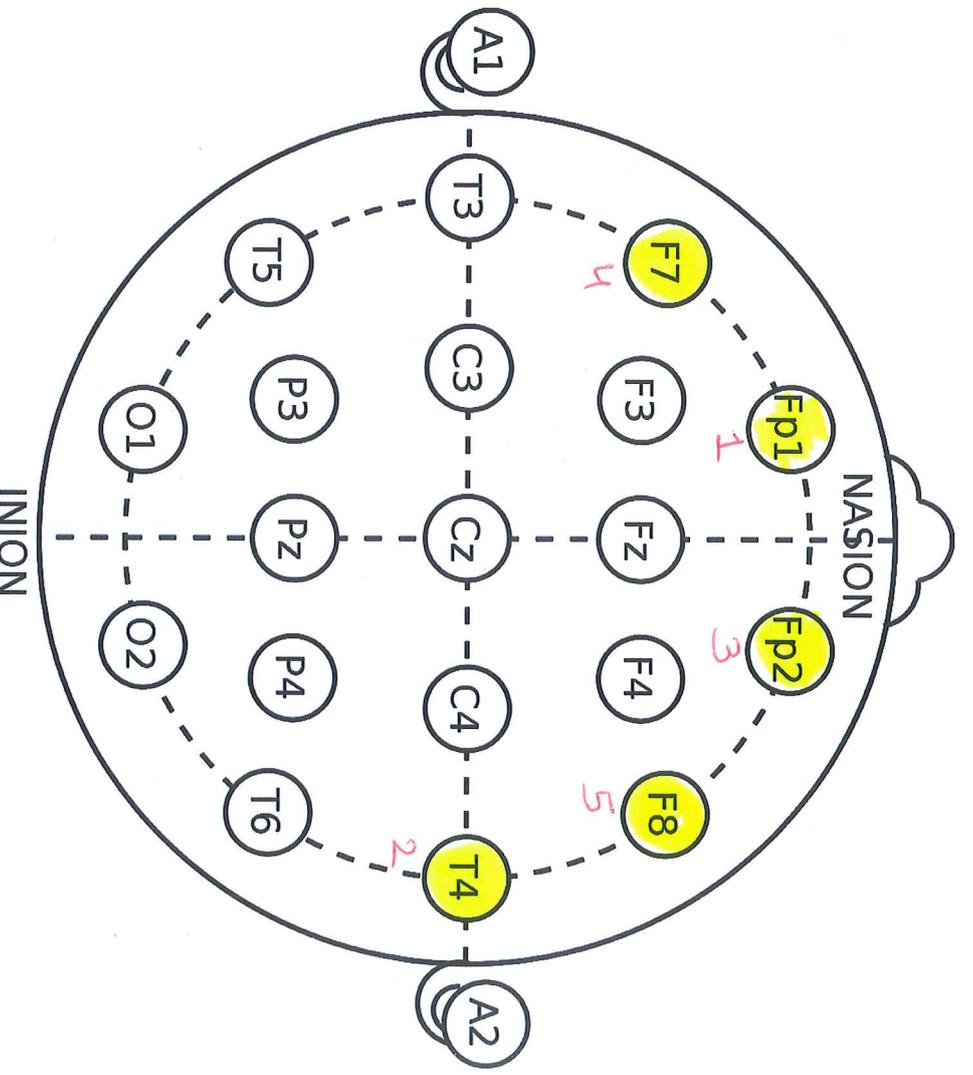
HF

Each 5 min Data



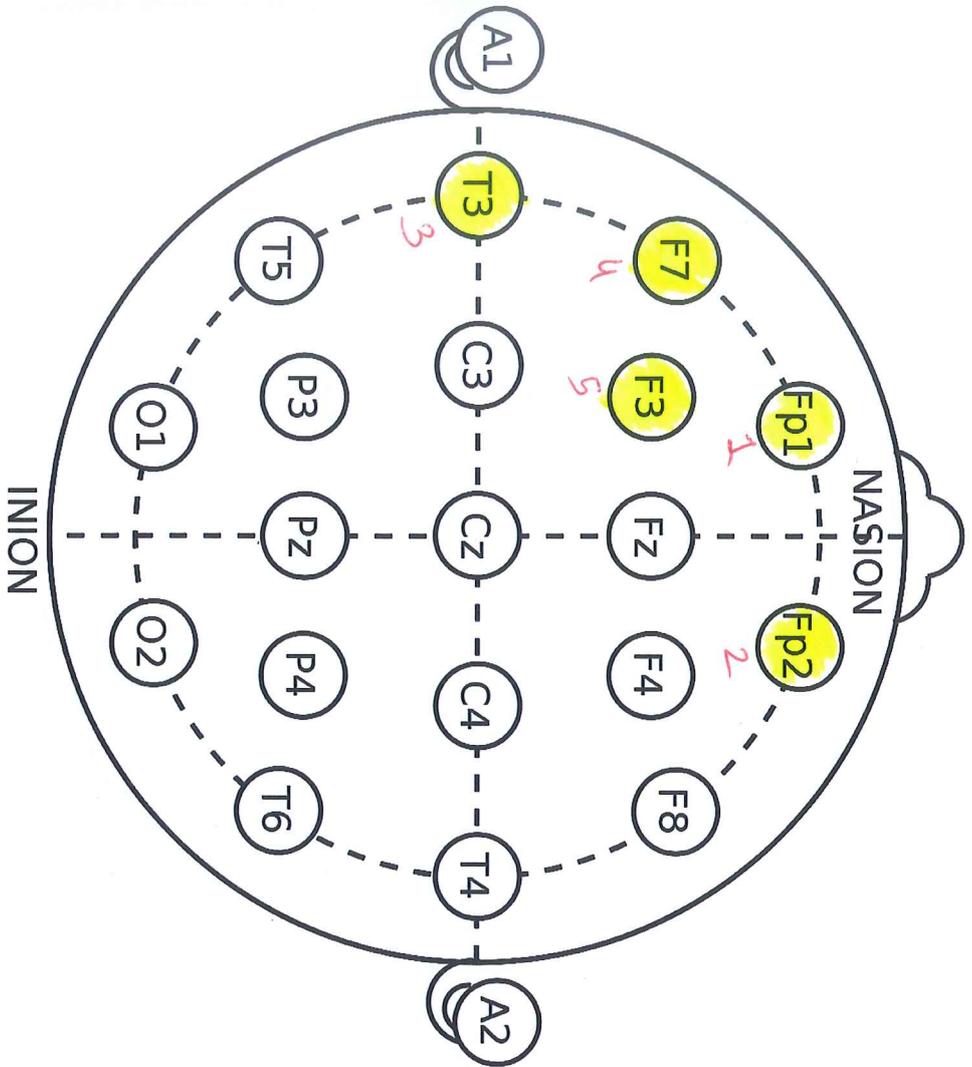
LF

Theta  
(CEEG)



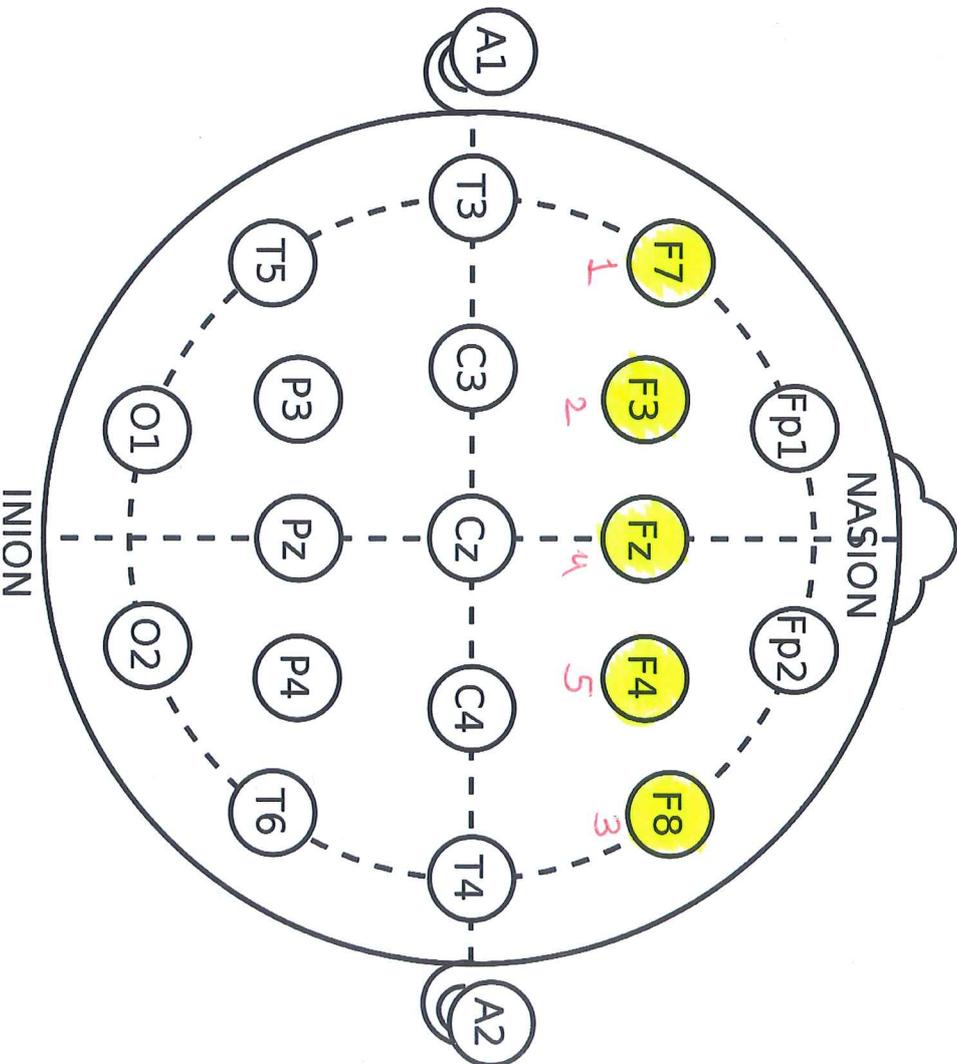
HF

Each 2.5 min Data



LF

Theta (EEG)

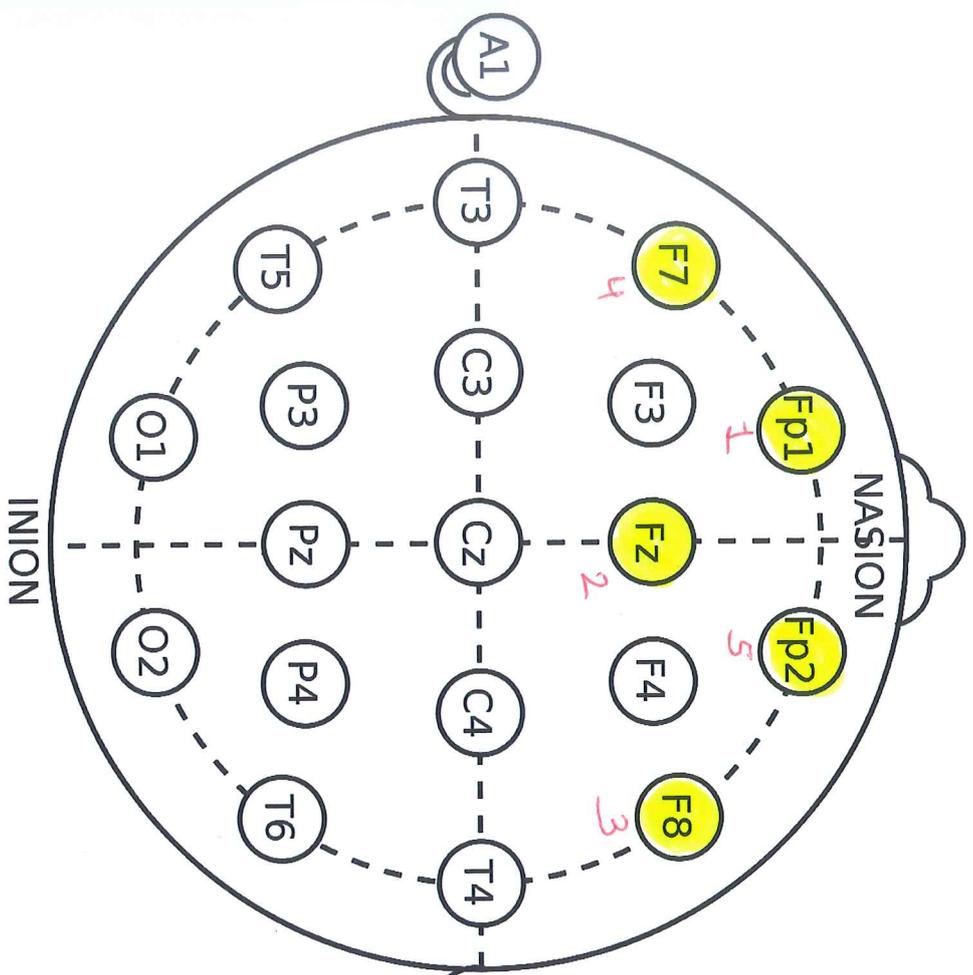


HF

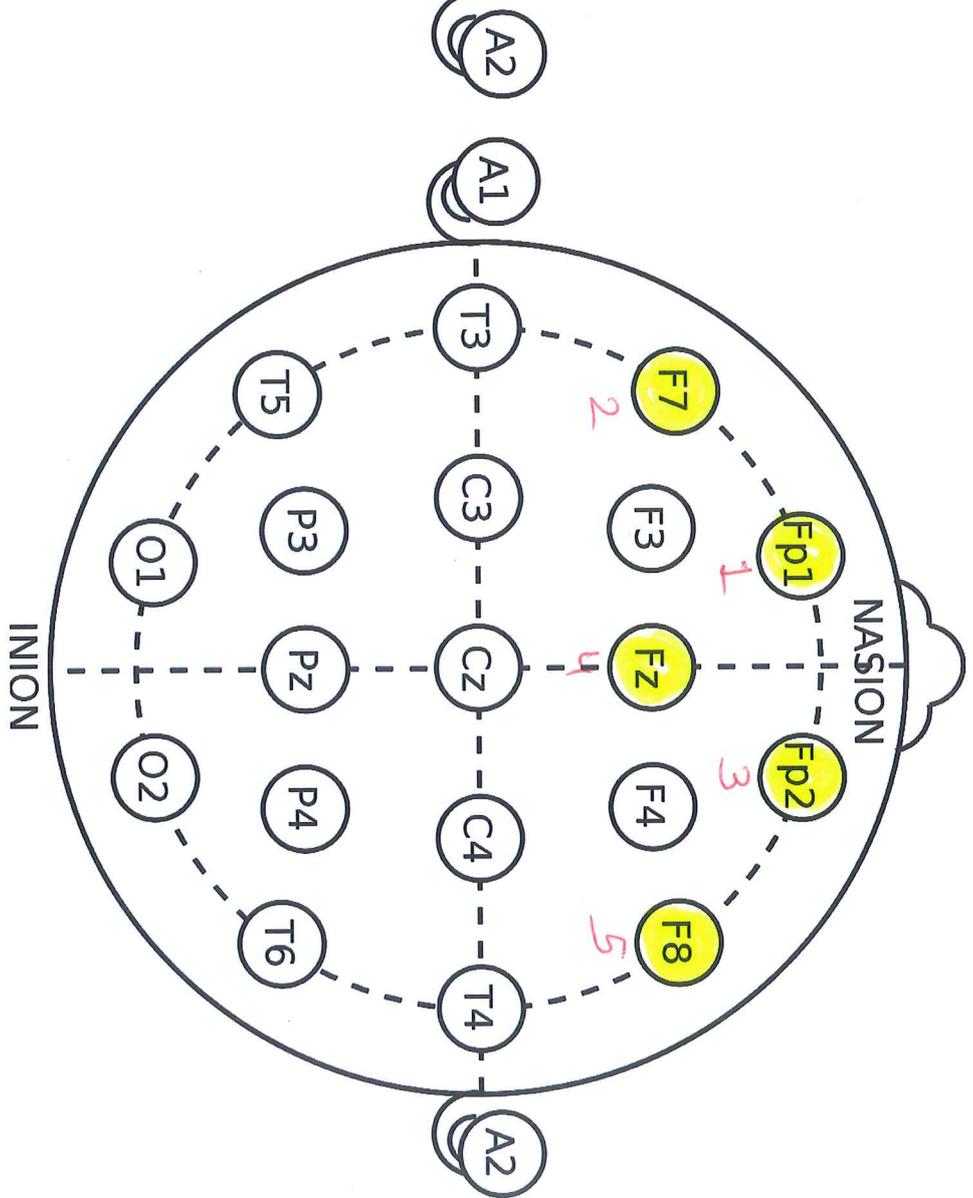
Even 1 min data

Each 5 min data

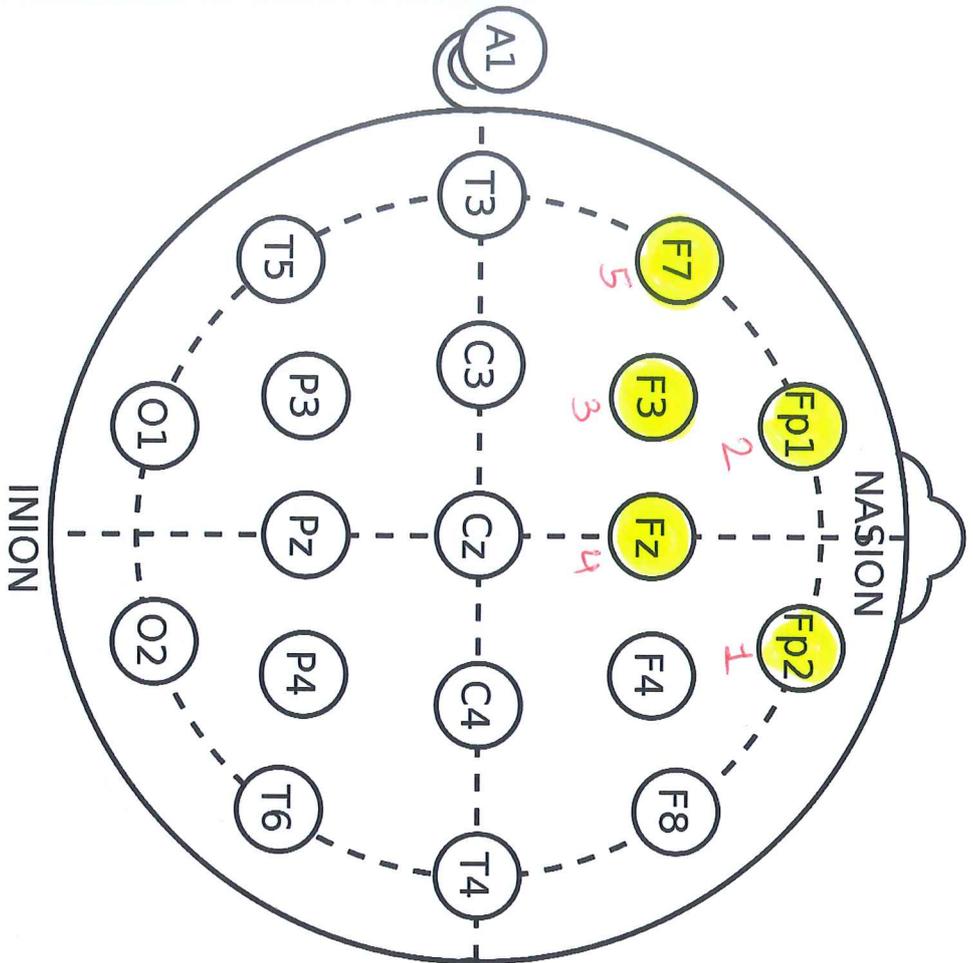
Alpha (EEG)



LF

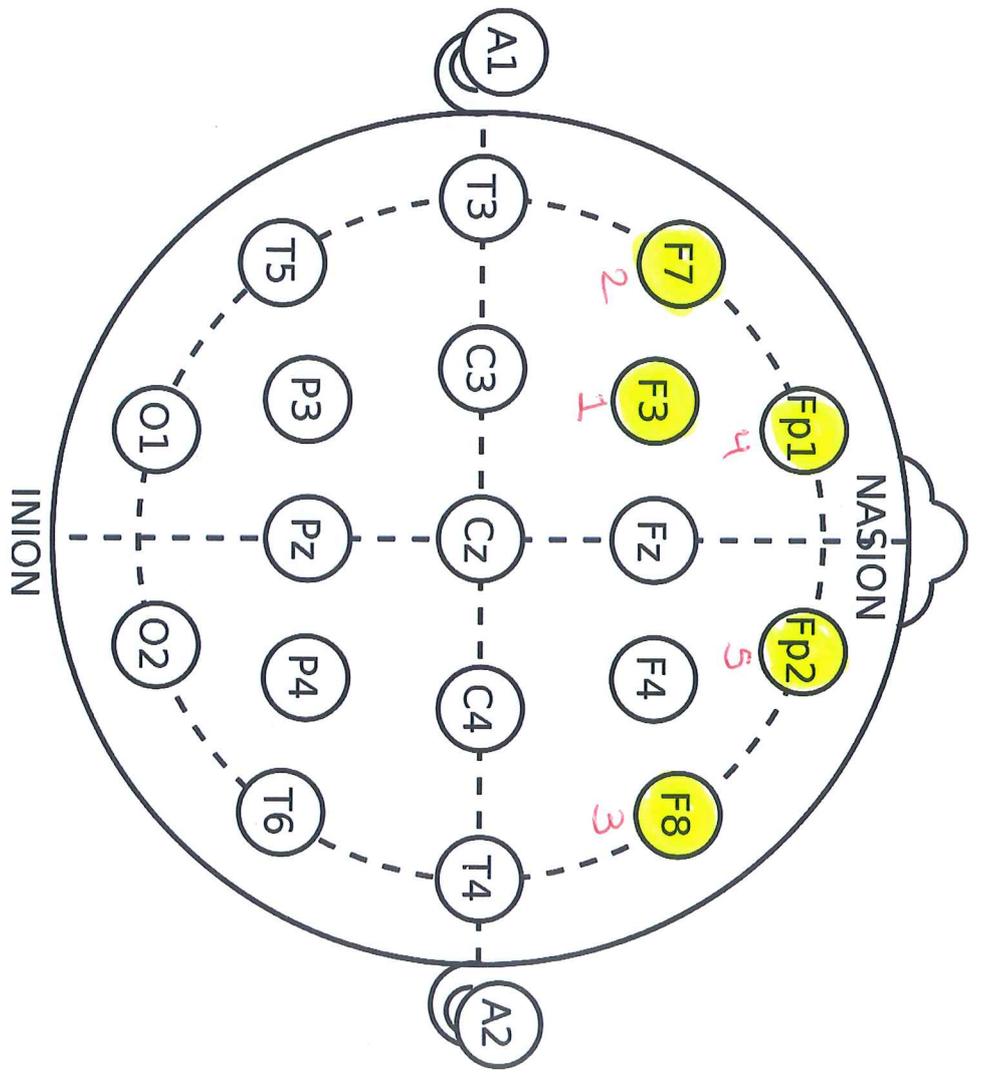


HF



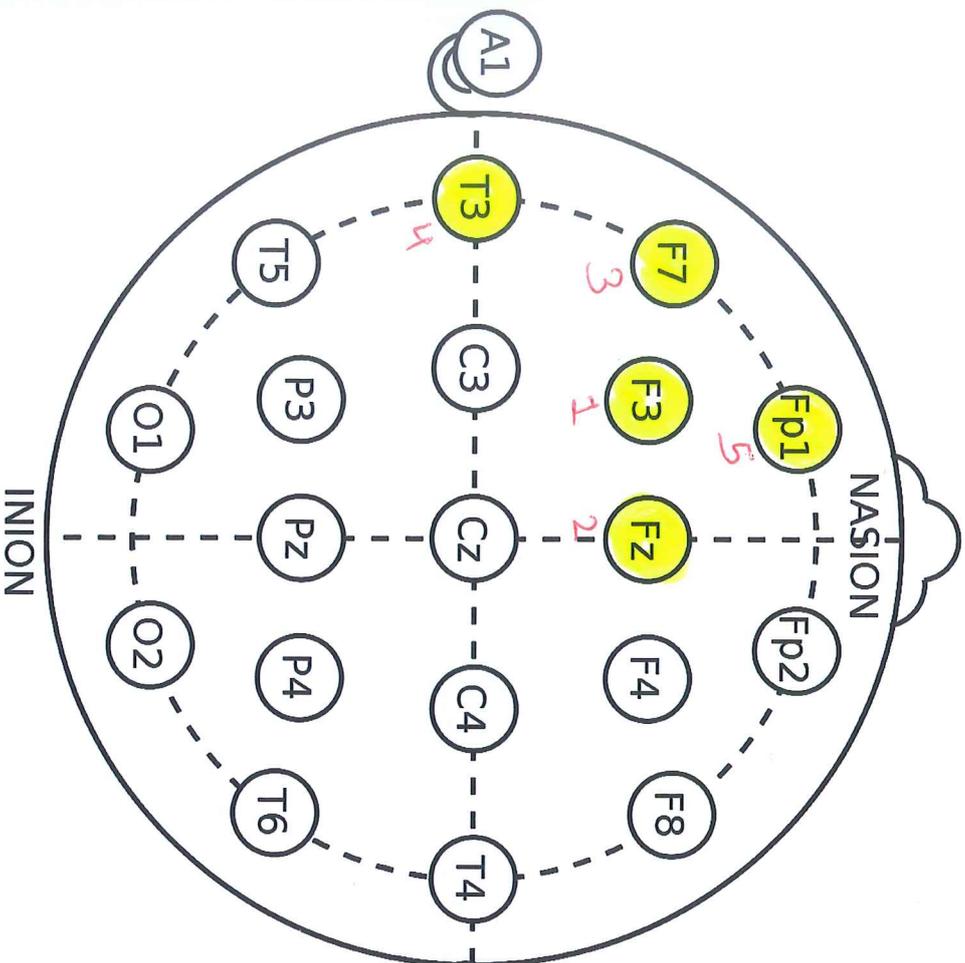
LF

Alpha  
(EEG)

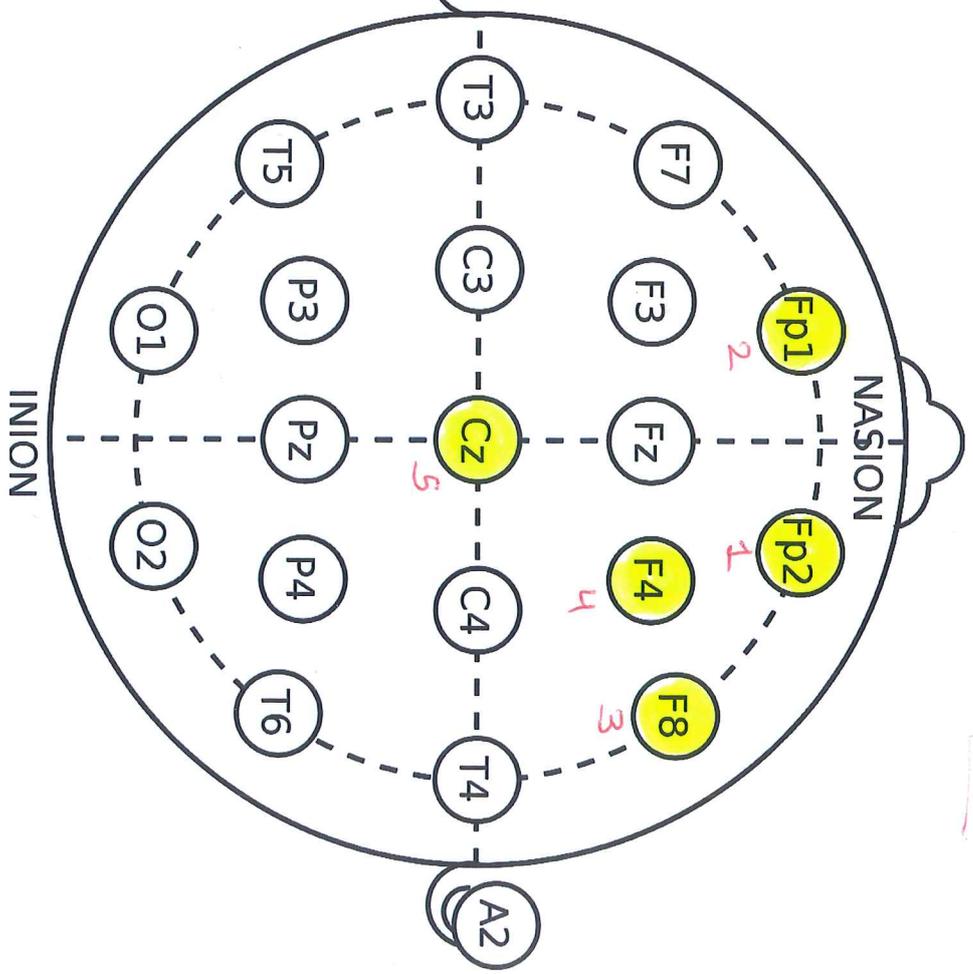


HF

Each 2.5 min Data



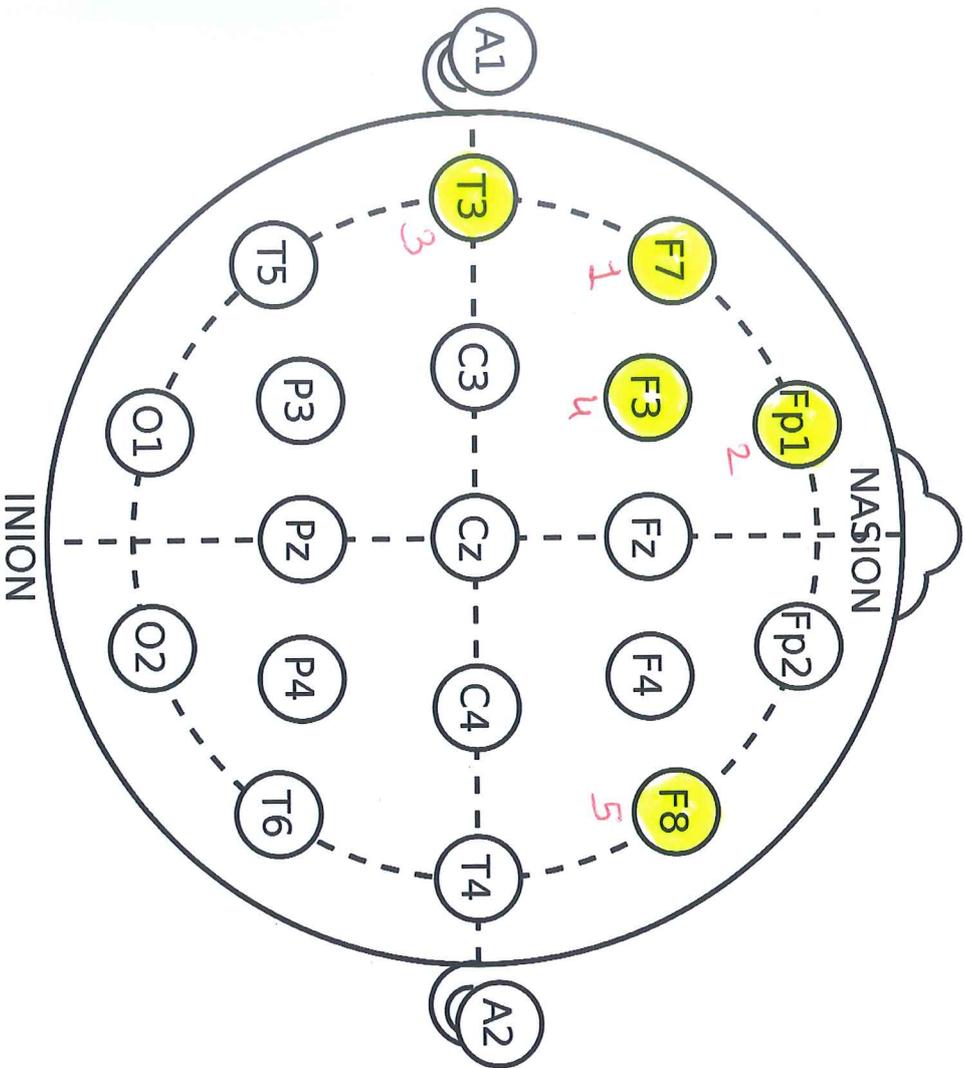
LF



HF

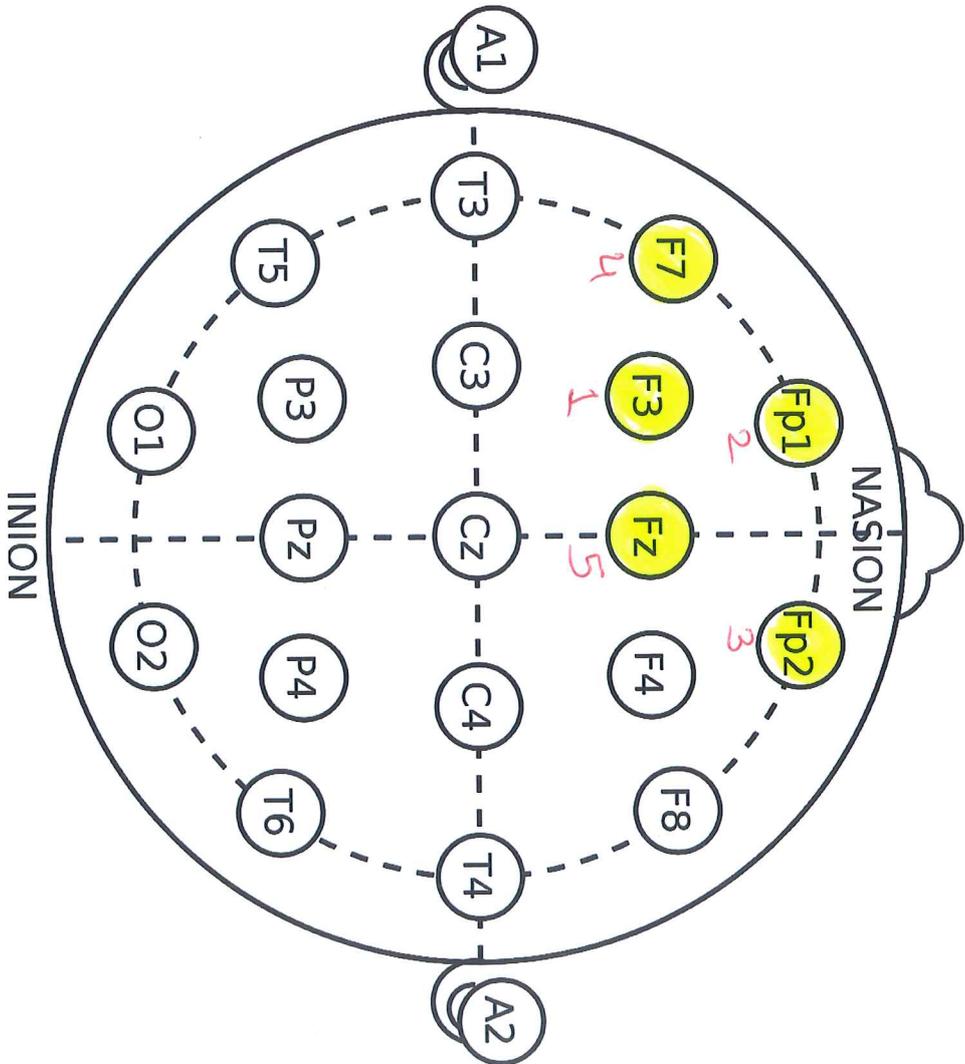
Alpha  
(EEG)

Each 1 min Data



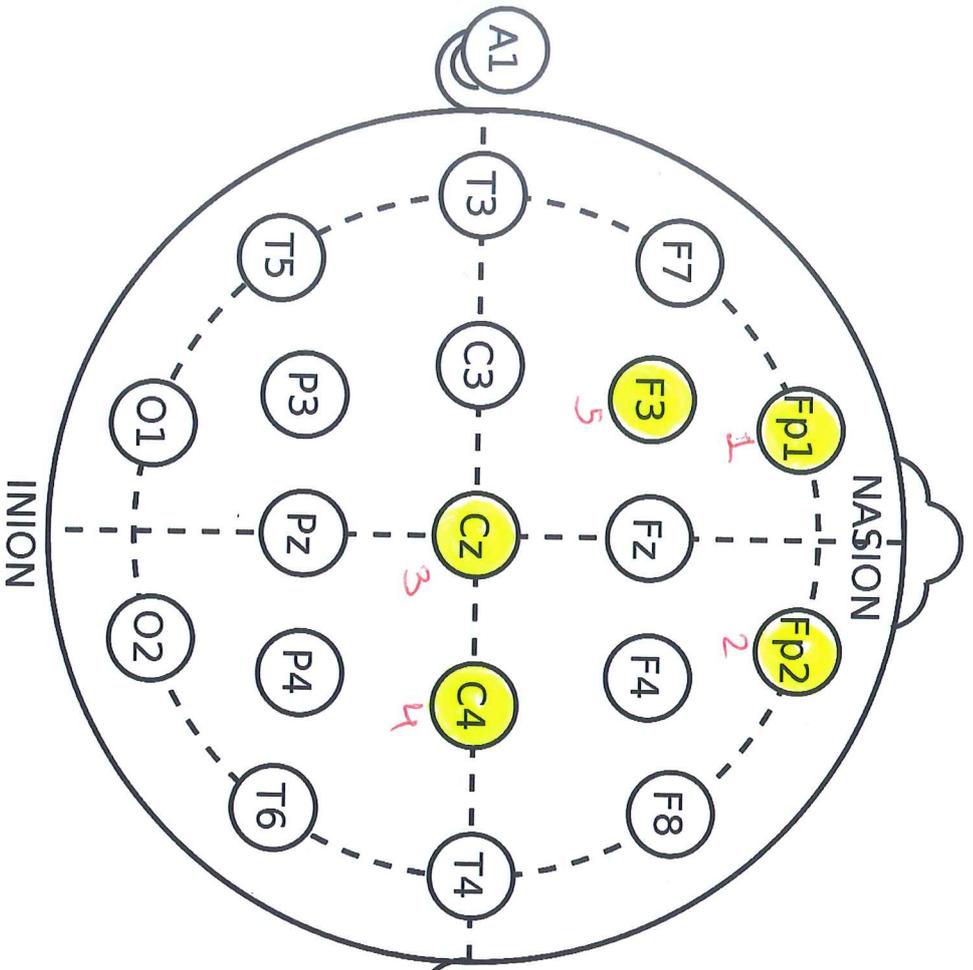
LF

Beta  
(EEG)



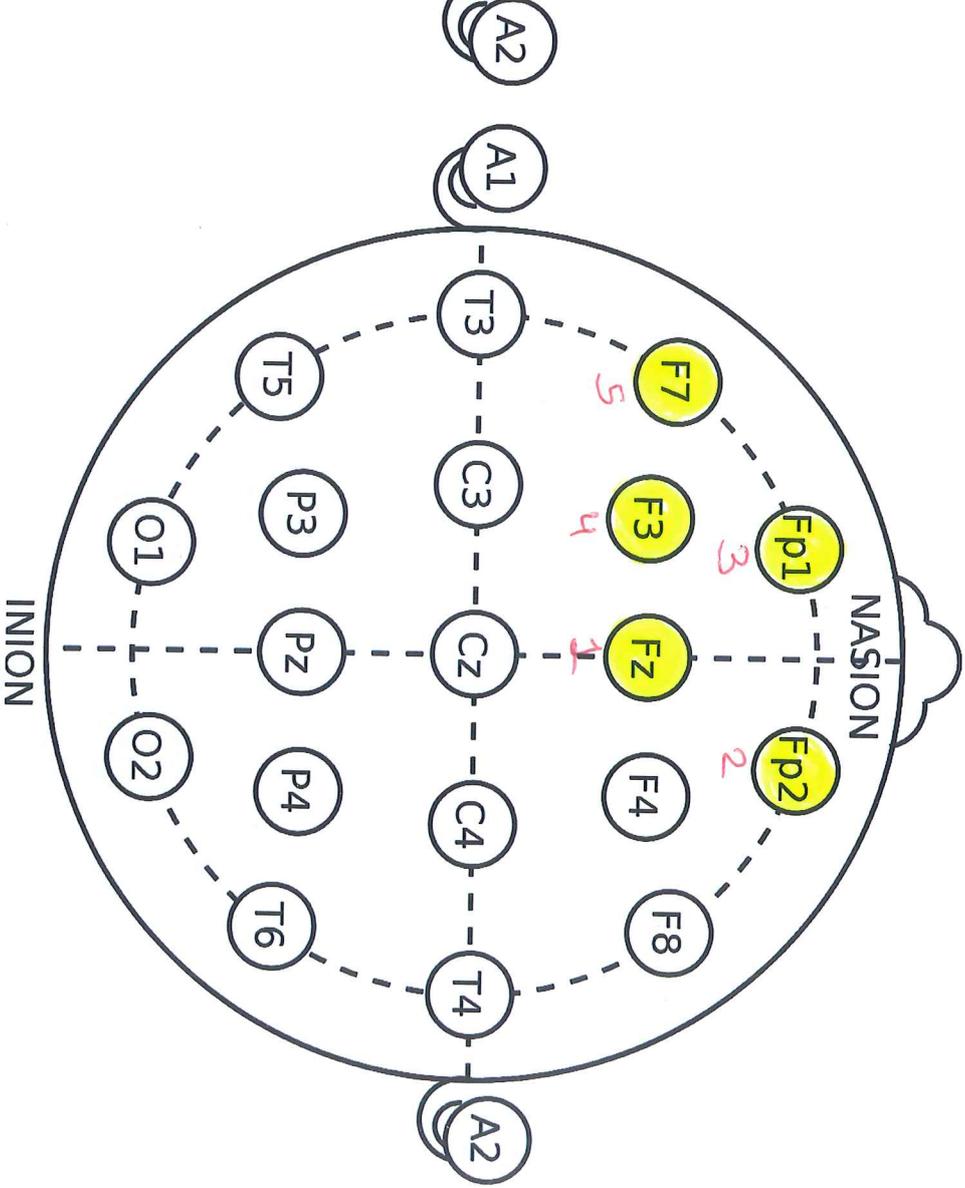
HF

Each 5 min data



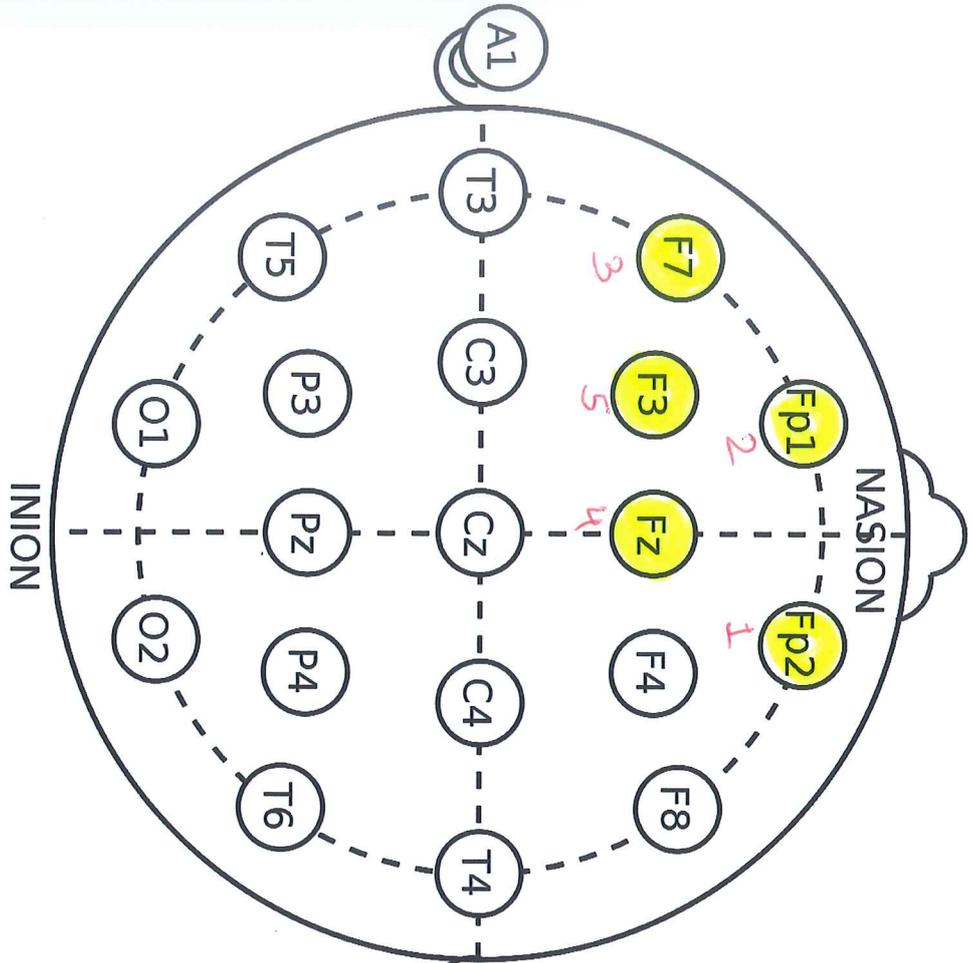
LF

Beta  
(EEG)

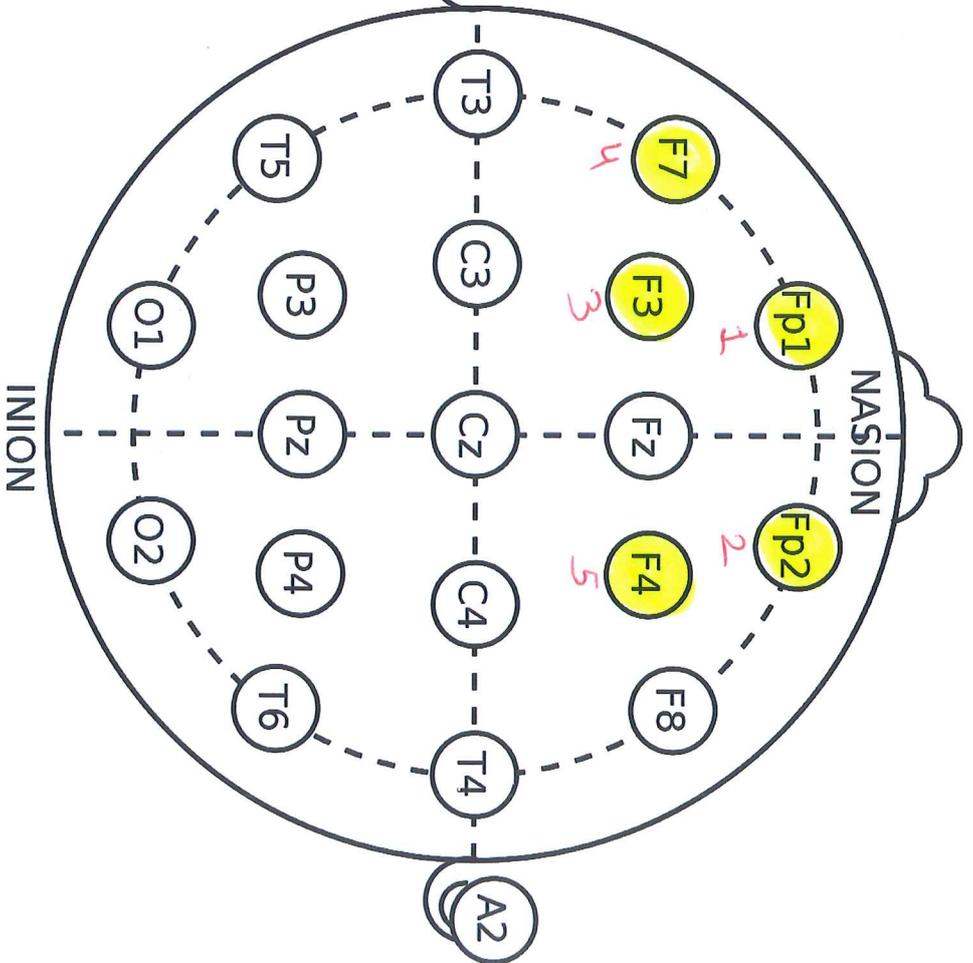


HF

Each 2.5 min data



LF



HF

Beta  
(EEG)

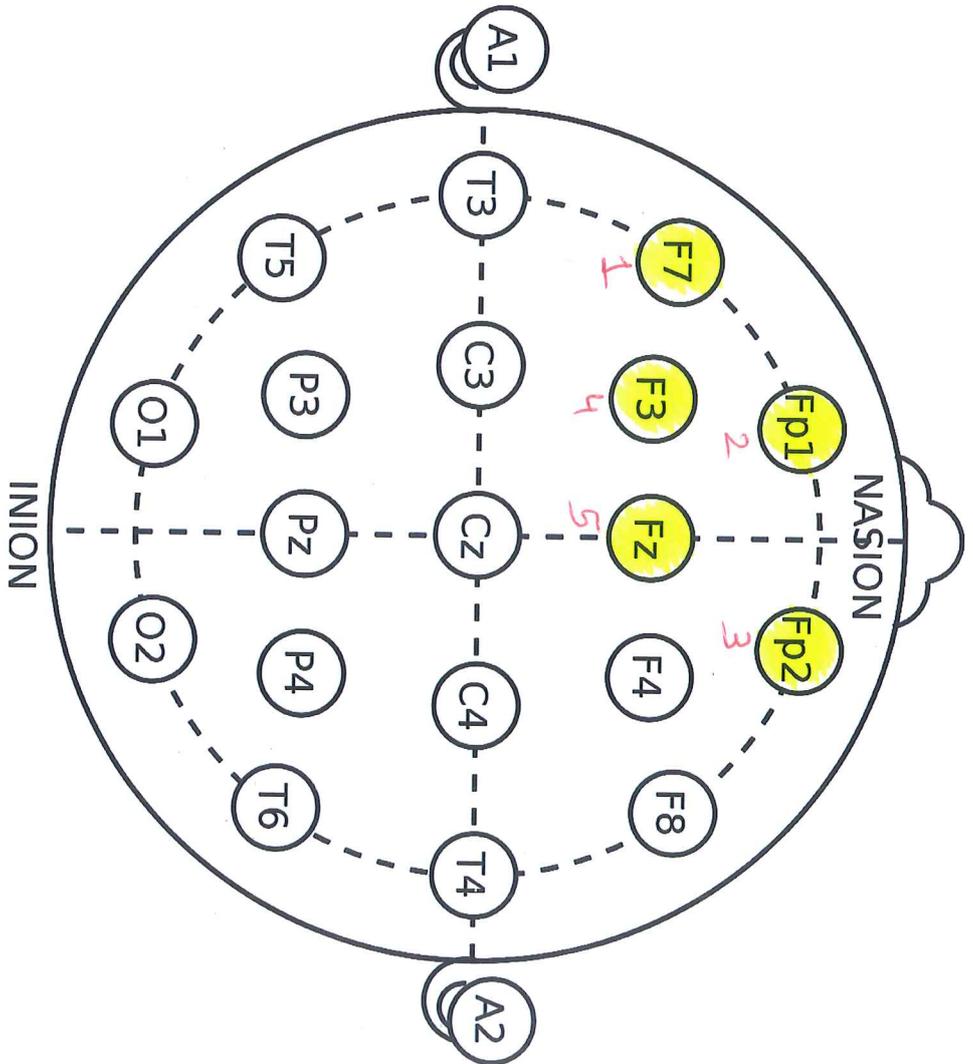
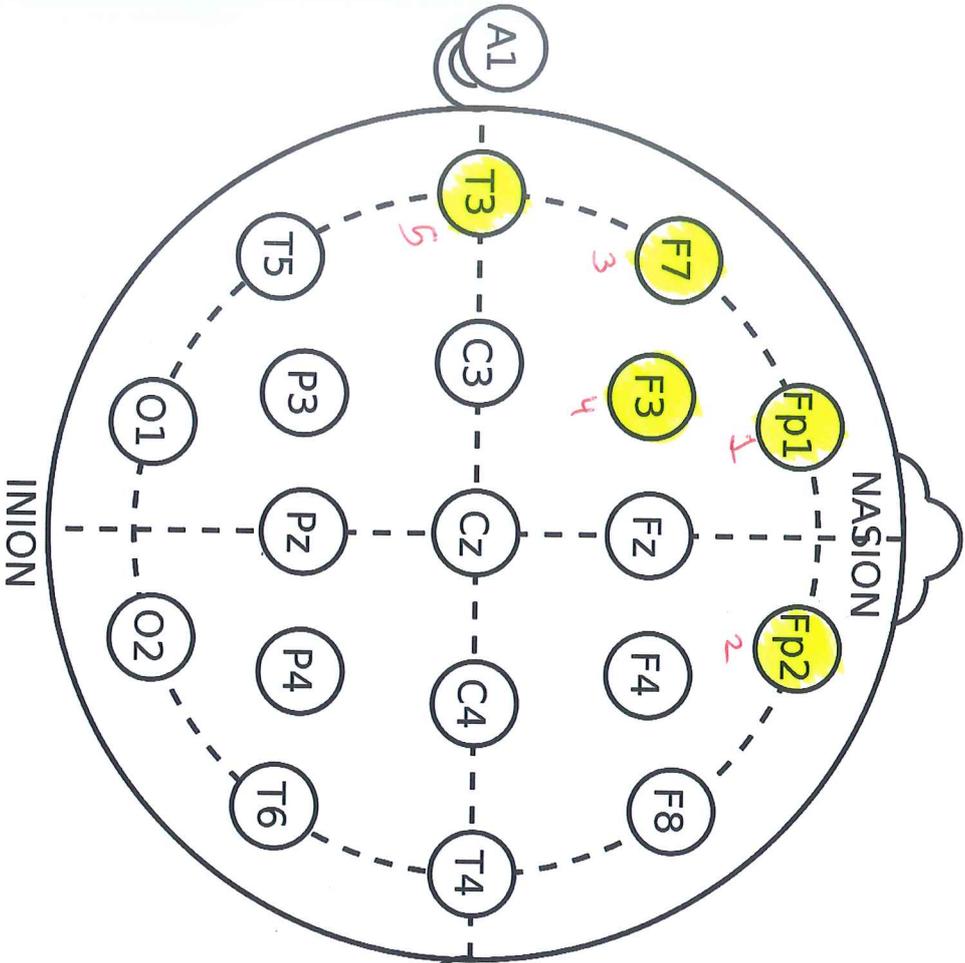
Each 1 min Data

Delta (EEG)

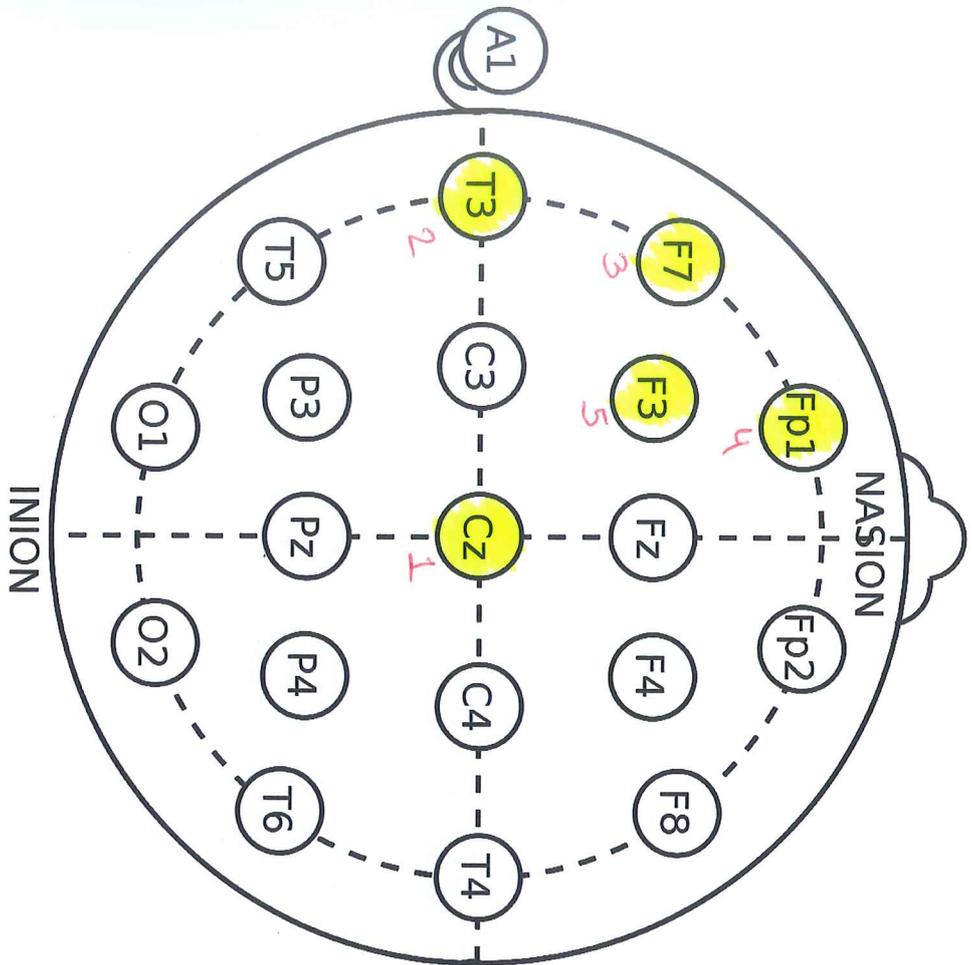
Each 5 min Data

LF

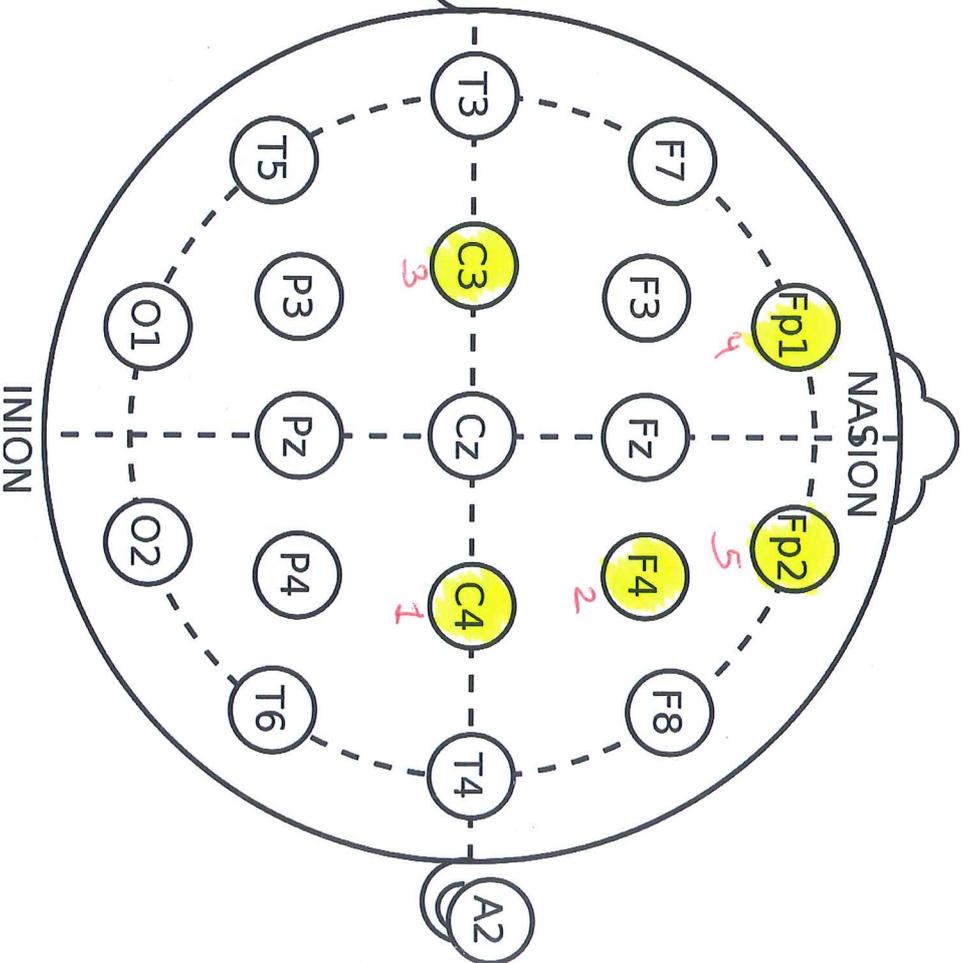
HF



(Average Result)



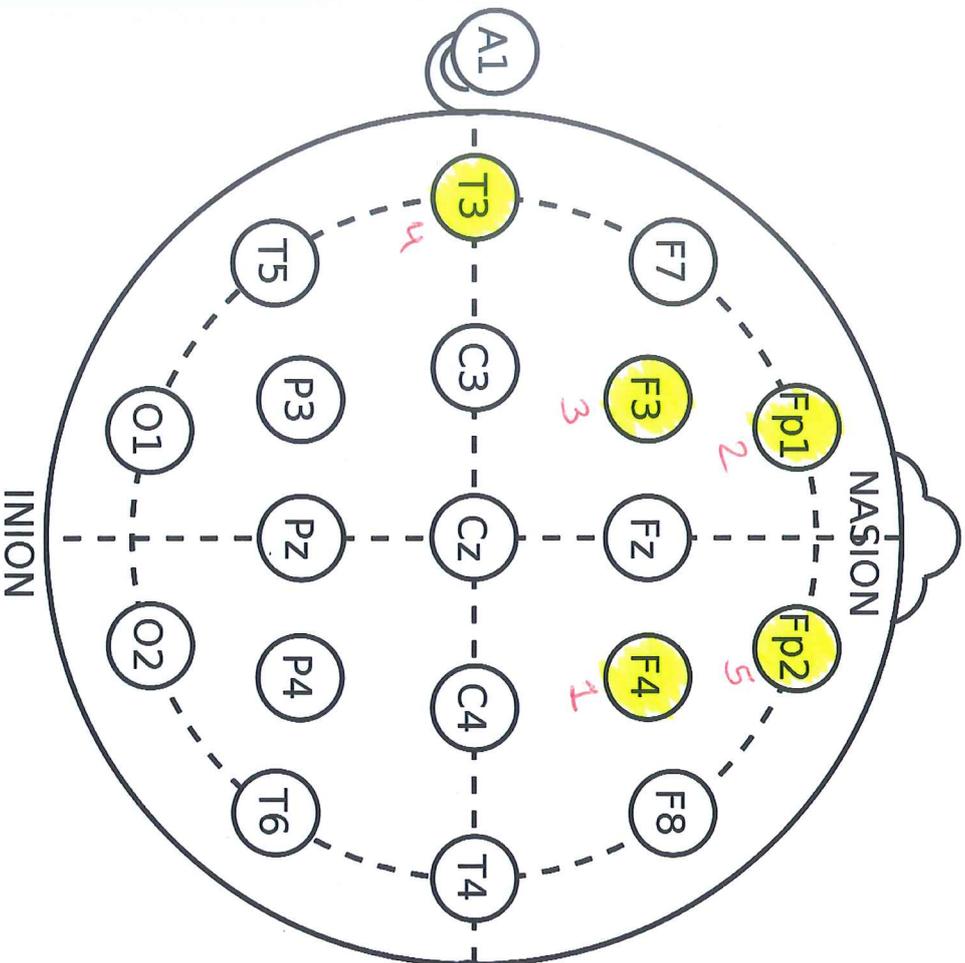
LF



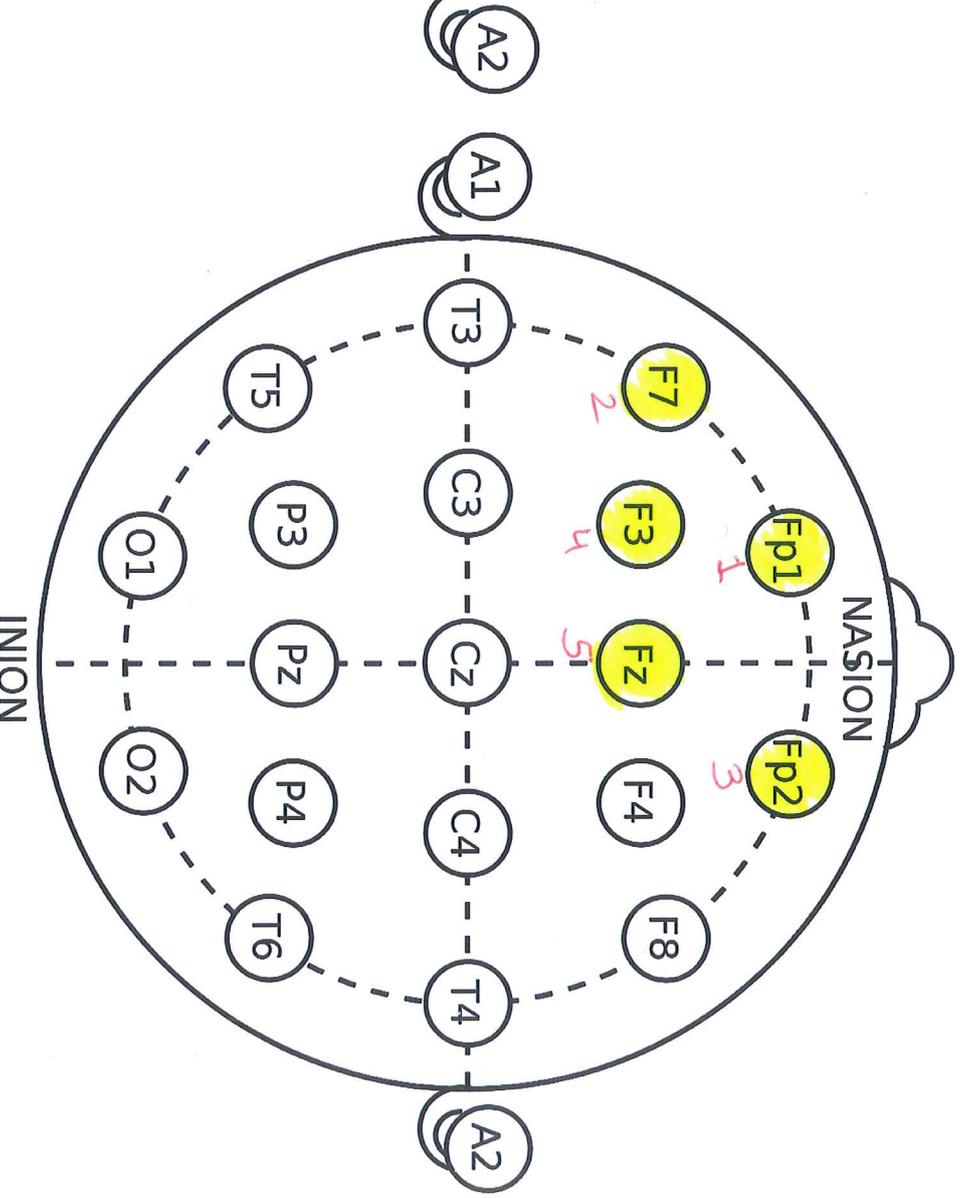
HF

Delta (EEG)

Each 2.5 min Data



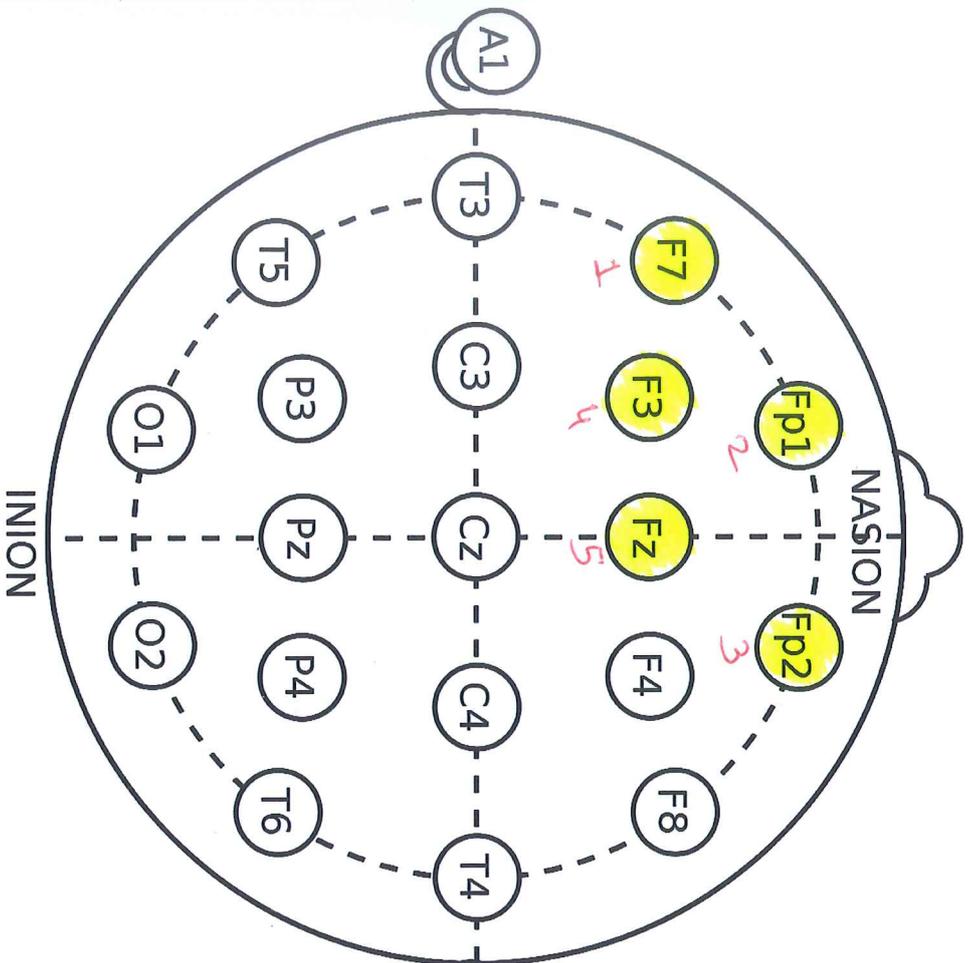
LF



HF

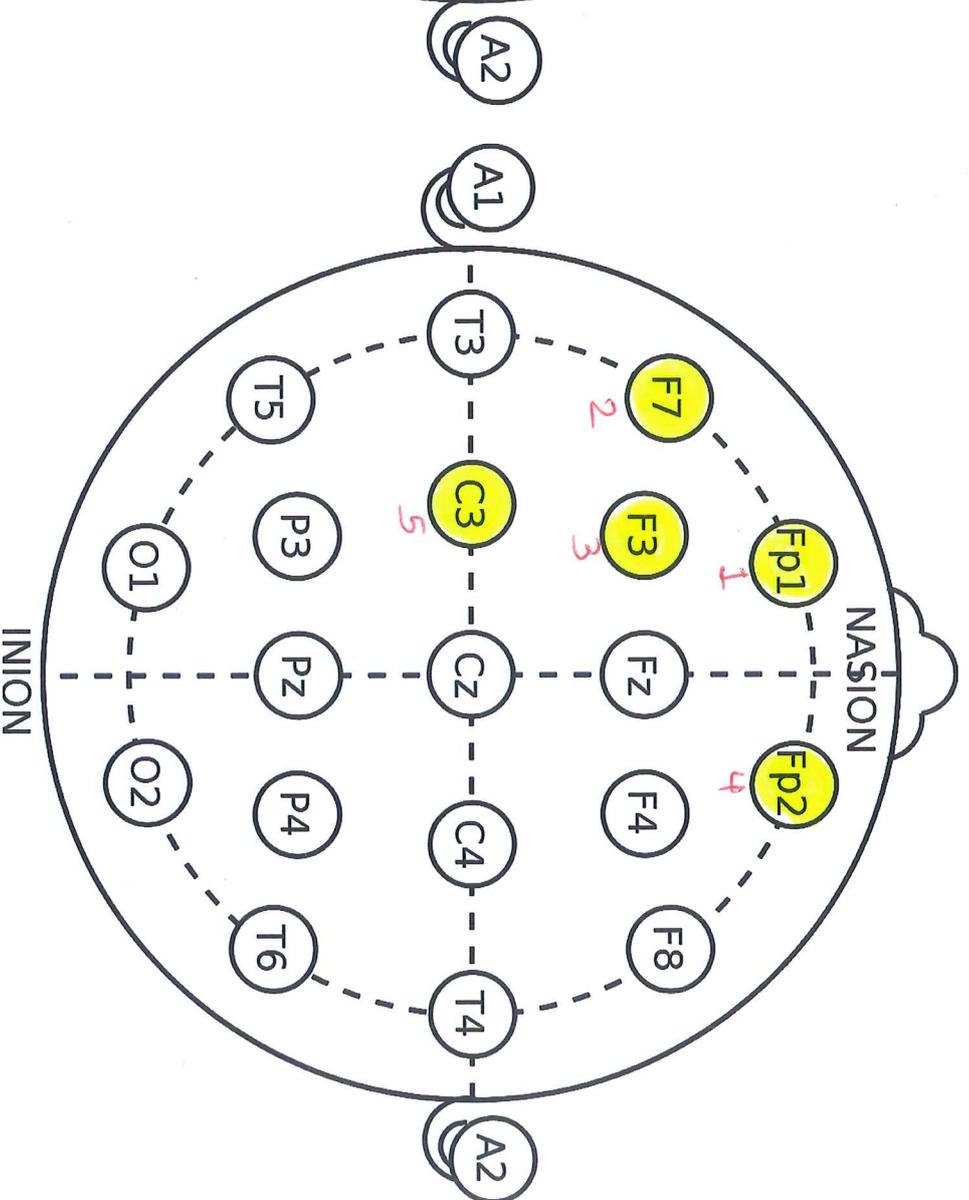
Delta (EEG)

Each 1 min Data



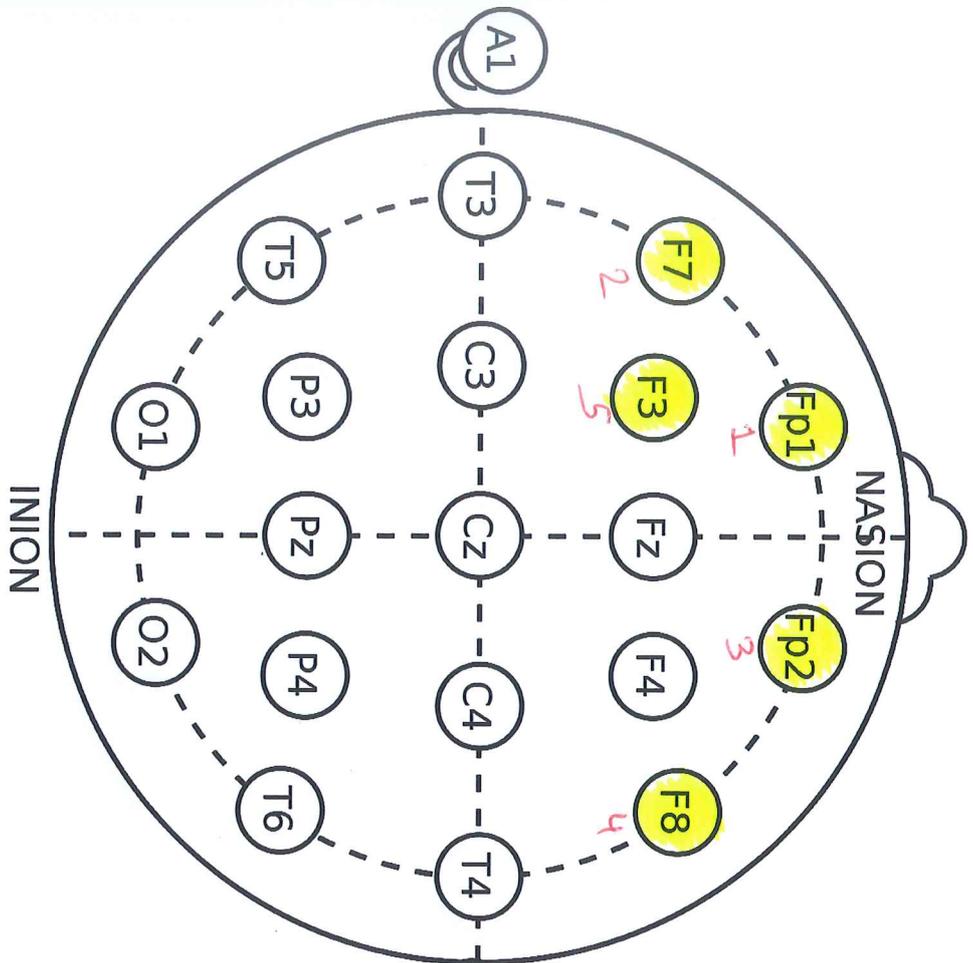
LF

Theta  
(EEG)



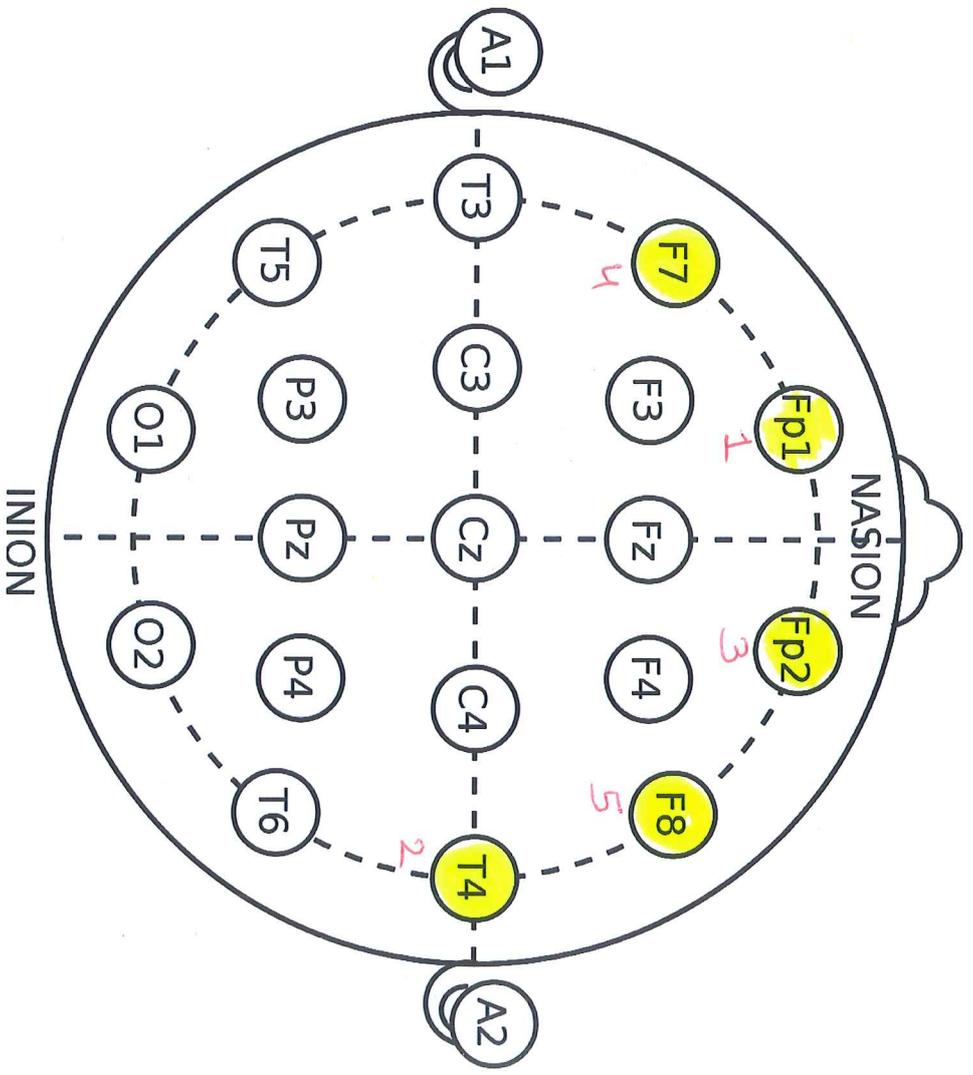
HF

Each 5 min Data



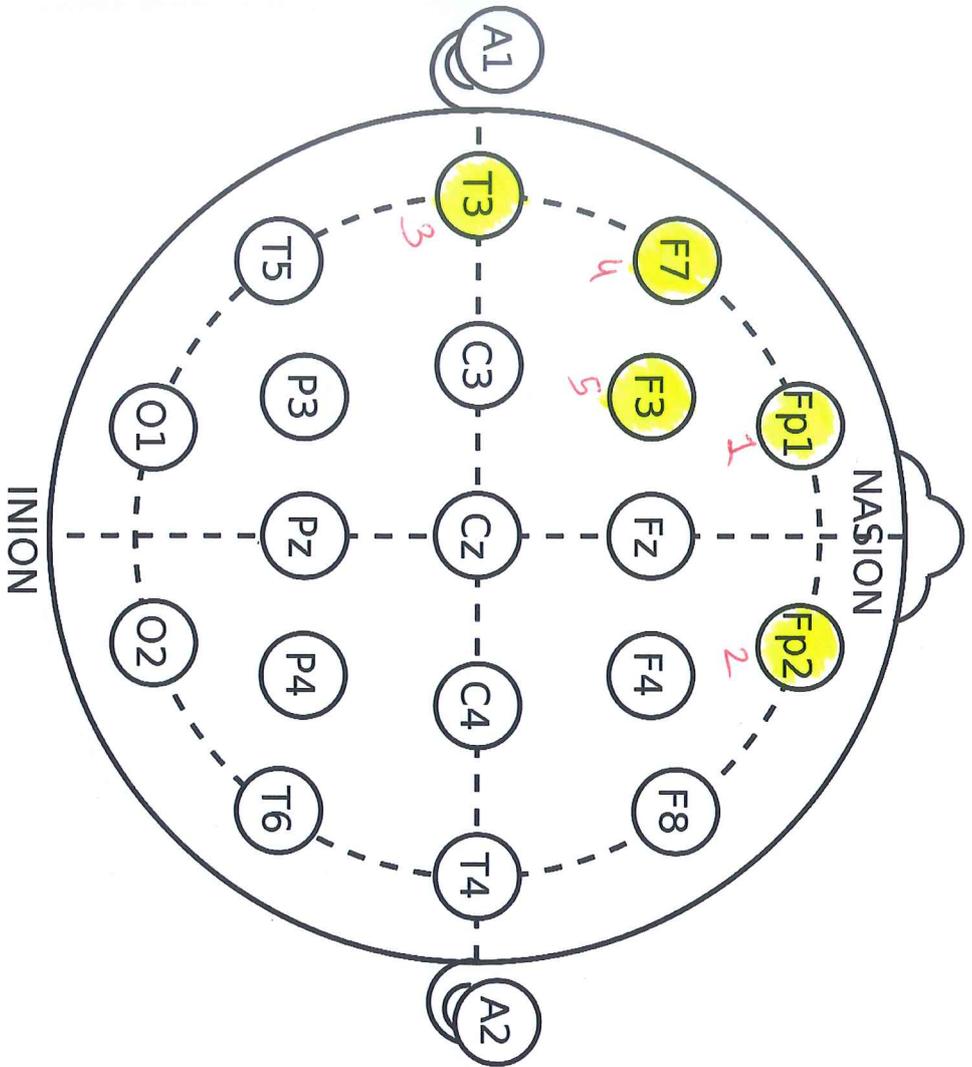
LF

Theta  
(CEEG)



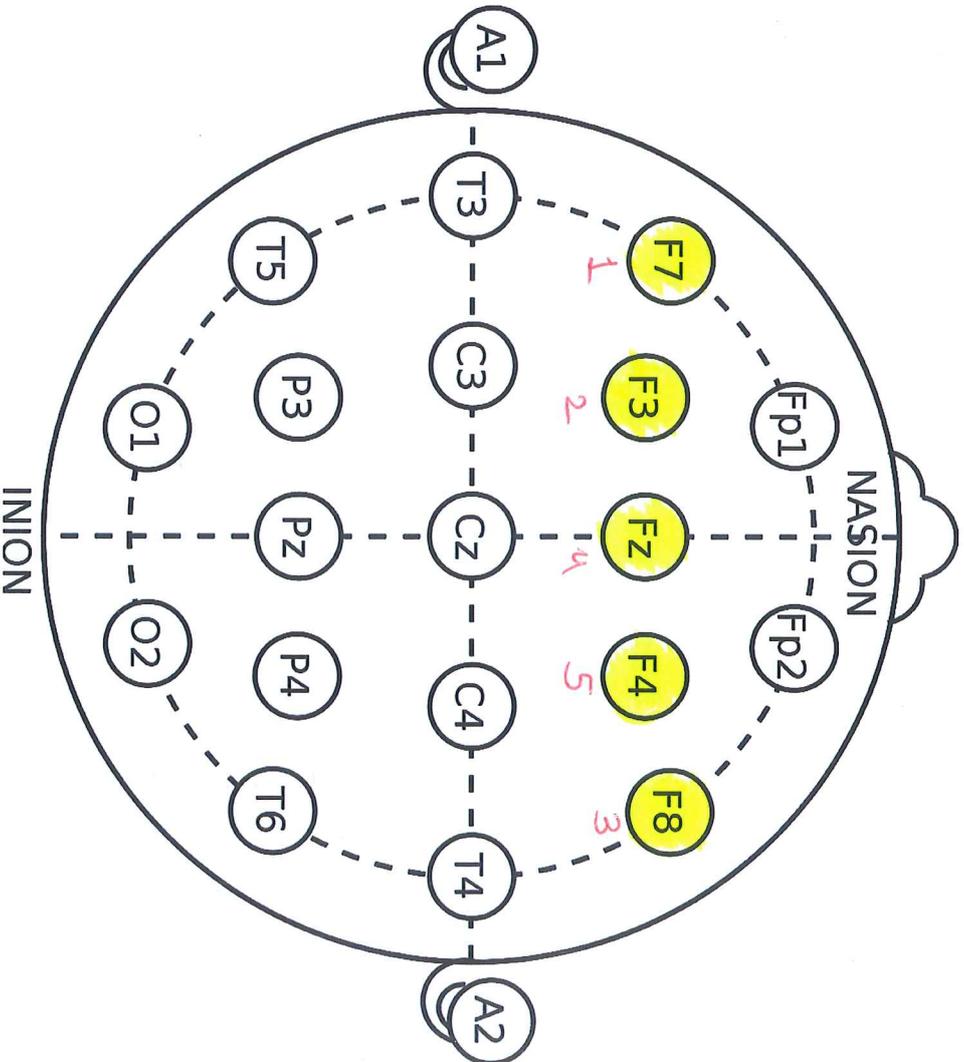
HF

Each 2.5 min Data



LF

Theta (EEG)

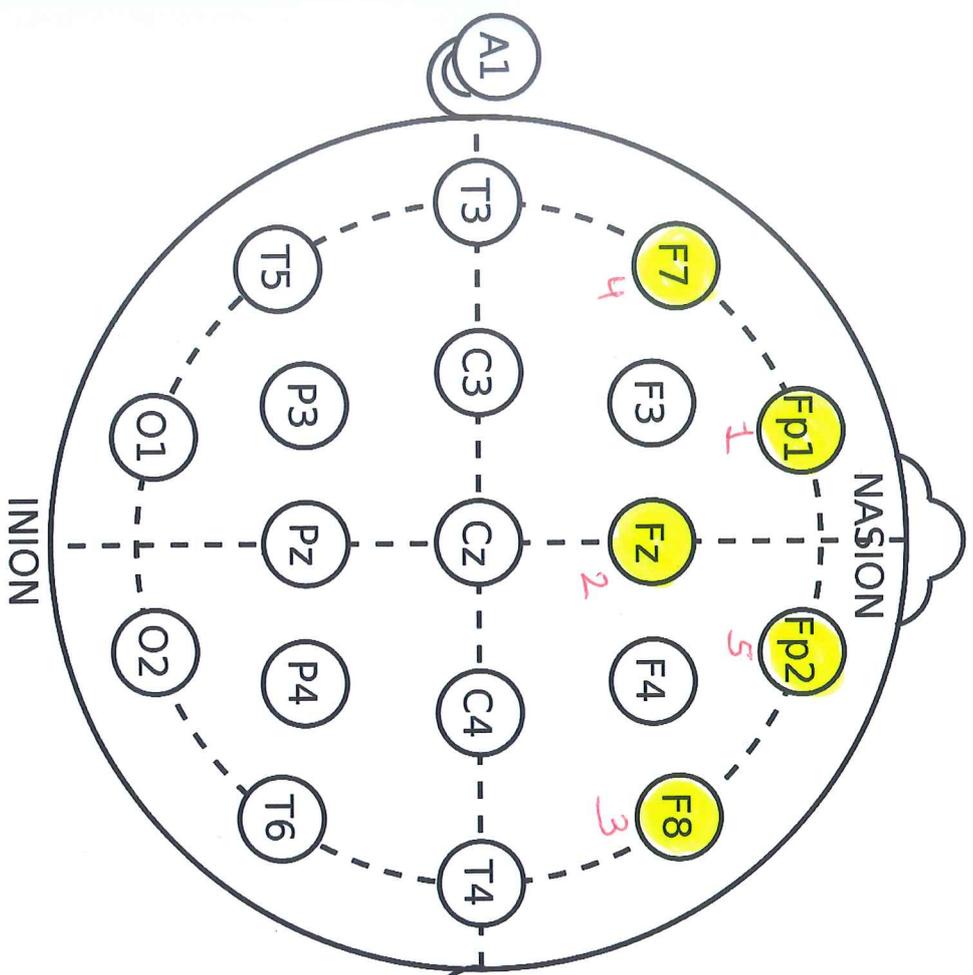


HF

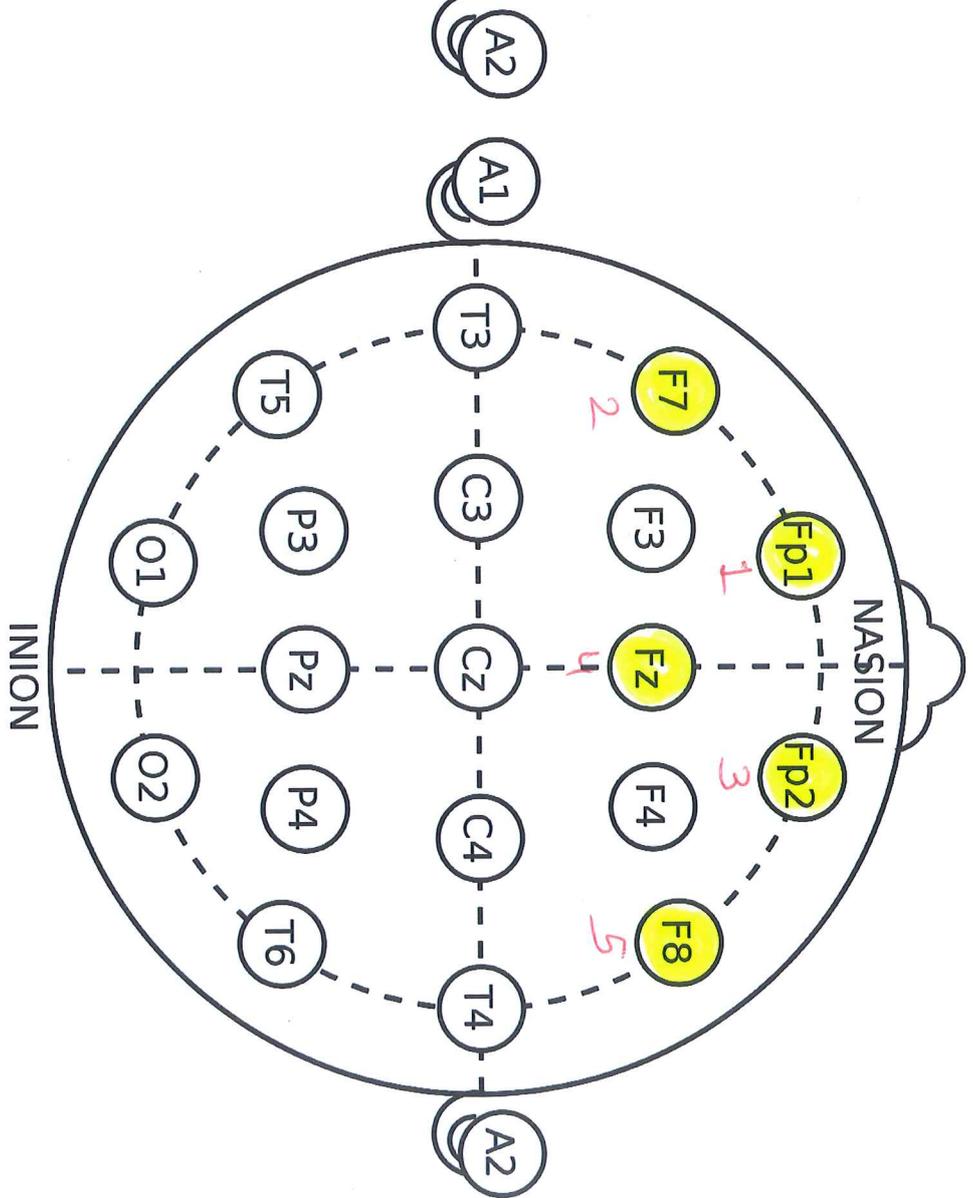
Even 1 min data

Each 5 min data

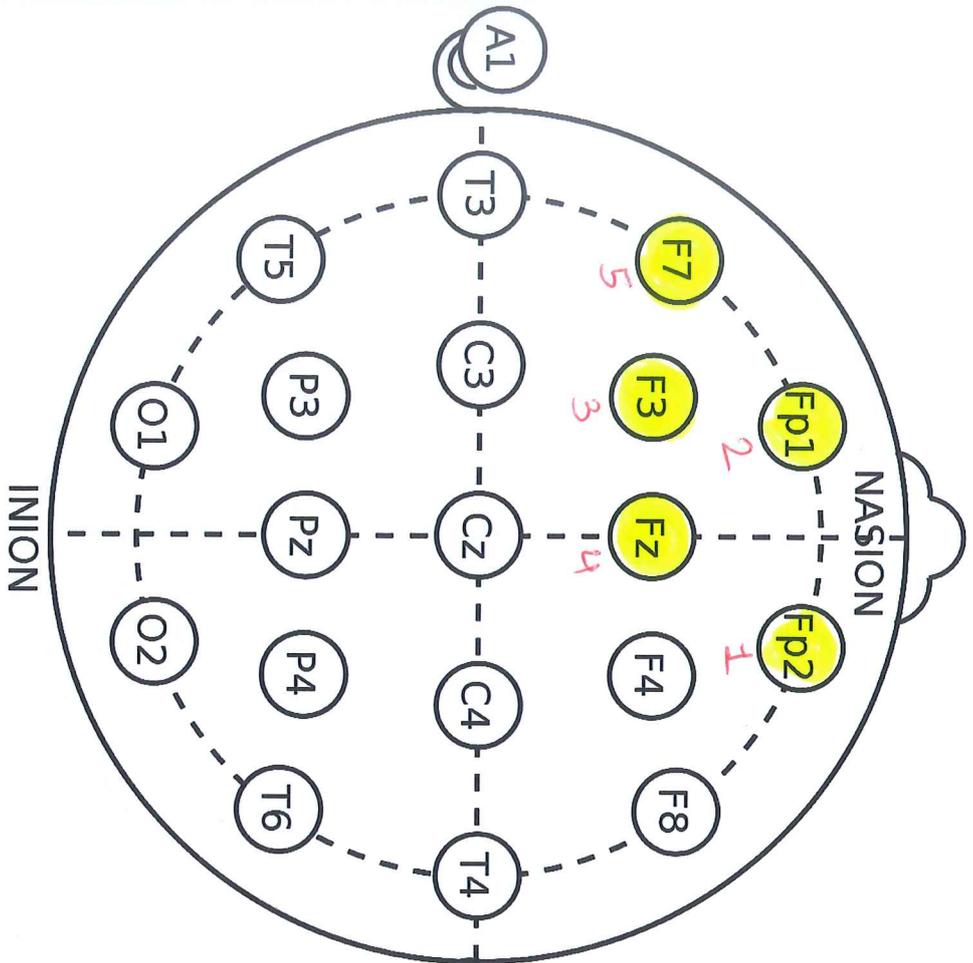
Alpha (EEG)



LF

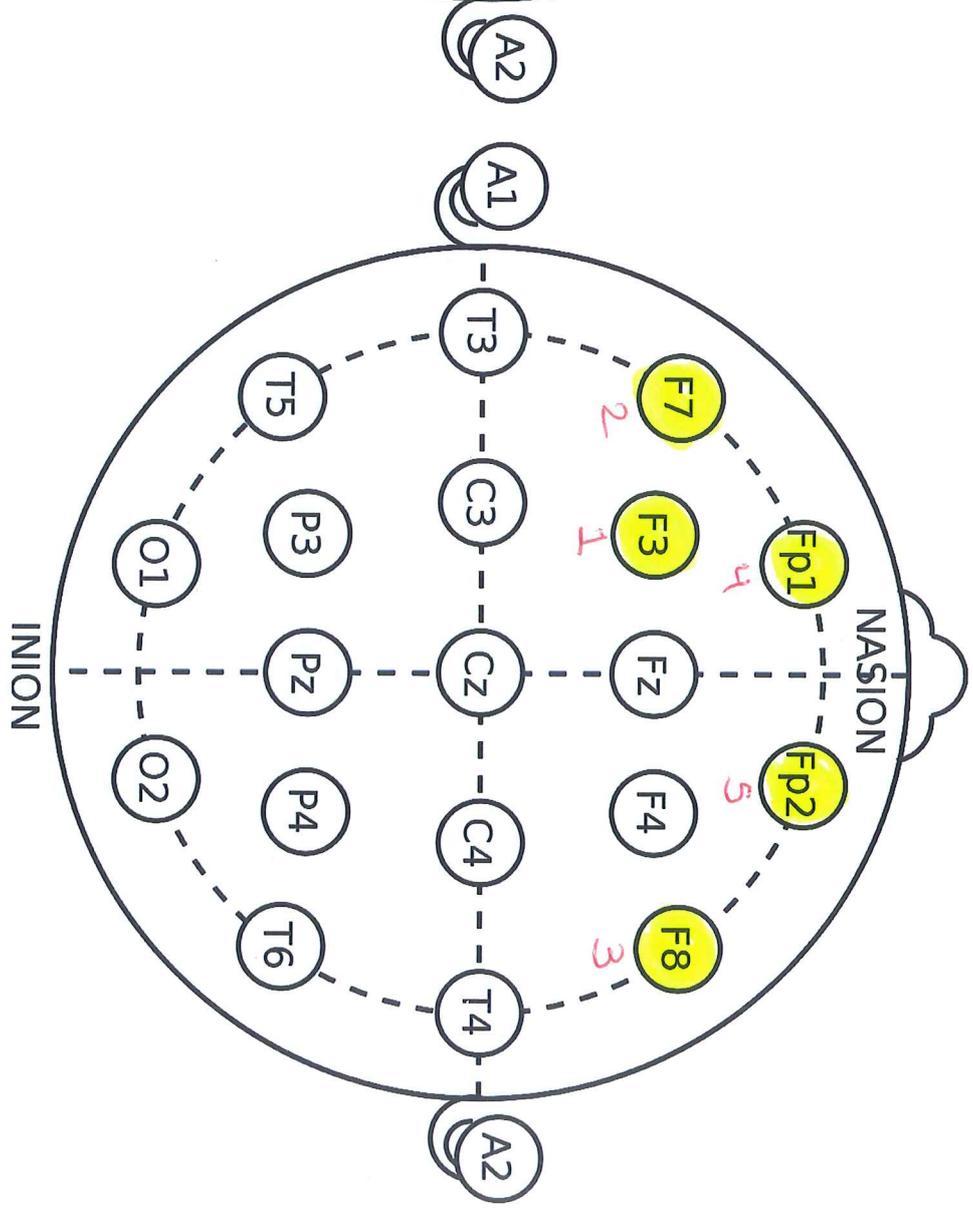


HF



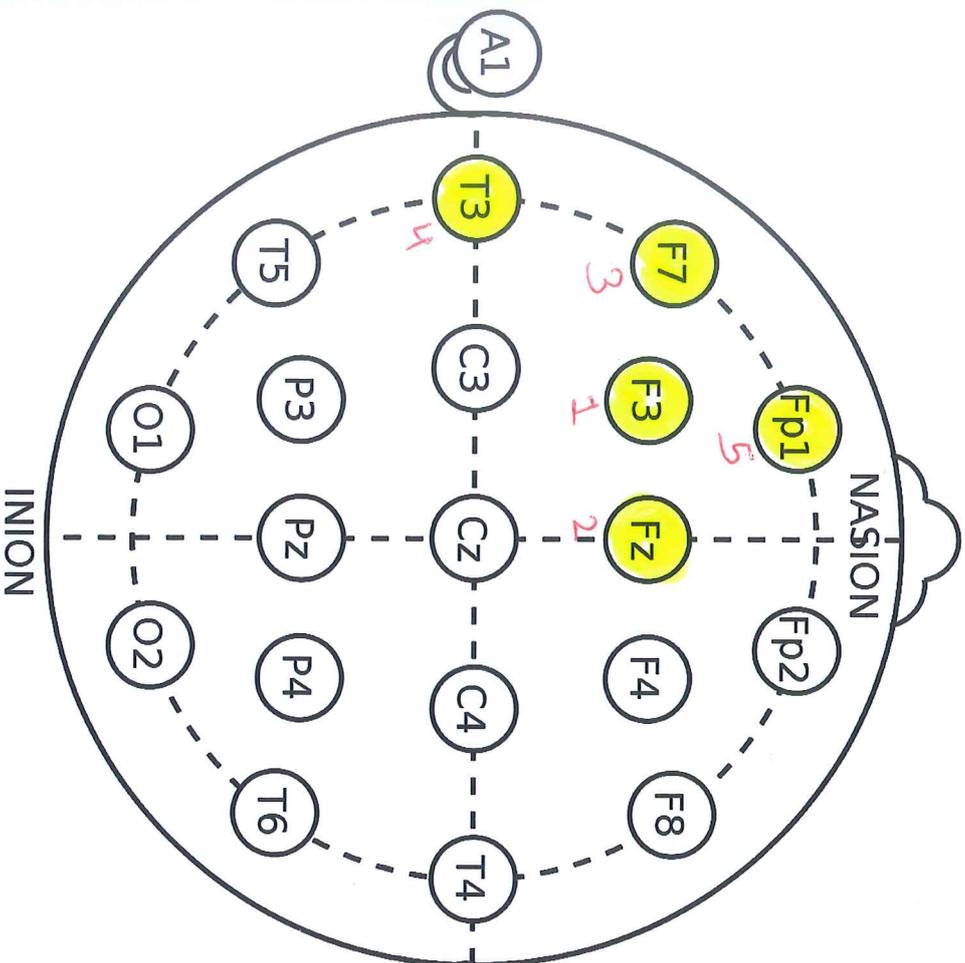
LF

Alpha  
(EEG)



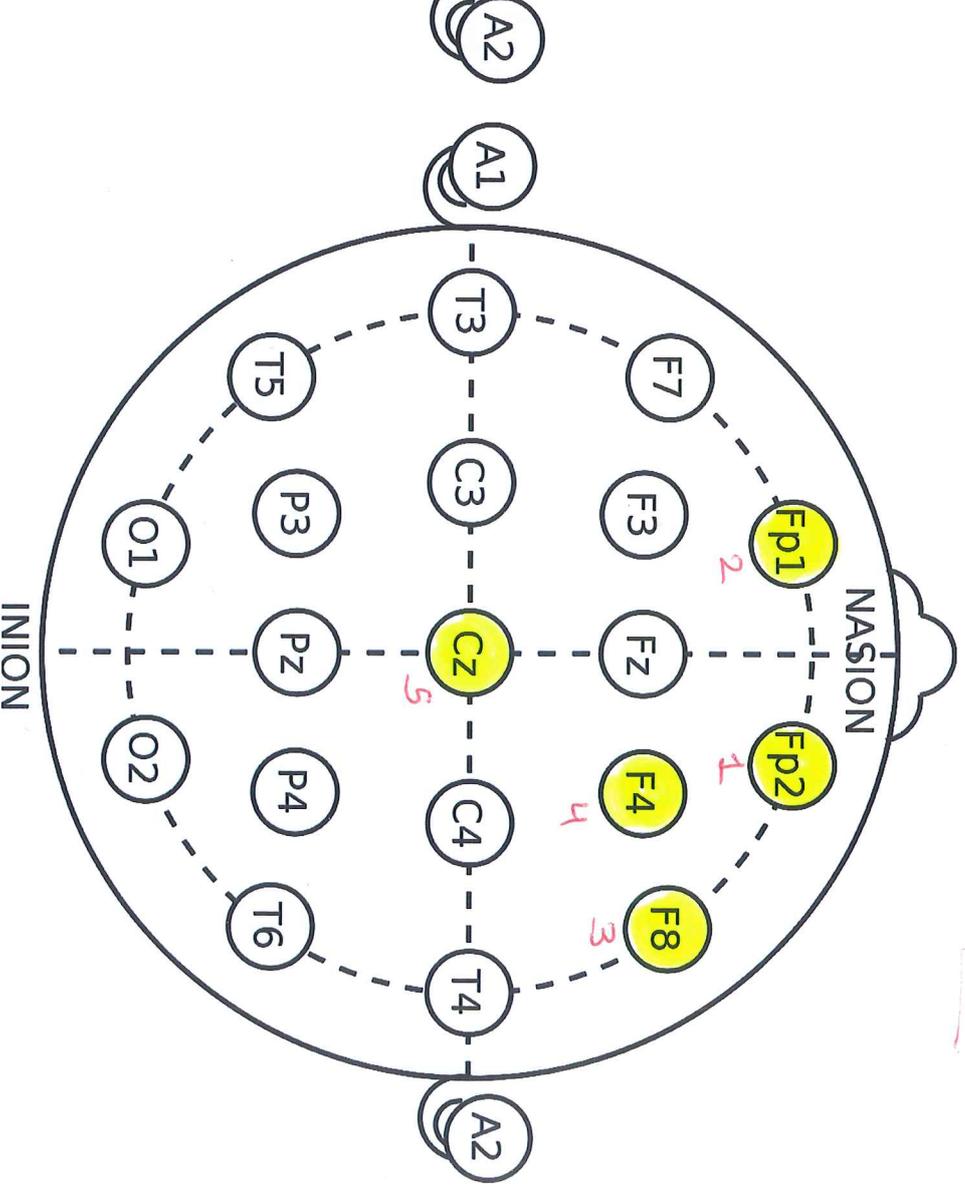
HF

Each 2.5 min Data



LF

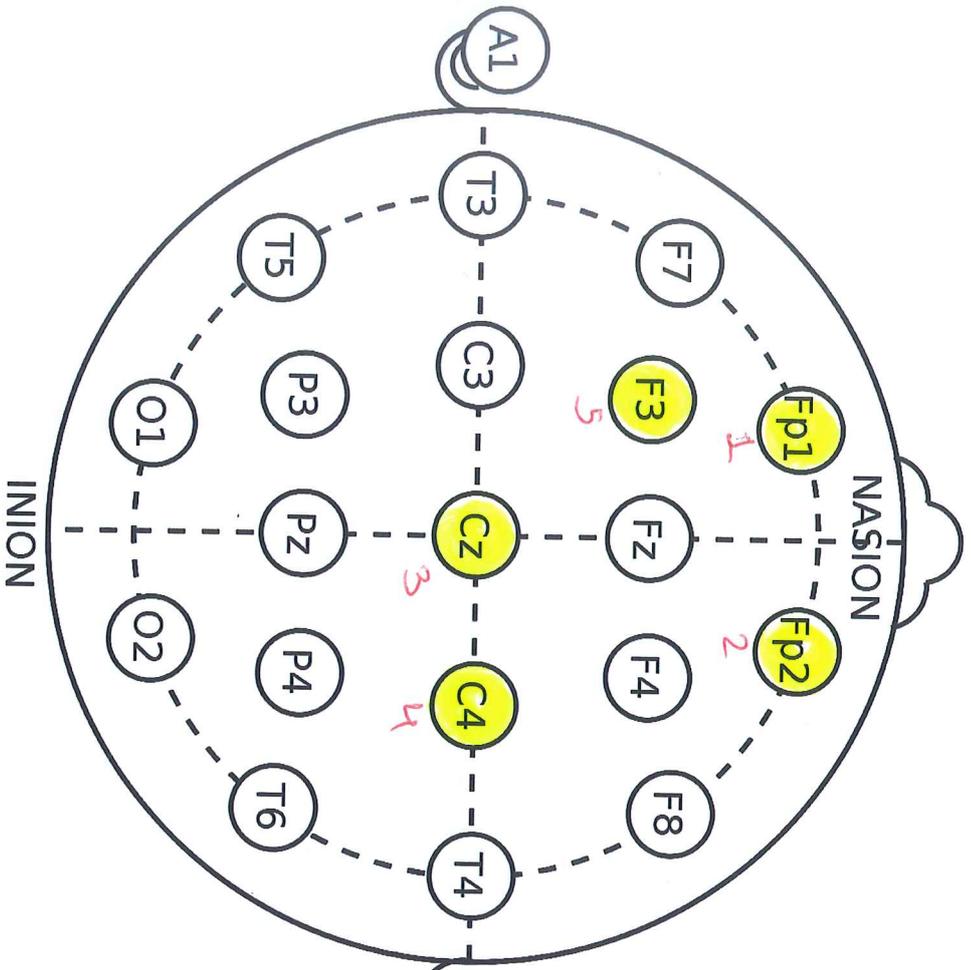
Alpha  
(EEG)



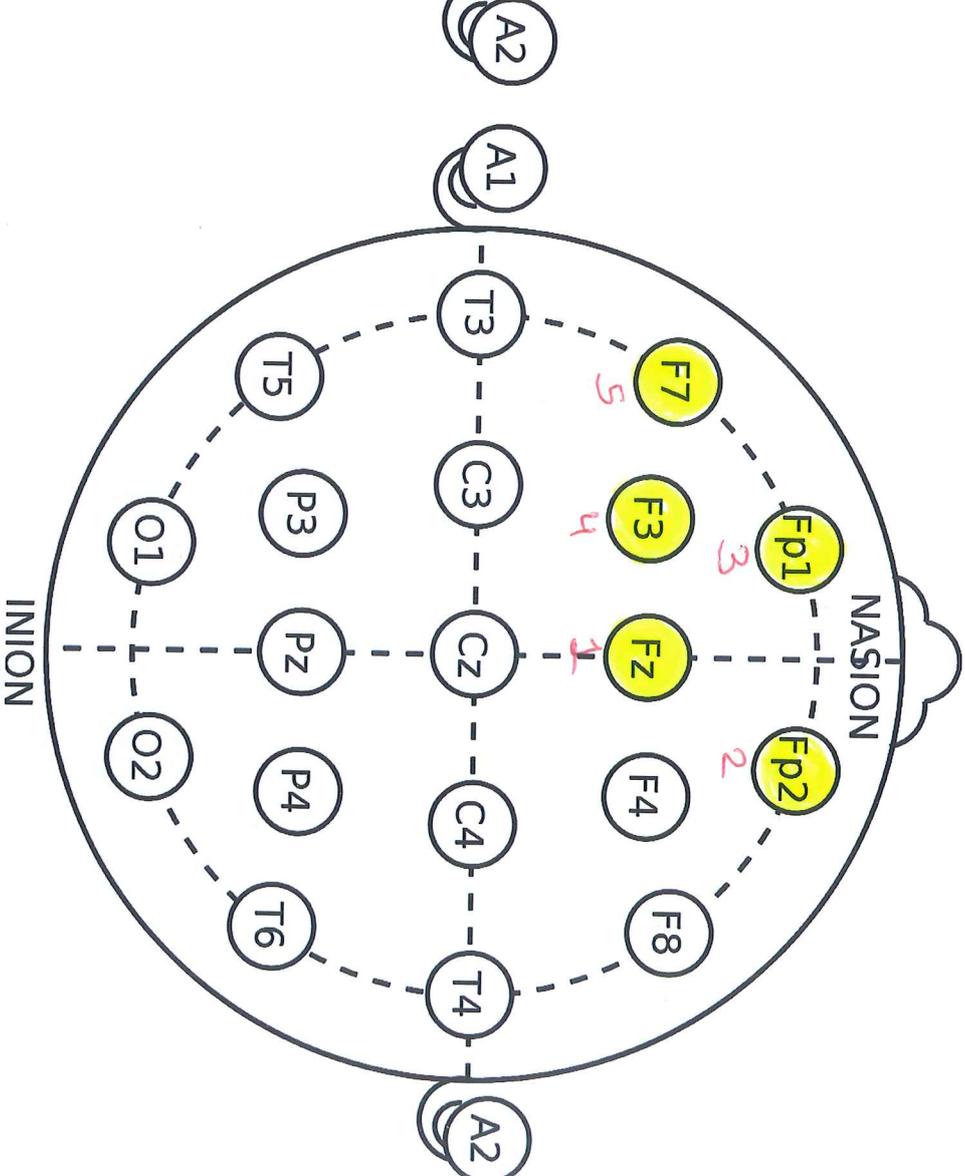
HF

Each 1 min Data





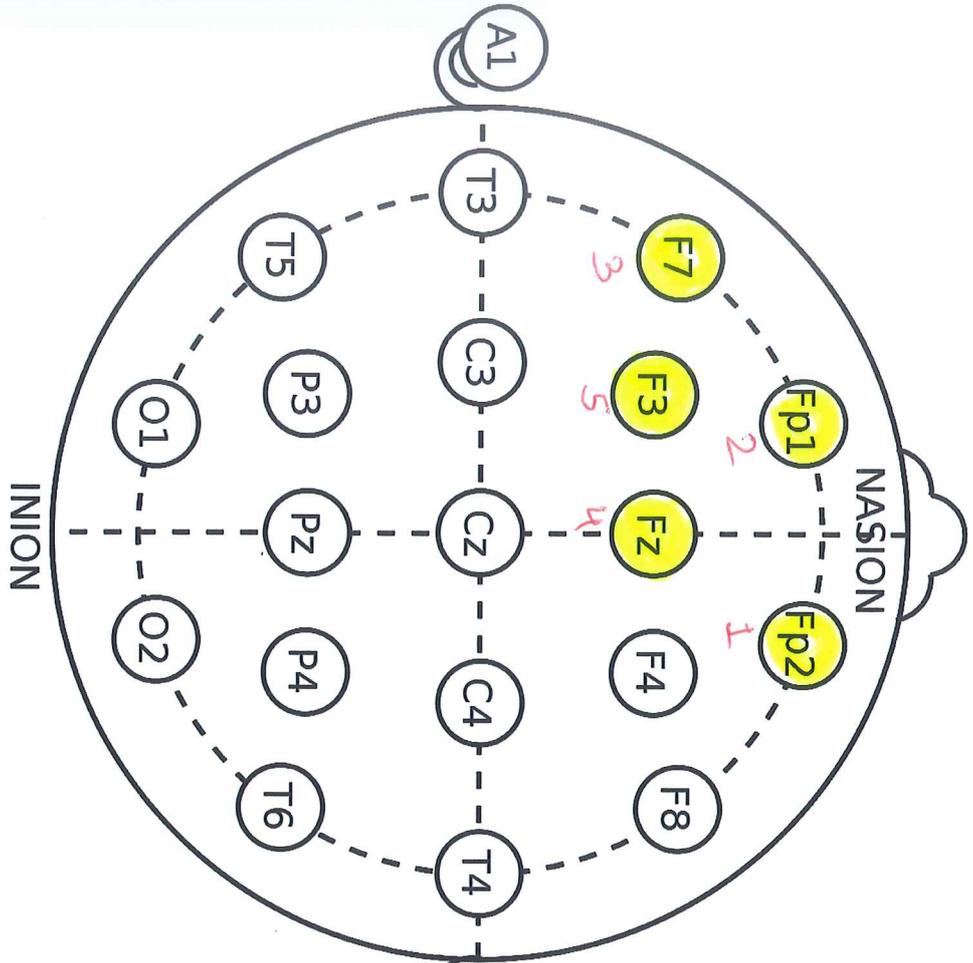
LF



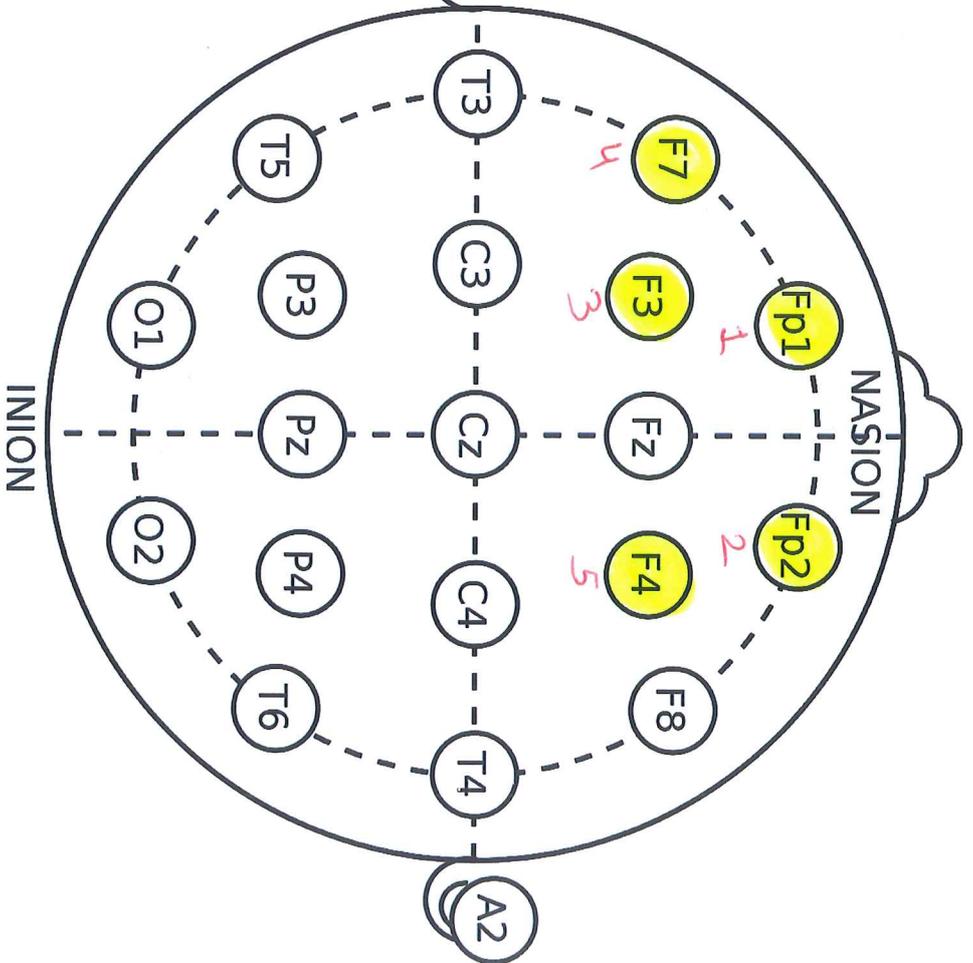
HF

Beta  
(EEG)

Each 2.5 min data



LF



HF

Beta  
(EEG)

Each 1 min Data

EEG	HRV	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14	E15	E16	E17	E18	E19	Average
DELTA	LF	0.31	0.34	0.41	0.22	0.19	0.17	0.41	0.28	0.43	0.33	0.27	0.32	0.27	0.28	0.27	0.23	0.28	0.30	0.27	<b>0.29</b>
	HF	0.25	0.24	0.35	0.29	0.22	0.36	0.39	0.23	0.47	0.39	0.34	0.23	0.28	0.32	0.30	0.28	0.23	0.28	0.31	<b>0.30</b>
THETA	LF	0.34	0.31	0.36	0.22	0.26	0.26	0.32	0.17	0.33	0.33	0.24	0.23	0.30	0.28	0.30	0.21	0.33	0.23	0.25	<b>0.28</b>
	HF	0.28	0.19	0.31	0.30	0.24	0.27	0.29	0.17	0.34	0.29	0.24	0.26	0.32	0.25	0.28	0.24	0.25	0.31	0.37	<b>0.27</b>
ALPHA	LF	0.28	0.20	0.26	0.20	0.28	0.22	0.31	0.29	0.23	0.29	0.28	0.22	0.25	0.30	0.29	0.26	0.33	0.19	0.15	<b>0.25</b>
	HF	0.29	0.22	0.20	0.19	0.23	0.24	0.34	0.28	0.39	0.22	0.19	0.32	0.21	0.30	0.28	0.32	0.19	0.35	0.29	<b>0.26</b>
BETA	LF	0.22	0.24	0.25	0.19	0.26	0.24	0.16	0.31	0.31	0.22	0.27	0.32	0.25	0.21	0.24	0.16	0.28	0.21	0.27	<b>0.24</b>
	HF	0.22	0.25	0.18	0.30	0.30	0.26	0.26	0.28	0.25	0.23	0.24	0.32	0.22	0.30	0.29	0.26	0.30	0.32	0.24	<b>0.27</b>

(a)

EEG	HRV	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14	E15	E16	E17	E18	E19	Average
DELTA	LF	0.32	0.37	0.38	0.23	0.15	0.22	0.38	0.29	0.41	0.34	0.25	0.35	0.28	0.30	0.27	0.26	0.30	0.29	0.26	<b>0.30</b>
	HF	0.24	0.26	0.37	0.30	0.29	0.37	0.36	0.23	0.47	0.34	0.30	0.30	0.26	0.36	0.27	0.29	0.29	0.24	0.28	<b>0.31</b>
THETA	LF	0.38	0.29	0.38	0.25	0.33	0.27	0.29	0.21	0.35	0.29	0.24	0.24	0.31	0.27	0.31	0.24	0.33	0.25	0.25	<b>0.29</b>
	HF	0.28	0.26	0.33	0.33	0.24	0.30	0.33	0.13	0.33	0.27	0.22	0.27	0.33	0.28	0.31	0.25	0.35	0.34	0.38	<b>0.29</b>
ALPHA	LF	0.24	0.17	0.24	0.23	0.24	0.25	0.30	0.30	0.21	0.32	0.27	0.18	0.24	0.30	0.22	0.21	0.33	0.18	0.15	<b>0.24</b>
	HF	0.27	0.19	0.18	0.24	0.24	0.23	0.29	0.33	0.31	0.18	0.17	0.25	0.15	0.27	0.18	0.33	0.25	0.37	0.24	<b>0.25</b>
BETA	LF	0.21	0.20	0.24	0.14	0.23	0.23	0.14	0.41	0.33	0.23	0.28	0.23	0.21	0.18	0.24	0.22	0.23	0.25	0.16	<b>0.23</b>
	HF	0.16	0.31	0.20	0.22	0.18	0.24	0.25	0.28	0.23	0.28	0.31	0.20	0.17	0.34	0.28	0.28	0.24	0.27	0.18	<b>0.24</b>

(b)

EEG	HRV	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14	E15	E16	E17	E18	E19	Average
DELTA	LF	0.34	0.31	0.35	0.24	0.17	0.21	0.39	0.26	0.45	0.28	0.29	0.31	0.29	0.25	0.22	0.25	0.33	0.27	0.25	<b>0.29</b>
	HF	0.27	0.24	0.36	0.29	0.17	0.37	0.41	0.21	0.45	0.35	0.36	0.26	0.26	0.31	0.29	0.27	0.31	0.24	0.25	<b>0.30</b>
THETA	LF	0.35	0.30	0.30	0.23	0.29	0.29	0.27	0.20	0.34	0.33	0.30	0.15	0.30	0.30	0.26	0.20	0.26	0.28	0.22	<b>0.27</b>
	HF	0.31	0.23	0.31	0.28	0.23	0.29	0.31	0.16	0.33	0.28	0.24	0.20	0.31	0.33	0.27	0.27	0.36	0.35	0.36	<b>0.29</b>
ALPHA	LF	0.23	0.25	0.33	0.27	0.25	0.17	0.28	0.35	0.24	0.32	0.32	0.22	0.28	0.27	0.27	0.24	0.27	0.27	0.23	<b>0.27</b>
	HF	0.33	0.33	0.20	0.27	0.26	0.25	0.25	0.25	0.34	0.27	0.29	0.24	0.20	0.19	0.22	0.27	0.20	0.38	0.34	<b>0.27</b>
BETA	LF	0.16	0.23	0.35	0.22	0.27	0.21	0.26	0.38	0.31	0.34	0.34	0.27	0.27	0.24	0.31	0.25	0.21	0.24	0.28	<b>0.27</b>
	HF	0.24	0.25	0.22	0.22	0.27	0.28	0.27	0.29	0.36	0.21	0.30	0.29	0.17	0.19	0.26	0.23	0.26	0.31	0.34	<b>0.26</b>

(c)

EEG	HRV	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14	E15	E16	E17	E18	E19	Average
DELTA	LF	0.36	0.26	0.31	0.31	0.22	0.25	0.35	0.28	0.38	0.30	0.27	0.30	0.32	0.21	0.31	0.21	0.31	0.26	0.29	<b>0.29</b>
	HF	0.29	0.20	0.38	0.33	0.18	0.34	0.39	0.24	0.45	0.35	0.37	0.25	0.22	0.27	0.29	0.30	0.32	0.28	0.29	<b>0.30</b>
THETA	LF	0.37	0.31	0.27	0.26	0.31	0.23	0.26	0.23	0.26	0.40	0.35	0.27	0.33	0.22	0.29	0.23	0.24	0.28	0.27	<b>0.28</b>
	HF	0.29	0.22	0.36	0.28	0.25	0.31	0.29	0.22	0.30	0.39	0.20	0.26	0.27	0.30	0.31	0.27	0.28	0.43	0.35	<b>0.29</b>
ALPHA	LF	0.43	0.39	0.36	0.29	0.24	0.20	0.38	0.19	0.32	0.38	0.32	0.26	0.34	0.29	0.36	0.25	0.22	0.28	0.33	<b>0.31</b>
	HF	0.27	0.30	0.37	0.28	0.33	0.31	0.38	0.19	0.36	0.40	0.18	0.31	0.33	0.35	0.35	0.28	0.28	0.48	0.41	<b>0.32</b>
BETA	LF	0.38	0.34	0.30	0.26	0.30	0.23	0.27	0.20	0.25	0.42	0.35	0.28	0.36	0.22	0.30	0.20	0.25	0.28	0.29	<b>0.29</b>
	HF	0.29	0.24	0.35	0.27	0.24	0.31	0.28	0.23	0.30	0.40	0.18	0.24	0.25	0.31	0.32	0.22	0.26	0.40	0.37	<b>0.29</b>

(d)

Table F.1: Dataset 1- Method 1 - Averaged participants' correlation performance: (a) Level 0 (No WT), (b) Level 1 WT, (c) Level 2 WT, (d) Level 3 WT)

EEG	HRV	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14	E15	E16	E17	E18	E19	Average	
DELTA	LF	0.20	0.24	0.35	0.29	0.34	0.24	0.26	0.27	0.17	0.22	0.23	0.25	0.37	0.38	0.30	0.30	0.38	0.33	0.28	0.28	<b>0.28</b>
	HF	0.16	0.35	0.32	0.28	0.35	0.31	0.28	0.30	0.12	0.27	0.16	0.22	0.16	0.28	0.24	0.27	0.41	0.26	0.19	0.26	<b>0.26</b>
THETA	LF	0.23	0.31	0.34	0.31	0.34	0.26	0.24	0.27	0.21	0.31	0.37	0.37	0.29	0.36	0.28	0.29	0.35	0.24	0.27	0.30	<b>0.30</b>
	HF	0.16	0.24	0.25	0.24	0.29	0.28	0.23	0.29	0.22	0.37	0.24	0.24	0.21	0.23	0.30	0.16	0.24	0.30	0.11	0.24	<b>0.24</b>
ALPHA	LF	0.22	0.30	0.32	0.34	0.26	0.27	0.26	0.27	0.23	0.21	0.34	0.40	0.34	0.39	0.22	0.27	0.26	0.22	0.29	0.28	<b>0.28</b>
	HF	0.20	0.21	0.19	0.28	0.28	0.26	0.18	0.29	0.15	0.24	0.22	0.28	0.14	0.21	0.17	0.16	0.22	0.22	0.13	0.21	<b>0.21</b>
BETA	LF	0.25	0.30	0.30	0.30	0.29	0.32	0.27	0.29	0.23	0.25	0.29	0.36	0.33	0.40	0.23	0.33	0.37	0.33	0.27	0.30	<b>0.30</b>
	HF	0.27	0.26	0.22	0.33	0.27	0.30	0.16	0.28	0.12	0.24	0.31	0.35	0.19	0.22	0.17	0.31	0.38	0.28	0.24	0.26	<b>0.26</b>

(a)

EEG	HRV	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14	E15	E16	E17	E18	E19	Average	
DELTA	LF	0.21	0.20	0.27	0.28	0.28	0.19	0.20	0.31	0.13	0.19	0.16	0.22	0.34	0.32	0.26	0.27	0.38	0.28	0.29	0.29	<b>0.25</b>
	HF	0.18	0.37	0.31	0.30	0.31	0.34	0.35	0.39	0.15	0.28	0.26	0.24	0.25	0.28	0.29	0.33	0.41	0.25	0.25	0.25	<b>0.29</b>
THETA	LF	0.26	0.28	0.33	0.32	0.32	0.27	0.28	0.26	0.23	0.31	0.37	0.30	0.29	0.36	0.29	0.30	0.39	0.27	0.27	0.27	<b>0.30</b>
	HF	0.13	0.24	0.25	0.24	0.22	0.26	0.22	0.28	0.21	0.35	0.24	0.17	0.22	0.23	0.29	0.13	0.19	0.32	0.11	0.22	<b>0.22</b>
ALPHA	LF	0.22	0.32	0.31	0.34	0.29	0.24	0.27	0.26	0.22	0.21	0.28	0.40	0.32	0.39	0.21	0.24	0.24	0.21	0.27	0.27	<b>0.28</b>
	HF	0.17	0.19	0.19	0.28	0.22	0.25	0.16	0.31	0.15	0.26	0.19	0.28	0.14	0.20	0.15	0.12	0.21	0.20	0.13	0.20	<b>0.20</b>
BETA	LF	0.23	0.31	0.30	0.33	0.26	0.28	0.27	0.25	0.23	0.22	0.26	0.32	0.33	0.42	0.23	0.33	0.35	0.29	0.30	0.30	<b>0.29</b>
	HF	0.23	0.22	0.23	0.34	0.25	0.30	0.14	0.26	0.12	0.22	0.30	0.33	0.18	0.22	0.17	0.34	0.38	0.27	0.28	0.28	<b>0.25</b>

(b)

EEG	HRV	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14	E15	E16	E17	E18	E19	Average	
DELTA	LF	0.24	0.27	0.34	0.19	0.35	0.28	0.28	0.31	0.16	0.17	0.20	0.24	0.31	0.32	0.27	0.31	0.36	0.24	0.32	0.27	<b>0.27</b>
	HF	0.29	0.35	0.40	0.33	0.33	0.39	0.24	0.38	0.16	0.36	0.28	0.22	0.28	0.18	0.30	0.32	0.44	0.23	0.27	0.27	<b>0.30</b>
THETA	LF	0.27	0.28	0.32	0.33	0.34	0.28	0.27	0.27	0.22	0.29	0.37	0.31	0.29	0.37	0.29	0.29	0.38	0.27	0.27	0.27	<b>0.30</b>
	HF	0.13	0.24	0.23	0.23	0.22	0.24	0.22	0.29	0.20	0.34	0.24	0.17	0.22	0.23	0.28	0.13	0.19	0.31	0.10	0.10	<b>0.22</b>
ALPHA	LF	0.25	0.33	0.29	0.30	0.36	0.26	0.20	0.22	0.23	0.22	0.28	0.38	0.33	0.41	0.18	0.26	0.25	0.23	0.29	0.29	<b>0.28</b>
	HF	0.16	0.19	0.22	0.27	0.21	0.30	0.25	0.30	0.13	0.27	0.19	0.28	0.15	0.23	0.15	0.20	0.25	0.23	0.16	0.22	<b>0.22</b>
BETA	LF	0.24	0.32	0.27	0.31	0.33	0.24	0.22	0.20	0.24	0.25	0.24	0.34	0.31	0.40	0.23	0.22	0.29	0.21	0.22	0.22	<b>0.27</b>
	HF	0.18	0.19	0.20	0.30	0.28	0.26	0.18	0.26	0.12	0.21	0.20	0.30	0.22	0.22	0.19	0.28	0.30	0.32	0.18	0.18	<b>0.23</b>

(c)

EEG	HRV	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14	E15	E16	E17	E18	E19	Average	
DELTA	LF	0.25	0.29	0.36	0.17	0.37	0.29	0.32	0.31	0.17	0.20	0.21	0.23	0.28	0.32	0.27	0.33	0.32	0.21	0.33	0.27	<b>0.27</b>
	HF	0.30	0.34	0.43	0.32	0.34	0.39	0.28	0.36	0.16	0.39	0.27	0.21	0.30	0.14	0.28	0.34	0.40	0.21	0.28	0.28	<b>0.30</b>
THETA	LF	0.29	0.26	0.32	0.32	0.29	0.25	0.25	0.32	0.22	0.29	0.34	0.37	0.30	0.32	0.38	0.39	0.35	0.26	0.28	0.28	<b>0.31</b>
	HF	0.14	0.25	0.21	0.27	0.26	0.25	0.31	0.30	0.13	0.34	0.25	0.22	0.21	0.24	0.31	0.21	0.22	0.27	0.08	0.24	<b>0.24</b>
ALPHA	LF	0.29	0.25	0.40	0.34	0.28	0.24	0.22	0.30	0.20	0.28	0.29	0.40	0.30	0.30	0.35	0.34	0.32	0.37	0.28	0.28	<b>0.30</b>
	HF	0.16	0.24	0.22	0.29	0.31	0.23	0.30	0.18	0.15	0.33	0.12	0.23	0.20	0.20	0.27	0.19	0.22	0.29	0.17	0.23	<b>0.23</b>
BETA	LF	0.28	0.28	0.35	0.33	0.27	0.26	0.25	0.24	0.22	0.35	0.23	0.38	0.30	0.33	0.40	0.33	0.31	0.28	0.31	0.30	<b>0.30</b>
	HF	0.15	0.23	0.23	0.25	0.36	0.25	0.33	0.23	0.16	0.41	0.14	0.20	0.20	0.25	0.32	0.17	0.25	0.26	0.21	0.24	<b>0.24</b>

(d)

Table F.2: Dataset 2- Method 1 - Averaged participants' correlation performance: (a) Level 0 (No WT), (b) Level 1 WT, (c) Level 2 WT, (d) Level 3 WT

EEG	HRV	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14	E15	E16	E17	E18	E19	Average
DELTA	LF	0.25	0.29	0.27	0.29	0.29	0.29	0.28	0.31	0.27	0.27	0.29	0.25	0.26	0.27	0.16	0.26	0.34	0.27	0.25	<b>0.27</b>
	HF	0.23	0.24	0.19	0.22	0.20	0.22	0.19	0.25	0.21	0.26	0.16	0.26	0.26	0.22	0.28	0.29	0.24	0.35	0.33	<b>0.24</b>
THETA	LF	0.33	0.29	0.28	0.27	0.26	0.27	0.26	0.28	0.24	0.37	0.35	0.27	0.28	0.40	0.38	0.34	0.29	0.29	0.26	<b>0.30</b>
	HF	0.29	0.34	0.20	0.19	0.18	0.18	0.19	0.23	0.21	0.29	0.31	0.32	0.27	0.31	0.35	0.36	0.34	0.21	0.31	<b>0.27</b>
ALPHA	LF	0.19	0.28	0.31	0.26	0.27	0.24	0.32	0.28	0.29	0.29	0.34	0.26	0.35	0.39	0.31	0.28	0.32	0.26	0.26	<b>0.29</b>
	HF	0.23	0.25	0.25	0.21	0.18	0.22	0.25	0.21	0.27	0.21	0.27	0.25	0.30	0.31	0.31	0.27	0.21	0.33	0.34	<b>0.26</b>
BETA	LF	0.26	0.33	0.28	0.27	0.27	0.20	0.32	0.26	0.33	0.27	0.30	0.25	0.32	0.30	0.31	0.27	0.34	0.25	0.30	<b>0.29</b>
	HF	0.23	0.26	0.25	0.28	0.19	0.22	0.24	0.18	0.29	0.21	0.18	0.22	0.29	0.22	0.31	0.33	0.20	0.33	0.36	<b>0.25</b>

(a)

EEG	HRV	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14	E15	E16	E17	E18	E19	Average
DELTA	LF	0.25	0.29	0.27	0.29	0.29	0.29	0.28	0.31	0.27	0.27	0.29	0.25	0.26	0.27	0.17	0.26	0.34	0.27	0.25	<b>0.27</b>
	HF	0.23	0.24	0.19	0.22	0.20	0.22	0.18	0.25	0.21	0.26	0.16	0.26	0.26	0.22	0.28	0.29	0.24	0.35	0.33	<b>0.24</b>
THETA	LF	0.33	0.31	0.28	0.26	0.26	0.27	0.26	0.27	0.24	0.38	0.35	0.28	0.28	0.40	0.37	0.35	0.30	0.29	0.26	<b>0.30</b>
	HF	0.30	0.34	0.20	0.19	0.18	0.18	0.20	0.23	0.21	0.29	0.32	0.36	0.27	0.31	0.34	0.35	0.34	0.20	0.31	<b>0.27</b>
ALPHA	LF	0.20	0.28	0.31	0.26	0.27	0.24	0.33	0.28	0.30	0.29	0.34	0.27	0.35	0.38	0.31	0.28	0.32	0.26	0.26	<b>0.29</b>
	HF	0.22	0.25	0.25	0.21	0.18	0.22	0.25	0.21	0.28	0.20	0.25	0.24	0.30	0.31	0.31	0.27	0.21	0.33	0.34	<b>0.25</b>
BETA	LF	0.26	0.32	0.28	0.27	0.27	0.20	0.32	0.26	0.33	0.26	0.30	0.25	0.32	0.30	0.31	0.27	0.34	0.25	0.30	<b>0.28</b>
	HF	0.23	0.26	0.25	0.28	0.18	0.22	0.24	0.17	0.29	0.20	0.18	0.21	0.30	0.23	0.31	0.34	0.20	0.34	0.37	<b>0.25</b>

(b)

EEG	HRV	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14	E15	E16	E17	E18	E19	Average
DELTA	LF	0.25	0.27	0.27	0.29	0.29	0.29	0.26	0.31	0.29	0.26	0.28	0.25	0.24	0.25	0.18	0.26	0.35	0.26	0.24	<b>0.27</b>
	HF	0.25	0.24	0.19	0.21	0.21	0.22	0.18	0.25	0.22	0.30	0.14	0.27	0.26	0.22	0.27	0.27	0.24	0.33	0.32	<b>0.24</b>
THETA	LF	0.33	0.36	0.28	0.25	0.25	0.26	0.26	0.26	0.24	0.39	0.32	0.34	0.29	0.39	0.37	0.36	0.33	0.30	0.26	<b>0.31</b>
	HF	0.34	0.29	0.21	0.20	0.18	0.17	0.21	0.22	0.20	0.27	0.27	0.40	0.27	0.31	0.32	0.35	0.32	0.19	0.32	<b>0.26</b>
ALPHA	LF	0.21	0.29	0.30	0.26	0.26	0.23	0.33	0.28	0.31	0.28	0.33	0.22	0.35	0.38	0.30	0.27	0.34	0.25	0.26	<b>0.29</b>
	HF	0.20	0.26	0.26	0.23	0.18	0.22	0.25	0.21	0.30	0.19	0.21	0.20	0.31	0.31	0.31	0.29	0.22	0.33	0.34	<b>0.25</b>
BETA	LF	0.23	0.30	0.31	0.25	0.26	0.22	0.33	0.27	0.31	0.25	0.30	0.22	0.35	0.35	0.26	0.25	0.35	0.25	0.27	<b>0.28</b>
	HF	0.23	0.25	0.27	0.22	0.17	0.22	0.25	0.21	0.30	0.20	0.19	0.20	0.30	0.30	0.28	0.27	0.23	0.33	0.35	<b>0.25</b>

(c)

EEG	HRV	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14	E15	E16	E17	E18	E19	Average
DELTA	LF	0.27	0.19	0.27	0.29	0.29	0.29	0.27	0.31	0.31	0.20	0.21	0.28	0.19	0.27	0.24	0.24	0.33	0.23	0.24	<b>0.26</b>
	HF	0.28	0.29	0.20	0.21	0.21	0.21	0.19	0.22	0.22	0.34	0.19	0.27	0.25	0.28	0.27	0.24	0.29	0.31	0.32	<b>0.25</b>
THETA	LF	0.27	0.33	0.28	0.24	0.25	0.25	0.29	0.24	0.24	0.38	0.21	0.25	0.30	0.37	0.38	0.37	0.37	0.29	0.26	<b>0.29</b>
	HF	0.32	0.23	0.21	0.21	0.18	0.17	0.24	0.20	0.22	0.28	0.18	0.23	0.29	0.29	0.33	0.34	0.30	0.23	0.32	<b>0.25</b>
ALPHA	LF	0.29	0.35	0.26	0.24	0.25	0.25	0.28	0.24	0.23	0.35	0.27	0.19	0.29	0.37	0.37	0.36	0.35	0.26	0.24	<b>0.29</b>
	HF	0.30	0.26	0.20	0.21	0.18	0.18	0.24	0.19	0.24	0.26	0.21	0.23	0.25	0.30	0.32	0.33	0.31	0.18	0.29	<b>0.25</b>
BETA	LF	0.28	0.32	0.28	0.24	0.25	0.25	0.29	0.24	0.25	0.37	0.20	0.25	0.28	0.37	0.36	0.36	0.38	0.27	0.25	<b>0.29</b>
	HF	0.33	0.22	0.21	0.21	0.18	0.17	0.23	0.19	0.22	0.27	0.18	0.22	0.28	0.30	0.34	0.36	0.31	0.22	0.32	<b>0.25</b>

(d)

Table F.3: Dataset 1- Method 2 - Averaged participants' correlation performance: (a) Level 0 (No WT), (b) Level 1 WT, (c) Level 2 WT, (d) Level 3 WT

EEG	HRV	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14	E15	E16	E17	E18	E19	Average
DELTA	LF	0.23	0.22	0.21	0.21	0.23	0.24	0.33	0.19	0.22	0.23	0.22	0.17	0.23	0.21	0.24	0.22	0.24	0.24	0.26	<b>0.23</b>
	HF	0.30	0.30	0.29	0.30	0.30	0.29	0.30	0.30	0.29	0.28	0.29	0.24	0.28	0.29	0.27	0.29	0.29	0.29	0.28	<b>0.29</b>
THETA	LF	0.16	0.18	0.14	0.17	0.19	0.15	0.23	0.25	0.18	0.21	0.19	0.33	0.15	0.25	0.23	0.19	0.14	0.15	0.15	<b>0.19</b>
	HF	0.32	0.36	0.29	0.29	0.30	0.33	0.36	0.31	0.29	0.31	0.30	0.33	0.31	0.32	0.32	0.31	0.30	0.35	0.34	<b>0.32</b>
ALPHA	LF	0.21	0.22	0.21	0.19	0.32	0.23	0.29	0.18	0.21	0.21	0.21	0.19	0.22	0.21	0.21	0.21	0.23	0.22	0.23	<b>0.22</b>
	HF	0.28	0.29	0.28	0.28	0.33	0.28	0.31	0.31	0.28	0.28	0.27	0.25	0.28	0.28	0.28	0.28	0.29	0.28	0.28	<b>0.29</b>
BETA	LF	0.25	0.23	0.22	0.20	0.21	0.21	0.22	0.19	0.22	0.22	0.22	0.17	0.22	0.22	0.23	0.22	0.22	0.21	0.21	<b>0.21</b>
	HF	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.28	0.29	0.29	0.29	0.28	0.29	0.29	0.28	0.29	0.29	0.29	0.29	<b>0.29</b>

(a)

EEG	HRV	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14	E15	E16	E17	E18	E19	Average
DELTA	LF	0.23	0.22	0.21	0.22	0.23	0.25	0.33	0.19	0.22	0.23	0.22	0.17	0.23	0.21	0.24	0.22	0.24	0.25	0.26	<b>0.23</b>
	HF	0.30	0.30	0.29	0.30	0.30	0.29	0.30	0.30	0.29	0.28	0.29	0.24	0.28	0.29	0.28	0.29	0.29	0.29	0.28	<b>0.29</b>
THETA	LF	0.16	0.18	0.14	0.18	0.19	0.15	0.23	0.25	0.18	0.21	0.19	0.33	0.15	0.25	0.23	0.19	0.14	0.15	0.15	<b>0.19</b>
	HF	0.32	0.36	0.29	0.29	0.29	0.33	0.36	0.31	0.29	0.32	0.30	0.33	0.31	0.32	0.31	0.31	0.30	0.35	0.34	<b>0.32</b>
ALPHA	LF	0.21	0.22	0.21	0.19	0.28	0.23	0.28	0.17	0.21	0.21	0.21	0.17	0.22	0.21	0.21	0.21	0.23	0.22	0.23	<b>0.22</b>
	HF	0.28	0.29	0.28	0.28	0.30	0.28	0.31	0.31	0.28	0.28	0.27	0.26	0.28	0.28	0.28	0.28	0.29	0.28	0.28	<b>0.29</b>
BETA	LF	0.25	0.23	0.22	0.20	0.21	0.21	0.22	0.19	0.22	0.22	0.22	0.16	0.22	0.22	0.23	0.22	0.22	0.21	0.21	<b>0.21</b>
	HF	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.28	0.29	0.29	0.29	0.28	0.29	0.29	0.28	0.29	0.29	0.29	0.29	<b>0.29</b>

(b)

EEG	HRV	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14	E15	E16	E17	E18	E19	Average
DELTA	LF	0.23	0.22	0.21	0.23	0.33	0.28	0.34	0.18	0.22	0.24	0.23	0.16	0.23	0.21	0.23	0.22	0.24	0.26	0.27	<b>0.24</b>
	HF	0.30	0.30	0.29	0.30	0.31	0.29	0.30	0.29	0.30	0.28	0.29	0.24	0.28	0.29	0.28	0.29	0.29	0.29	0.28	<b>0.29</b>
THETA	LF	0.17	0.19	0.14	0.19	0.30	0.16	0.24	0.25	0.18	0.21	0.19	0.34	0.15	0.26	0.24	0.19	0.14	0.15	0.16	<b>0.20</b>
	HF	0.34	0.36	0.29	0.30	0.30	0.33	0.36	0.32	0.29	0.32	0.31	0.35	0.31	0.33	0.30	0.31	0.31	0.36	0.35	<b>0.32</b>
ALPHA	LF	0.22	0.22	0.22	0.21	0.18	0.22	0.25	0.18	0.22	0.22	0.22	0.13	0.22	0.22	0.22	0.22	0.22	0.22	0.22	<b>0.21</b>
	HF	0.28	0.29	0.29	0.28	0.29	0.29	0.29	0.29	0.28	0.29	0.28	0.27	0.28	0.29	0.29	0.29	0.29	0.28	0.28	<b>0.29</b>
BETA	LF	0.22	0.22	0.22	0.21	0.20	0.22	0.23	0.19	0.22	0.22	0.22	0.16	0.22	0.22	0.22	0.22	0.22	0.22	0.22	<b>0.21</b>
	HF	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.28	0.28	0.29	0.28	0.28	0.29	0.29	0.29	0.29	0.29	0.28	0.29	<b>0.29</b>

(c)

EEG	HRV	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14	E15	E16	E17	E18	E19	Average
DELTA	LF	0.35	0.39	0.34	0.24	0.27	0.34	0.41	0.26	0.15	0.25	0.30	0.36	0.34	0.27	0.33	0.35	0.24	0.36	0.38	<b>0.31</b>
	HF	0.30	0.38	0.25	0.38	0.31	0.48	0.40	0.34	0.29	0.34	0.39	0.34	0.26	0.35	0.39	0.33	0.23	0.43	0.51	<b>0.35</b>
THETA	LF	0.27	0.30	0.33	0.22	0.33	0.21	0.29	0.25	0.19	0.28	0.28	0.36	0.40	0.44	0.32	0.38	0.42	0.31	0.44	<b>0.32</b>
	HF	0.38	0.44	0.34	0.30	0.29	0.36	0.36	0.32	0.27	0.32	0.38	0.38	0.36	0.36	0.27	0.33	0.39	0.31	0.37	<b>0.34</b>
ALPHA	LF	0.37	0.35	0.37	0.24	0.27	0.23	0.31	0.26	0.19	0.25	0.29	0.34	0.35	0.37	0.32	0.37	0.39	0.34	0.43	<b>0.32</b>
	HF	0.42	0.45	0.33	0.27	0.24	0.37	0.36	0.30	0.25	0.31	0.35	0.36	0.30	0.34	0.28	0.36	0.37	0.30	0.36	<b>0.33</b>
BETA	LF	0.27	0.29	0.32	0.20	0.27	0.22	0.29	0.25	0.17	0.25	0.24	0.37	0.43	0.34	0.28	0.35	0.44	0.34	0.43	<b>0.30</b>
	HF	0.39	0.43	0.35	0.28	0.26	0.36	0.36	0.32	0.26	0.31	0.36	0.39	0.38	0.33	0.27	0.36	0.45	0.32	0.40	<b>0.35</b>

(d)

Table F.4: Dataset 2- Method 2 - Averaged participants' correlation performance: (a) Level 0 (No WT), (b) Level 1 WT, (c) Level 2 WT, (d) Level 3 WT

# Bibliography

- H. D. Abarbanel, R. Brown, J. J. Sidorowich, and L. S. Tsimring. The analysis of observed chaotic data in physical systems. *Reviews of modern physics*, 65(4):1331, 1993.
- D. Abásolo, R. Hornero, P. Espino, D. Alvarez, and J. Poza. Entropy analysis of the eeg background activity in alzheimer’s disease patients. *Physiological measurement*, 27(3): 241, 2006.
- S. Abdulla, M. Diykh, R. L. Laft, K. Saleh, and R. C. Deo. Sleep eeg signal analysis based on correlation graph similarity coupled with an ensemble extreme machine learning algorithm. *Expert Systems With Applications*, 138:112790, 2019.
- H. Abdullah, G. Holland, I. Cosic, and D. Cvetkovic. Correlation of sleep eeg frequency bands and heart rate variability. In *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE*, pages 5014–5017. IEEE, 2009.
- H. Abdullah, N. C. Maddage, I. Cosic, and D. Cvetkovic. Cross-correlation of eeg frequency bands and heart rate variability for sleep apnoea classification. *Medical & biological engineering & computing*, 48(12):1261–1269, 2010.
- R. Acharya, O. Faust, N. Kannathal, T. Chua, and S. Laxminarayan. Non-linear analysis of eeg signals at various sleep stages. *Computer methods and programs in biomedicine*, 80(1):37–45, 2005.
- U. R. Acharya, K. P. Joseph, N. Kannathal, C. M. Lim, and J. S. Suri. Heart rate variability: a review. *Medical and biological engineering and computing*, 44(12):1031–1051, 2006.
- U. R. Acharya, H. Fujita, M. Adam, O. S. Lih, V. K. Sudarshan, T. J. Hong, J. E. Koh, Y. Hagiwara, C. K. Chua, C. K. Poo, et al. Automated characterization and classification of coronary artery disease and myocardial infarction by decomposition of eeg signals: a comparative study. *Information Sciences*, 377:17–29, 2017.

- P. S. Addison. *The illustrated wavelet transform handbook: introductory theory and applications in science, engineering, medicine and finance*. CRC press, 2017.
- V. X. Afonso, W. J. Tompkins, T. Q. Nguyen, and S. Luo. Ecg beat detection using filter banks. *IEEE transactions on biomedical engineering*, 46(2):192–202, 1999.
- M. Akay. Wavelet applications in medicine. *IEEE spectrum*, 34(5):50–56, 1997.
- M. Akin. Comparison of wavelet transform and fft methods in the analysis of eeg signals. *Journal of medical systems*, 26(3):241–247, 2002.
- M. Akin, M. Kiymik, M. Arserim, and I. Turkoglu. Separation of brain signals using fft and neural networks. *Proc. of Biyomut 2000*, pages 161–164, 2000.
- M. Ako, T. Kawara, S. Uchida, S. Miyazaki, K. Nishihara, J. Mukai, K. Hirao, J. Ako, and Y. Okubo. Correlation between electroencephalography and heart rate variability during sleep. *Psychiatry and clinical neurosciences*, 57(1):59–65, 2003.
- H. M. Al-Angari and A. V. Sahakian. Use of sample entropy approach to study heart rate variability in obstructive sleep apnea syndrome. *Biomedical Engineering, IEEE Transactions on*, 54(10):1900–1904, 2007.
- G. Alba, E. Pereda, S. Mañas, L. D. Méndez, M. R. Duque, A. González, and J. J. González. The variability of eeg functional connectivity of young adhd subjects in different resting states. *Clinical Neurophysiology*, 127(2):1321–1330, 2016.
- R. Alcaraz and J. J. Rieta. A review on sample entropy applications for the non-invasive analysis of atrial fibrillation electrocardiograms. *Biomedical Signal Processing and Control*, 5(1):1–14, 2010.
- M. L. Almanza-Sepúlveda, M. Hernández-González, J. C. Hevia-Orozco, C. Amezcua-Gutiérrez, and M. A. Guevara. Verbal and visuospatial working memory during pregnancy: Eeg correlation between the prefrontal and parietal cortices. *Neurobiology of learning and memory*, 148:1–7, 2018.
- K. Amaratunga, J. R. Williams, S. Qian, and J. Weiss. Wavelet–galerkin solutions for one-dimensional partial differential equations. *International Journal for Numerical Methods in Engineering*, 37(16):2703–2716, 1994.
- I. Amato. Chaos breaks out at nih, but order may come of it. *Science*, 256(5065):1763–1765, 1992.

- R. G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, and C. E. Elger. Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state. *Physical Review E*, 64(6):061907, 2001.
- G. Antonioli and P. Tonella. Eeg data compression techniques. *IEEE Transactions on Biomedical engineering*, 44(2):105–114, 1997.
- F. Araújo, I. Antelmi, A. C. Pereira, D. Maria do Rosário, C. J. Grupi, J. E. Krieger, and A. J. Mansur. Lower heart rate variability is associated with higher serum high-sensitivity c-reactive protein concentration in healthy individuals aged 46 years or more. *International journal of cardiology*, 107(3):333–337, 2006.
- M. ATKINSON and R. T. BRADLEY. Electrophysiological evidence of intuition: Part 2. a system-wide process? *THE JOURNAL OF ALTERNATIVE AND COMPLEMENTARY MEDICINE*, 10(2):325–336, 2004.
- S. Ayers. The application of chaos theory to psychology. *Theory & Psychology*, 7(3): 373–398, 1997.
- J. K. Beard. *The FFT in the 21st century: Eigenspace processing*. Springer Science & Business Media, 2013.
- T. P. Beauchaine and J. F. Thayer. Heart rate variability as a transdiagnostic biomarker of psychopathology. *International Journal of Psychophysiology*, 98(2):338–350, 2015.
- C. F. Beckmann, M. DeLuca, J. T. Devlin, and S. M. Smith. Investigations into resting-state connectivity using independent component analysis. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 360(1457):1001–1013, 2005.
- J. Benesty, J. Chen, Y. Huang, and I. Cohen. Pearson correlation coefficient. *Noise reduction in speech processing*, pages 1–4, 2009.
- J. Berg, G. Neely, U. Wiklund, and U. Landström. Heart rate variability during sedentary work and sleep in normal and sleep-deprived states. *Clinical physiology and functional imaging*, 25(1):51–57, 2005.
- G. G. Berntson, J. T. Cacioppo, and K. S. Quigley. Autonomic determinism: the modes of autonomic control, the doctrine of autonomic space, and the laws of autonomic constraint. *Psychological review*, 98(4):459, 1991.

- G. G. Berntson, J. T. Cacioppo, and K. S. Quigley. Autonomic cardiac control. i. estimation and validation from pharmacological blockades. *Psychophysiology*, 31(6):572–585, 1994.
- L. Biel, O. Pettersson, L. Philipson, and P. Wide. Ecg analysis: a new approach in human identification. In *IMTC/99. Proceedings of the 16th IEEE Instrumentation and Measurement Technology Conference (Cat. No. 99CH36309)*, volume 1, pages 557–561. IEEE, 1999.
- L. Biel, O. Pettersson, L. Philipson, and P. Wide. Ecg analysis: a new approach in human identification. *IEEE Transactions on Instrumentation and Measurement*, 50(3):808–812, 2001.
- J. T. Bigger Jr, R. E. Kleiger, J. L. Fleiss, L. M. Rolnitzky, R. C. Steinman, and J. P. Miller. Components of heart rate variability measured during healing of acute myocardial infarction. *The American journal of cardiology*, 61(4):208–215, 1988.
- J. T. Bigger Jr, J. L. Fleiss, R. C. Steinman, L. M. Rolnitzky, R. E. Kleiger, and J. N. Rottman. Frequency domain measures of heart period variability and mortality after myocardial infarction. *Circulation*, 85(1):164–171, 1992.
- M. Bilucaglia, L. Pederzoli, W. Giroladini, E. Prati, and P. Tressoldi. Eeg correlation at a distance: A re-analysis of two studies using a machine learning approach. *F1000Research*, 8, 2019.
- A. Bissacco, A. Chiuso, Y. Ma, and S. Soatto. Recognition of human gaits. In *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, volume 2, pages II–II. IEEE, 2001.
- P. Bob, M. Susta, K. Glaslova, and N. N. Boutros. Dissociative symptoms and interregional eeg cross-correlations in paranoid schizophrenia. *Psychiatry research*, 177(1):37–40, 2010.
- M. Botcharova, S. F. Farmer, and L. Berthouze. Markers of criticality in phase synchronization. *Frontiers in systems neuroscience*, 8:176, 2014.
- G. E. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung. *Time series analysis: forecasting and control*. John Wiley & Sons, 2015.
- M. Breakspear and J. Terry. Nonlinear interdependence in neural systems: motivation, theory, and relevance. *International Journal of Neuroscience*, 112(10):1263–1284, 2002.

- C. Bregler. Learning and recognizing human dynamics in video sequences. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 568–574. IEEE, 1997.
- M. Brennan, M. Palaniswami, and P. Kamen. Do existing measures of poincare plot geometry reflect nonlinear features of heart rate variability? *Biomedical Engineering, IEEE Transactions on*, 48(11):1342–1347, 2001.
- E. O. Brigham and E. O. Brigham. *The fast Fourier transform and its applications*, volume 448. prentice Hall Englewood Cliffs, NJ, 1988.
- P. J. Brockwell, R. A. Davis, and M. V. Calder. *Introduction to time series and forecasting*, volume 2. Springer, 2002.
- D. J. Brotman, L. D. Bash, R. Qayyum, D. Crews, E. A. Whitsel, B. C. Astor, and J. Coresh. Heart rate variability predicts esrd and ckd-related hospitalization. *Journal of the American Society of Nephrology*, 21(9):1560–1570, 2010.
- E. Buccelletti, E. Gilardi, E. Scaini, L. Galiuto, R. Persiani, A. Biondi, F. Basile, and N. G. Silveri. Heart rate variability and myocardial infarction: systematic literature review and metanalysis. *Eur Rev Med Pharmacol Sci*, 13(4):299–307, 2009.
- T. G. Buchman, J. P. Cobb, A. S. Lapedes, and T. B. Kepler. Complex systems analysis: a tool for shock research. *SHOCK-AUGUSTA-*, 16(4):248–251, 2001.
- J. R. Buck, M. M. Daniel, and A. Singer. *Computer explorations in signals and systems using MATLAB*. Prentice Hall Upper Saddle River, NJ, 2002.
- A. Bultheel et al. Learning to swim in a sea of wavelets. *Bulletin of the Belgian Mathematical society-simon stevin*, 2(1):1–45, 1995.
- G. E. Burch and N. P. DePasquale. *A history of electrocardiography*. Number 1. Norman Publishing, 1990.
- C. Burrus and T. W. Parks. *DFT/FFT and Convolution Algorithms: theory and Implementation*. John Wiley & Sons, Inc., 1991.
- H. Busby and D. Trujillo. Smoothing noisy data using dynamic programming and generalized cross-validation. *Journal of biomechanical engineering*, 110:37, 1988.

- U. H. Buzzi, N. Stergiou, M. J. Kurz, P. A. Hageman, and J. Heidel. Nonlinear dynamics indicates aging affects variability during gait. *Clinical biomechanics*, 18(5):435–443, 2003.
- A. J. Camm, C. M. Pratt, P. J. Schwartz, H. R. Al-Khalidi, M. J. Spyt, M. J. Holroyde, R. Karam, E. H. Sonnenblick, and J. M. Brum. Mortality in patients after a recent myocardial infarction: a randomized, placebo-controlled trial of azimilide using heart rate variability for risk stratification. *Circulation*, 109(8):990–996, 2004.
- H. Cao, D. E. Lake, M. P. Griffin, and J. R. Moorman. Increased nonstationarity of neonatal heart rate before the clinical diagnosis of sepsis. *Annals of biomedical engineering*, 32(2):233–244, 2004.
- D. W. Carmichael, S. Vulliemoz, R. Rodionov, J. S. Thornton, A. W. McEvoy, and L. Lemieux. Simultaneous intracranial eeg–fmri in humans: protocol considerations and data quality. *Neuroimage*, 63(1):301–309, 2012.
- S. Carnot. Reflections on the motive power of fire, and on machines fitted to develop that power. *Paris: Bachelier*, 1824.
- M. Casdagli. Nonlinear prediction of chaotic time series. *Physica D: Nonlinear Phenomena*, 35(3):335–356, 1989.
- A. Catarino, O. Churches, S. Baron-Cohen, A. Andrade, and H. Ring. Atypical eeg complexity in autism spectrum conditions: a multiscale entropy analysis. *Clinical neurophysiology*, 122(12):2375–2383, 2011.
- J. T. Cavanaugh, N. Kochi, and N. Stergiou. Nonlinear analysis of ambulatory activity patterns in community-dwelling older adults. *Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences*, 65(2):197–203, 2009.
- J. S. Chall and A. F. Mirsky. *Education and the Brain: The Seventy-seventh Yearbook of the National Society for the Study of Education*. National Society for the Study of Education, 1978.
- A. B. Chan and N. Vasconcelos. Classifying video with kernel dynamic textures. In *2007 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–6. IEEE, 2007.
- S. Chandra, G. Sharma, M. Sharma, D. Jha, and A. P. Mittal. Workload regulation by sudarshan kriya: an eeg and ecg perspective. *Brain informatics*, 4(1):13, 2017.

- C. Chang, C. D. Metzger, G. H. Glover, J. H. Duyn, H.-J. Heinze, and M. Walter. Association between heart rate variability and fluctuations in resting-state functional connectivity. *Neuroimage*, 68:93–104, 2013.
- C.-S. Chang, C.-W. Ko, H.-C. Lien, and M.-C. Chou. Effect of electroacupuncture on st. 36 (zusanli) and li. 10 (shousanli) acupuncture points on heart rate variability. *The American journal of Chinese medicine*, 38(02):231–239, 2010.
- T. Chang, S. J. Schiff, T. Sauer, J.-P. Gossard, and R. E. Burke. Stochastic versus deterministic variability in simple neuronal circuits: I. monosynaptic spinal cord reflexes. *Biophysical journal*, 67(2):671–683, 1994.
- C. Chatfield. *The analysis of time series: an introduction*. Chapman and Hall/CRC, 2016.
- W. Chen, Z. Wang, H. Xie, and W. Yu. Characterization of surface emg signal based on fuzzy entropy. *IEEE Transactions on neural systems and rehabilitation engineering*, 15(2):266–272, 2007.
- Y. Chen, W. Wang, X. Zhao, M. Sha, Y. Liu, X. Zhang, J. Ma, H. Ni, and D. Ming. Age-related decline in the variation of dynamic functional connectivity: a resting state analysis. *Frontiers in aging neuroscience*, 9:203, 2017.
- Y.-J. Choi and S.-H. Cho. Effect of electroacupuncture stimulation at different frequencies on brain waves. *Acupuncture & Electro-Therapeutics Research*, 44(1):11–22, 2019.
- E. Chowdhury and L. Ludeman. Discrimination of cardiac arrhythmias using a fuzzy rule-based method. In *Computers in Cardiology 1994*, pages 549–552. IEEE, 1994.
- E. C.-P. Chua, W.-Q. Tan, S.-C. Yeo, P. Lau, I. Lee, I. H. Mien, K. Puvanendran, and J. J. Gooley. Heart rate variability can be used to estimate sleepiness-related decrements in psychomotor vigilance during total sleep deprivation. *Sleep*, 35(3):325–334, 2012.
- B. Chun-Hua and N. Xin-Bao. Determining the minimum embedding dimension of non-linear time series based on prediction method. *Chinese Physics*, 13(5):633, 2004.
- F. Cignetti, F. Schena, and A. Rouard. Effects of fatigue on inter-cycle variability in cross-country skiing. *Journal of biomechanics*, 42(10):1452–1459, 2009.
- R. Clausius. *The mechanical theory of heat: with its applications to the steam-engine and to the physical properties of bodies*. J. van Voorst, 1867.

- D. A. Coast, R. M. Stern, G. G. Cano, and S. A. Brillner. An approach to cardiac arrhythmia analysis using hidden markov models. *IEEE Transactions on biomedical Engineering*, 37(9):826–836, 1990.
- J. Collins and C. De Luca. The effects of visual input on open-loop and closed-loop postural control mechanisms. *Experimental Brain Research*, 103(1):151–163, 1995.
- P. Comon. Independent component analysis, a new concept? *Signal processing*, 36(3):287–314, 1994.
- J. W. Cooley and J. W. Tukey. An algorithm for the machine calculation of complex fourier series. *Mathematics of computation*, 19(90):297–301, 1965.
- L. T. Cooper, K. L. Baughman, A. M. Feldman, A. Frustaci, M. Jessup, U. Kuhl, G. N. Levine, J. Narula, R. C. Starling, J. Towbin, et al. The role of endomyocardial biopsy in the management of cardiovascular disease: a scientific statement from the american heart association, the american college of cardiology, and the european society of cardiology endorsed by the heart failure society of america and the heart failure association of the european society of cardiology. *Journal of the American College of Cardiology*, 50(19):1914–1931, 2007.
- M. D. Costa, T. Henriques, M. N. Munshi, A. R. Segal, and A. L. Goldberger. Dynamical glucometry: use of multiscale entropy analysis in diabetes. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 24(3):033139, 2014.
- J. D. Cryer and N. Kellert. *Time series analysis*. Springer, 1991.
- K. Cuevas and M. A. Bell. Eeg and ecg from 5 to 10 months of age: Developmental changes in baseline activation and cognitive processing during a working memory task. *International Journal of Psychophysiology*, 80(2):119–128, 2011.
- D. Cvetkovic, E. D. Übeyli, and I. Cosic. Wavelet transform feature extraction from human ppg, ecg, and eeg signal responses to elf pemf exposures: A pilot study. *Digital signal processing*, 18(5):861–874, 2008.
- F. H. L. da Silva, J.-P. Pijn, and W. J. Wadman. Dynamics of local neuronal networks: control parameters and state bifurcations in epileptogenesis. In *Progress in brain research*, volume 102, pages 359–370. Elsevier, 1994.
- I. Daskalov and I. Christov. Improvement of resolution in measurement of electrocardiogram rr intervals by interpolation. *Medical engineering & physics*, 19(4):375–379, 1997.

- I. Daubechies. The wavelet transform, time-frequency localization and signal analysis. *IEEE transactions on information theory*, 36(5):961–1005, 1990.
- E. de Azevedo Botter, C. L. Nascimento, and T. Yoneyama. A neural network with asymmetric basis functions for feature extraction of eeg p waves. *IEEE Transactions on Neural Networks*, 12(5):1252–1255, 2001.
- D. Dheeru and E. Karra Taniskidou. UCI machine learning repository, 2017. URL <http://archive.ics.uci.edu/ml>.
- J. B. Dingwell and J. P. Cusumano. Nonlinear time series analysis of normal and pathological human walking. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 10(4):848–863, 2000.
- A. D. Dolatabadi, S. E. Z. Khadem, and B. M. Asl. Automated diagnosis of coronary artery disease (cad) patients using optimized svm. *Computer methods and programs in biomedicine*, 138:117–126, 2017.
- G. Doretto, A. Chiuso, Y. N. Wu, and S. Soatto. Dynamic textures. *International Journal of Computer Vision*, 51(2):91–109, 2003.
- H. Doufesh, F. Ibrahim, N. A. Ismail, and W. A. Wan Ahmad. Application of self organizing map for correlation hunting between alpha band power of eeg signals and other physiological parameters during muslim prayer (salat). *Biomedical Engineering: Applications, Basis and Communications*, 30(04):1850029, 2018.
- Dr.Evans. Acupuncture with electric stimulation, 1998-2019. URL <http://www.evansacupuncture.com/acupuncturewes.html>.
- Dr.Xie. Acupuncture and chinese herbel medicines for pain and health, 1998-2019. URL <https://acupunctureforhealthcare.com/services/manual-acupuncture/>.
- P. Duhamel and M. Vetterli. Fast fourier transforms: a tutorial review and a state of the art. *Signal processing*, 19(4):259–299, 1990.
- D. Durante, D. B. Dunson, et al. Bayesian inference and testing of group differences in brain networks. *Bayesian Analysis*, 13(1):29–58, 2018.
- Z. Duszak and W. Koczkodaj. Using principal component transformation in machine learning.

- G. Edlinger and C. Guger. Correlation changes of eeg and ecg after fast cable car ascents. In *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the*, pages 5540–5543. IEEE, 2006.
- C. L. Ehlers, J. Havstad, D. Prichard, and J. Theiler. Low doses of ethanol reduce evidence for nonlinear structure in brain activity. *Journal of Neuroscience*, 18(18):7474–7486, 1998.
- W. Einthoven. The different forms of the human electrocardiogram and their signification. *The Lancet*, 179(4622):853–861, 1912.
- M. Elad, J.-L. Starck, P. Querre, and D. L. Donoho. Simultaneous cartoon and texture image inpainting using morphological component analysis (mca). *Applied and Computational Harmonic Analysis*, 19(3):340–358, 2005.
- T. F. o. t. E. S. o. C. t. N. A. S. o. P. Electrophysiology. Heart rate variability: standards of measurement, physiological interpretation, and clinical use. *Circulation*, 93(5):1043–1065, 1996.
- J. B. Elsner and A. A. Tsonis. *Singular spectrum analysis: a new tool in time series analysis*. Springer Science & Business Media, 2013.
- O. Faust, U. R. Acharya, H. Adeli, and A. Adeli. Wavelet-based eeg processing for computer-aided seizure detection and epilepsy diagnosis. *Seizure*, 26:56–64, 2015.
- R. D. Fields. White matter matters. *Scientific American*, 298(3):54–61, 2008.
- M. D. Fox, A. Z. Snyder, J. L. Vincent, M. Corbetta, D. C. Van Essen, and M. E. Raichle. The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences*, 102(27):9673–9678, 2005.
- M. D. Fox, M. Corbetta, A. Z. Snyder, J. L. Vincent, and M. E. Raichle. Spontaneous neuronal activity distinguishes human dorsal and ventral attention systems. *Proceedings of the National Academy of Sciences*, 103(26):10046–10051, 2006.
- R. J. Frank, N. Davey, and S. P. Hunt. Time series prediction and neural networks. *Journal of Intelligent and Robotic Systems*, 31(1-3):91–103, 2001.
- B. L. Fredrickson and M. F. Losada. Positive affect and the complex dynamics of human flourishing. *American Psychologist*, 60(7):678–686, 2005.

- N. P. Friedman, A. Miyake, R. P. Corley, S. E. Young, J. C. DeFries, and J. K. Hewitt. Not all executive functions are related to intelligence. *Psychological science*, 17(2):172–179, 2006.
- U. Frith. *Autism: Explaining the enigma*. Blackwell Publishing, 2003.
- X. Gao, Y. Wang, X. Li, and D. Tao. On combining morphological component analysis and concentric morphology model for mammographic mass detection. *IEEE Transactions on Information Technology in Biomedicine*, 14(2):266–273, 2010.
- A. Garfinkel, M. Spano, and W. Ditto. Weiss. *J. N., Controlling cardiac chaos, Science*, 257:12, 1992.
- M. A. Gartstein, G. R. Hancock, N. V. Potapova, S. D. Calkins, and M. A. Bell. Modeling development of frontal electroencephalogram (eeg) asymmetry: Sex differences and links with temperament. *Developmental science*, 23(1):e12891, 2020.
- H. Gastaut. Electrocorticographic study of the reactivity of rolandic rhythm. *Revue neurologique*, 87(2):176–182, 1952.
- B. Ghanem and N. Ahuja. Phase based modelling of dynamic textures. In *2007 IEEE 11th International Conference on Computer Vision*, pages 1–8. IEEE, 2007.
- G. Giakas. Power spectrum analysis and filtering. *Innovative analyses of human movement*, 1:223–223, 2004.
- G. Giakas and V. Baltzopoulos. Time and frequency domain analysis of ground reaction forces during walking: an investigation of variability and symmetry. *Gait & Posture*, 5(3):189–197, 1997.
- G. Giakas, V. Baltzopoulos, P. H. Dangerfield, J. C. Dorgan, and S. Dalmira. Comparison of gait patterns between healthy and scoliotic patients using time and frequency domain analysis of ground reaction forces. *Spine*, 21(19):2235–2242, 1996.
- P. J. Gianaros and K. S. Quigley. Autonomic origins of a nonsignal stimulus-elicited bradycardia and its habituation in humans. *Psychophysiology*, 38(3):540–547, 2001.
- D. Giri, U. R. Acharya, R. J. Martis, S. V. Sree, T.-C. Lim, T. A. VI, and J. S. Suri. Automated diagnosis of coronary artery disease affected patients using lda, pca, ica and discrete wavelet transform. *Knowledge-Based Systems*, 37:274–282, 2013.

- A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley. Physiobank, physiotookit, and physionet: components of a new research resource for complex physiologic signals. *Circulation*, 101(23):e215–e220, 2000.
- B. Goldstein, D. Towell, S. Lai, K. Sonnenthal, and B. Kimberly. Uncoupling of the autonomic and cardiovascular systems in acute brain injury. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, 275(4):R1287–R1292, 1998.
- N. Golyandina, V. Nekrutkin, and A. A. Zhigljavsky. *Analysis of time series structure: SSA and related techniques*. Chapman and Hall/CRC, 2001.
- B. Gourévitch and J. J. Eggermont. A simple indicator of nonstationarity of firing rate in spike trains. *Journal of neuroscience methods*, 163(1):181–187, 2007.
- R. Govindan, K. Narayanan, and M. Gopinathan. On the evidence of deterministic chaos in ecg: Surrogate and predictability analysis. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 8(2):495–502, 1998.
- P. Graziano and K. Derefinko. Cardiac vagal control and children’s adaptive functioning: A meta-analysis. *Biological psychology*, 94(1):22–37, 2013.
- J. F. Grcar. Matrix stretching for linear equations. *arXiv preprint arXiv:1203.2377*, 2012.
- F. Grouiller, L. Vercueil, A. Krainik, C. Segebarth, P. Kahane, and O. David. A comparative study of different artefact removal algorithms for eeg signals acquired during functional mri. *Neuroimage*, 38(1):124–137, 2007.
- J. Hevia-Orozco, A. Sanz-Martin, M. Guevara, and M. Hernández-Gonzalez. Eeg correlation during social decision-making in institutionalized adolescents. *Abnorm Behav Psychol*, 3(131):2472–0496, 2017.
- R. C. Hilborn et al. *Chaos and nonlinear dynamics: an introduction for scientists and engineers*. Oxford University Press on Demand, 2000.
- S. L. Horowitz. A syntactic algorithm for peak detection in waveforms with applications to cardiography. *Communications of the ACM*, 18(5):281–285, 1975.
- Y. H. Hu, W. J. Tompkins, J. L. Urrusti, and V. X. Afonso. Applications of artificial neural networks for ecg signal detection and classification. *Journal of electrocardiology*, 26: 66–73, 1993.

- Y. H. Hu, S. Palreddy, and W. J. Tompkins. A patient-adaptable ecg beat classifier using a mixture of experts approach. *IEEE transactions on biomedical engineering*, 44(9): 891–900, 1997.
- N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu. The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 454(1971):903–995, 1998.
- H. V. Huikuri and P. K. Stein. Heart rate variability in risk stratification of cardiac patients. *Progress in cardiovascular diseases*, 56(2):153–159, 2013.
- C. M. Hurvich and C.-L. Tsai. Regression and time series model selection in small samples. *Biometrika*, 76(2):297–307, 1989.
- A. Hyvärinen and E. Oja. Independent component analysis: algorithms and applications. *Neural networks*, 13(4):411–430, 2000.
- T. Ince, S. Kiranyaz, and M. Gabbouj. Automated patient-specific classification of premature ventricular contractions. In *2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 5474–5477. IEEE, 2008.
- Y. İşler and M. Kuntalp. Combining classical hrv indices with wavelet entropy measures improves to performance in diagnosing congestive heart failure. *Computers in biology and medicine*, 37(10):1502–1510, 2007.
- D. Ivanov, H. Posch, and C. Stumpf. Statistical measures derived from the correlation integrals of physiological time series. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 6(2):243–253, 1996.
- S. James and J. Walker. A primer on wavelets and scientific applications, 1999.
- S. Janjaraşjitt, M. Scher, and K. Loparo. Nonlinear dynamical analysis of the neonatal eeg time series: the relationship between sleep state and complexity. *Clinical Neurophysiology*, 119(8):1812–1823, 2008.
- I. Janszky, M. Ericson, M. Lekander, M. Blom, K. Buhlin, A. Georgiades, and S. Ahnve. Inflammatory markers and heart rate variability in women with coronary heart disease. *Journal of internal medicine*, 256(5):421–428, 2004.

- C.-W. Jao, Y.-L. Chen, T. H. Huang, C.-T. Tseng, C.-S. Yang, C. Y. Lin, S. J. Tsai, P.-S. Wang, and Y.-T. Wu. Status change revealed by electrocardiography (ecg) and electroencephalography (eeg) during cycling exercise. In *2017 International Automatic Control Conference (CACCS)*, pages 1–5. IEEE, 2017.
- M. Jarczok, J. Koenig, D. Mauss, J. Fischer, and J. Thayer. Lower heart rate variability predicts increased level of c-reactive protein 4 years later in healthy, nonsmoking adults. *Journal of internal medicine*, 276(6):667–671, 2014.
- E. Jayachandran et al. Analysis of myocardial infarction using discrete wavelet transform. *Journal of medical systems*, 34(6):985–992, 2010.
- J. R. Jennings, L. K. Sheu, D. C.-H. Kuan, S. B. Manuck, and P. J. Gianaros. Resting state connectivity of the medial prefrontal cortex covaries with individual differences in high-frequency heart rate variability. *Psychophysiology*, 53(4):444–454, 2016.
- D.-H. Jeong, Y.-D. Kim, I.-U. Song, Y.-A. Chung, and J. Jeong. Wavelet energy and wavelet coherence as eeg biomarkers for the diagnosis of parkinson’s disease-related dementia and alzheimer’s disease. *Entropy*, 18(1):8, 2015.
- M. I. Johnson and J. M. Bjordal. Transcutaneous electrical nerve stimulation for the management of painful conditions: focus on neuropathic pain. 2011.
- V. Jurcak, D. Tsuzuki, and I. Dan. 10/20, 10/10, and 10/5 systems revisited: their validity as relative head-surface-based positioning systems. *Neuroimage*, 34(4):1600–1611, 2007.
- F. Jurysta, P. Van De Borne, P.-F. Migeotte, M. Dumont, J.-P. Lanquart, J.-P. Degaute, and P. Linkowski. A study of the dynamic interactions between sleep eeg and heart rate variability in healthy young men. *Clinical neurophysiology*, 114(11):2146–2155, 2003.
- C. Jutten and J. Herault. Blind separation of sources, part i: An adaptive algorithm based on neuromimetic architecture. *Signal processing*, 24(1):1–10, 1991.
- N. Kannathal, M. L. Choo, U. R. Acharya, and P. Sadasivan. Entropies for detection of epilepsy in eeg. *Computer methods and programs in biomedicine*, 80(3):187–194, 2005.
- H. Kantz and T. Schreiber. *Nonlinear time series analysis*, volume 7. Cambridge university press, 2004.
- H. Kantz, J. Kurths, and G. Mayer-Kress. *Nonlinear analysis of physiological data*. Springer Science & Business Media, 2012.

- D. Kaplan and L. Glass. Understanding nonlinear dynamics. 1995.
- D. Kaplan and L. Glass. *Understanding nonlinear dynamics*. Springer Science & Business Media, 2012.
- C. K. Karmakar, A. H. Khandoker, A. Voss, and M. Palaniswami. Sensitivity of temporal heart rate variability in poincaré plot to changes in parasympathetic nervous system activity. *Biomedical engineering online*, 10(1):17, 2011.
- O. Kärner. On nonstationarity and antipersistency in global temperature series. *Journal of Geophysical Research: Atmospheres*, 107(D20), 2002.
- M. B. Kennel, R. Brown, and H. D. Abarbanel. Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Physical review A*, 45(6): 3403, 1992.
- J. Kevric and A. Subasi. Comparison of signal decomposition methods in classification of eeg signals for motor-imagery bci system. *Biomedical Signal Processing and Control*, 31:398–406, 2017.
- L. Khadra, A. Al-Fahoum, and H. Al-Nashash. Detection of life-threatening cardiac arrhythmias using the wavelet transformation. *Medical and Biological Engineering and Computing*, 35(6):626–632, 1997.
- R. M. Kil, S. H. Park, and S. Kim. Optimum window size for time series prediction. In *Engineering in Medicine and Biology Society, 1997. Proceedings of the 19th Annual International Conference of the IEEE*, volume 4, pages 1421–1424. IEEE, 1997.
- D.-K. Kim, K.-M. Lee, J. Kim, M.-C. Whang, and S. W. Kang. Dynamic correlations between heart and brain rhythm during autogenic meditation. *Frontiers in human neuroscience*, 7, 2013.
- K. Kircanski, L. M. Williams, and I. H. Gotlib. Heart rate variability as a biomarker of anxious depression response to antidepressant medication. *Depression and anxiety*, 36(1):63–71, 2019.
- M. G. Kitzbichler, M. L. Smith, S. R. Christensen, and E. Bullmore. Broadband criticality of human brain network synchronization. *PLoS computational biology*, 5(3):e1000314, 2009.

- R. E. Kleiger, J. P. Miller, J. T. Bigger Jr, and A. J. Moss. Decreased heart rate variability and its association with increased mortality after acute myocardial infarction. *The American journal of cardiology*, 59(4):256–262, 1987.
- S. B. Klein and B. M. Thorne. *Biological psychology*. Macmillan, 2006.
- G. H. Klem, H. O. Lüders, H. Jasper, C. Elger, et al. The ten-twenty electrode system of the international federation. *Electroencephalogr Clin Neurophysiol*, 52(3):3–6, 1999.
- D. E. Knuth. Sorting and searching, 2nd edn. the art of computer programming, vol. 3, 1998.
- S. Koelstra, C. Muhl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras. Deap: A database for emotion analysis; using physiological signals. *IEEE Transactions on Affective Computing*, 3(1):18–31, 2012.
- J. B. Kruskal. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 29(1):1–27, 1964.
- P. Kumari and A. Vaish. Brainwave’s energy feature extraction using wavelet transform. In *Electrical, Electronics and Computer Science (SCEECS), 2014 IEEE Students’ Conference on*, pages 1–5. IEEE, 2014.
- P. Kumari, S. Kumar, and A. Vaish. Feature extraction using empirical mode decomposition for biometric system. In *Signal Propagation and Computer Technology (ICSPCT), 2014 International Conference on*, pages 283–287. IEEE, 2014.
- R. Kunhimangalam, P. K. Joseph, and O. Sujith. Nonlinear analysis of eeg signals: Surrogate data analysis. *IRBM*, 29(4):239–244, 2008.
- H. Ku’nkkel and G. Dolce. Clean-computerized eeg analysis. *Gustav Fisher, Stuttgart*, pages 365–383, 1975.
- A. Kushki, J. Fairley, S. Merja, G. King, and T. Chau. Comparison of blood volume pulse and skin conductance responses to mental and affective stimuli at different anatomical sites. *Physiological measurement*, 32(10):1529, 2011.
- Y. Kutlu and D. Kuntalp. Feature extraction for eeg heartbeats using higher order statistics of wpd coefficients. *Computer methods and programs in biomedicine*, 105(3):257–267, 2012.

- L. Ladislao and S. Fioretti. Nonlinear analysis of posturographic data. *Medical & biological engineering & computing*, 45(7):679–688, 2007.
- D. Lake. Improved entropy rate estimation in physiological data. In *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, pages 1463–1466. IEEE, 2011.
- D. E. Lake, J. S. Richman, M. P. Griffin, and J. R. Moorman. Sample entropy analysis of neonatal heart rate variability. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, 283(3):R789–R797, 2002.
- R. Lane, E. Reiman, G. Ahern, and J. Thayer.  $\alpha_1$  activity in medial prefrontal cortex correlates with vagal component of heart rate variability during emotion. *Brain and Cognition*, 47(1-2):97–100, 2001.
- D. T. Lee and A. Yamamoto. Wavelet analysis: theory and applications. *Hewlett Packard journal*, 45:44–44, 1994.
- S.-H. Lee and J. S. Lim. Parkinson’s disease classification using gait characteristics and wavelet-based feature extraction. *Expert Systems with Applications*, 39(8):7338–7344, 2012.
- W. G. Lennox. *Epilepsy and related disorders*, volume 2. Little, Brown, 1960.
- J. P. Lewis. Fast normalized cross-correlation. In *Vision interface*, volume 10, pages 120–123, 1995.
- N. G. Lewis, J. B. McGovern, J. C. Miller, D. R. Eddy, and E. M. Forster. Eeg indices of g-induced loss of consciousness (g-loc). Technical report, SCHOOL OF AEROSPACE MEDICINE BROOKS AFB TX, 1988.
- Y. Li, T. Zhang, L. Deng, B. Wang, and M. Nakamura. Eeg physiological signals correlation under condition of+ gz accelerations. *Journal of Multimedia*, 8(1):64–71, 2013.
- C.-W. Lin, J.-S. Wang, and P.-C. Chung. Mining physiological conditions from heart rate variability analysis. *IEEE Computational Intelligence Magazine*, 5(1):50–58, 2010.
- Y.-D. Lin and Y. H. Hu. Power-line interference detection and suppression in ecg signal processing. *IEEE Transactions on Biomedical Engineering*, 55(1):354–357, 2007.

- L.-M. Liou, D. Ruge, M.-C. Kuo, J.-C. Tsai, C.-W. Lin, M.-N. Wu, C.-Y. Hsu, and C.-L. Lai. Functional connectivity between parietal cortex and the cardiac autonomic system in uremics. *The Kaohsiung journal of medical sciences*, 30(3):125–132, 2014.
- C.-B. Liu, R.-S. Lin, N. Ahuja, and M.-H. Yang. Dynamic textures synthesis as nonlinear manifold learning and traversing. In *BMVC*, volume 2, pages 859–868, 2006.
- S. J. Luck. *An introduction to the event-related potential technique*. MIT press, 2014.
- A. P. Mackey, A. T. M. Singley, and S. A. Bunge. Intensive reasoning training alters patterns of brain connectivity at rest. *Journal of Neuroscience*, 33(11):4796–4803, 2013.
- S. Makeig, A. J. Bell, T.-P. Jung, and T. J. Sejnowski. Independent component analysis of electroencephalographic data. In *Advances in neural information processing systems*, pages 145–151, 1996.
- V. T. Mäkinen, P. J. May, and H. Tiitinen. The use of stationarity and nonstationarity in the detection and analysis of neural oscillations. *NeuroImage*, 28(2):389–400, 2005.
- M. Malik and A. J. Camm. Heart rate variability. *Clinical cardiology*, 13(8):570–576, 1990.
- S. Malik and V. Verma. Comparative analysis of dct, haar and daubechies wavelet for image compression. *International Journal of Applied Engineering Research*, 7(11):2012, 2012.
- G. Manis. Fast computation of approximate entropy. *Computer methods and programs in biomedicine*, 91(1):48–54, 2008.
- G. J. Martin, N. M. Magid, G. Myers, P. S. Barnett, J. W. Schaad, J. S. Weiss, M. Lesch, and D. H. Singer. Heart rate variability and sudden death secondary to coronary artery disease during ambulatory electrocardiographic monitoring. *The American journal of cardiology*, 60(1):86–89, 1987.
- J. Martinerie, C. Adam, M. Le Van Quyen, M. Baulac, S. Clemenceau, B. Renault, and F. J. Varela. Epileptic seizures can be anticipated by non-linear analysis. *Nature medicine*, 4(10):1173, 1998.
- K. Martínez, A. B. Solana, M. Burgaleta, J. A. Hernández-Tamames, J. Álvarez-Linera, F. J. Román, E. Alfayate, J. Privado, S. Escorial, M. A. Quiroga, et al. Changes in resting-state functionally connected parietofrontal networks after videogame practice. *Human brain mapping*, 34(12):3143–3157, 2013.

- A. Matonia, J. Jezewski, T. Kupka, K. Horoba, J. Wrobel, and A. Gacek. The influence of coincidence of fetal and maternal qrs complexes on fetal heart rate reliability. *Medical and Biological Engineering and Computing*, 44(5):393–403, 2006.
- MayoClinic. Seizures, 1998-2018. URL <https://www.mayoclinic.org/diseases-conditions/seizure/symptoms-causes/syc-20365711>.
- D. MAYoR. The technical acupuncturist—where did that week go, and was it healthy?!
- D. Mayor and T. Steffert. “i think you’ll find it’s the actual point that’s causing me to blink—every time it fires off in colon 4”“i felt the blinking started when the electrical stimulation started”(study participants).
- D. Mayor and T. Steffert. Eeg and eyeblink response to different acupuncture modalities: preliminary results from four pilot studies. *F1000Research*, 4, 2013.
- D. Mayor and T. Steffert. Measuring mood—relative sensitivity of numerical rating and likert scales in the context of teaching electroacupuncture. initial findings and the influence of response style on results. In *Proceedings of the 18th ARRC International Acupuncture Research Symposium, London, UK*, volume 19, 2016.
- A. K. McAllister, D. C. Lo, and L. C. Katz. Neurotrophins regulate dendritic growth in developing visual cortex. *Neuron*, 15(4):791–803, 1995.
- R. McCraty, M. Atkinson, D. Tomasino, and W. P. Stuppy. Analysis of twenty-four hour heart rate variability in patients with panic disorder. *Biological psychology*, 56(2):131–150, 2001.
- R. McCraty, M. Atkinson, D. Tomasino, and R. T. Bradley. The coherent heart heart-brain interactions, psychophysiological coherence, and the emergence of system-wide order. *Integral Review: A Transdisciplinary & Transcultural Journal for New Thought, Research, & Praxis*, 5(2), 2009.
- D. McGrath, T. N. Judkins, I. I. Pipinos, J. M. Johanning, and S. A. Myers. Peripheral arterial disease affects the frequency response of ground reaction forces during walking. *Clinical Biomechanics*, 27(10):1058–1063, 2012.
- W. Men, D. Falk, T. Sun, W. Chen, J. Li, D. Yin, L. Zang, and M. Fan. The corpus callosum of albert einstein’s brain: another clue to his high intelligence? *Brain*, 137(4): e268–e268, 2014.

- M. Mennes, C. Kelly, X.-N. Zuo, A. Di Martino, B. B. Biswal, F. X. Castellanos, and M. P. Milham. Inter-individual differences in resting-state functional connectivity predict task-induced bold activity. *Neuroimage*, 50(4):1690–1701, 2010.
- D. J. Miller, N. Stergiou, and M. J. Kurz. An improved surrogate method for detecting the presence of chaos in gait. *Journal of biomechanics*, 39(15):2873–2876, 2006.
- K.-i. Minami, H. Nakajima, and T. Toyoshima. Real-time discrimination of ventricular tachyarrhythmia with fourier-transform neural network. *IEEE transactions on Biomedical Engineering*, 46(2):179–185, 1999.
- M. Mirsadeghi, H. Behnam, R. Shalhaf, and H. J. Moghadam. Characterizing awake and anesthetized states using a dimensionality reduction method. *Journal of Medical Systems*, 40(1):1, 2016.
- Mirwais. Finds minimum Embedding Dimension with false nearest neighbours method. <https://uk.mathworks.com/matlabcentral/fileexchange/37239-minimum-embedding-dimension>, 2012. [Online; accessed 18-May-2018].
- S. Mitra, M. A. Riley, and M. T. Turvey. Chaos in human rhythmic movement. *Journal of Motor Behavior*, 29(3):195–198, 1997.
- T. Miyashita, K. Ogawa, H. Itoh, Y. Arai, M. Ashidagawa, M. Uchiyama, Y. Koide, T. Andoh, and Y. Yamada. Spectral analyses of electroencephalography and heart rate variability during sleep in normal subjects. *Autonomic Neuroscience*, 103(1):114–120, 2003.
- T. Mizuno, T. Takahashi, R. Y. Cho, M. Kikuchi, T. Murata, K. Takahashi, and Y. Wada. Assessment of eeg dynamical complexity in alzheimer’s disease using multiscale entropy. *Clinical Neurophysiology*, 121(9):1438–1446, 2010.
- P. Moeynoi and Y. Kitjaidure. Dimension reduction based on canonical correlation analysis technique to classify sleep stages of sleep apnea disorder using eeg and ecg signals. In *2017 14th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, pages 455–458. IEEE, 2017.
- B. Moon, H. Lee, Y. Lee, J. Park, I. Oh, and J. Lee. Fuzzy systems to process ecg and eeg signals for quantification of the mental workload. *Information Sciences*, 142(1-4): 23–35, 2002.

- F. Moon and P. J. Holmes. A magnetoelastic strange attractor. *Journal of Sound and Vibration*, 65(2):275–296, 1979.
- B. R. Moreira, A. P. Duque, C. S. Massolar, R. de Lima Pimentel, M. F. Mediano, T. C. Guimarães, and L. F. Rodrigues Jr. Transcutaneous electrical stimulation of pc5 and pc6 acupoints modulates autonomic balance in heart transplant patients: A pilot study. *Journal of acupuncture and meridian studies*, 2019.
- G. J. Mpitsos and S. Soinila. In search of a unified theory of biological organization: What does the motor system of a sea slug tell us about human motor integration? Technical report, OREGON STATE UNIV NEWPORT HATFIELD MARINE SCIENCE CENTER, 1992.
- I. Mporas, V. Tsirka, E. I. Zacharaki, M. Koutroumanidis, M. Richardson, and V. Megalooikonomou. Seizure detection using eeg and ecg signals for computer-based monitoring, analysis and management of epileptic patients. *Expert Systems with Applications*, 42(6):3227–3233, 2015.
- W. Mumtaz, L. Xia, M. A. M. Yasin, S. S. A. Ali, and A. S. Malik. A wavelet-based technique to predict treatment outcome for major depressive disorder. *PloS one*, 12(2): e0171409, 2017.
- R. Mutihac. Wavelets, fourier transform, and fractals. In *Proceedings of the 6th WSEAS international conference on Wavelet analysis & multirate systems*, pages 49–58. World Scientific and Engineering Academy and Society (WSEAS), 2006.
- S. H. Na, S.-H. Jin, S. Y. Kim, and B.-J. Ham. Eeg in schizophrenic patients: mutual information analysis. *Clinical Neurophysiology*, 113(12):1954–1960, 2002.
- S. Nasehi and H. Pourghassem. Real-time seizure detection based on eeg and ecg fused features using gabor functions. In *Intelligent Computation and Bio-Medical Instrumentation (ICBMI), 2011 International Conference on*, pages 204–207. IEEE, 2011.
- E. Niedermeyer and F. L. da Silva. *Electroencephalography: basic principles, clinical applications, and related fields*. Lippincott Williams & Wilkins, 2005.
- J. Nolan, P. D. Batin, R. Andrews, S. J. Lindsay, P. Brooksby, M. Mullen, W. Baig, A. D. Flapan, A. Cowley, R. J. Prescott, et al. Prospective study of heart rate variability and mortality in chronic heart failure: results of the united kingdom heart failure evaluation and assessment of risk trial (uk-heart). *Circulation*, 98(15):1510–1516, 1998.

- G. Nolte, O. Bai, L. Wheaton, Z. Mari, S. Vorbach, and M. Hallett. Identifying true brain interaction from eeg data using the imaginary part of coherency. *Clinical neurophysiology*, 115(10):2292–2307, 2004.
- M. Nurujjaman, R. Narayanan, and A. S. Iyengar. Comparative study of nonlinear properties of eeg signals of normal persons and epileptic patients. *Nonlinear biomedical physics*, 3(1):6, 2009.
- L. M. Oberman, E. M. Hubbard, J. P. McCleery, E. L. Altschuler, V. S. Ramachandran, and J. A. Pineda. Eeg evidence for mirror neuron dysfunction in autism spectrum disorders. *Cognitive brain research*, 24(2):190–198, 2005.
- S. O’Regan, S. Faul, and W. Marnane. Automatic detection of eeg artefacts arising from head movements. In *2010 annual international conference of the ieee engineering in medicine and biology*, pages 6353–6356. IEEE, 2010.
- F. F. Orsucci. The paradigm of complexity in clinical neurocognitive science. *The Neuroscientist*, 12(5):390–397, 2006.
- S. Osowski and T. H. Linh. Ecg beat recognition using fuzzy hybrid neural network. *IEEE Transactions on Biomedical Engineering*, 48(11):1265–1271, 2001.
- S. Osowski, L. T. Hoai, and T. Markiewicz. Support vector machine-based expert system for reliable heartbeat recognition. *IEEE transactions on biomedical engineering*, 51(4):582–589, 2004.
- T. E. Ozkurt, M. Sun, and R. J. Scwabassi. Decomposition of meg signals with sparse representations. In *2007 IEEE 33rd Annual Northeast Bioengineering Conference*, pages 112–113. IEEE, 2007.
- M. Paluš. Testing for nonlinearity using redundancies: Quantitative and qualitative aspects. *Physica D: Nonlinear Phenomena*, 80(1-2):186–205, 1995.
- M. Paluš. Nonlinearity in normal human eeg: cycles, temporal asymmetry, nonstationarity and randomness, not chaos. *Biological cybernetics*, 75(5):389–396, 1996.
- J. M. Palva, A. Zhigalov, J. Hirvonen, O. Korhonen, K. Linkenkaer-Hansen, and S. Palva. Neuronal long-range temporal correlations and avalanche dynamics are correlated with behavioral scaling laws. *Proceedings of the National Academy of Sciences*, 110(9):3585–3590, 2013.

- J. Pan and W. J. Tompkins. A real-time qrs detection algorithm. *IEEE Trans. Biomed. Eng.*, 32(3):230–236, 1985.
- G. Papakonstantinou and F. Gritzali. Syntactic filtering of eeg waveforms. *Computers and Biomedical Research*, 14(2):158–167, 1981.
- D. B. Percival, A. T. Walden, et al. *Spectral analysis for physical applications*. cambridge university press, 1993.
- V. Periwal and D. Shevitz. Unitary-matrix models as exactly solvable string theories. *Physical review letters*, 64(12):1326, 1990.
- E. E. Peters, E. R. Peters, and D. Peters. *Fractal market analysis: applying chaos theory to investment and economics*, volume 24. John Wiley & Sons, 1994.
- M. Petretta, D. Bonaduce, S. Themistoclakis, A. Ianniciello, L. Scalfi, E. De Filippo, F. Contaldo, F. Marciano, and M.-L. Migaux. Heart rate variability as a measure of autonomic nervous system function in anorexia nervosa. *Clinical cardiology*, 20(3): 219–224, 1997.
- S. M. Pincus. Approximate entropy as a measure of system complexity. *Proceedings of the National Academy of Sciences*, 88(6):2297–2301, 1991.
- S. Pissanetzky. *Sparse Matrix Technology-electronic edition*. Academic Press, 1984.
- B. Pompe. Measuring statistical dependences in a time series. *Journal of Statistical Physics*, 73(3):587–610, 1993.
- A. Porta, S. Guzzetti, R. Furlan, T. Gnechi-Ruscione, N. Montano, and A. Malliani. Complexity and nonlinearity in short-term heart period variability: comparison of methods based on local nonlinear prediction. *IEEE Transactions on Biomedical Engineering*, 54(1):94–106, 2006.
- K. Prabhu. *Window functions and their applications in signal processing*. CRC press, 2013.
- D. Prichard and J. Theiler. Generalized redundancies for time series analysis. *Physica D: Nonlinear Phenomena*, 84(3-4):476–493, 1995.
- G. E. Prinsloo, H. L. Rauch, D. Karpul, and W. E. Derman. The effect of a single session of short duration heart rate variability biofeedback on eeg: a pilot study. *Applied psychophysiology and biofeedback*, 38(1):45–56, 2013.

- J. G. Proakis, D. G. Manolakis, et al. *Digital signal processing-principles, algorithms and applications* ., volume 2. Macmillan New York, 1992.
- R. Q. Quiroga. Quantitative analysis of eeg signals: time-frequency methods and chaos theory. *Institute of Physiology-Medical University Lubeck and Institute of Signal Processing-Medical University Lubeck*, 1998.
- H. Rajaguru and S. K. Prabhakar. Power spectral density with correlation dimension for epilepsy classification from eeg signals. In *2017 2nd International Conference on Communication and Electronics Systems (ICCES)*, pages 376–379. IEEE, 2017.
- L. Ralaivola et al. Dynamical modeling with kernels for nonlinear time series prediction. In *Advances in neural information processing systems*, pages 129–136, 2004.
- S. Ramdani, B. Seigle, J. Lagarde, F. Bouchara, and P. L. Bernard. On the use of sample entropy to analyze human postural sway data. *Medical engineering & physics*, 31(8): 1023–1031, 2009.
- C. Ramon and M. D. Holmes. Noninvasive localization of epileptic sites from stable phase synchronization patterns on different days derived from short duration interictal scalp deeg. *Brain topography*, 26(1):1–8, 2013.
- P. E. Rapp. A guide to dynamical analysis. *Integrative physiological and behavioral science*, 29(3):311–327, 1994.
- I. A. A. Rastiti, H.-L. Zheng, et al. Electroencephalogram brain connectome: An approach in research to identify the effect of acupuncture on human brain wave. *World Journal of Traditional Chinese Medicine*, 4(3):127, 2018.
- J. S. Richman and J. R. Moorman. Physiological time-series analysis using approximate entropy and sample entropy. *American Journal of Physiology-Heart and Circulatory Physiology*, 278(6):H2039–H2049, 2000.
- P. H. Richter and H.-J. Scholz. Chaos in classical mechanics: The double pendulum. In *Stochastic phenomena and chaotic behaviour in complex systems*, pages 86–97. Springer, 1984.
- C. Rieke, F. Mormann, R. G. Andrzejak, T. Kreuz, P. David, C. E. Elger, and K. Lehnertz. Discerning nonstationarity from nonlinearity in seizure-free and preseizure eeg recordings from epilepsy patients. *IEEE transactions on Biomedical Engineering*, 50 (5):634–639, 2003.

- B. F. Robinson, S. E. Epstein, G. D. Beiser, and E. Braunwald. Control of heart rate by the autonomic nervous system: studies in man on the interrelation between baroreceptor mechanisms and exercise. *Circulation Research*, 19(2):400–411, 1966.
- S. Rombouts, R. Keunen, and C. Stam. Investigation of nonlinear structure in multichannel eeg. *Physics Letters A*, 202(5-6):352–358, 1995.
- J. Rottenberg, A. Clift, S. Bolden, and K. Salomon. Rsa fluctuation in major depressive disorder. *Psychophysiology*, 44(3):450–458, 2007.
- P. Rubel, J. Fayn, P. Macfarlane, and J. Willems. Development of a conceptual reference model for digital eeg data storage. In *Computers in Cardiology 1991, Proceedings.*, pages 109–112. IEEE, 1991.
- S. Sakai, E. Hori, K. Umeno, N. Kitabayashi, T. Ono, and H. Nishijo. Specific acupuncture sensation correlates with eegs and autonomic changes in human subjects. *Autonomic Neuroscience*, 133(2):158–169, 2007.
- M. Sakaki, H. J. Yoo, L. Nga, T.-H. Lee, J. F. Thayer, and M. Mather. Heart rate variability is associated with amygdala functional connectivity with mpfc across younger and older adults. *Neuroimage*, 139:44–52, 2016.
- C. B. Saper, T. E. Scammell, and J. Lu. Hypothalamic regulation of sleep and circadian rhythms. *Nature*, 437(7063):1257, 2005.
- J. D. Scargle. Studies in astronomical time series analysis. ii-statistical aspects of spectral analysis of unevenly spaced data. *The Astrophysical Journal*, 263:835–853, 1982.
- K. Schiecke, B. Pester, D. Piper, F. Benninger, M. Feucht, L. Leistritz, and H. Witte. Non-linear directed interactions between hrv and eeg activity in children with tle. *IEEE Transactions on Biomedical Engineering*, 63(12):2497–2504, 2016.
- A. Schödl, R. Szeliski, D. H. Salesin, and I. Essa. Video textures. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, pages 489–498. ACM Press/Addison-Wesley Publishing Co., 2000.
- T. Schreiber and A. Schmitz. Surrogate time series. *Physica D: Nonlinear Phenomena*, 142(3-4):346–382, 2000.
- E. B. Schroeder, L. E. Chambless, D. Liao, R. J. Prineas, G. W. Evans, W. D. Rosamond, and G. Heiss. Diabetes, glucose, insulin, and heart rate variability: the atherosclerosis risk in communities (aric) study. *Diabetes care*, 28(3):668–674, 2005.

- W. W. Seeley, V. Menon, A. F. Schatzberg, J. Keller, G. H. Glover, H. Kenna, A. L. Reiss, and M. D. Greicius. Dissociable intrinsic connectivity networks for salience processing and executive control. *Journal of Neuroscience*, 27(9):2349–2356, 2007.
- F. Shaffer and J. Ginsberg. An overview of heart rate variability metrics and norms. *Frontiers in public health*, 5:258, 2017.
- F. Shaffer, R. McCraty, and C. L. Zerr. A healthy heart is not a metronome: an integrative review of the heart’s anatomy and heart rate variability. *Frontiers in psychology*, 5:1040, 2014.
- R. Shaw. The dripping faucet as a model chaotic system. 1984.
- T.-W. Shen, W. Tompkins, and Y. Hu. One-lead ecg for identity verification. In *Proceedings of the Second Joint 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society*[[*Engineering in Medicine and Biology*, volume 1, pages 62–63. IEEE, 2002.
- R. H. Shumway and D. S. Stoffer. Time series regression and exploratory data analysis. In *Time series analysis and its applications*, pages 47–82. Springer, 2011.
- L.-Y. Shyu, Y.-H. Wu, and W. Hu. Using wavelet transform and fuzzy neural network for vpc detection from the holter ecg. *IEEE Transactions on Biomedical Engineering*, 51(7):1269–1273, 2004.
- A. Siegel and H. N. Saprú. *Essential neuroscience*. Lippincott Williams & Wilkins, 2006.
- M. Sifuzzaman, M. Islam, and M. Ali. Application of wavelet transform and its advantages compared to fourier transform. 2009.
- F. d. Silva. Eeg analysis: theory and practice. *Electroencefalography*, 3:1097–1123, 1987.
- B. P. Simon and C. Eswaran. An ecg classifier designed using modified decision based neural networks. *Computers and Biomedical Research*, 30(4):257–272, 1997.
- J. P. Singh, M. G. Larson, H. Tsuji, J. C. Evans, C. J. O’Donnell, and D. Levy. Reduced heart rate variability and new-onset hypertension: insights into pathogenesis of hypertension: the framingham heart study. *Hypertension*, 32(2):293–297, 1998.
- B. Sivakumar. Chaos theory in hydrology: important issues and interpretations. *Journal of hydrology*, 227(1-4):1–20, 2000.

- E. Skordalakis. Syntactic eeg processing: A review. *Pattern recognition*, 19(4):305–313, 1986.
- M. W. Slutzky, P. Cvitanovic, and D. J. Mogul. Deterministic chaos and noise in three in vitro hippocampal models of epilepsy. *Annals of Biomedical Engineering*, 29(7):607–618, 2001.
- S. W. Smith. The breadth and depth of dsp. *The Scientist and Engineer’s Guide to Digital Signal Processing, Second Edition*, σελ, pages 9–10, 2002.
- L. Soares-Miranda, C. Negrao, L. Antunes-Correa, T. Nobre, P. Silva, R. Santos, S. Vale, and J. Mota. High levels of c-reactive protein are associated with reduced vagal modulation and low physical activity in young adults. *Scandinavian journal of medicine & science in sports*, 22(2):278–284, 2012.
- J. Song, C. Davey, C. Poulsen, P. Luu, S. Turovets, E. Anderson, K. Li, and D. Tucker. Eeg source localization: sensor density and head surface coverage. *Journal of neuroscience methods*, 256:9–21, 2015.
- L. Sörnmo and P. Laguna. *Bioelectrical signal processing in cardiac and neurological applications*, volume 8. Academic Press, 2005.
- V. Srinivasan, C. Eswaran, and N. Sriraam. Approximate entropy-based epileptic eeg detection using artificial neural networks. *Information Technology in Biomedicine, IEEE Transactions on*, 11(3):288–295, 2007.
- C. J. Stam, T. Van Woerkom, and R. Keunen. Non-linear analysis of the electroencephalogram in creutzfeldt-jakob disease. *Biological Cybernetics*, 77(4):247–256, 1997.
- T. Steffert. *Real-Time Electroencephalogram Sonification for Neurofeedback*. PhD thesis, The Open University, 2018.
- T. Steffert and D. Mayor. The fickleness of data: Estimating the effects of different aspects of acupuncture treatment on heart rate variability (hrv). initial findings from three pilot studies.
- T. Steffert and D. Mayor. Does the cortical response to electroacupuncture depend on stimulation frequency? results of a pilot eeg study first proposed at the aacp conference in 20011. In *AACP Annual Conference, Wokefield Park, Reading, UK*, pages 18–19, 2013.

- N. Stergiou. *Innovative analyses of human movement*. Human Kinetics Publishers, 2004.
- N. Stergiou. *Nonlinear analysis for human movement variability*. CRC Press, 2016.
- N. Stergiou, G. Giakas, J. E. Byrne, and V. Pomeroy. Frequency domain characteristics of ground reaction forces during walking of young and elderly females. *Clinical Biomechanics*, 17(8):615–617, 2002.
- P. Stoica and R. L. Moses. *Introduction to spectral analysis*, volume 1. Prentice hall Upper Saddle River, NJ, 1997.
- D. T. Stuss and D. F. Benson. *The frontal lobes*. Raven Pr, 1986.
- A. Subasi. Automatic recognition of alertness level from eeg by using neural network and wavelet coefficients. *Expert Systems with Applications*, 28(4):701–711, 2005.
- V. K. Sudarshan, U. R. Acharya, S. L. Oh, M. Adam, J. H. Tan, C. K. Chua, K. P. Chua, and R. San Tan. Automated diagnosis of congestive heart failure using dual tree complex wavelet transform and statistical features extracted from 2s of eeg signals. *Computers in Biology and Medicine*, 83:48–58, 2017.
- T. Sun, X. Zou, C. Ding, J. Dai, J. Li, J. Wang, and F. Hou. Analysis of epilepsy eeg and eeg correlation based on irc algorithm. In *2018 International Conference on Information Systems and Computer Aided Education (ICISCAE)*, pages 398–401. IEEE, 2018.
- T. Takahashi, T. Murata, T. Hamada, M. Omori, H. Kosaka, M. Kikuchi, H. Yoshida, and Y. Wada. Changes in eeg and autonomic nervous activity during meditation and their association with personality traits. *International Journal of Psychophysiology*, 55(2):199–207, 2005.
- T. Takahashi, R. Y. Cho, T. Mizuno, M. Kikuchi, T. Murata, K. Takahashi, and Y. Wada. Antipsychotics reverse abnormal eeg complexity in drug-naive schizophrenia: a multi-scale entropy analysis. *Neuroimage*, 51(1):173–182, 2010.
- F. Takens. Detecting strange attractors in turbulence. In *Dynamical systems and turbulence, Warwick 1980*, pages 366–381. Springer, 1981.
- H. Takeuchi, Y. Taki, H. Hashizume, Y. Sassa, T. Nagase, R. Nouchi, and R. Kawashima. Effects of training of processing speed on neural systems. *Journal of Neuroscience*, 31(34):12139–12148, 2011.

- J. L. Talmon. *Pattern recognition of the ECG: A structured analysis*. PhD thesis, Vrije Universiteit te Amsterdam, 1983.
- M. P. Tarvainen, J.-P. Niskanen, J. A. Lipponen, P. O. Ranta-Aho, and P. A. Karjalainen. Kubios hrv—heart rate variability analysis software. *Computer methods and programs in biomedicine*, 113(1):210–220, 2014.
- W. O. Tatum IV, S. R. Benbadis, A. Hussain, S. Al-Saadi, B. Kaminski, L. S. Heriaud, and F. L. Vale. Ictal eeg remains the prominent predictor of seizure-free outcome after temporal lobectomy in epileptic patients with normal brain mri. *Seizure*, 17(7):631–636, 2008.
- R. Taylor. Interpretation of the correlation coefficient: a basic review. *Journal of diagnostic medical sonography*, 6(1):35–39, 1990.
- M. Teplan. Fundamentals of eeg measurement. *Measurement science review*, 2(2):1–11, 2002.
- J. F. Thayer and R. D. Lane. A model of neurovisceral integration in emotion regulation and dysregulation. *Journal of affective disorders*, 61(3):201–216, 2000.
- J. F. Thayer, A. L. Hansen, E. Saus-Rose, and B. H. Johnsen. Heart rate variability, prefrontal neural function, and cognitive performance: the neurovisceral integration perspective on self-regulation, adaptation, and health. *Annals of Behavioral Medicine*, 37(2):141–153, 2009.
- J. Theiler and P. E. Rapp. Re-examination of the evidence for low-dimensional, nonlinear structure in the human electroencephalogram. *Electroencephalography and clinical Neurophysiology*, 98(3):213–222, 1996.
- J. Theiler, S. Eubank, A. Longtin, B. Galdrikian, and J. D. Farmer. Testing for nonlinearity in time series: the method of surrogate data. *Physica D: Nonlinear Phenomena*, 58(1-4):77–94, 1992.
- M. Thomas, M. K. Das, and S. Ari. Automatic eeg arrhythmia classification using dual tree complex wavelet based features. *AEU-International Journal of Electronics and Communications*, 69(4):715–721, 2015.
- P. Thomas and R. Moni. Methods for improving the classification accuracy of biomedical signals based on spectral features. *Technology*, 7(1):105–116, 2016.

- T. W. Thompson, M. L. Waskom, and J. D. Gabrieli. Intensive working memory training produces functional changes in large-scale frontoparietal networks. *Journal of cognitive neuroscience*, 28(4):575–588, 2016.
- L. Tong, R.-W. Liu, V. C. Soon, and Y.-F. Huang. Indeterminacy and identifiability of blind identification. *IEEE Transactions on circuits and systems*, 38(5):499–509, 1991a.
- L. Tong, G. Xu, and T. Kailath. A new approach to blind identification and equalization of multipath channels. In *Signals, Systems and Computers, 1991. 1991 Conference Record of the Twenty-Fifth Asilomar Conference on*, pages 856–860. IEEE, 1991b.
- S. Tong, Z. Li, Y. Zhu, and N. V. Thakor. Describing the nonstationarity level of neurological signals based on quantifications of time–frequency representation. *IEEE Transactions on Biomedical Engineering*, 54(10):1780–1785, 2007.
- P. Trahanias and E. Skordalakis. Syntactic pattern recognition of the eeg. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(7):648–657, 1990.
- A. I. Triggiani, A. Valenzano, C. Del Percio, N. Marzano, A. Soricelli, A. Petito, A. Bellomo, E. Başar, C. Mundi, G. Cibelli, et al. Resting state rolandic mu rhythms are related to activity of sympathetic component of autonomic nervous system in healthy humans. *International Journal of Psychophysiology*, 103:79–87, 2016.
- J. K. Udupa and I. S. Murthy. Syntactic approach to eeg rhythm analysis. *IEEE Transactions on Biomedical Engineering*, (7):370–375, 1980.
- M. Unser. Splines: A perfect fit for signal and image processing. *IEEE Signal processing magazine*, 16(ARTICLE):22–38, 1999.
- M. Unser and A. Aldroubi. A review of wavelets in biomedical applications. *Proceedings of the IEEE*, 84(4):626–638, 1996.
- V. A. Vakorin, S. Lippé, and A. R. McIntosh. Variability of brain signals processed locally transforms into higher connectivity with brain development. *Journal of neuroscience*, 31(17):6405–6413, 2011.
- M. Valderrama, C. Alvarado, S. Nikolopoulos, J. Martinerie, C. Adam, V. Navarro, and M. Le Van Quyen. Identifying an increased risk of epileptic seizures using a multi-feature eeg–ecg classification. *Biomedical Signal Processing and Control*, 7(3):237–244, 2012.

- C. L. Vaughan, B. L. Davis, C. Jeremy, et al. Dynamics of human gait. 1999.
- R. N. Vigário. Extraction of ocular artefacts from eeg using independent component analysis. *Electroencephalography and clinical neurophysiology*, 103(3):395–404, 1997.
- S. Vikman, T. H. Mäkikallio, S. Yli-Mäyry, M. Nurmi, K. Juhani Airaksinen, and H. V. Huikuri. Heart rate variability and recurrence of atrial fibrillation after electrical cardioversion. *Annals of medicine*, 35(1):36–42, 2003.
- V.Kaye. Transcutaneous electro acupuncture system, 1998-2019. URL <http://emedicine.medscape.com/article/325107-overview>.
- C. Wagner, B. Nafz, and P. Persson. Chaos in blood pressure control. *Cardiovascular research*, 31(3):380–387, 1996.
- J. Wang, A. Sehmi, N. Jones, and D. De Bono. A knowledge-based system for qualitative ecg simulation and ecg analysis. In *[1991] Proceedings Computers in Cardiology*, pages 733–736. IEEE, 1991.
- J. M. Wang, D. J. Fleet, and A. Hertzmann. Gaussian process dynamical models for human motion. *IEEE transactions on pattern analysis and machine intelligence*, 30(2):283–298, 2007.
- L.-Z. Wang, Y.-L. Pei, and J. Leng. Correlation between vigilance level and driving performance: Influence of the driving duration and circadian rhythm. In *CICTP 2019*, pages 331–343. 2019.
- Y. Wang and S.-C. Zhu. A generative method for textured motion: Analysis and synthesis. In *European Conference on Computer Vision*, pages 583–598. Springer, 2002.
- Y. Wang, Y.-S. Zhu, N. V. Thakor, and Y.-H. Xu. A short-time multifractal approach for arrhythmia detection based on fuzzy neural network. *IEEE Transactions on Biomedical Engineering*, 48(9):989–995, 2001.
- R. M. Warner. *Spectral analysis of time-series data*. Guilford Press, 1998.
- D. Wendi, N. Marwan, and B. Merz. In search of determinism-sensitive region to avoid artefacts in recurrence plots. *International Journal of Bifurcation and Chaos*, 28(01): 1850007, 2018.
- J. L. Willems and E. Lesaffre. Comparison of multigroup logistic and linear discriminant ecg and veg classification. *Journal of electrocardiology*, 20(2):83–92, 1987.

- A. Wolf, J. B. Swift, H. L. Swinney, and J. A. Vastano. Determining lyapunov exponents from a time series. *Physica D: Nonlinear Phenomena*, 16(3):285–317, 1985.
- H. J. Woltring. On optimal smoothing and derivative estimation from noisy displacement data in biomechanics. *Human Movement Science*, 4(3):229–245, 1985.
- S. R. Wurdeman, J. M. Huisinga, M. Filipi, and N. Stergiou. Multiple sclerosis affects the frequency content in the vertical ground reaction forces during walking. *Clinical biomechanics*, 26(2):207–212, 2011.
- C. C. Yang, C.-W. Lai, H. Y. Lai, and T. B. Kuo. Relationship between electroencephalogram slow-wave magnitude and heart rate variability during sleep in humans. *Neuroscience letters*, 329(2):213–216, 2002.
- S. Yarling. Time series modeling as an approach to automatic feedback control of robotic positioning errors. In *Electronic Manufacturing Technology Symposium, 1993, Fifteenth IEEE/CHMT International*, pages 443–449. IEEE, 1993.
- X. Yong, R. K. Ward, and G. E. Birch. Artifact removal in eeg using morphological component analysis. In *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 345–348. IEEE, 2009.
- C. S. Yoo, D. C. Jung, Y. M. Ahn, Y. S. Kim, S.-G. Kim, H. Yoon, Y. J. Lim, and S. H. Yi. Automatic detection of seizure termination during electroconvulsive therapy using sample entropy of the electroencephalogram. *Psychiatry research*, 195(1):76–82, 2012.
- L. C. Young, M. P. Roediger, G. Grandits, J. Baker, C. Somboonwit, I. Williams, J. D. Lundgren, J. D. Neaton, and E. Z. Soliman. Relationship between inflammatory and coagulation biomarkers and cardiac autonomic function in hiv-infected individuals. *Biomarkers in medicine*, 8(9):1073–1083, 2014.
- L. Yuan, F. Wen, C. Liu, and H.-Y. Shum. Synthesizing dynamic texture with closed-loop linear dynamic system. In *European Conference on Computer Vision*, pages 603–616. Springer, 2004.
- Y. Zhao, J. Sun, and M. Small. Evidence consistent with deterministic chaos in human cardiac data: surrogate and nonlinear dynamical modeling. *International Journal of Bifurcation and Chaos*, 18(01):141–160, 2008.

- Y. Zhong, H. Wang, K. H. Ju, K.-M. Jan, and K. H. Chon. Nonlinear analysis of the separate contributions of autonomic nervous systems to heart rate variability using principal dynamic modes. *Biomedical Engineering, IEEE Transactions on*, 51(2):255–262, 2004.
- K. H. Zou, K. Tuncali, and S. G. Silverman. Correlation and simple linear regression. *Radiology*, 227(3):617–628, 2003.