

# Machine Learning and Neuroinformatics/Brain-Computer Interfacing

Seminar Project

Project Report

Students:

Robert Reichel (01431168)

January 18, 2022

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Experimental Setup</b>	<b>2</b>
2.1	Paradigm . . . . .	2
<b>3</b>	<b>Methods</b>	<b>3</b>
3.1	Conventional processing in the frequency domain . . . . .	4
<b>4</b>	<b>Results</b>	<b>5</b>
4.1	Patient AC21 . . . . .	5
4.2	Patient AC22 . . . . .	11
4.3	Patient AC23 . . . . .	17
<b>5</b>	<b>Discussion</b>	<b>23</b>

## 1 Introduction

This report details the analysis of prerecorded EEG data from patients in a minimally conscious state (MCS) to investigate the possibility of constructing a Brain-Computer Interface (BCI) to help the patients in communication with people. As people in a MCS are not able to communicate and it is currently unknown to what degree they are able to perceive their surroundings, it is unknown if they are able to follow the instructions given during the experiment or if they are able to carry them out.

The data was gathered ahead of this analysis by performing experiments described in section 2 by other researchers, thus, the experimental setup will be described according to the documentation.

## 2 Experimental Setup

### 2.1 Paradigm

The aim of the experiment was to investigate the possibility of constructing a BCI to be used by people in a MCS. Since people in a MCS are not able to consciously control the movements of their eyes and it is unknown if they actively perceive the visual information captured through their eyes, a auditory paradigm was chosen. While it is uncertain if the instructions given in the experiment were actually picked up and followed by the patients, the use of an auditory paradigm removes the need for visual attention towards a display.

The patients were sat in a chair with their arms resting on a table in front of them. The room was shielded electromagnetically and from noise to reduce the amount of noise picked up by the EEG and to ensure the same conditions for all participants. They had an 16 channel EEG cap placed on their heads according to figure 2. They received auditory instructions to imagine either clenching both hands to fists or to lift both feet of the ground with the heels remaining on the ground and then dropping the feet again. One patient, AC23, additionally had trials where he was instructed to imagine no movement and thus, had rest periods. The timings of the paradigm are depicted in figure 1.

As can be inferred from figure 1, the trial starts and after one second, a warning is given to the subject to draw attention to concentrate and not imagine any movement. One and a half second after this, instructions are given whether to imagine hand or foot movement. After these

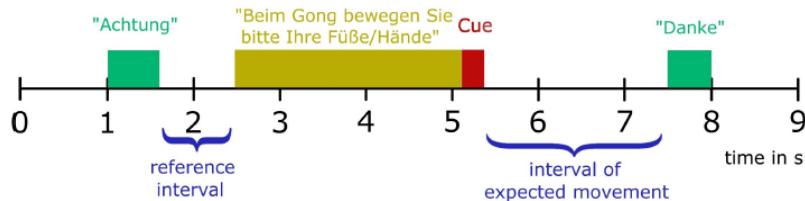


Figure 1: Timescale of the paradigm and the timings

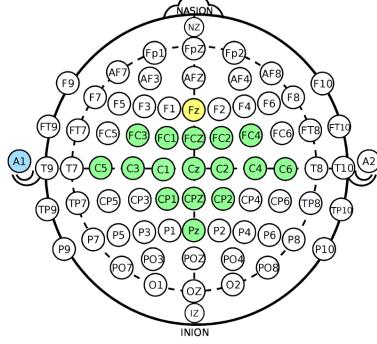


Figure 2: Placement of the electrodes

instructions, a gong is used to signal the start of the imagination period. This period is ended with another cue, after which a short rest period is placed. Overall, 4 trials with 20 trials per trial type were recorded per patient, which resulted in 80 trials per patient and movement type. In the data, the end of the instructions are signaled with the trial type (60 for hand movement, 61 for foot movement and 62 for no movement) and have a duration of 3.6 seconds.

### 3 Methods

The preprocessing was done according to current literature and research. After reading in all data and removing invalid values such as NaN, the signals were filtered using a 4<sup>th</sup> order Butterworth bandpass Filter with the cutoff-frequencies of either 0.1 and 40 Hz for analysis in the frequency domain or 0.3 and 3 Hz for analysing the movement-related cortical potentials (MRCP) of the signal.

After filtering the data, artifact rejection was performed using the data filtered for MRCP usage to make sure the analysed data was valid and any possible patterns or properties found in the data correspond with movement imagination and are not due to other co-occurrences of additional phenomena. The data was then used to perform simple analysis using Topoplots to infer activation patterns, event-related desynchronisation/synchronisation (ERDS) maps to investigate any possible differences between classes in frequency bands between 1 and 40 Hz. Additionally, the band power was plotted both using a linear and logarithmic y-axis to investigate any possible information encoded signal which could be made visible in the frequency domain. To further investigate the data in the time-domain, the MRCP of the data was plotted for the classes of hand and foot movement with the mean potential across the channels and the standard error of the mean.

After these steps, the decision was made to further evaluate the possibility of using the data for classification. To this end, features were extracted from both the time domain and the frequency domain and used to train a shrinkage linear discriminant analysis (LDA) classifier.

The three types of features extracted from the frequency domain were:

- The mean power spectral density over the whole frequency range of 0.1 to 40 Hz.

- The mean power spectral density over the whole frequency range of 0.1 to 40 Hz, after filtering the signal using a common spatial pattern (CSP) filter trained on the data
- The mean power spectral density over frequency bins of 8 Hz to 12 Hz, 10 Hz to 14 Hz, 14 Hz to 19 Hz, 17 Hz to 22 Hz, 20 Hz to 25 Hz, 23 Hz to 28 Hz and 26 Hz to 31 Hz. This results in a division of the power spectral density into bins representing the frequency ranges, where movement imagination usually is represented.

The power spectral density is a measure of the signal's energy and has been successfully used to construct features for classification purposes in the application of Brain-Computer Interfaces.

The features from the time domain are all based on MRCP and are described below:

- Method Robert: All significantly different time points in the interval between trials: The data was split into different trials and the mean MRCP for all classes were calculated. Using the Wilcoxon-rank-sum test, all points that show significant differences between movement types were identified and the amplitude of those points were used as features.
- Method Valeria: All time points between the left-outermost and right-outermost significantly different time points: a window of significance was constructed by downsampling the signal to 8 Hz and performing a Wilcoxon-rank-sum test to investigate time points with a statistically significant difference between movement types. The outermost time points were then used to define a window in which all data points were used as features.
- Method Paper: All time points in a one-second window that performs best for classification: A one-second sliding window across the downsampled data was used to define the features in the time domain. For all possible feature sets, a shrinkage LDA classifier was trained and the dataset that performed best was used for further evaluation.

The windows that were defined for each method are plotted in section ?? as a sanity check. From this, it is visible that while the features from the Paper method are predefined as the one-second window, the features for the Robert method are spanning most of the MRCP interval and thus lead to an increase in feature size. The biggest problem however lies withing the Valeria method, where a poor choice of the p-value results in either no reference window at all or a very narrow window. It has to be said that both the Robert and Valeria method suffer from this problem, which necessitated the usage of a p value of 0.5 for the Wilcoxon-rank-sum test since lower p-values corrupted the feature windows. This is a first indicator showing that the methods using features from the time domain might not be usable for this dataset.

### 3.1 Conventional processing in the frequency domain

To create a reference point to indicate the level of information within the data, the data was analysed by filtering the data with bandpass filters into 10 frequency bands with a width of 4 Hz between 0.1 and 40 Hz. The signals were then used to train a CSP filter using pairs of bands in the range from 2 to 16 CSP bands and the band power of each band was calculated. The calculated values were used to train an LDA classifier for each frequency band. Using the band power from the two best performing frequency bands, a final shrinkage LDA classifier was trained and the performance was evaluated using the accuracy of the predictions. It has to be noted that while the ERDS maps

show high levels of synchronization and desynchronization in multiple frequency bands up to 30 Hz, the frequency bands that perform best in differentiating between the movement classes are the bands between 0.01 and 4 Hz as well as 4 and 8 Hz.

## 4 Results

### 4.1 Patient AC21

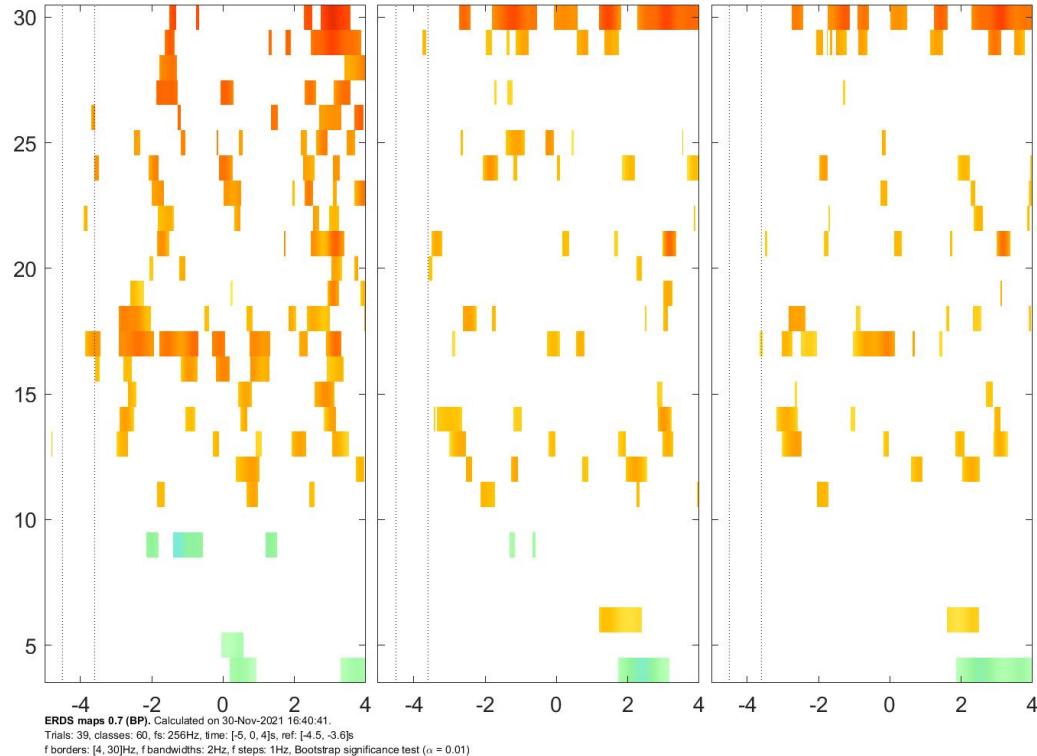


Figure 3: ERDS Map for patient AC 21, hand movement

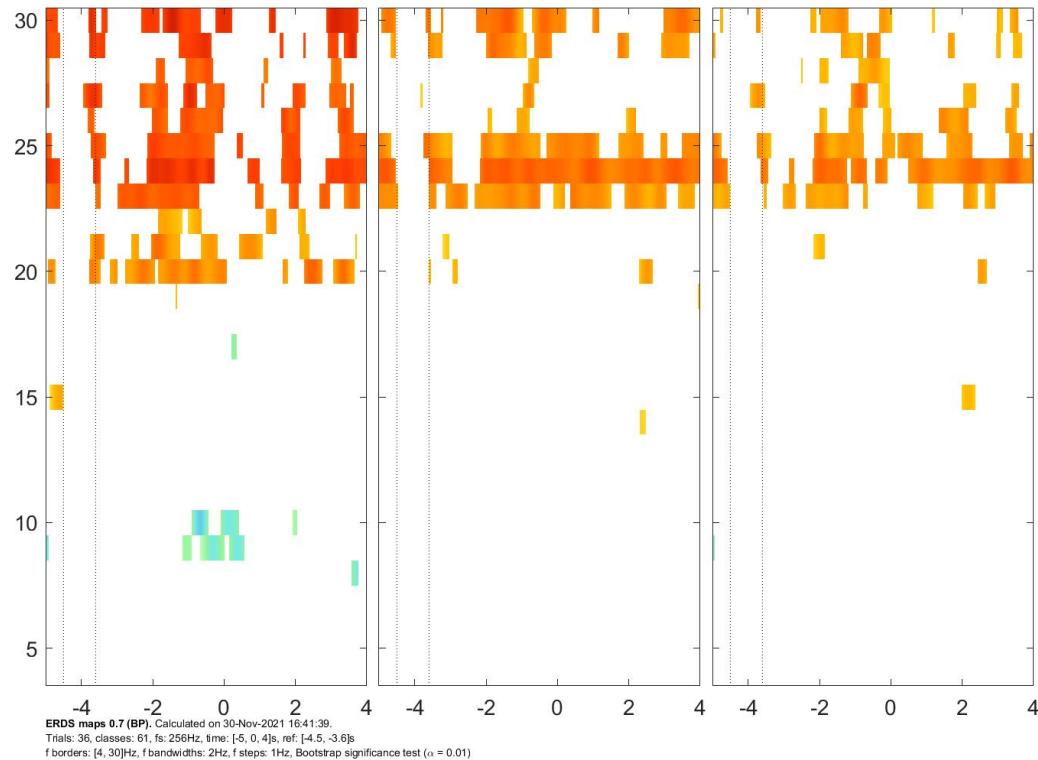


Figure 4: ERDS Map for patient AC 21, foot movement

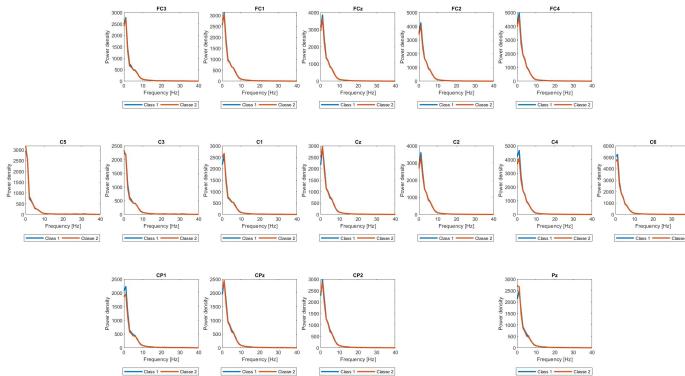


Figure 5: Bandpower for Patient AC21, linear plotted

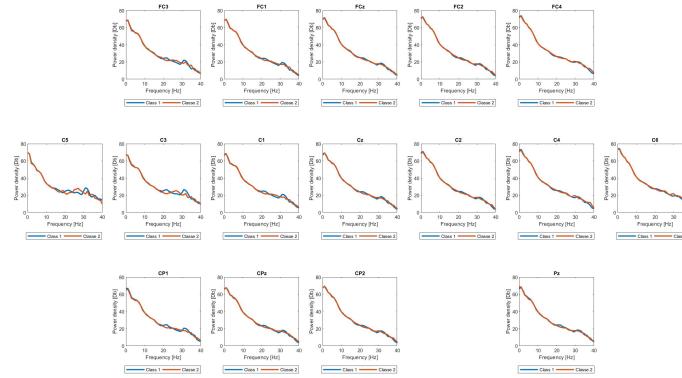


Figure 6: Bandpower for Patient AC21, logarithmic plotted

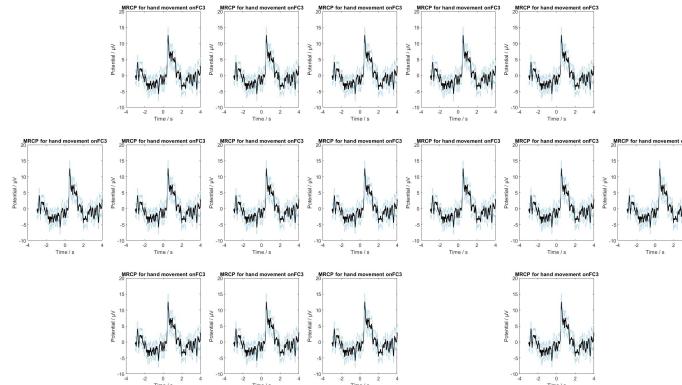


Figure 7: MRCP Patient AC21 for class 1

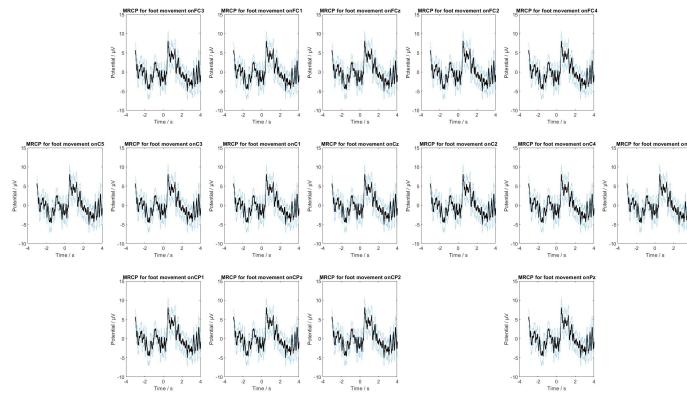


Figure 8: MRCP Patient AC21 for class 2

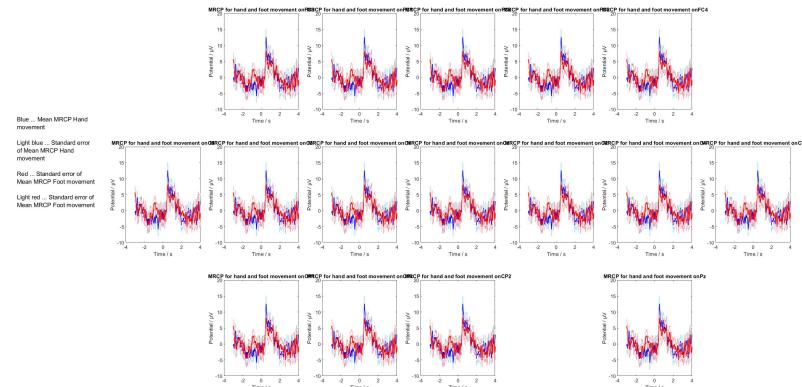


Figure 9: MRCP Patient AC21 for two classes

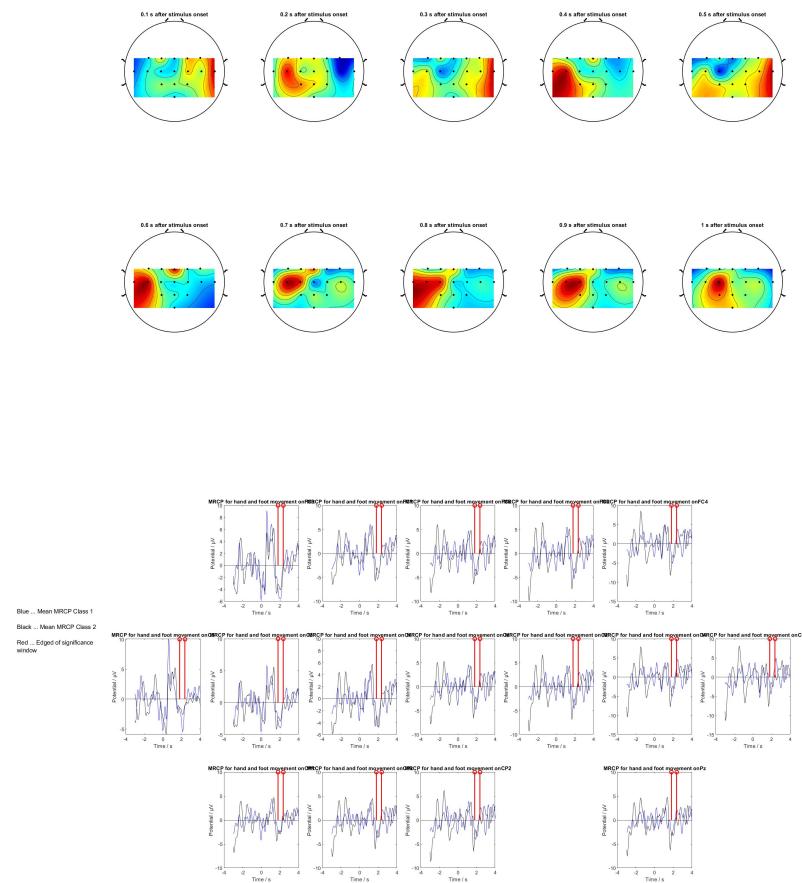


Figure 10: Time domain features, Method from Paper, Patient AC21

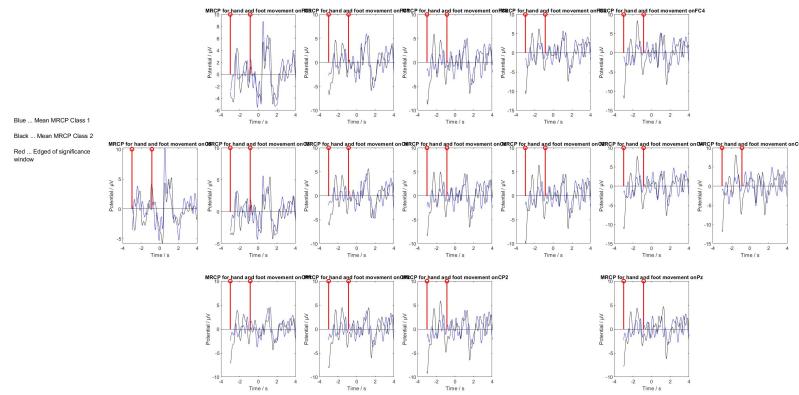


Figure 11: Time domain features, Method from Robert, Patient AC21

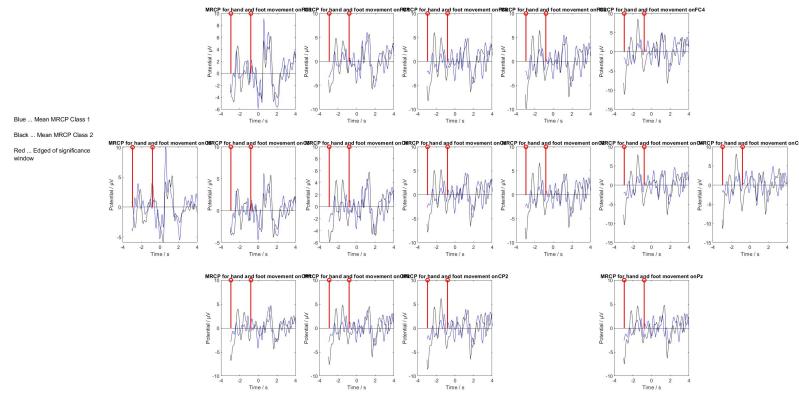


Figure 12: Time domain features, Method from Valeria, Patient AC21

## 4.2 Patient AC22

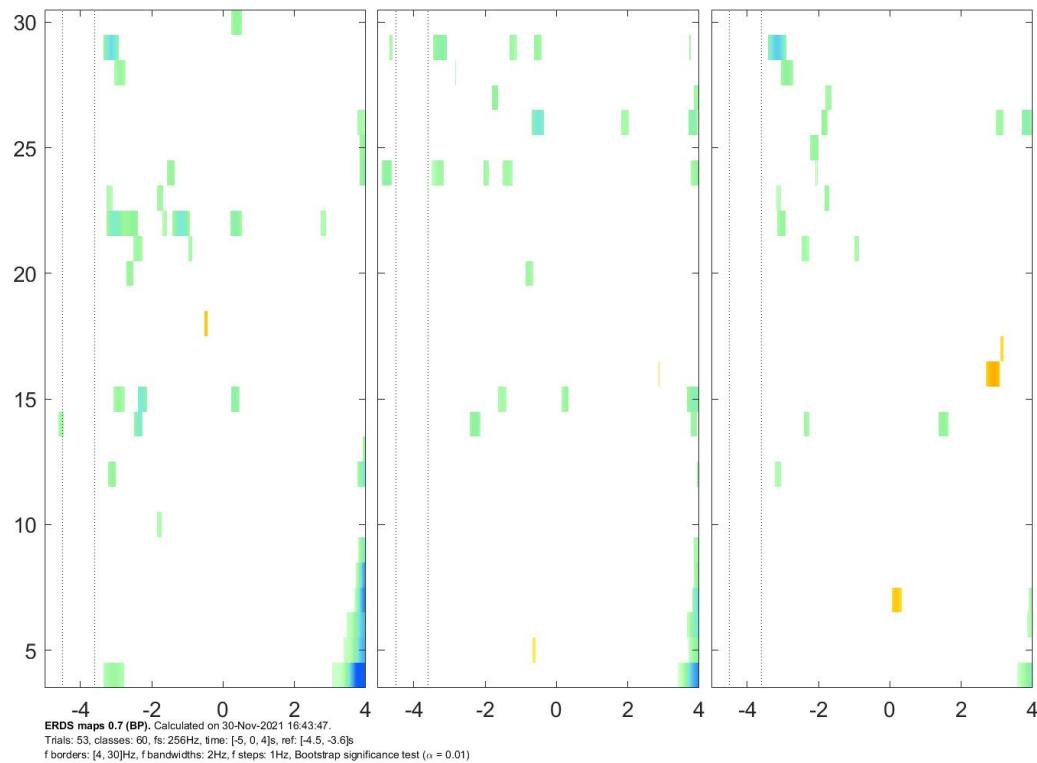


Figure 13: ERDS Map for patient AC 22, hand movement

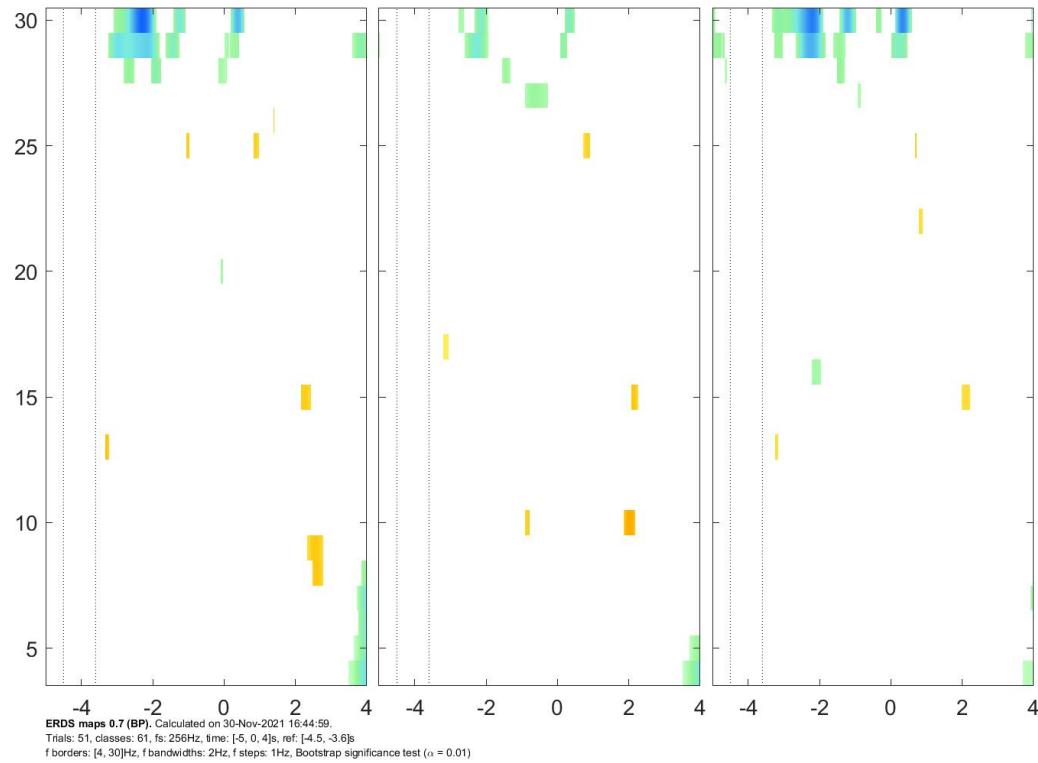


Figure 14: ERDS Map for patient AC 22, foot movement

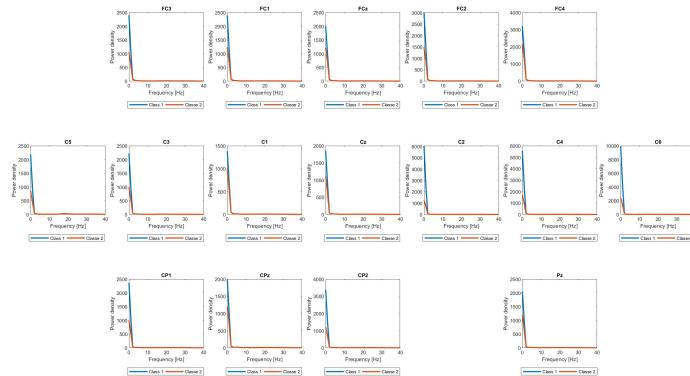


Figure 15: Bandpower for Patient AC22, linear plotted

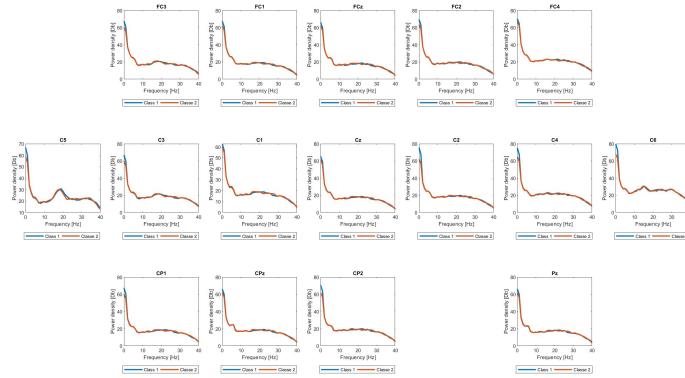


Figure 16: Bandpower for Patient AC22, logarithmic plotted

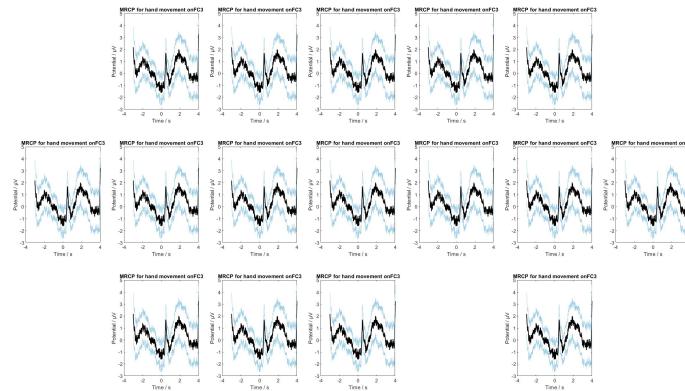


Figure 17: MRCP Patient AC22 for class 1

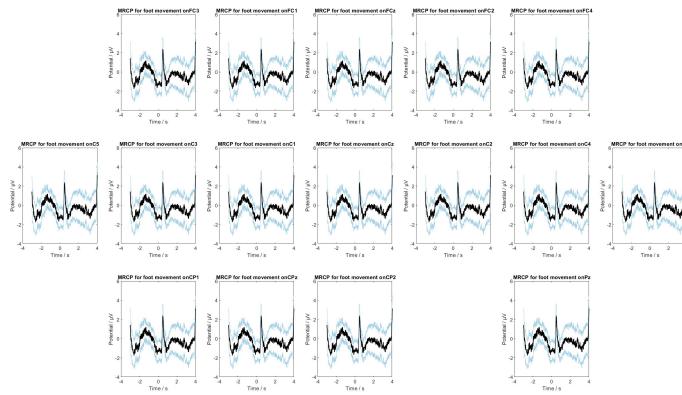


Figure 18: MRCP Patient AC22 for class 2

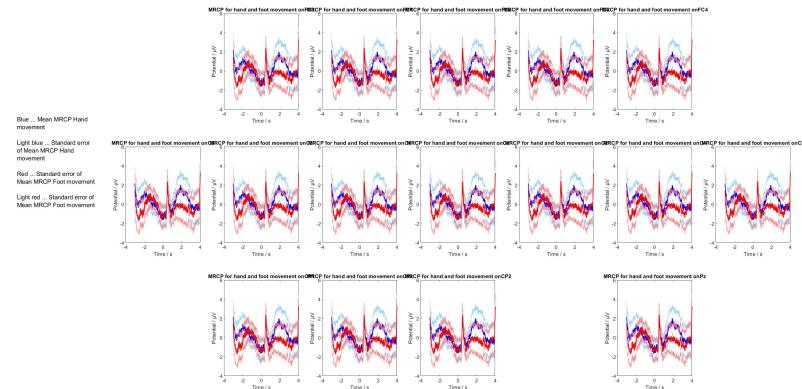


Figure 19: MRCP Patient AC22 for two classes

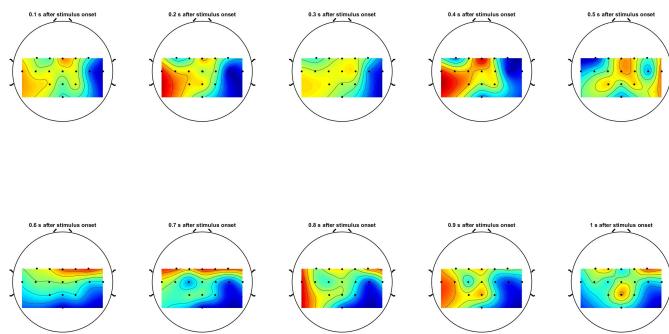


Figure 20: Topoplot for Patient AC22

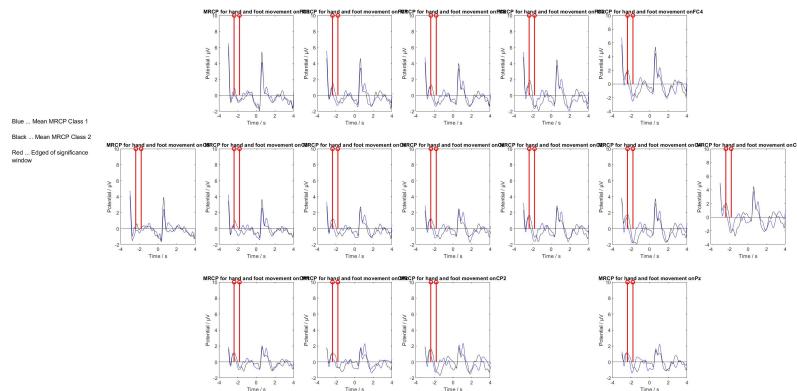


Figure 21: Time domain features, Method from Paper, Patient AC22

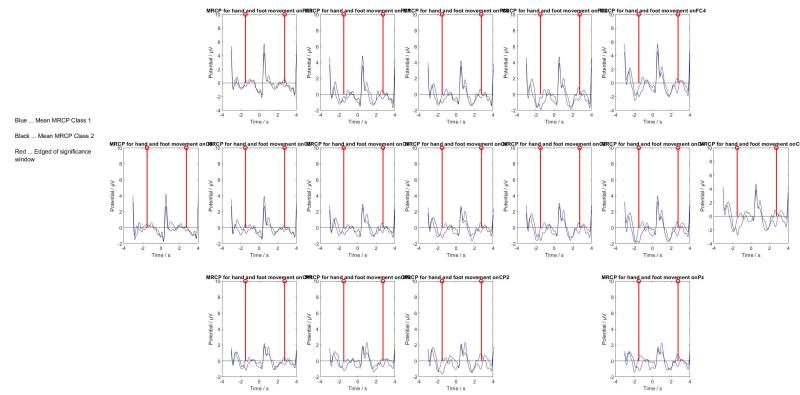


Figure 22: Time domain features, Method from Robert, Patient AC22

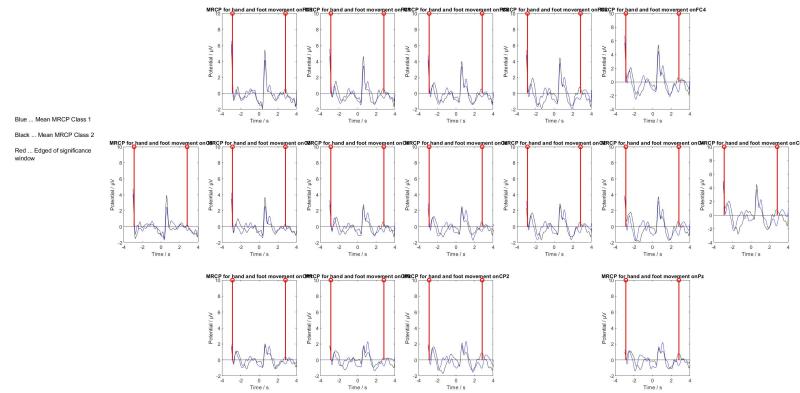


Figure 23: Time domain features, Method from Valeria, Patient AC22

### 4.3 Patient AC23

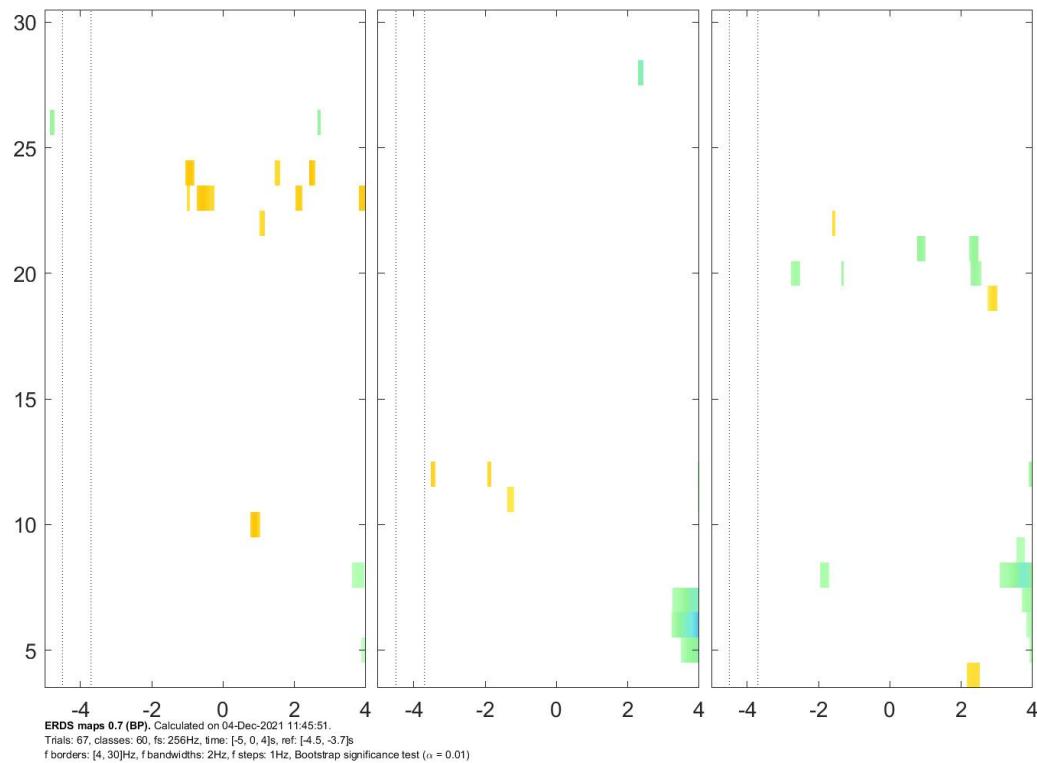


Figure 24: ERDS Map for patient AC 23, hand movement

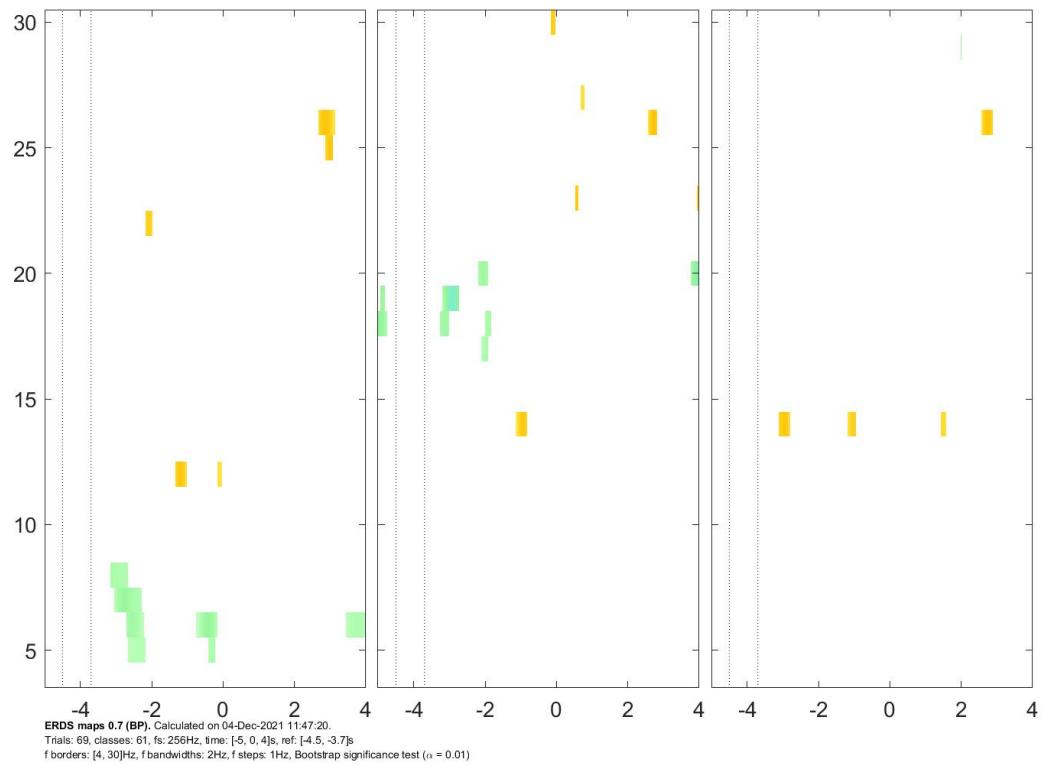


Figure 25: ERDS Map for patient AC 23, foot movement

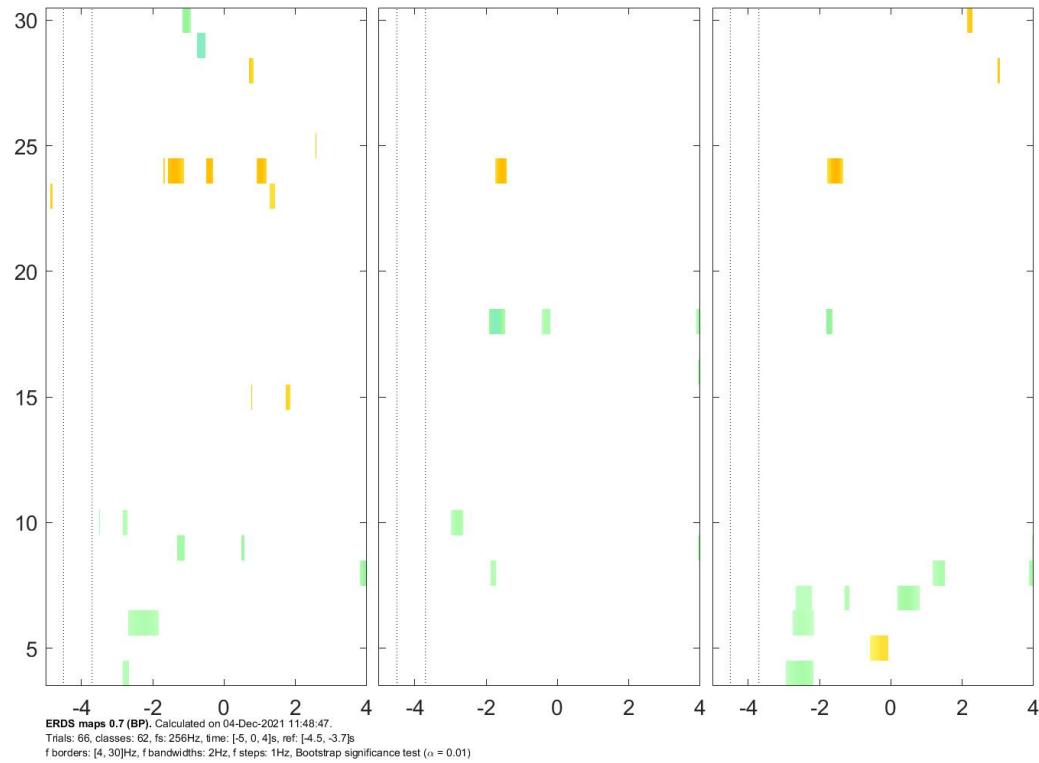
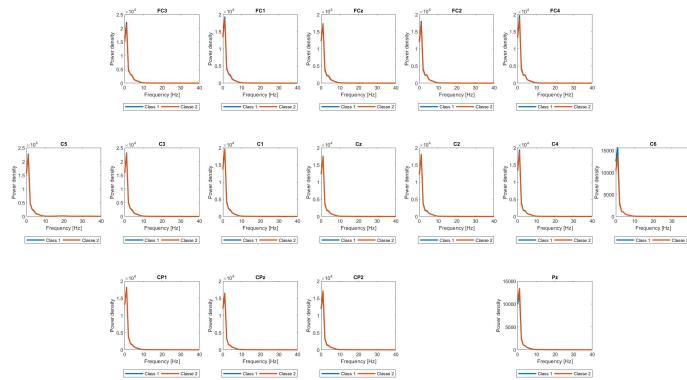


Figure 26: ERDS Map for patient AC 23, rest condition



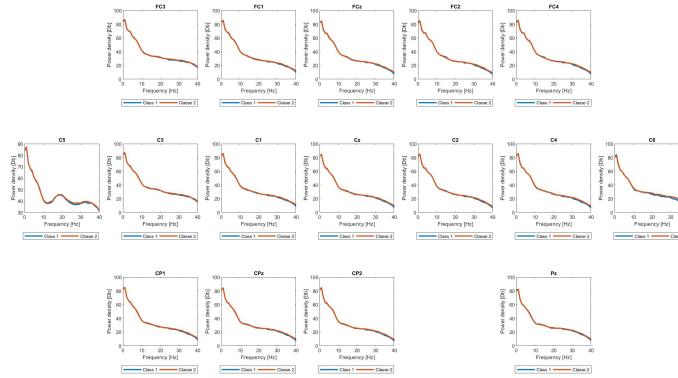
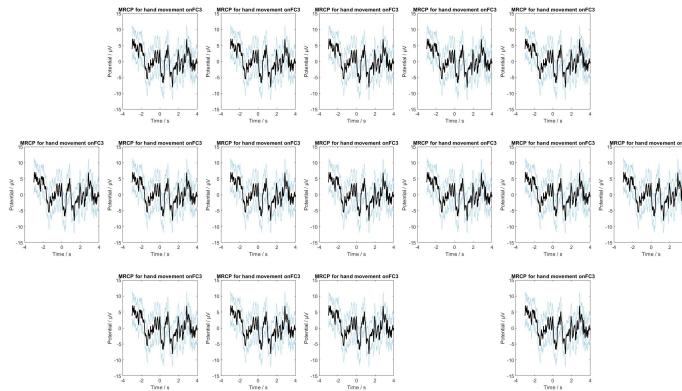
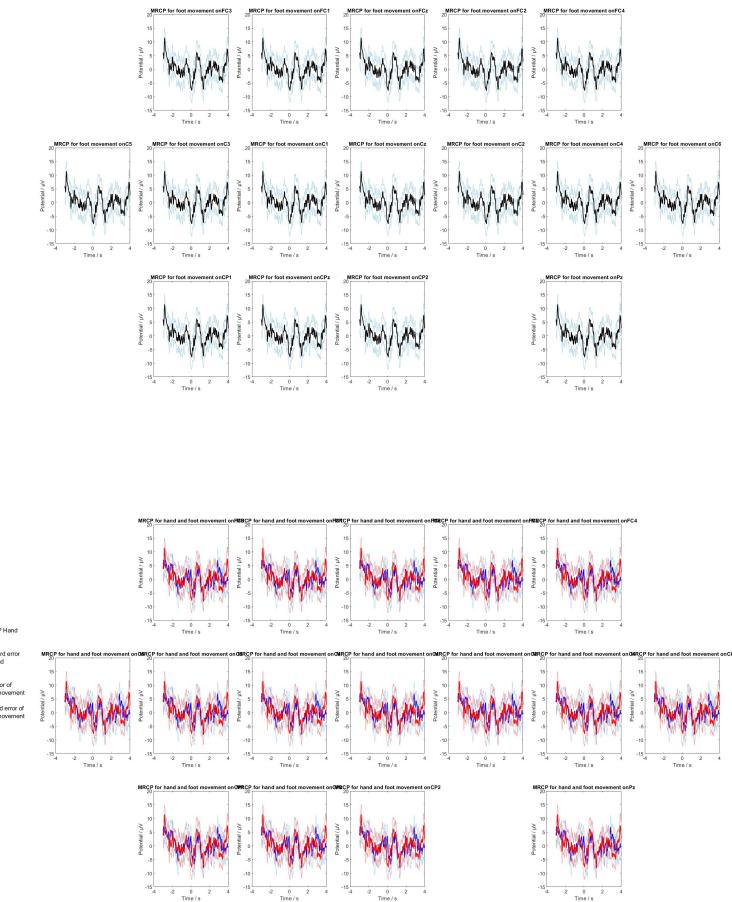


Figure 27: Bandpower for Patient AC23, logarithmic plotted





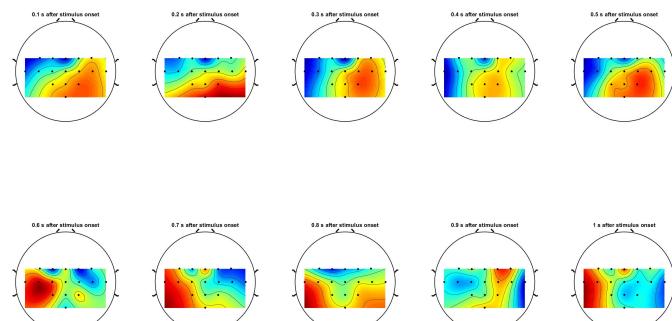


Figure 28: Topoplot for Patient AC23

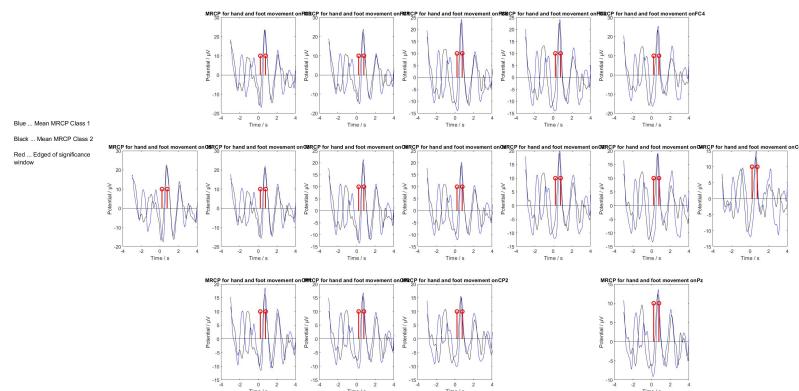


Figure 29: Time domain features, Method from Paper, Patient AC23, Hand and Foot Movement

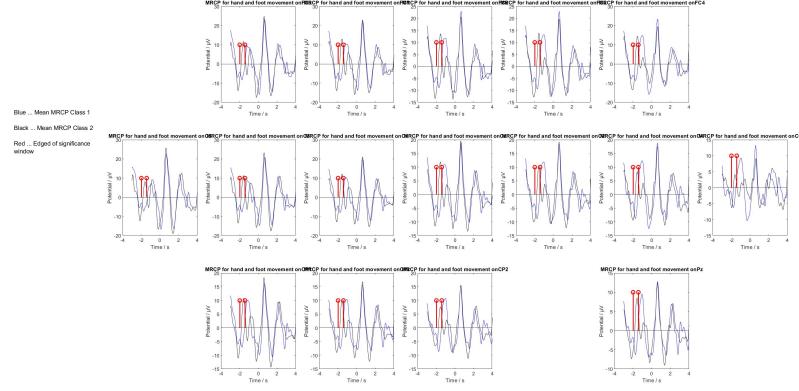


Figure 30: Time domain features, Method from Paper, Patient AC23, Foot Movement and Rest

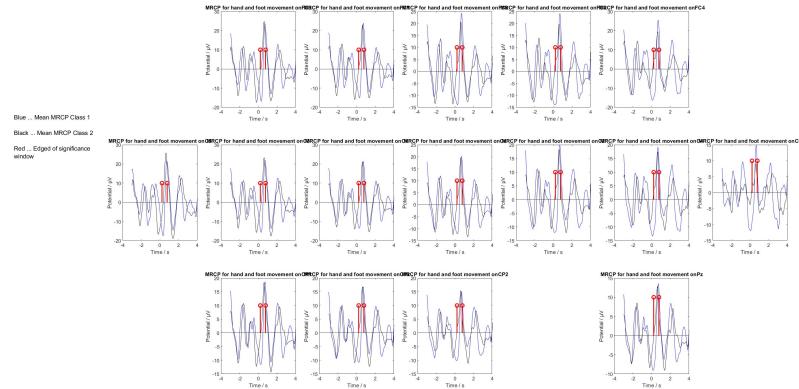


Figure 31: Time domain features, Method from Paper, Patient AC23, Hand Movement and Rest

## 5 Discussion

The results from the analysis of the data are promising. The ERDS maps hint at information encoded in the frequency bands, thus indicating the possibility of using the data for binary decision making. While the linear band power plots do not show any difference in band power, the use of a logarithmic y-axis highlights the possible differences between the two classes, thus validating the approach of training a binary classification through shrinkage LDA for the patients.

The analysis of the data in the time domain yields no significant information. The MRCP itself does not show any interesting characteristics.

Another interesting point of discussion are the achieved accuracies. For the time domain features, the classification accuracy, even for the best methods, is around 60 percent with an exception for the Valeria method and Patient AC22, where an accuracy of nearly 80 percent was reached. When considering the confidence intervals for the chance level of this 80 trials, two-class classification task, an accuracy of 60 percent does not provide any significant difference from the chance level. Through this, it can be concluded that either the method of analysing the data in the time domain, or the information encoded in the time domain of the signal itself is not suited for classification.

A much more pronounced difference can be seen when performing classification using the frequency domain features. Through this, the method described in the paper, where the mean power spectral density (PSD) across certain frequency bins were used as features, all three patients achieved an accuracy of over 80 percent. For patient AC23, the two other methods performed equally well, while they underperformed for the other two patients.

Through this, the idea was hatched to perform classification with features from both the time and the frequency domain, which has been proven to increase performance for Brain-Computer Interfaces. To this end, the best performing features from the time domain and the frequency domain were combined and used as new dataset for classification. From the results, it can be seen that the prognosis of increasing the accuracy through this step did not hold true. Through introducing a higher level of randomness into the classification by adding time domain features to the frequency domain features, the performance of the overall system was dragged down.

The data was analyzed using the classical approach, during which the signals were filtered using CSP filters. For each band, classification was done using a shrinkage LDA algorithm and the two best performing frequency bands were used to construct the final dataset for classification. This approach is robuster considering variations in the time domain and thus may be more usable with patients in a minimally conscious state.

From these results, it can be concluded that frequency domain features are better suited for the analysis of the data. This is supported by the fact that it is unclear if coma patients are able to understand the instructions and follow them. It may be that they understand the paradigm but can not follow it exactly, leading to trigger misalignments. Since time domain features display time- and phase-locked properties, these trigger misalignments can lead to inaccuracies in the feature extraction and classification process. Therefore, for situations in which the correct execution of the paradigm can not be guaranteed, the usage of frequency domain features offers a much more robust approach for building brain-computer interfaces.

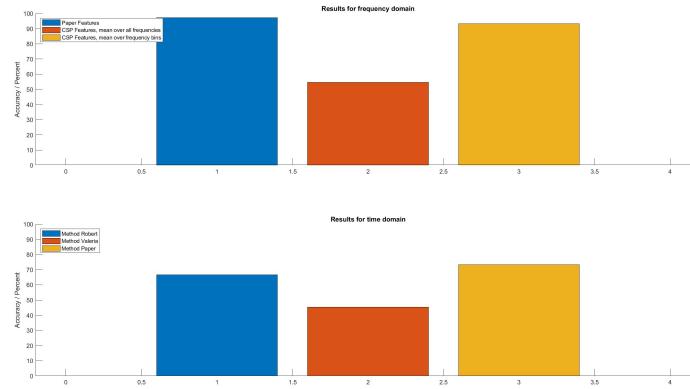


Figure 32: Classification results for patient AC21

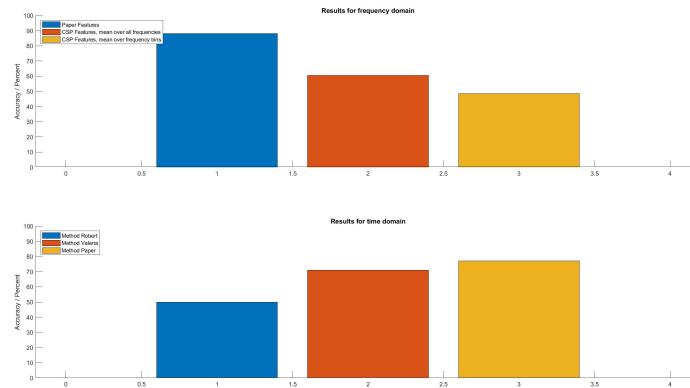


Figure 33: Classification results for patient AC22

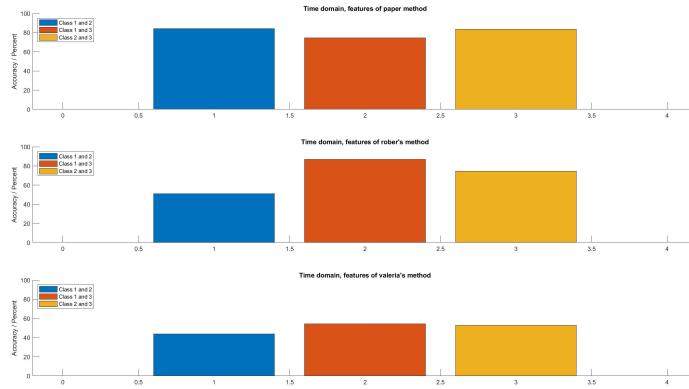


Figure 34: Classification results for patient AC23 in the time domain

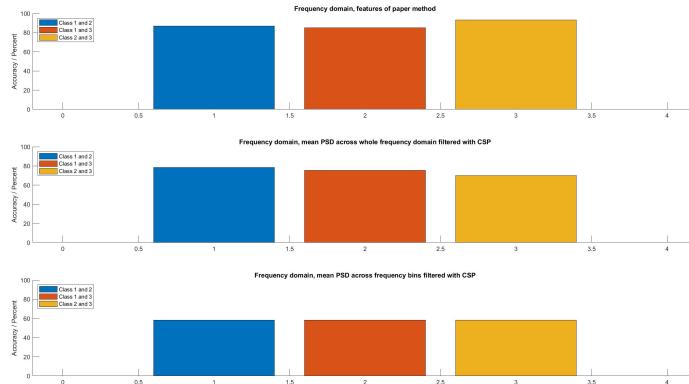


Figure 35: Classification results for patient AC23 in the frequency domain

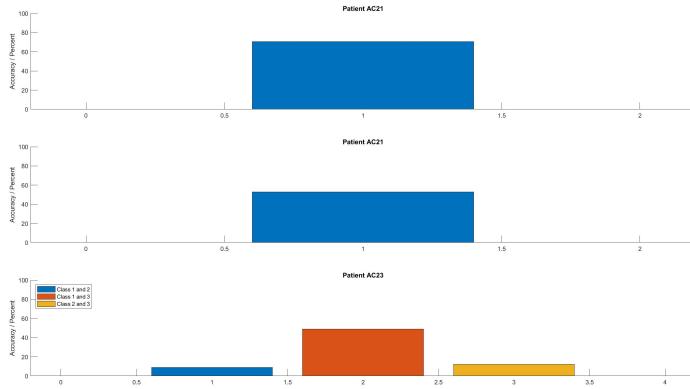


Figure 36: Classification results for combined features

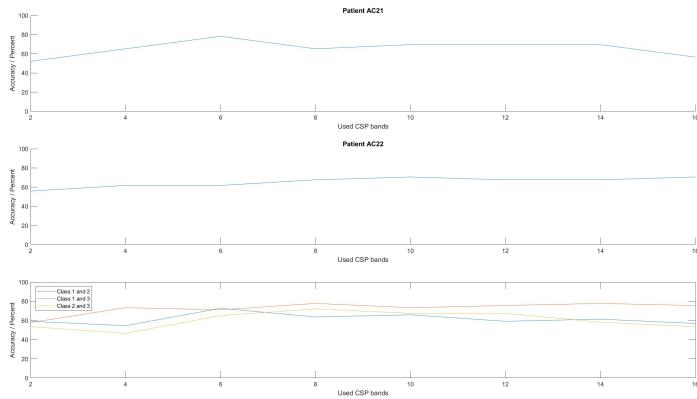


Figure 37: Classification results for CSP filtered band power features