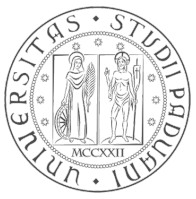


Università degli Studi di Padova

Dipartimento di Ingegneria dell'Informazione

Laurea Magistrale in Bioingegneria

Longitudinal ALS study: research of EEG biomarkers during the progression of the disease

Relatrice

Prof.ssa Alessandra Bertoldo

Correlatori

Dr. Ujwal Chaudhary

Prof. Niels Birbaumer

Institute of Medical Psychology and Behavioural Neurobiology, Tübingen

Laureanda

Arianna Secco

9 Dicembre 2019

Anno Accademico 2018/2019

Abstract

Amyotrophic lateral sclerosis is a neurodegenerative disorder that involves both upper motor neurons and lower motor neurons, progressively bringing patients to a state of complete paralysis. There is evidence of physiological changes happening at a neural level but, nowadays, the diagnosis and the progression of the disease are still based on clinical symptoms. Electroencephalography (EEG) is currently used to extract features that differentiate patients from control groups, but examining how the brain signals change during the evolution of the disease is still challenging due to the different diagnosis types and disease courses.

In this thesis a longitudinal analysis of EEG resting state data is performed for three patients, setting a signal processing pipeline to detect which features of the signal are significantly changing over the observation period, and which of them are possibly changing with a monotonous trend. What has been found is a substantial difference in the spectral content of EEG signal between late-stage patients and the patient observed during his transition from locked-in to completely locked-in state (CLIS), demonstrating that EEG can to represent the neurodegeneration in its different stages. The patient during the transition still exhibits a clear peak in alpha frequency band, that is missing in the other patients. Those, on the other hand, present a shift of alpha rhythm to lower frequencies. These findings suggest a probable gradual decreasing of alpha activity after the transition to CLIS, which could be investigated in a larger dataset covering a wider observation period.

Contents

List of Abbreviations

ALS Amyotrophic Lateral Sclerosis

EEG Electroencephalogram

EOG Electrooculogram

MEG Magnetoencephalogram

TMS Transcranial magnetic stimulation

fNIRS Functional near-infrared spectroscopy

fMRI Functional magnetic resonance

UMN Upper motor neurons

LMN Lower motor neurons

LIS Locked-in state

CLIS Completely locked-in state

DOC Disorders of consciousness

MCS Minimally conscious state

VS Vegetative state

BCI Brain Computer Interface

ICA Independent component analysis

PSD Power spectral density

ANOVA One way analysis of variance

List of Figures

List of Tables

Chapter 1

Introduction

Amyotrophic lateral sclerosis (ALS) is a neurodegenerative disorder that progressively brings patients to a state of complete paralysis and impossibility to interact with the external environment. Many studies investigated whether brain signals can be useful to discriminate between patients and control groups, but monitoring the progression of the disease is still challenging.
In this thesis EEG signals have been analysed to answer some questions: is EEG signal changing in some of its features during the evolution of the disease? If yes, is it evolving in the same features that differentiate patients from healthy subjects?

According to those questions, this work aims to perform a longitudinal analysis of EEG resting states recorded for three different ALS patients, to investigate whether it is possible to find some features in EEG signal that change during the evolution and degeneration of the disease.
The final objective of this study is to propose a solid method to approach this kind of analysis, focusing on data preprocessing and on the statistic validation of the results obtained after data processing and features extraction.

Since considering the nature of the data and its biological correlation with patients' condition is of crucial importance, the analysis is approached with a brief overview of ALS disease in Section 1.1, giving a quick analysis of the physiological factors assumed to have a role in the onset of this syndrome and discussing the implications for patients at a late stage of the disease. One of the aspects that is covered is the transition from locked-in state (LIS) to complete locked-in state (CLIS).
Next to that, in Section 1.2, the choice of EEG as source of information is explained and in Section 1.3 the current state of the art in the analysis of brain signals for ALS

patients is presented, resuming the results obtained in the literature in the research of markers able to differentiate ALS from healthy patients. Successively Brain-Computer Interfaces (BCI) application to communication is reviewed, being the only way that is left to those patients to continue interacting with the external environment even in the locked-in state stage. Finally are presented the current results for longitudinal studies, that are still few in the literature but that would have an important role for a better understanding of ALS disease and for improving patients' living conditions.
This thesis project is part of this last type of studies, considering and facing the limits of available patients' data.
In Chapter 2 is going to be provided a detailed explanation of data and the preprocessing and processing methods used in the analysis.
Finally, in Chapters 3 and 4 all the results are presented and discussed, giving the final conclusions in Chapter 5.

Data were provided by Professor Niels Birbaumer's BCI research group, from the Institute of Neural Psychology of the University of Tuebingen.

1.1 Amyotrophic Lateral Sclerosis

Amyotrophic lateral sclerosis (ASL) is a syndrome referring to a specific form of motor neuron disease that involves both upper motor neurons (UMN) and lower motor neurons (LMN) [1].
UMNs are found in the cerebral cortex and brainstem and are responsible for carrying information down to activate interneurons and LMNs, which innervate skeletal muscle fibers and turn directly the signal to muscles, to contract or relax. While UMNs in the cerebral cortex are the main source of voluntary movement, LMNs are the actuators.
"Amyotrophic" relates to muscle atrophy, weakness and fasciculation (involuntary muscle contraction and relaxation), characteristic of a LMN disease. "Lateral sclerosis" refers to the hardness to palpation of the lateral columns of the spinal cord found in autopsy specimens, due to gliosis following degeneration of cortical spinal tract; this is related to signs of UMN disease [1].
UMN disease signs are the main elements to differentiate ALS from other types of motor neuropathies.
The disease consists in a progressive degeneration of the corticospinal tract, brainstem and spinal anterior horn neurons, with a heterogeneous clinical presentation and course.

This syndrome affects people worldwide and establishing an exact incidence rate is problematic in epidemiologic studies for ALS, due to difficulties in the determination of a specific date for disease onset and the variable latency for symptoms to manifest. Some studies have established that the incidence of ALS in Europe is almost uniformly around 2.16 per 100000 person-years [2].
According to these studies men have a slightly higher incidence of disease with respect to women (3.0 versus 2.4 per 100000 person-years), and the peak age at onset is 58-63 years for sporadic disease and 47-52 years for familial disease, with the incidence increasing rapidly after 80 years [3].

1.1.1 Clinical phenotypes and progression

A first distinction between ALS patients is related to the underlying causes of the disease: about 5-10% of the cases are familial, with a Mendelian pattern of inheritance, while 90% of cases are classified as sporadic [3] [1]. For the familial ALS some genes have been identified as possible causes for the presentation of typical clinical phenotypes, while for the sporadic ALS the pathological path of the disease is still mostly unclear.
Some studies proposed also environmental factors competing in the incidence of the disease, as the exposure to heavy metals, persistent viral infection [1], a prolonged and intensive physical exertion or active service in the armed force [3] .

The main presentations of the syndrome consist of:

- limb-onset ALS, with UMN and LMN signs in the limbs (about 70% of patients);
- bulbar-onset ALS, starting with speech and swallowing difficulties and limb features developing later (25%);
- ALS with pure UMN involvement;
- progressive muscular atrophy, with pure LMN involvement [4].

Generally, the disease signs involving the limbs are spasticity, weakness and brisk deep tendon reflexes concerning UMN disturbance, while common LMNs' limb symptoms are fasciculation, wasting and weakness.
Regarding bulbar features the most spread is spastic dysarthria for UMN dysfunction, while bulbar LMN disorder can be identified by tongue wasting, weakness, fasciculations, flaccid dysarthria and later dysphagia [3]. The disease and the presentation of its symptoms are progressive, and the rate of the progression depends both on the age at

symptoms onset and on onset symptoms themselves.
Reduced survival rates are related with older age at symptoms onset, early respiratory muscle dysfunction and bulbar onset disease, while on the other hand limb-onset disease and younger age at presentation can be some predictors for prolonged survival [3].

The transition from LIS to CLIS

The advanced stage of ALS is identified as locked-in state (LIS), a term denoting a neurological condition consisting of tetraplegia and paralysis of all cranial nerves except vertical eye movements [5]. A further progression of the disease leads to a total incapability of moving all muscles, including the residual eye movement, bringing the patients to a complete locked-in state (CLIS).
This step in the advancement of the disease is of particular interest and importance for the quality of life of patients, since it excludes them from the possibility to communicate through eye movement. Communication using eye movement is demonstrated to be effective for LIS patients through neuroprosthethic devices as Brain-Computer Interfaces (BCIs), explained in Section 1.3.2. BCIs usage for CLIS patients implies the analysis of other brain signals as electroencephalogram (EEG) and functional near-infra-red spectroscopy (fNIRS), but currently there are no proves for stable and lasting communication using those signals, probably due to the lack of knowledge of this particular brain state and its implication to brain's electrophysiological activity.
Understanding a possible correlation between the progression of the disease and related brain activity would help in improving communication techniques and consequently the quality of life for this late stage of the disease. Deeper considerations on the topic are postponed to Section 1.2.

1.1.2 Pathophysiology of the disease

The pathophysiological mechanisms underlying ALS development have to be investigated in a complex interaction between genetic and molecular pathways.
In particular, the mutation of SOD1 seems one of the elements more involved in the concurrent and auto-reinforcing process that leads to the the death of motor neurons, and it represents also a possible connection in the pathogenesis of both familial and sporadic ALS.
SOD1 is an enzyme that catalyzes the convertion of toxic superoxid radicals to hydrogen peroxide and oxygen [1]; its mutation is currently thought to lead to a toxic gain-of-function. Moreover, SOD1's mutation induces instability and misfolding of the SOD1 peptide, leading to the formation of intracellular aggregates that disrupts axonal

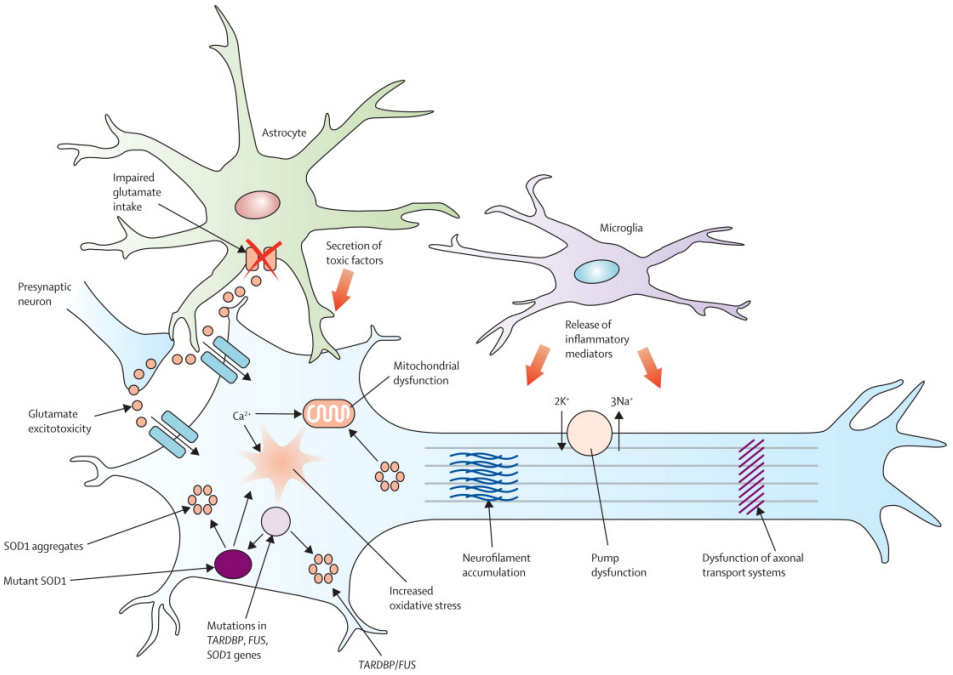


FIGURE 1.1: Cellular and molecular processes involved in ALS neurodegeneration [3].

transport systems and vital cellular functions [3].

Another factor that seems to be connected to ALS pathogenesis is glutamate-induced excitotoxicity, consisting in an excessive entry of extracellular calcium through the inappropriate activation of glutamate receptors. This theory is supported by findings of increased glutamate levels in cerebrospynal fluid of patients with sporadic ALS [1]. Another effect of this excitotoxicity is the generation of free radicals, and both these elements are shown to lead to neurodegeneration and cellular death. In addition to these two processes other anomalous neural conditions are implicated in the generation of the disease (as structural abnormalities of mitochondria, dysfunction of sodium/potassium pump, autophagy and disrupted axonal transport systems) together with non-regular behaviour neural cells, as astrocytes and microglia, that could generate an insufficient release of neurotrophic factors, secretion of neurotoxic mediators and modulation of glutamate receptor expression. Figure 1.1 offers a schematic view of these processes.

1.1.3 Diagnosis and treatments

Currently there is no diagnostic test for ALS, and diagnosis are made relying on the indentification of UMN and LMN signs, and their combination in the limb or bulbar territories. These features are usually examinated through El Escorial criteria [6].

Generally diagnostic certainty entails a delay of one year from the onset of symptoms to diagnosis [7], due to the need of observing the possible progression and the spread of the symptoms to other regions, and for the possible overlap with other neuropathies

and disorders of motor neurons that present similar features. Anyway, this delay implies a late start for the treatments that could slow the progression of early symptoms, and that is the reason why an early diagnostic biomarker would be clinically useful.
Electromyography can be used to count the number of surviving neurons, becoming an help for diagnosis and for assessing the efficacy of treatments.
A great contribution to the diagnostic process comes from neuroimaging techniques, specially for their ability to exclude other pathological causes. In particular magnetic resonance voxel-based morphometry allows to quantify grey and white matter volumes. Those are important features in diagnosis, since an extensive decrease in frontotemporal white matter volume has been documented for ALS patients [7].
Functional magnetic resonance (fMRI) is used to measure regional changes in the blood-oxygenation-level-dependent response to stimuli, given that it has been proved that there is a decrease in regional pattern activation during motor tasks in ALS patients, and a parallel activation of other regions that is correlated with UMNs disturbances.
Moreover there are other methods that helps in documenting the involment of the UMNs, as magnetic resonance spectroscopy and magnetic stimulation of the motor cortex [1]. The first is used to measure the number of surviving neurons in the motor cortex, and the second is used to assess the conduction of the corticospinal tract.

The only one drug approved by the Food and Drug Administration (FDA) for treating ALS is riluozole, considered a neuroprotective therapy as it is a glutamate antagonist. Some studies and retrospective analysis proved that patients who received riluozole had a slower degeneration with respect to a control group [8] and extendend survival of patient of 3-6 months [9], supporting the theory of the excitotoxicity of glutamate in the pathogenesis of ALS.
A key role in the maintainance of patients' good quality of life is played by the symptomatic treatment, provided by the cooperation of physioterapists, speech therapists, respiratory physicians, gastroenterologists and care takers. The biggest issues in the treatment of the symptoms are related to respiratory insufficiency, that is the main cause of death [3]. Respiratory failure indicates the degeneration of both central respiratory centres and motor neurons involved in the activity of the phrenic nerve.
Patients and their families have to face the need of artificial ventilation as soon as respiratory symptoms appear, and most of the patients improve their quality of life with the use of non-invasive ventilation; patients that are intolerant to this form of ventilation, or when this help is no more sufficient, can choose the option of invasive ventilation via tracheostomy, keeping them alive but implying a profound decrease in their quality of life.

1.2 EEG in ALS

Electroencephalography (EEG) refers to the recording of the oscillations of brain electric potentials, and it is acquired through electrodes placed on the scalp with high time resolution. On the other hand, each scalp electrode records electrical activity at very large scales, recording electric potentials generated in the underlying tissue from a very large neural population in the cortical layer. For this reason the spatial resolution is not optimal but EEG signals can still detect synaptic actions that are related to specific brain states.
EEG is analysed both in time and frequency domain. Frequency domain and spectral analysis are of particular interest since the frequency component of EEG, considering separated frequency bands, is correlated to different brain conditions. Usually frequency ranges are categorized as [10]:

- delta (1 to 4 Hz), the highest in amplitude and slowest in wave, represents the grey matter of the brain, it is found in all sleeping stages. It is normal and dominant rhythm in infants and it is abnormal for adults that are awake.
- theta (4 to 8 Hz), related to subconscious activity, observed in deep relaxation and meditation. It is normal for children under 13 years and abnormal for adults in awake state.
- alpha (8 to 13 Hz), found on relaxation state in both sides of the head. It represents the white matter of the brain.
- beta (13 to 30 Hz), concerned with behaviour and actions, related to sensorial perceptions. These waves are typically related to conscious states like talking, problem solving and decision making.
- gamma (higher than 30 Hz), associated with perception and consciousness and occurring during hyper-alertness and integration of sensory inputs.

EEG is applied both to resting state and task-related studies, due to its correlation with vigilance and thinking activities. EEG experiments are also suitable for neurological disorders as Alzheimer, seizure disorders, attention deficit, autism, Parkinson, ALS and many others due to the correlation of these disorders to brain signals with a specific frequency. Moreover, the correlation of EEG with brain physiological changes can be investigated.

In particular, the literature some EEG studies are focusing on ALS disorders, both for resting state analysis [11] [12] [13] and for task related application as BCI [14] [15].

Their aim is mainly to compare EEG signal for ALS patients and controls, to find some features of the signal that can be used as biomarkers for the disease, possibly to improve diagnosis and to assess treatment efficacy.
Their analysis is made through quantitative electroencephalography (QEEG), and the state of the art concerning these studies is reported in Section 1.3.1.
Generally, it has been found an overall power decrease for power spectral densities and an apparent shift of power to low frequencies.

1.2.1 Motivation: why EEG

EEG is particularly elected for resting state analysis because it closely reflects real-time neural activity with high temporal resolution, being at the same time widely accessible and portable.
Besides the possibility of extracting time domain and spectral features from EEG, a multichannel acquisition allows retrieving synchrony and connectivity measures between brain regions, which are strongly correlated with magnetic resonance imaging (MRI) features and are demonstrated to have similar discriminating power [16]. The great advantage of EEG with respect to MRI is in terms of cost and accessibility.
Moreover, quantitative EEG has the potential to capture upper motor system changes in ALS [17], a determinant aspect in the disease's diagnosis.
For this reason EEG application to diagnosis or communication through EEG-BCI would be applicable to patients easily and at their home. EEG-BCI approaches are still unable to maintain a stable rate of communication with those patients [14], so it would be useful to understand better the characteristics of this signal and how it modifies during the progression of the disease, to improve communication performances and consequently quality of life for these people. In particular EEG-BCIs could possibly focus on different bands that are more significant or more stable during the progression of the disease, allowing to keep communicating also for CLIS patients [18].
Furthermore, there are a lot of studies investigating the relation of EEG signal with different states of consciousness in disorders of consciousness (DOC), vegetative state and coma[19] [20] [21]. The same features can be applied to ALS patients to help in understanding better their condition at late-stages of the disease, as they differentiate from DOC patients. Indeed, for both LIS and CLIS patients consciousness is thought to be fully preserved, but in LIS patients it can be demonstrated through voluntary blinking while for CLIS patients the total loss of muscular control prevents from having an objective assessment of their state of consciousness.

1.3 State of the art

Currently the literature offers several studies and reviews analysing EEG signals to investigate ALS disease. In this section different aspects in the state of the art of the research are reported.

1.3.1 Control vs patients studies and biomarkers

The majority of the studies focus on the comparison between ALS patients and controls, to use EEG features measured in these subjects as a tool to help in diagnosis and in the assessment of treatments' effectiveness.
R. Mai et al. in a qEEG study [12] found a significant difference between ALS group and controls in relative alpha power in central regions, while the other frequency bands could not separate the groups, result that is in agreement with the site of the pathological changes (primary motor cortex). The separation occurred for the "mu rhythm", that is a frequency component with the same range of classical alpha band but, contrarily, is not located in the posterior regions and is independent from those oscillations. Moreover, the absence of separation for theta and delta power, which are the ones thought to be the most reliable electrophysiological correlate of cognitive decline, indicates that EEG signs characteristics of dementia were not present in their results.
Hohmann et al. documented a shift in alpha peak frequency comparing ALS late-stage patients and controls [18], and in addition to that they found a significant difference in the location of alpha peak with respect to other neurodegenerative syndromes as Alzheimer's disease and schizophrenia. That aspect can be of help in specifying the diagnosis.
Jayaram et al. investigated the spectral features of cortical processes after applying independent component analysis to EEG signals from patients and healthy subjects, founding a global bandpower enhancement in the high-γ range [14]. Moreover, as in previously mentioned studies, they found a peak in the lower-α range in the central area for ALS patients.
Nasseroleslami et al. conducted a complex study that combines EEG characteristics of ALS and control subjects with MRI [16], to correlate electrophysiological changes to physiological and functional evidence from brain imaging techniques. In particular, they considered both spectral power and connectivity measures, finding a decrease in θ-band power spreading to adjacent frequency bands in ALS patients, together with a general increase in average connectivity, particularly significant over bilateral motor-regions of the scalp for θ-band and over parietal and frontal scalp regions for γ-band. The location of their results is in accordance with the one that is specific for the degeneration of

ALS disease. Since the syndrome is known to cause a degeneration of white matter, the expected electrophysiological evidence would be an attenuation of communication and connectivity across the brain; the increase in connectivity could therefore be interpreted as a reflection of extra compensatory activity within and outside areas associated with white matter degeneration [16].
Another useful result from their study is the absence of difference in the EEG indices between patients treated with Riluozole and those not on Riluozole ALS therapy [16], implying the possibility to assume that this pharmacological treatment is not affecting EEG signal and consequently is not interfering with qEEG results.
Duckick et al. in their study [17] documented a widespread spectral power decrease from δ to β bands, with a parallel increase in amplitude envelope correlation (a connectivity measure) and decrease in synchrony (through the analysis of imaginary coherence) over the same frequency ranges, comparing ALS patients with controls. The recognized discriminant measures are correlated with structural degeneration as reported from MRI results. In particular, they found a strong correlation between imaginary coherence and the average cortical volume, while the decrease in spectral power can be attributed to structural degeneration of pyramidal cells.

EEG in disorders of consciousness

Other studies focus on EEG signatures correlated to the state of consciousness of patients, differentiating patients in a vegetative state (VS) from those in a minimally conscious state (MCS). The research of biomarkers related to this topic has several implications, from diagnosis to prognosis and evaluation of brain intervention [20].
Many studies are resumed in reviews that found out spectrum power, coherence and entropy as the most frequently used features in differentiating consciousness levels. In particular the increase in low band power (δ and θ) and the decrease in mid-high band power (α) resulted being a common characteristics of patients with DOC, while comparing MCS with VS patients the second ones resulted having increased δ and decreased α power with respect to the first ones [20]. Additionally, according to the hypothesis that a brain in a lower consciousness level would present a decrease in brain activity, EEG complexity measures (both in time domain and in frequency domain) showed a congruent decrease.
These results are confirmed by the study of Sitt et al. [19], that identifies also a greater fluctuation over spectral features values across MCS patients, in agreement to the definition of that state as a fluctuating one. On the other hand, they found a stable state of increased δ and reduced θ-α power as a solid sign of unconsciousness. Looking at

complexity metrics, they noticed a increase in spectral entropy in MCS with respect to VS patients.
An interesting aspect of their study is the evaluation of fluctuations and stability of the indexes.

Considering ALS syndrome as a disorder of consciousness is incorrect since late-stage ALS patients, which state is recognized as LIS or CLIS, are still completely conscious, as by definition these forms are characterized by preserved consciousness with the only difference that CLIS patients are almost or completely unable to interact with the external environment [11].
Comparing EEG signatures for ALS disease and progression to the ones usually assessing a state of decreased/lack of consciousness could be additional evidence of the preserved consciousness in those patients.

1.3.2 BCI: communication

Brain-computer-interfaces (BCIs) allows LIS patients to efficiently communicate through eye-movement, while for CLIS patients, who completely lost eye-movement as last muscle under volitional control, other sources to use as controllable neural signals are needed.
Kubler et al. in their study [22] investigated whether there is a relationship between physical impairment and BCI performance, trying to answer to what prevents CLIS patients from learning BCI-control paradigm. They demonstrated that CLIS patients still have intact cognitive processing analysing their response to an event-related brain potential test [22], hypothesizing that, differently from other states of physical impairment, people with longer time period in CLIS may have lost their voluntary cognitive activity, goal-directed thinking and imagery. Since BCI paradigms require the user to learn how to regulate the target EEG response by means of online feedback (e.g. auditory, visual), the failure of BCI with CLIS patients can be attributed to the loss of the perception of contingency between the required physiological behaviour and its consequences (represented by some kind of reward in this case).
They concluded leaving an open question about the possibility of transferring the pre learned BCI-control from LIS to CLIS.
Other peripheral autonomic psycohphysiological measures could be used for communication, like skin conductance response (SCR), heart rate (HR) and respiration [18]; anyway this is not possible in advanced ALS patients that are possibly artificially ventilated, may not have sufficient residual muscle activity to learn and control these physiopathological mechanisms, and could have had a degeneration of sympathetic nerve fibers too.

Jaymaran et al., after finding a global increase in γ-power range in ALS and a parallel decrease in lower frequency oscillation [14], suggested that a possible reason for the decrease in BCI performance was related to the increasing inefficiency of low-frequency bands for these patients. Indeed, generally BCI paradigms focus on low-frequency ranges of EEG signals, while the power increase in other oscillations can make them good candidates for BCIs usage.
For this reason understanding the electrophysiological implication of neurodegeneration in the transition from LIS to CLIS could open new possibilities for BCI communication, exploring new features and exploiting different characteristics of the signal.

1.3.3 Longitudinal studies

Currently in the literature there is a lack of longitudinal studies focusing on the analysis of the progression of the disease, and a parallel failure in the adoption of the same EEG biomarkers used to separate patients from controls in the separation of different ALS phenotypes. Anyway, this aspect is of crucial importance to monitor patients' state through the disease development, possibly assessing some changes or improvements related to pharmaceutical treatments or environmental conditions.
Murguialday et al. investigated the transition from LIS to CLIS from the physiological point of view [18] demonstrating that eye muscle is the very last muscle group under volitional control in the transition.
Moreover, an EEG biomarker able to quantify the degree of locked-in state, and thus to differentiate general LIS from CLIS, would help in monitoring also other neuropathies with the same implications for late-stage patients [18].
Hohmann et al., after finding evidence of a slowing down of alpha peak frequency in CLIS patients with respect to controls and LIS patients, suggested that this shift might arise gradually in the transition from LIS to CLIS [13], and so it could be therefore investigated in longitudinal studies monitoring the transition.
The same hypothesis is supported comparing two studies, the first investigating abnormal EEG patterns in LIS [11] and the second focused on CLIS patients, specifically related to ALS disease [23]. The LIS study documented a significant difference between α and θ in controls and patients, with a slight shift from 9.8 Hz for the individual α frequency (IAF) peak to 9.4 Hz, in controls and patients respectively, while the second study found a greater shift to lower frequency ranges in peaks for CLIS patients. This result in [23] was interpreted as a general EEG slowing, with the hypothesis that this slowing of neural activity could be progressive during the advancement of the disease.

Moreover the CLIS study added other considerations on the hypothesis of cognitive impairment for these patients on the base of γ reduction, that is usually associated with a reduction in cognitive performances, but being those results are not univocal among studies it is premature to make a conclusion on their base. The progressive loss of muscular activity and volitional control on movements could also be the cause, since usually high frequency oscillations are affected by muscular activity.
In support of this, Jayaram et al. hypothesized a non linear relationship between γ power and disease progression [14].
Nasseroleslami et al. examined longitudinal changes in EEG's spectral content, noticing significant progressive changes in connectivity measures [16]. The connectivity increase in their analysis was particularly evident in θ and γ-band coherences.

Chapter 2

Materials and Methods

2.1 Patients and visits

The following longitudinal analysis considers data acquired from three patients, further referred to as Patient 4, Patient 6 and Patient 11.

Patient 4: Female, 29 years old, CLIS, was diagnosed with juvenile ALS in December 2012. She was completely paralyzed within half a year after diagnosis and has been artificially ventilated since March 2013, fed through a percutaneous endoscopic gastrostomy tube since April 2013, and is in home care. She was able to communicate with the eye-tracking device from early 2013 to August 2014 but was unable to use the eye-tracking device after the loss of eye control in August 2014. After August 2014 family members were able to communicate with her by training her to move her eyes to the right to answer "yes" and to the left to answer "no" questions until December 2014. In January 2015 eye control was completely lost and she tried to answer "yes" by twitching the right corner of her mouth and that too varied considerably and parents lost reliable communication contact. From 2015 she is using EEG- and fNIRS-based BCI for communication. In mid-2017 the patient had a collapsed lung that damaged her condition more.

Patient 6: Male, 40 years old, CLIS, was diagnosed with bulbar ALS in 2009. He has been artificially ventilated and fed through a percutaneous endoscopic gastrostomy tube since September 2010 and is in home care. He lost speech and capability to move by 2010. After that, he was trained to move his chin for "yes". No communication with eye movements, other muscles, or assistive communication devices was possible since 2012.

Patient 11: Male, 34 years old, CLIS, was diagnosed with not-bulbar ALS in August 2015. He lost speech and capability to walk by the end of 2015. He has been fed through a percutaneous endoscopic gastrostomy tube and artificially ventilated since July 2016 and is in home care. He started used an eye-tracking-based communication system (ECTS) from August 2016 to August 2017, when he was not able to fixate his gaze anymore. After that, the family could find a way to communicate observing patient's eye movement until he lost definitively the capability to control his eyes in March 2019.

Tables 2.1 2.2 and 2.3 provide the details related to the acquisition of EEG resting state recordings for each patient.
EEG recordings were acquired during visits to patients, with a variable duration going from one to six days. In the Tables the field *Visit* refers to different visits, reported with their date in *Dates*, while *Sessions* refers to the number of EEG resting state recordings that were acquired on different days in the correspondent visit. In the following is sometimes referred to "sessions" as "days".

EEG resting state recordings for Patient 4 have been extracted from the first 10 minutes of three EEG sleep recordings, acquired in 2016, 2017 and 2019. Details on recordings and visits dates are reported in Table 2.1.
EEG resting state recordings for Patient 6 cover the period going from May 2017 to January 2019, details are reported in Table 2.2.
Available EEG resting state recordings for Patient 11 cover the period going from May 2018 to March 2019, details of the visits and acquisition setup are reported in Table 2.3.

Considering patients' information related to the onset of symptoms and the evolution of the disease is important since EEG resting state recordings and their frequency profiles are not consistent from patient to patient, and taking into account where the data analyzed are temporally located within the anamnesis of each patient can provide a further comprehension of the results.

What is going to be analysed is the evolution of EEG resting states through the progression of the disease, with the objective of investigating the existence of a biomarker that correlates EEG signal with the neural degeneration that occurs during the advancement of ALS syndrome.

Visit number	Date	Sessions	Acquisition channels
1	January 2016	1	CF3,CF4,C3,C4,CP3,CP4 EOGV,EOGH
2	March 2017	1	F1,F2,C1,C2,O1,O2 EOGL,EOGR
3	March 2019	1	Cz,C3,C4,FZ,F3,F4,AF3,AF4 EOGU,EOGD,EOGR,EOGL

TABLE 2.1: Details of visits for Patient 4. "Sessions" refers to the number of EEG resting state recordings for each visit.

Visit number	Date	Sessions	Acquisition channels
1	May 2017	6	FC5,FC6,C5,C6,Cz EOGU,EOGD,EOGR,EOGL
2	September 2017	5	FC2,FC4,FC1,FC3,Cz EOGU,EOGD,EOGR,EOGL (session 1)
			P7,P4,Cz,PZ,P3,P8 O1,O2,T8,C4,F4 (other sessions)
3	April 2018	3	F3,FC3,F4,FC4,Cz EOGU,EOGD,EOGR,EOGL
4	May 2018	4	F4,FC4,C2,Cz,C1,FZ,F3,FC3,P4,PZ,P3 EOGU,EOGR,EOGL
5	January 2019	3	AF3,F3,F5,FC3,FC5,C5,C3,T7,CP5,CP3,CP1 C1,Fz,FCz,F4,Cz,FC4,C4,C2,CP2,CP4,CPz EOGU,EOGR,EOGL

TABLE 2.2: Details of visits for Patient 6. "Sessions" refers to the number of EEG resting state recordings for each visit.

2.2 EEG acquisition

Resting states recordings are acquired from patients while they are lying in bed, with the instruction of staying relaxed trying not to think anything and refrain from sleeping. EEG electrodes were attached according to the 10-5 system (setup is shown in Figure 2.1) with reference and ground channels placed respectively to their right mastoid and to the forehead. EOG electrodes were placed as it is shown in Figure 2.2.
EEG and EOG signals were recorded using V-Amp amplifier and active electrodes (Brain

Visit number	Date	Sessions	Acquisition channels
1	May 2018	1	F4,FC4,F3,FC3,Cz,C1,C2
			EOGU,EOGD,EOGR,EOGL
2	August 2018	1	AF3,AF4,F3,F1,FZ,F2,F4,FC3,FC1,FC2,FC4,Cz
			EOGU,EOGD,EOGR,EOGL
3	September 2018	4	AF3,AF4,F3,F1,FZ,F2,F4,FC3,FC1,FC2,FC4,Cz
			EOGU,EOGD,EOGR,,EOGL
4	November 2018	3	F3,F1,F2,F4,FC3,FC1,FC2,FC4,Cz,C1,C2,C4
			EOGU,EOGD,EOGR,EOGL
5	December 2018	3	F3,F1,F2,F4,FC3,FC1,FC2,FC4,Cz,C1,C2
			EOGU,EOGD,EOGR,EOGL
6	January 2019	4	F3,F1,FZ,F2,F4,FC1,FC2,FC4,Cz,C1,C2,C4
			EOGU,EOGD,EOGR,EOGL
7	February 2019	2	F3,F4,FC3,FC4,Cz,C1,C2
			EOGU,EOGD,EOGR,EOGL
8	March 2019	1	F3,F4,FC3,FC1,FC2,FC4,Cz,C1,C2
			EOGU,EOGD,EOGR,EOGL

TABLE 2.3: Details of visits for Patient 11. "Sessions" refers to the number of EEG resting state recordings for each visit.

Products, Germany) with a sampling frequency rate of 500 Hz.
The length of the acquisitions is variable, almost 5 minutes for Patient 6 and 10 minutes for Patient 11. Patient 4's recordings have a duration of exactly 10 minutes since they are manually extracted from the first minutes of sleeping recordings.
The number and position of electrodes varies between visits due to different conditions and patients' clinical needs.

2.3 Data preprocessing

All data analysis was performed using Matlab 2018b, [24] and in particular EEG recordings were stored and processed with EEGLAB Matlab toolbox [25].
The preprocessing pipeline for EEG signals consisted in:

- Filtering
- Resampling

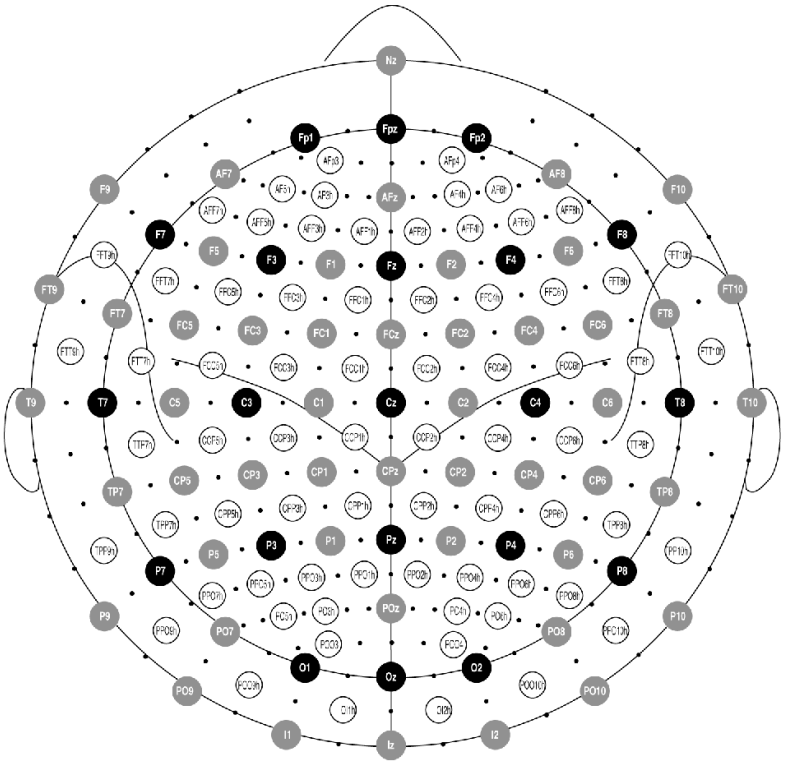


FIGURE 2.1: EEG electrodes' configuration in 10-5 system.

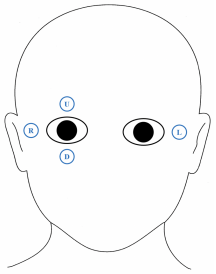


FIGURE 2.2: EOG electrodes' placement. Simbols R, L, U, D indicate positions of EOG electrodes EOGR, EOGL, EOGU, EOGD respectively.

- Independent component analysis
- Component removal
- Channel(s) interpolation

First of all data were filtered through a windowed sinc passband filter, with bandpass set to [0.5 45] Hz, through the function `pop firws` provided by EEGLAB. That function realizes a symmetric FIR filter with Hamming window type, and the order can be automatically set by defining the transition band. For this particular application to EEG data sampled at 500 Hz, specifying a transition band of 1 Hz, the order used was 1650. After that, the signals were downsampled to 128 Hz using EEGLAB function `pop resample`, to save computational cost on the next preprocessing steps.

Moreover, information about channel localization on the scalp are added to EEGLAB dataset through the functions `pop chanedit` and `readlocs`, taking the information about electrode localization in the standard 10-5, corresponding to the one adopted for the acquisition. The file containing channels' spatial coordinates is provided by EEGLAB toolbox [26].

2.3.1 Independent Component Analysis

Independent component analysis (ICA) is a powerful mathematical tool for separating a signal into its independent sources.
EEG multichannel recording allows to obtain a set of highly correlated signals, due to the distance between the scalp (where electrodes are applied) and the sources of neural activities, and the different resistivities of tissues that brain's electric potentials crosses before they are acquired; moreover EEG signal does not have a high spatial resolution, and is easily corrupted by artifacts (as muscular activity, eye movements).
The ICA algorithm is able to identify independent sources from EEG signal employing statistical tools, without using information about the physical location or configuration of source generators. In its mathematical formulation, the aim of ICA is to find a matrix W and a vector w so that the elements $u = [u_1...u_N]^T$ of the linear transform $u = Wx+w$ of the random vector $x = [x_1...x_N]^T$ are statistically independent.
Applied to EEG analysis, the rows of the input matrix x are the EEG signals recorded at different electrodes, the rows of the output data matrix u are the time courses of activation of the ICA components, and the columns of the inverse matrix W give the projection strengths of the respective components onto the scalp sensors [27]. The scalp topographies of the components provide information about the location of the sources.
ICA technique is particularly suited for domains where sources are independent, the propagation delays are negligible, the sources are analog and with a probability density function similar to the gradient of a logistic sigmoid (plausible for EEG signals) and the number of independent signals sources is equal to the number of sensors [28]. These requirements are satisfied by the characteristics of EEG signals, indeed:

1. EEG dynamics can be modeled as a collection of statistically independent brain processes

2. Volume conduction in the brain is instantaneous, ensuring no delays in propagation

The assumption regarding the equality between the number of sources and the number of sensors is more uncertain, since the exact number of independent sources of brain

activity is unknown; however, with the objective of artifact rejection, it is possible to consider just the first components as the most independent signals, that can be easily related to sources that are external with respect to brain activity, which is more correlated.

In this particular application, ICA is used to isolate and remove eye movement's components from EEG channels, since it is admissible to neglect other sources of muscular activity because of CLIS condition. To this purpose, ICA algorithm is applied on all EEG and EOG channels.
Removing the eye movement's components from EEG data is of crucial importance for the analysis of those patients' recordings, since they can be characterized by a different level in the ability to control eye muscles over the progression of the disease. Without removing this "noisy" component there is the risk of interpreting a decrease in low-frequency range power as a signature of the progression, while it could be caused simply by a decrease in eye movements, that affect the same frequency range.
The application of this preprocessing technique, as was documented in multiple studies [16], allows ensuring that results are not due to artifact components.

The ICA algorithm is applied using `runica` function provided by EEGLAB. After that, the components are selected through the analysis of components' power spectral densities, selecting the ones with the highest and most deviating power at low frequencies, together with the ones corresponding to unexpected high power peaks in high frequencies. Both of these spectral features can be related to slow and fast eye movements. Selecting components to remove from data is still done manually, but is less demanding than selecting and rejecting continuous data by eye.

2.3.2 Channels interpolation

It is clear from Tables 2.1 2.2 and 2.3 the EEG electrodes placed on the scalp for resting state recordings acquisition are different between patients and within different visits to the same patient.
EEG channels do not have a high spatial resolution and reflect the electrical activity of a huge and spread population of neurons, and thus doesn't reveal exactly the activity corresponding to the brain region where the electrode is located. In addition to that, volume conduction effects are predominant. Anyway, to compare different patients is necessary to use a common channel to have a more systematic analysis. Since the electrode *Cz* is present in all recordings for patients 6 and 11 it was chosen for spectral analysis. For patient 4 Cz electrode is missing in the first two visits (see Table 2.1) and so it was obtained through the interpolation of neighbor channels.

Channel interpolation was performed through `eeg interp` function of EEGLAB toolbox [25], that recalls Matlab `interp` function. This tool exploits the spread nature of EEG and the known standard electrode location to obtain the activity of a channel when is missing or badly recorded. Channel interpolation needs a high density of electrodes over the scalp, but in this specific case Cz interpolation is admitted since the signals from neighbors channels are acquired.
In Appendix A a comparison between real Cz activity and the one reconstructed by interpolation is shown for one record as an example.

For Patient 11, since he is the one with the more stable set of electrodes within visits, a fixed set of channels was obtained by interpolation. The final set of channels obtained for all recordings is:

$$channels = [F2, F4, FC2, FC4, C4, C2, Cz, C1, FC1, F3, F1, Fz]$$

which is equal to the available channels for visit 6 (see Table 2.3).

2.3.3 Visit wise data management

Since the aim of this longitudinal-single patient analysis is to observe the evolution of EEG characteristics through time, the objective is to compare data visit wise, referring to the visits mentioned in Table 2.3. As reported in the column *sessions* of the same table, the number of EEG resting state recordings acquired in different visits is not the same for all the visits. To consider and measure all the visits equally it was decided to combine day wise resting state recordings belonging to the same visit, also to reduce the influence of outlier data and to improve the overall significance of the features extracted.

Single sessions combination for days of the same visit is done differently for time and frequency domains.
Visit wise - time domain features are obtained averaging the values of the features obtained from day wise analysis in time domain.
Before frequency domain processing, EEG data is normalized channel by channel across sessions related to the same visit. The normalization is performed through the `zscore` function provided by Matlab in this way:

$$data_{norm} = zscore(data, 0, 3)$$

where *data* is the *samples* × *channels* × *sessions* referring to resting states recordings of the same visit and the third parameter indicates the dimension on which to apply the normalization, which is the one corresponding to the sessions. The output matrix $data_{norm}$ has the same dimensions of *data* but, for each sample, values were z-scored across sessions of the same visit (subtracted by mean across trials for each channel and divided by the standard deviation across sessions for each channel). This procedure was applied to remove the offset deviation and noise in EEG samples related to setting conditions at the acquisition time, assuming that EEG recordings are approximatively homogeneous between days of the same visit.
After that, visit wise features are extracted from visit wise power spectral density (PSD), obtained by averaging the PSDs computed on correspondent electrodes from resting state recordings acquired in sessions of the same visit, obtaining a single PSD representative of frequency domain behavior for each visit. The resulting power spectral densities are dimensionless due to the normalization applied.
Z scoring EEG data across sessions of the same visit is applied before power spectral density computation, to reduce the standard deviation of day wise PSDs grouped by visits, details on the validation of this choice are reported in Section 3.1.3.

2.4 Time domain analysis

Time domain analysis and feature extraction were performed from day wise EEG recordings, preprocessed as described in Section 2.3.
The features ectracted from EEG signal in time domain are two complexity metrics, since it is known from the literature [29] that brain complexity is able to distinguish healthy subjects from patients in many neuropathies and disorders of consciousness [19].

2.4.1 Features

- *Kolmogorov complexity index*
- *Permutation entropy*

These two features are described in the successive paragraphs.

Kolmogorov complexity index

Measuring EEG's compressibility has been found to be an effective way to assess its complexity and redundancy, and for this reason has been related to consciousness state

in many studies [19].

Kolmogorov index gives a measure of the complexity of a given binary string, computed as the number of bits of the shortest computer program that can generate this string [30]. The choice of the program language to compress the string is not unique and here the adopted compression algorithm corresponds to Lempel-Ziv compression technique [31], where they employ the concept of encoding future segments of the source-output via maximum-length copying from a dictionary containing the recent past output. The idea of this compression algorithm is the following: as the input data is being processed, a dictionary keeps a correspondence between the longest encountered words and a list of code values. The words are replaced by their corresponding codes and so the input file is compressed. Therefore, the efficiency of the algorithm increases as the number of long, repetitive words in the input data increases.
In this study Kolmogorov complexity is computed through Matlab function `kolmogorov` following the algorithm described in [30]. It receives as input a sequence of binary simbols $S = (S_1, S_2, ..., S_n)$ obtained from the original EEG sequence $X = (X_1, X_2, ..., X_n)$ as:

$$S_i = 1 \; if \; X_i > A$$

$$S_i = 0 \; if \; X_i < A$$

where $A = (X_1 + X_2 + ... + X_n)/n$. The conversion of data into a digital stream is taken from [32], where this complexity index is applied to EEG for assessment of mental fatigue.

Permutation entropy

Permutation entropy (PE) is an effective method to compare time series and distinguish different types of behaviors (e.g. periodic, chaotic or random); for an EEG epoch it's a measure of its distance from white noise [33]. One of its advantages is the one of being robust to low signal to noise ratios and of not requiring strong stationary assumptions. [29].
Its definition starts from the one of an ordinal pattern of an m-tuple of real numbers $(x_1, x_2, ..., x_m)$, that describes how its elements relate to one another in terms of position and value. The parameter m is called the *order* of the pattern, and for each tuple of length m exists a set of $m!$ ordinal patterns. Figure 2.3 shows an example of the 6 ordinal patterns for a tuple of length 3 [33].

The EEG signal is transformed into a sequence of symbols before estimating entropy, considering consecutive sub-vectors of the signal of size m. These sub-vectors can be

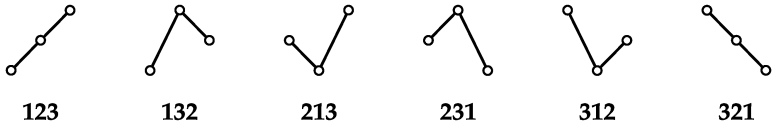


FIGURE 2.3: Example of an ordinal pattern of length 3

made of either consecutive elements or of elements separated by τ samples (where τ is an integer). Each sub-vector of length m is associated with a unique symbol, based solely on the ordering of its m signal amplitudes. After the symbolic transform, the probability of each symbol is estimated, and PE is computed by applying Shannon's classical formula to the probability distribution of the symbols:

$$PE = -\frac{1}{log(m!)} \sum_{i=1}^{m!} p_i log(p_i)$$

where p_i is the probability of one of the $m!$ symbols, and $log(m!)$ is a normalization factor.

In this analysis PE is applied through Matlab function `pec` from [34] where m is set equal to 3 and τ equal to 1, as it is one of the most common adoption for these parameters [19] to avoid aliasing and an excessive computation complexity (resulting from increasing m and τ).

2.5 Frequency domain analysis

Frequency domain analysis implies the extraction of features from EEG's power spectral density (PSD).The power spectrum is computed separately for each electrode of each EEG record using Matlab's `pwelch` function, which returns the power spectral density estimate of the input signal using Welch's overlapped segment averaging estimator. Window length has been set to 5 seconds while the number of overlapping points corresponds to 2 seconds.

The classic PSD computation through periodogram requires the spectral content of the signal to be stationary over the time period considered. Because it is not the case for EEG signal, averaging the periodograms obtained over short segments of the windows, Welch's method allows to drastically reduce this variance.

A comparison of different choices for `pwelch` parameters is shown in Appendix A.1. Since EEG signal is not stationary is important to choose a window length that is long enough, otherwise important components of the spectral activity would be neglected.

After power spectral density is computed separately for each channel, each PSD is normalized by its median to reduce the effect of different offsets between the recordings, related to subject condition during the specific acquisition (for example a change in impedance due to skin humidity) or to the equipment [35]. Normalization can lead to the loss of amplitude information of the signal, but is acceptable considering relative features.
Power spectral densities are visualized in logarithmic scale to overcome power law scaling, for which power amplitude of frequency bins generally decreases with increasing frequency [35]. Logarithmic scaling allows to obtain a more interpretable power spectrum and simplifies quantitative comparisons of power across frequency bands.

2.5.1 Features

After single-channel PSD (Px) computation the following features are extracted:

- *Absolute band power*: integral of the power spectral density (Px) within each frequency band, summing for each band the frequency bins corresponding to the defined bands.
- *Relative band power*: percentage of the total power of the signal represented by specified frequency band. It is computed dividing the power in each frequency band by the total power (sum of Px over all the considered frequencies).
- *Spectral range*: represents the difference between the maximum and the minimum values of Px, considering Px over a specific frequency range.
- *Spectral entropy*: as entropy measures in time domain, spectral entropy (SE) is a measure of the uncertainty and organization of the signal in the frequency domain [36]. The SE treats the signal's normalized power distribution in the frequency domain (Px_{norm}) as a probability distribution, and calculates the Shannon entropy of it. It is computed as:

  $$SE = -\frac{1}{logN}\sum_{f} Px_{norm}[f]log(Px_{norm}[f])$$

  where f varies over the frequencies in the specified bands, and N is the length of the considered segment of Px.
- *Spectral flatness (Wiener Entropy)*: is a measure of the width and uniformity of the power spectrum. It is expressed on a scale of 0-1, where white noise has an

entropy value of 1. It is computed as:

$$WE = \frac{exp(1/N \sum_f logPx[f])}{\frac{1}{N} \sum_f logPx[f]}$$

2.6 Statistical analysis

The objectives of statistical tests for a longitudinal analysis of features' values are to:

- assess if features are significantly different between different visits
- investigate the possible presence of a trending behavior (monotonously increasing or decreasing) for features that are significantly changing during time

Having a small number of visits and sessions, statistically assess the presence of a trend in the change of features' values is difficult and not completely reliable. Specially for Patient 4 (see 2.1), for which data are available for just three visits, the analysis can only be limited to a visual comparison of power spectral densities.
Anyway, since the objective of this work is to propose a pipeline to use for longitudinal comparison of features, a statistical approach to verify the previous points is tested.

One way analysis of variance

First of all, to check if a significant change over features' values is revealed by the processing of EEG signal, one way analysis of variance (ANOVA) is applied to all features obtained from day wise analysis. Using metrics obtained from day wise analysis allows to consider also the variability of the considered indexes whithin the same visit.
This test is performed through Matlab function `anova1` as `anova1(y,group)` to test the equality of group means, specified in *group*, for the data in vector or matrix y, using the grouping of days in visits as groups. The aim is to check if features were significantly different when grouped visit wise from day wise analysis.
ANOVA tests the hypothesis that all group means are equal versus the alternative hypothesis that at least one group is different from the others. A low p-value resulting from this test means a rejection of the null hypothesis.

Mann Kendall trend test

Then, Mann Kendall trend test is applied to investigate the existence of trends over time for feature values. This test can be used for as few as four samples, but with only a few

data points the test has a high probability of not finding a trend. The more data points you have the more likely the test is going to find a true trend. The minimum number of recommended measurements is therefore at least 8 to 10.
For this reason trend analysis is applied to Patient 6 and Patient 11, considering however the low reliability of its results given by the limited number of samples.
Mann Kendall test [37] is a non-parametric form of monotonic trend regression analysis, that analyzes the sign of the difference between subsequent values. Each later-measured value is compared to all previous values, resulting in a total of $n(n-1)/2$ possible pairs where n is the number observations. This test admits missing values and doesn't require data to fit any particular distribution.
The test statistic S is computed as:

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} sign(y_j - y_i)$$

where y_i and y_j are subsequent measures [38].
The key assumption is that, uder the null hypotesis (absence of trend), S is approximately normally distributed with expectation $n(n-1)/4$ and variance $(2n^3+3n^2-5n)/72$ [39].
The test statistic τ is computed as:

$$\tau = \frac{S}{n(n-1)/2}$$

assuming values in the range $[-1+1]$ and is analogous to the correlation coefficient in regression analysis, meaning that for $\tau > 0$ the trend is increasing and for $\tau < 0$ the trend is decreasing.
The null hypotesis is rejected for values of S and τ significantly different from zero. When a significant tren is found (p-value < 0.05), the rate of change is computed using the Sen slope estimator as [38]:

$$\beta = median(\frac{y_j - y_i}{x_j - x_i})$$

for all $i < j, i = 1, 2, ..., n-1$ and $j = 2, 3..., n$, where x_i and x_j refers to the time labels of data. In this way the overall trend slope is computed as the median of all pairs of data used for S computation.
In this specific case time labels are obtained scaling the intervals between EEG recordings. The obtained slope is used to plot features' values' regression in Section 3.

Simple linear regression was applied to Patient 4's features values, since the number of sessions and visits was too low to perform a statistical test.
The goodness of the regression is assessed through the coefficient of determination, or

R^2. This coefficients corresponds to

$$R^2 = 1 - \frac{SS_{resid}}{SS_{total}}$$

where SS_{resid} is the sum of the squared residuals of the regression and SS_{total} is the sum of the squared differences from the mean of the dependent variable (total sum of squares). This statistic indicates how closely values you obtain from fitting a model match the dependent variable the model is intended to predict, and thus gives a measure of the percentage of variance of the data that the model is able to predict. For Patient 4's features analysis most significant features were identified as the ones resulting in a R^2 value higher than 0.9. The correspondent regression is shown in Section 3.

Chapter 3

Results

Since the objectives of this work include to propose and validate a possible method to perform efficiently a single patient-longitudinal data analysis, in the following sections the results are presented step by step, following the principal stages of the pipeline executed.

3.1 Preprocessing

3.1.1 Independent Component Analysis and component rejection

After the application of ICA to EEG data, the signal has an equivalent representation in a number of independent sources that is equal to the number of channels. ICA is performed on all channels including EOG channels, to capture artifacts due to eye movements from their activity. In this preprocessing step the objective is to remove eye movements' influence from EEG channels as much as possible.
In the following are shown the successive steps towards component rejection:

- visualization of components' spectrum and localization
- identification of the components mostly affected by eye moments
- reprojection of the signal into EEG channels' space to check the effect of those components' removal
- definitive removal of the components from the EEG record

All these steps are performed through EEGLAB toolbox and, in particular , with the pop-up tools for components visualization and removal. This toolbox offers the possibility of selecting which is the frequency of highest interest for component analysis, to

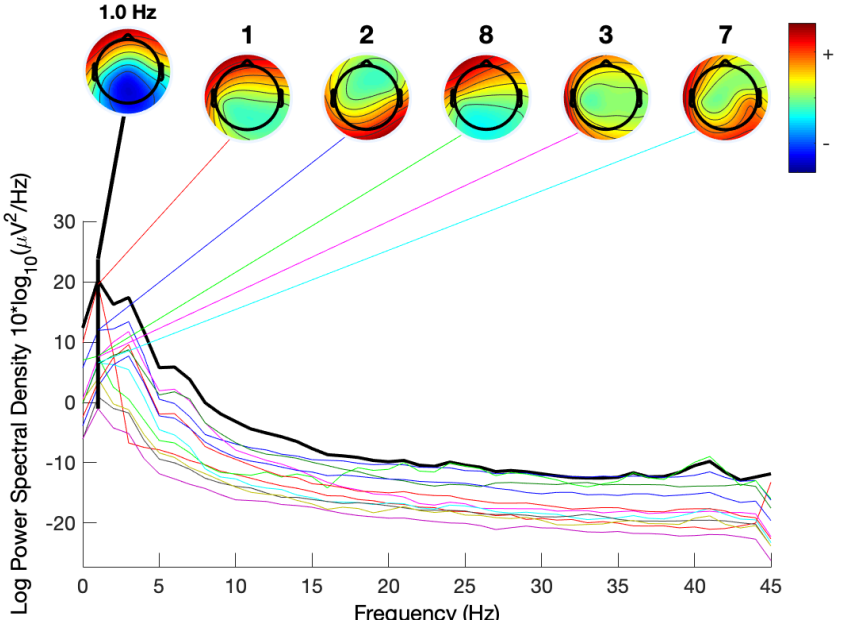
visualize in the plot of the components spectrum which components are affecting the most that particular frequency. Then, once the components to be removed have been selected, the software gives the option of reprojecting the data into the original channels space to visualize the effect of component removal on EEG channels. In this way is possible to check not to remove low frequency components that are related to brain activity and not to EOG artifacts. This issue is of particular importance in the analysis of those patients' EEG since, as shown by many studies mentioned in Section 1.3.1, they present a general slowing in the EEG frequency activity implying that the components related to their brain activity and the ones related to eye movements possibly overlap. Two examples of this procedure applied to two recordings are presented:

1. Patient 6, visit 1 day 2; it shows the rejection of EOG channels.
2. Patient 11, visit 3 day 2; it shows how ICA can be used to remove a component related to a specific noisy channel through source localization.

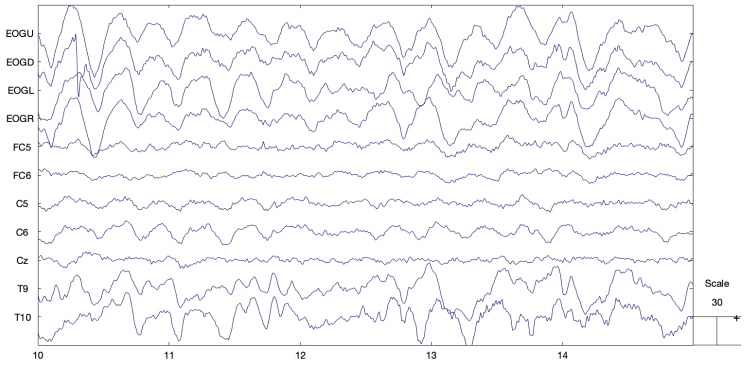
In Figures 3.1 and 3.2 ICA processing is shown in its different steps. With the objective of eye movement's artifacts rejection from EEG data, the frequency range of interest for components' spectral analysis is mainly the one corresponding to δ range, (low frequencies); indeed, EOG artifacts usually manifest their influence in EEG channels through the spread of low oscillations over the scalp, especially in the frontal electrodes, closer to eyes position. Artifacts due to eye movements, together with other unspecified sources of artifacts, can affect also high frequencies. For this reason in the spectral analysis of independent components is necessary to pay attention to both these features to efficiently apply artifact rejection.

An example showing the effect of EOG's component removal is shown in Figure 3.1. In the upper part of Figure 3.1a are displayed the 5 components affecting more the frequency of interest, with their location on the scalp. It is clear how with a first observation no components show an anomalous behavior at high frequencies. The components are shown from the most to the least independent source. Based on a visual analysis the components elected for the removal could be 1 2 3 7 and 8. Figure 3.1b shows the original EEG data scroll and Figure 3.1c shows a projection of the remaining components after rejection on original data's space (in red) over the original data (in blue), useful to check if the removal of the chosen components is admissible. In this particular example is clear how the rejection of the previously mentioned components is effective to reduce EOG artifacts. Indeed, from Figure 3.1c, is evident how the channels most affected by the rejection are EOG channels, while the EEG channels are intact excluding the attenuation due to artifact rejection.

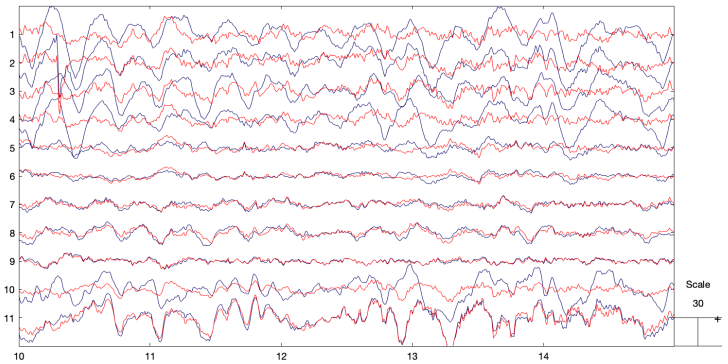
Figure 3.2 is another example of ICA for general artifact rejection. As in the previous example, in 3.2a are shown the five components related to 1 Hz activity reported with their location on the scalp. In this example component 5 is highly related to a precise location, correspondent to channel F1 (see Figure 2.1), which is suspicious considering that EEG activity is highly correlated between neighbors electrodes. This is one of the methods to spot noisy channels. From Figure 3.2b, showing EEG data scroll, emerges how channel F1 is actually a noisy channel.
Finally Figure 3.2c, that consists in a preview of the results of component rejection through the reprojection of the "new" EEG data (in red) over the original data (in blue), shows how the rejection of component 5 is effective in the correction of channel F1, eliminating the noise component from its activity.



(A) *"component spectra and maps"* plot obtained from EEGLAB popup windows.

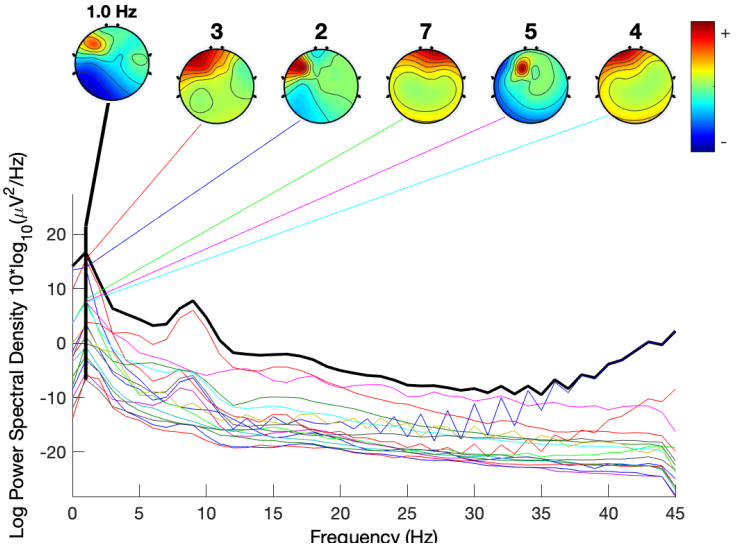


(B) EEG data before EOG's components rejection.

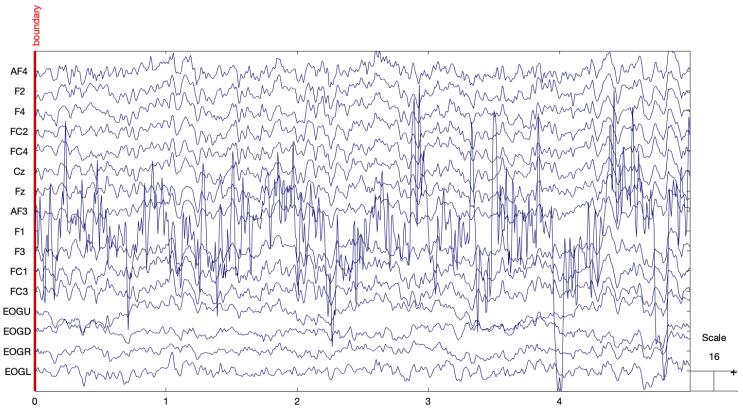


(C) EEG data after EOG's components rejection.

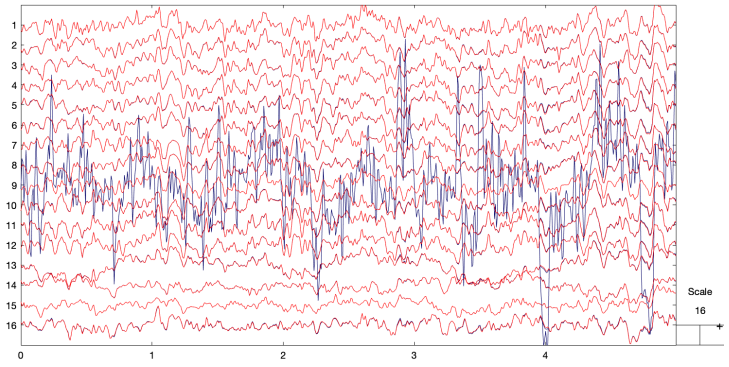
FIGURE 3.1: Example of ICA processing applied to Patient's 6 data from visit 1 day 2.



(A) *"component spectra and maps"* plot obtained from EEGLAB popup windows.



(B) EEG data before component 5 rejection.



(C) EEG data after component 5 rejection.

FIGURE 3.2: Example of ICA processing applied to Patient's 11 data from visit 3 day 2.

3.1.2 Channel Interpolation

In this section the results of channels interpolation are presented. As mentioned in Section 2.3.2 for Patients 4 and 11 channel interpolation technique was adopted to obtain a homogeneous set of EEG channels available for the processing.
For Patient 4 the analysis was performed just on channel Cz, which was obtained by interpolation for both visit 1 and 2 (see Table 3.5). In this case interpolation of Cz was admissible because in both sessions the other central electrodes were acquired, representing reliable neighbors sources for interpolation.
For Patient 6 too, just Cz channel was used for the analysis due to the relevant difference between channels sets between the visits. In this case Cz was already available for all the sessions, so interpolation was not carried on.
For Patient 11, since channels set were more uniform within different sessions, it was decided to interpolate the electrodes missing to each recording to obtain a fixed set of channels : [$F2$ $F4$ $FC2$ $FC4$ Fz $F1$ $F3$ $FC1$ $C4$ $C2$ $C1$ Cz].
To check the reliability of this method, a comparison between the "original" Cz channel and the reconstructed version through interpolation was made, using Patient's 11 data from visit 7 day 1. The results of the comparison are shown in Figure 3.3, giving the evidence of how interpolation is able to reproduce the activity of missing channels exploiting the activity of the neighbors, generating a signal with comparable amplitude and time course.
Figures 3.4a 3.4b and 3.4c present the final channel set obtained for Patient 11, showing the results of two, four and six channels interpolation respectively.

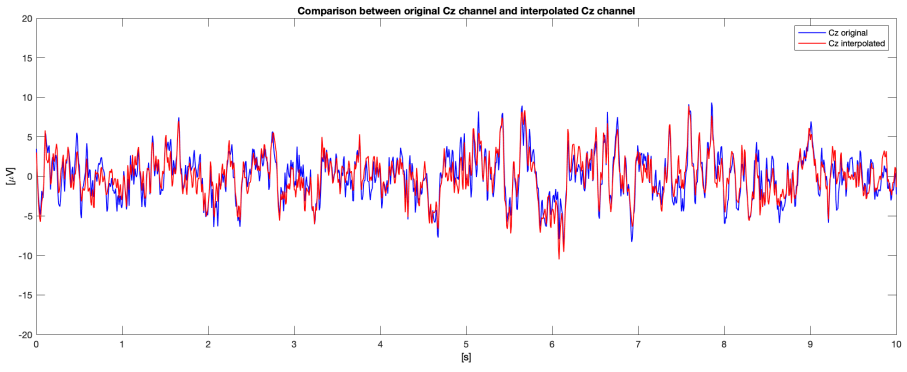
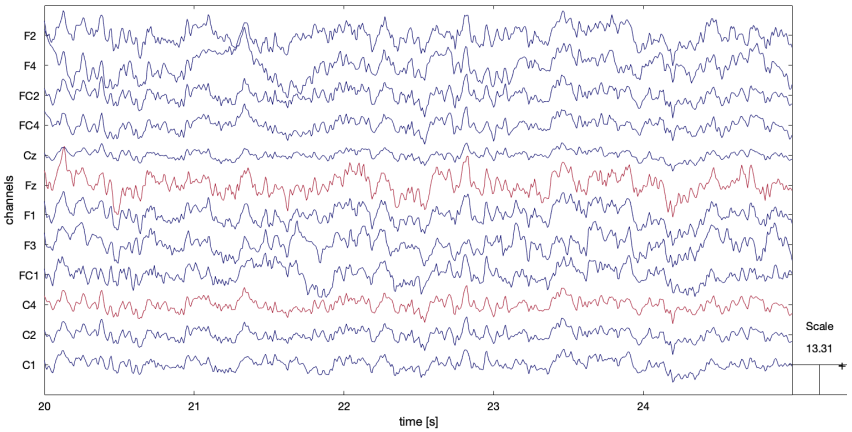
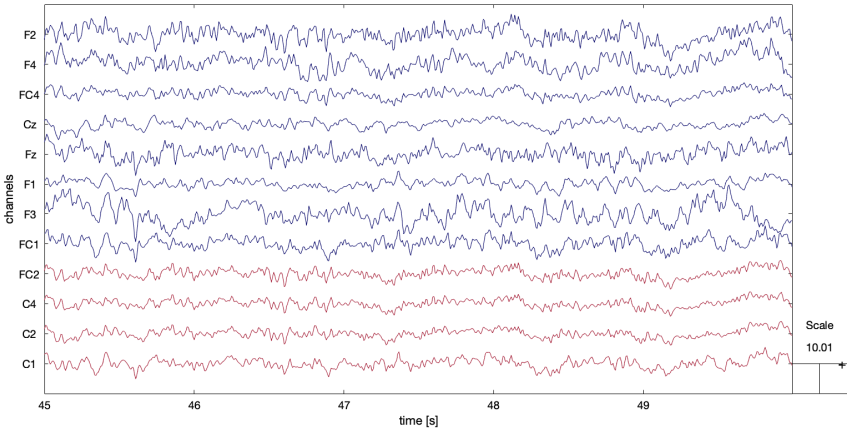


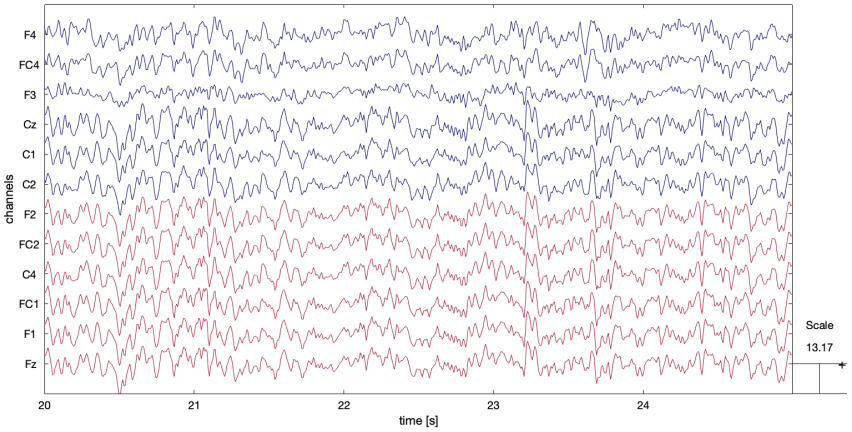
FIGURE 3.3: Comparison between original Cz channel and interpolated Cz channel.



(A) EEG channel set after channels interpolation - data from Patient 11 Visit 5 day 2.



(B) EEG channel set after channels interpolation - data from Patient 11 Visit 3 day 3.



(C) EEG channel set after channels interpolation - data from Patient 11 Visit 1 day 1.

3.1.3 Visit wise power spectral densities

Concerning the time domain analysis, the obtainment of a single description for data of the same visit is done after data processing at the features level, taking the median of the obtained features values in days of the same session.

For the analysis in frequency domain, after data preprocessing applied singularly to channels and recordings, the longitudinal analysis implies to group data of different sessions of the same visit into a singular representation, to allow a comparison across visits. The first step towards this objective is the application of `zscore` function to EEG data grouped by visits, that standardize data belonging to the same visit by the subtraction of the overall mean and division for their standard deviation (applied to correspondent-single channels across sessions of the same visit).
After that normalization, a representation of the spectral activity of each visit is obtained taking the mean of single sessions' power spectral densities, previously normalized for their respective median over the frequency range of interest ([0.5 45] Hz). To have a measure of how much this derived PSD is reliable in the representation of the single sessions, and to check whether those PSDs referring to EEG acquisition of subsequent days are similar, ANOVA test was applied to day-wise PSDs grouped by visits, having as null hypothesis the belonging of all the PSDs to the same distribution.
Figures 3.5, 3.6 and 3.7 show day-wise PSDs and visit-wise PSD superposed. In each subfigure is added the p-value resulting from the ANOVA statistics; a p-value higher than the threshold (0.05) state for the acceptance of the null hypothesis and so indicates a more reliable representation of day-wise EEG frequency behavior through the comprehensive visit-wise PSD.

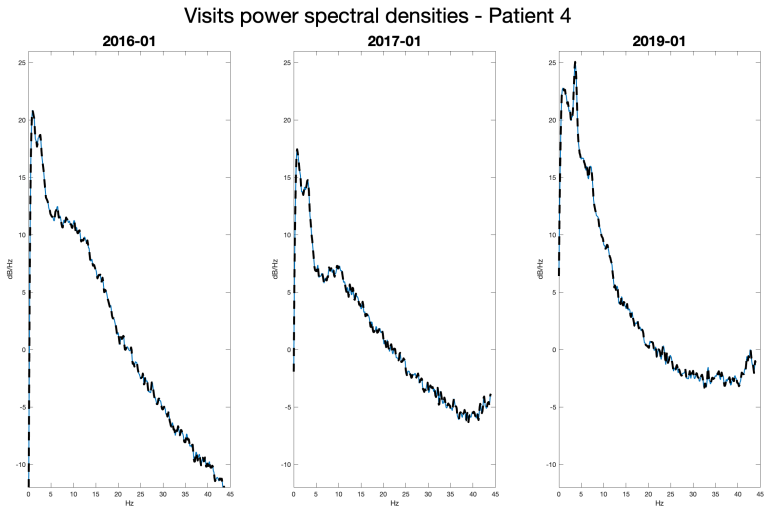


FIGURE 3.5: Visits' PSDs for Patient 4.
Concerning Patient 4's, since the available number of EEG recordings is equal to the number of visits, no sessions' grouping within visits is necessary.

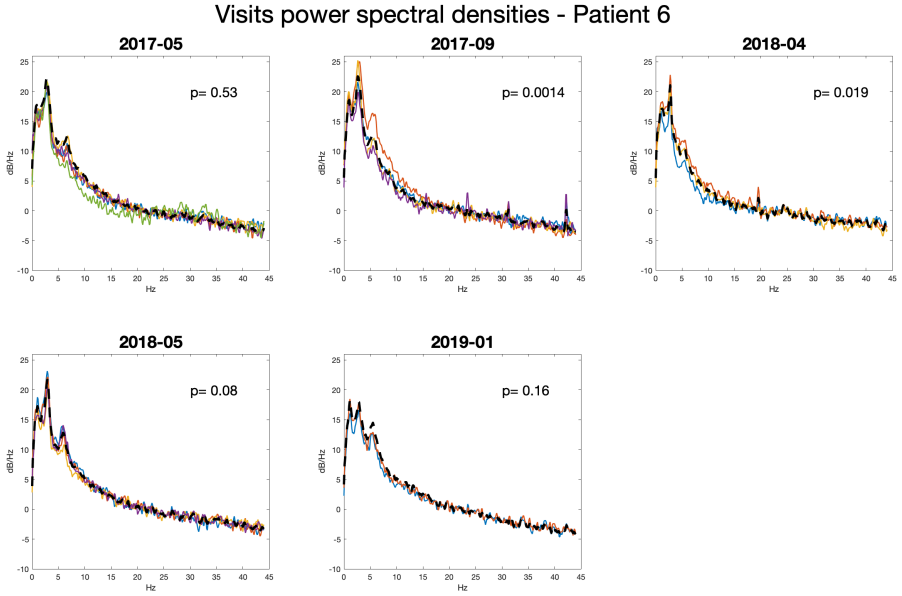


FIGURE 3.6: Visits' PSDs for Patient 6, in comparison with single session PSD after visit normalization.
Patient 6's results of ANOVA test reveals a positive outcome for the acceptance of the null hypothesis for visits 1 4 and 5(p-values higher than 0.05), while for visits 2 and 3 day-wise PSDs results having a less uniform spectral behavior across within the respective visit.

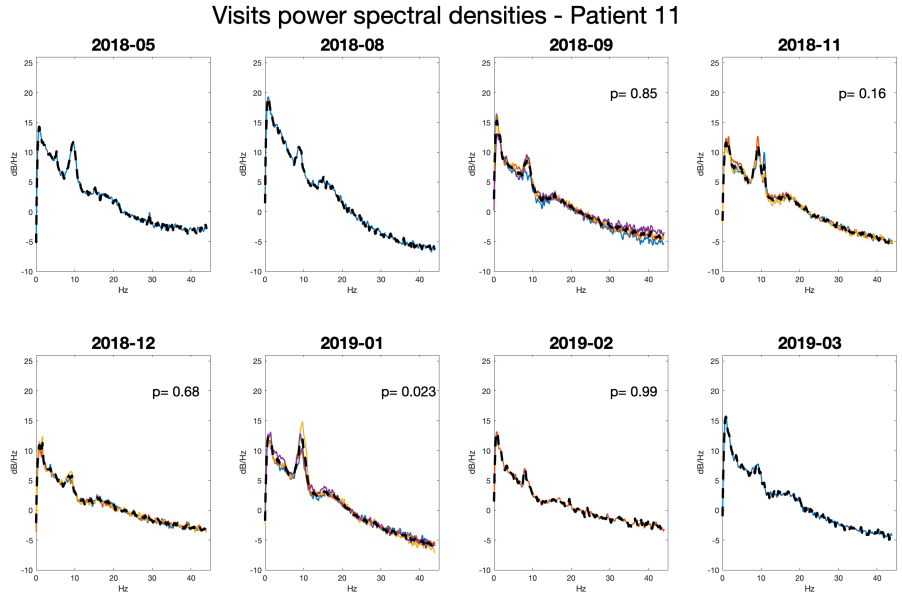


FIGURE 3.7: Visits' PSDs for Patient 11, in comparison with single session PSD after visit normalization.

Patient 11's has a better response to the visit grouping of day-wise PSDs, showing ANOVA's p-values lower than 0.05 just for the sixth visit. However, it has to be considered that for three visits a single EEG acquisition is available (see Table 2.3), while an optimal condition for the analysis would be the one of having a number of sessions equal for all the visits.

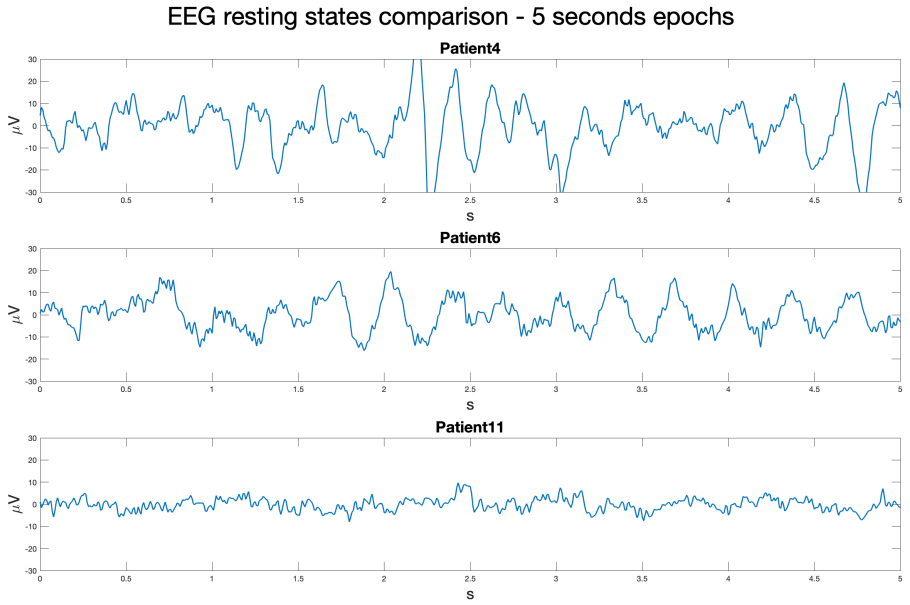


FIGURE 3.8: Comparison of EEG time course in Patients 4, 6 and 11. The figure shows three samples of 5 seconds EEG epochs in Patients 4, 6 and 11 respectively, giving the evidence of the slowing of alpha rhythm occurring in CLIS patients.

3.2 EEG overview in time domain

A starting point in the evaluation of the results is the observation of EEG time course in the three patients under analysis. In Figure 3.8 the comparison of samples of 5 seconds epochs for each patient is presented, to add evidence to the results obtained for power spectral densities analysis in frequency domain.

It is clear from Figure 3.8 that stable CLIS patients (Patient 4 and Patient 6) present an EEG time course with slower oscillations and higher amplitude with respect to the patient analysed during the transition (Patient 11), which EEG is comparable to the one of a healthy patient. The waveforms in the first two subfigures in Figure 3.8 are different from ones of the usual low frequency oscillations, showing a regular sinusoidal behavior resembling alpha activity rather than theta activity. This result is in accordance with the slowing of the alpha rhythm occurring in CLIS patients found in many studies in the literature [13] [12].

3.3 EEG's power spectral density evolution

In this section an overview of the evolution of EEG's power spectral densities is presented.
Data can be qualitative and quantitative compared just within the same patient.

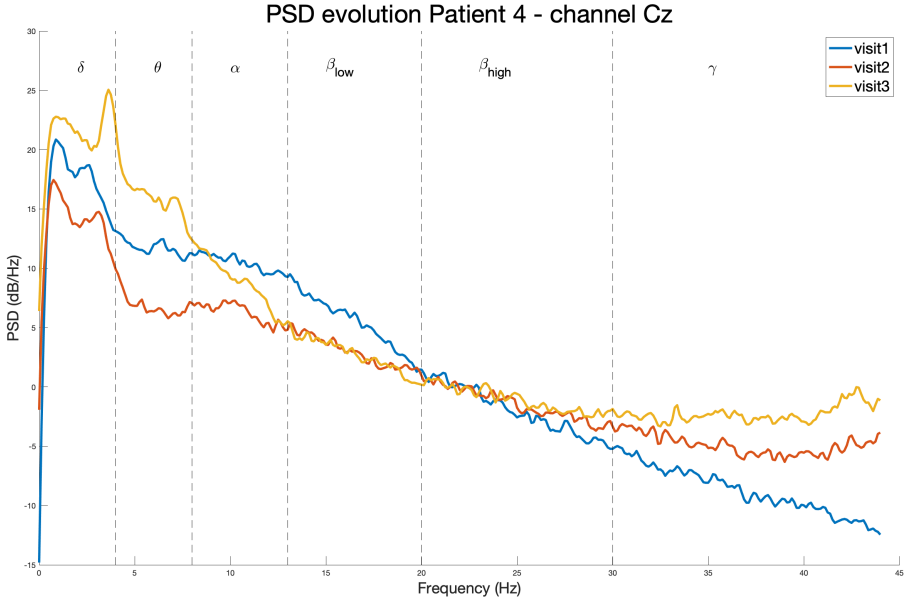


FIGURE 3.9: PSDs plots over subsequent visits for Patient 4.

Patient 4's data are the most spread in time, since they cover a period of three years (from 2016 to 2019, see Table 2.1). However, it has to be considered that just a single EEG resting state recording is available for each visit, and these recordings are extracted from the first 10 minutes of EEG sleeping recordings. It is though admissible to consider these data as "resting states", but their acquisition is different from the standard resting state one, since EEG electrodes used for sleep recordings have a different impedance with respect to the usual experimental ones. That's the main reason why comparing directly data with the ones of other patients, acquired with different settings, would be imprecise.
From Figure 3.9 is evident how the frequency components of EEG signal change between recordings belonging to different years, showing in particular a decrease in spectral power in the range [12 22] Hz (β frequency band) and an opposite increase for frequencies higher than 22 Hz(γ frequency band).
The quantitative analysis through the extraction of features in time and frequency domain is discussed in Section 3.4.1.

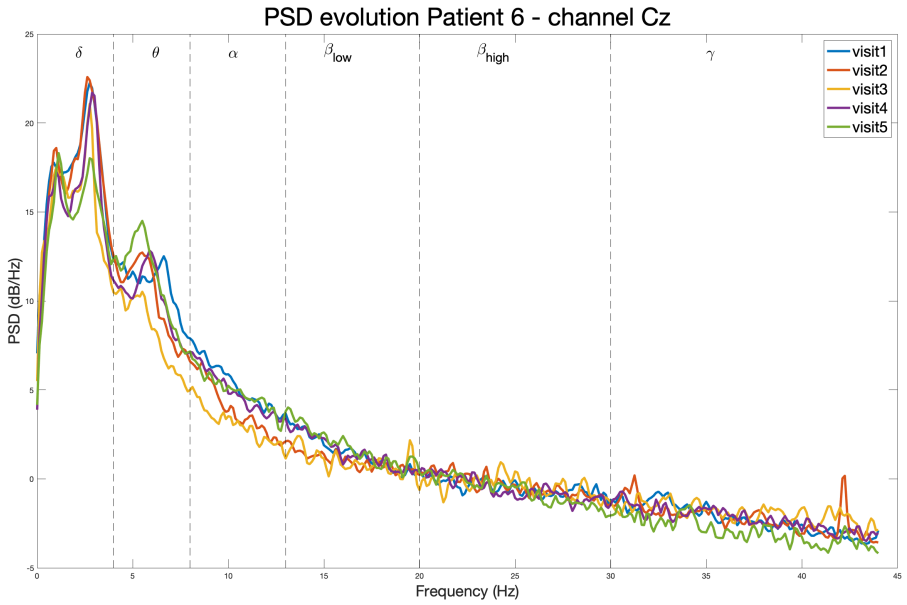


FIGURE 3.10: PSDs plots over subsequent visits for Patient 6.

Patient 6's PSD evolution across the visits is shown in Figure 3.10.
For this patient the analysis has been performed on data covering a period of almost 2 years (20 months), and their EEG's frequency content exhibit a lower variability with respect to the one of Patient 4. From a qualitative analysis it is clear how the frequency peaks in PSDs are almost stable with respect to their frequency localization, and also spectral power's amplitude doesn't show evident variations in subsequent visits. These observations, together with Patient 6's condition (CLIS, see 2.1), lead us to assume that there will be no significant difference from the quantitative analysis on features extracted.

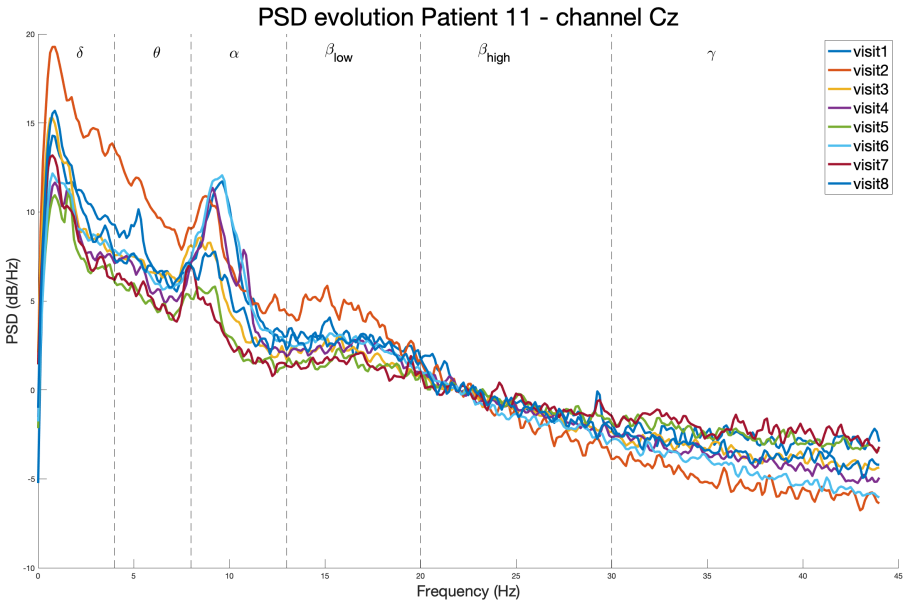


FIGURE 3.11: PSDs plots over subsequent visits for Patient 11.

Patient 11's PSD evolution is displayed in Figure 3.11.
EEG longitudinal analysis for this patient is of particular interest since the patient is supposed to have a major transition from LIS to CLIS during the year period in which data was acquired.
A qualitative analysis of EEG through the overview of PSD evolution is difficult in this case, since from Figure 3.11 can be just hypothesized that EEG's spectral content is significantly changing across visits. The results for features' quantitative and statistical analysis are reported in Section 3.4.3.

In agreement with what can be noticed from the observation of EEG's time course in Section 3.2, both Patients 4 and 6 exhibit frequency peaks in $\delta - \theta$ range instead of the normal peak in α band that can be found in the PSD of a healthy subject. On the other hand Patient 11's, which EEG shown in Figure 3.8 is similar to the one usually found in controls, present a clear peak in α frequency band in all the observations.

3.4 Features and statistics

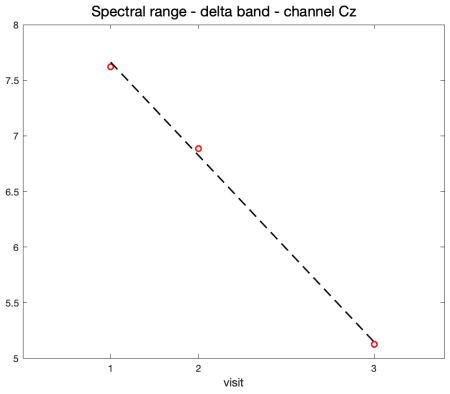
Data processing and statistical analysis have been performed to each patient's data as described in Section 2.4, 2.5 and 2.6.
The following sections present the results obtained and an explanation on how statistic analysis was performed considering the data available for each patient.

3.4.1 Patient 4

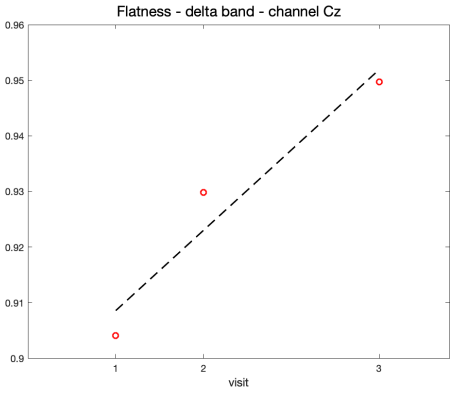
For Patient 4 features have been extracted separately for each EEG recording resembling a visit. The analysis has been performed on channel Cz.
Statistical analysis through one way ANOVA test and Mann Kendall test cannot be applied to this patient's data since the number of observations to consider is too low. Anyway, there are visible changes in the power spectral densities within different visits (see Figure 3.9), even though is difficult to quantify them in a reliable way.
The major changes visible comparing visit 2 and visit 3 in Figure 3.9 are probably due to the collapsed lung occurred in the patient in mid-2017, after visit 2. The oxygen shortage happened at the brain level is most probably the cause of the further slowing in the EEG, visible through the shift of the spectral power towards slower frequencies in visit 3's PSD.
Since the changes are clearly relevant it has been decided to apply a linear regression to features' values obtained after data processing, to select those features showing the highest consistency with respect to a linear regression through the three observations ($R^2 > 0.9$). The purpose of this analysis was just to select in an objective way what emerges from a qualitative visual analysis of power spectral densities.

Referring to Figures 3.12 3.13 3.14 3.15 and 3.16, after quantitative feature analysis the results can be resumed for each frequency band as:

- δ: decrease in spectral range and increase in flatness;
- θ: increase in spectral range and decrease in flatness;
- α: decrease in relative power, spectral entropy and flatness;
- β: increase in absolute power and decrease in spectral range;
- γ: increase in absolute power.

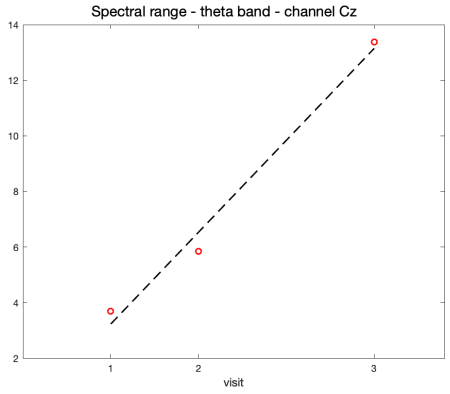


(A)

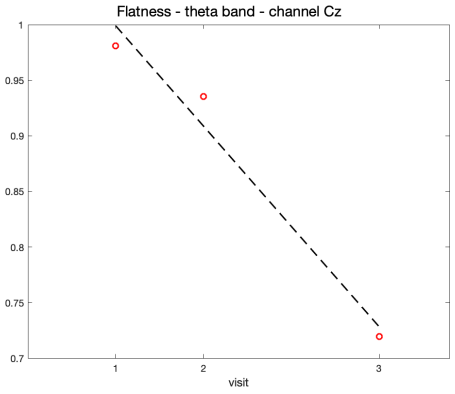


(B)

FIGURE 3.12: Features selected - delta band - Patient 4



(A)



(B)

FIGURE 3.13: Features selected - theta band - Patient 4

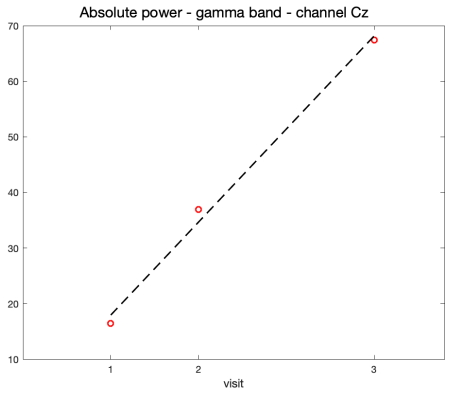
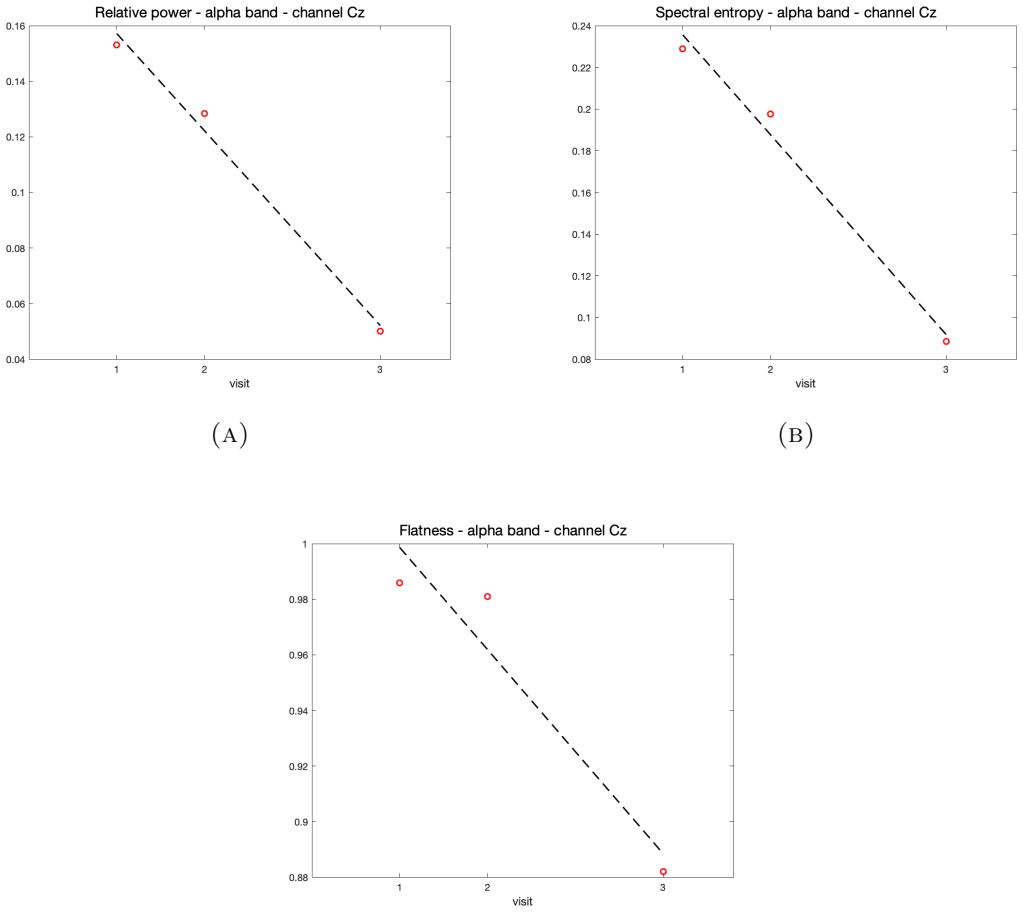


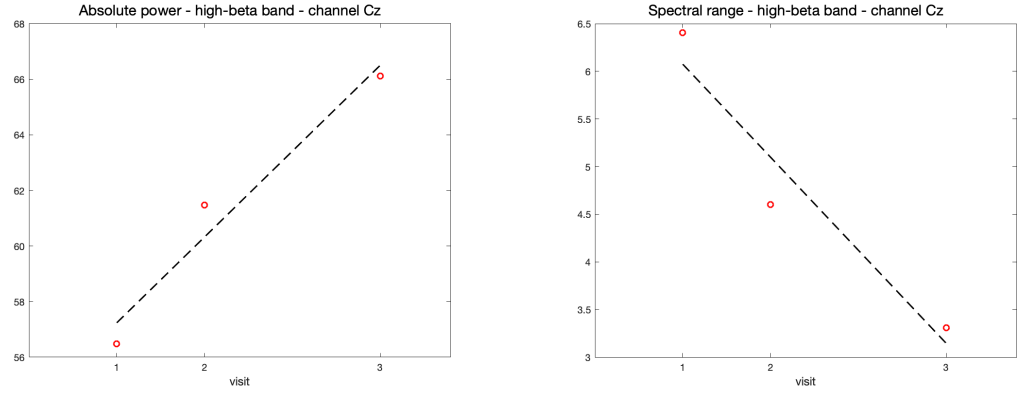
FIGURE 3.14: Features selected - gamma band - Patient 4



(A) (B)

(C)

FIGURE 3.15: Features selected - alpha band - Patient 4



(A) (B)

FIGURE 3.16: Features selected - beta band - Patient 4

3.4.2 Patient 6

Features analysis on Patient 6's data has been performed on channel Cz.
No features show a significant evolution over the period corresponding to the EEG acquisition, having all of them a p-value higher than 0.05 resulting from Mann Kendall trend test. One way ANOVA test confirms there is not even a statistically significant difference between features' values of different visits, excluding just two features. The results are shown in the following Table 3.1.
This result means that the variations registered on features values are not caused by a change within visits, and possibly the oscillations on values have the same variance within days.
ANOVA results are further clarified in Appendix A.2 Figure A.2, through the boxplots presenting the distribution of features values within days of the same visit.

Feature	**band**	**p-value**	**adjusted p-value**	**accepted**
Permutation entropy		0.8882	0.9135	false
Kolmogorov index		0.9802	0.9802	false
Absolute power	delta	0.4438	0.7099	false
	theta	0.8049	0.8522	false
	alpha	0.4754	0.7441	false
	low beta	0.3830	0.6566	false
	high beta	0.6036	0.8	false
	gamma	0.2067	0.4962	false
Relative power	delta	0.0034	0.0155	false
	theta	0.0012	0.0106	false
	alpha	0.0004	0.0094	true
	low beta	0.2649	0.5298	false
	high beta	0.7792	0.85	false
	gamma	0.7239	0.8406	false
Spectral entropy	delta	0.0013	0.0106	false

	theta	0.0015	0.0106	false
	alpha	0.0005	0.0094	true
	low beta	0.2632	0.5298	false
	high beta	0.7727	0.85	false
	gamma	0.7222	0.8406	false
Flatness	delta	0.0768	0.2303	false
	theta	0.0022	0.0132	false
	alpha	0.0749	0.2303	false
	low beta	0.3754	0.6566	false
	high beta	0.6061	0.8	false
	gamma	0.0713	0.2303	false
Range	delta	0.0326	0.1302	false
	theta	0.2232	0.5021	false
	alpha	0.0749	0.2303	false
	low beta	0.5088	0.7633	false
	high beta	0.6567	0.8152	false
	gamma	0.0028	0.0146	false
Total power		0.6222	0.8	false
Spectral entropy	broadband	0.1422	0.3656	false
Flatness	broadband	0.5897	0.8	false

TABLE 3.1: ANOVA results for Patient 6's features on Channel Cz. Both p-values and FDR corrected p-values are reported, and the last column indicates whether the correspondent feature had a positive outcome concerning changing significantly over the observation period (having a p-value lower than 0.01).

3.4.3 Patient 11

Applying the proposed pipeline to Patient 11's data, some features resulted in having a trending behavior over the period of observation.
While for the previous patients the analysis has been performed just on Cz channel, for Patient 11 EEG channels were interpolated to obtain a fixed set of twelve electrodes common to all visits (see Section 2.3.2). Having more channels available the analysis was performed channel wise, both to see if the eventually extracted features were consistent to all channels locations and to investigate if EEG signal presents different characteristics in different scalp regions.
In the following are reported:

- the results obtained from ANOVA test on features values related channel Cz, comparable to the results obtained for Patient 6 (Table 3.2)
- the results obtained from Mann Kendall trend test, reporting only the features correspondent to a p-value lower than 0.05 (chosen as significance threshold) together with the respective p-value obtained from the ANOVA test (Table 3.3)
- the plots of the features resulted in having a monotonous trending behavior across visits according to Mann-Kendall statistics (Figures 3.17 3.18 3.19 and 3.19)

ANOVA results are further clarified in Appendix A.2 Figure A.3, through the boxplots presenting the distribution of features values within days of the same visit.

Feature	**band**	**p-value**	**adjusted p-value**	**accepted**
Permutation entropy		0.1045	0.1107	false
Kolmogorov index		0.0591	0.0545	false
Absolute power	delta	6.15e-09	2.215e-07	true
	theta	1.148e-07	2.066e-06	true
	alpha	0.0186	0.0223	false
	low beta	8.187e-06	4.272e-05	true
	high beta	0.0009	0.0016	true
	gamma	5.208e-05	0.0001	true
Relative power	delta	0.0001	0.0004	true
	theta	0.6484	0.6669	false
	alpha	0.0029	0.0043	true
	low beta	0.0002	0.0004	true
	high beta	9.378e-06	4.272e-05	true
	gamma	5.061e-06	4.272e-05	true
Spectral entropy	delta	0.0002	0.0005	true
	theta	0.7075	0.7075	false
	alpha	0.0009	0.0016	true
	low beta	0.0001	0.0004	true
	high beta	9.493e-06	4.272e-05	true
	gamma	6.481e-06	4.272e-05	true
Flatness	delta	0.0166	0.0207	false
	theta	0.0007	0.0014	true
	alpha	0.0036	0.0052	true
	low beta	0.0001	0.0004	true

	high beta	0.0007	0.0014	true
	gamma	0.0006	0.0014	true
Range	delta	0.0317	0.0369	false
	theta	0.0394	0.0443	false
	alpha	0.0051	0.0070	true
	low beta	0.0056	0.0074	true
	high beta	0.0058	0.0075	true
	gamma	0.0011	0.0017	true
Total power		5.1556e-07	6.1867e-06	true
Spectral entropy	broadband	0.0006	0.0013	true
Flatness	broadband	5.1390e-05	0.0001	true

TABLE 3.2: ANOVA results for Patient 11's features on Channel Cz. Both p-value and FDR corrected p-values are reported, and the last column indicates whether the correspondent feature resulted in changing significantly over the observation period (p-value lower than 0.01).
Here are reported the results for channel Cz, to be consistent with the analysis made for Patient 6. ANOVA test was applied to features computed on all the EEG channels used in the processing for Patient 11.

Channel	Feature	Band	ANOVA p-value	Mann Kendall p-value	sign	accepted
Cz	relative power	theta	0.66689	0.0094	↓	false
Cz	spectral entropy	theta	0.7075	0.0187	↓	false
Cz	spectral range	low beta	0.0074	0.0354	↓	true
C1	relative power	low beta	0.0003	0.0187	↑	true
C1	spectral entropy	low beta	0.0002	0.0354	↑	true
F2	flatness	gamma	0.0005	0.0187	↓	true
F3	relative power	theta	0.0001	0.0354	↑	true
F3	relative power	low beta	0.0001	0.0354	↑	true
F3	spectral entropy	theta	7.80e-05	0.0354	↑	true
F3	spectral entropy	low beta	7.05e-05	0.0187	↑	true

TABLE 3.3: Results obtained after Mann Kendal trend test. The reported features are the ones having a p-value < 0.05 (adopted threshold to define significative features). For each selected feature, the p-value (FDR corrected) obtained from ANOVA test are reported too, to discard Mann Kendall trend test's false positives. Indeed, ANOVA should robustly reveal which features are significantly different between the visits, analysing their variance within the multiple sessions of the same visit.
It has to be considered the relative significance of the statistical results for this particular study due to the reduced number of observations.

Table 3.2 gives the evidence of how almost all features are significantly changing within visits, keeping a lower variability within days. This aspect is important since features values can change within days without being related to an overall change within visits, that is what resulted in Patient 6.

Figures 3.17 3.18 3.19 and 3.19 show the regression plots obtained from Mann Kendall slope (see 2.6) in the trend analysis of each feature in separated channels (on the left in each subfigure). Mann Kendall trend test is performed on features values extracted from the processing of visit wise PSDs (mean of day wise PSDs). In parallel to that, each feature is evaluated from day wise PSDs, and their variability within days of the same visit is shown in a boxplot (on the right in each subfigure).
The aim to that is the comparison of features values obtained from visit-wise PSDs and the ones obtained by the processing of day-wise PSDs to:

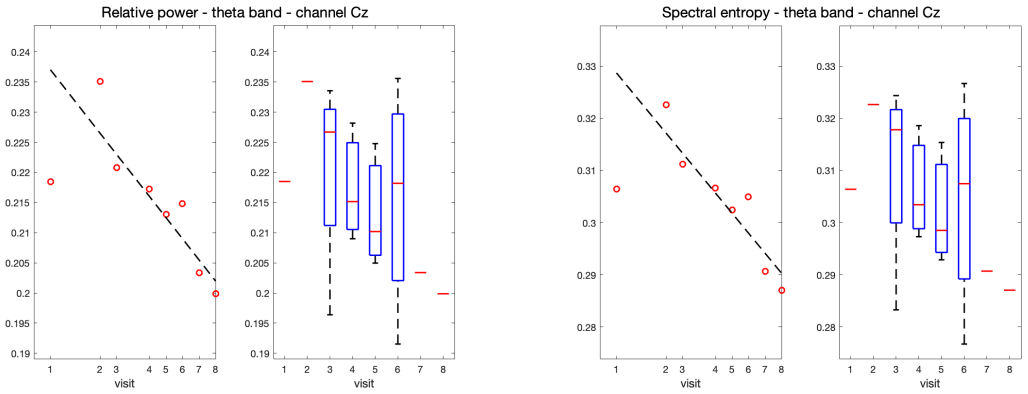
- validate the adoption of visit-wise PSD for extracting features, with respect to day-wise features values and their variance

- a better understanding of the importance of validation of Mann Kendall statistic results through ANOVA p-value

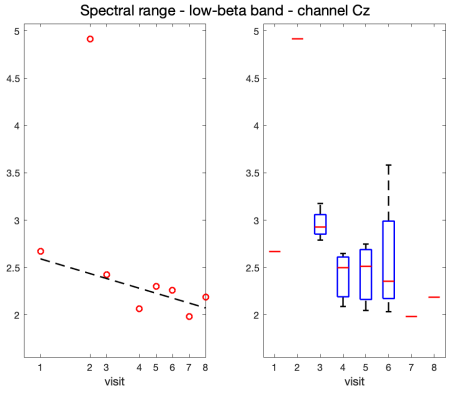
Indeed, from Table 3.3 we can see that two of the features selected by Mann Kendall test for trend analysis are not accepted; ANOVA test, analysing the variability of features values within days of the same visit, rejected them not recognizing a significative change within subsequent visits in day-wise features values grouped by visits.
The reason to that is clearly visible in Figure 3.17 (A) and (B), related to those "rejected" features; their values have a high variance within days of the same visit that is comparable to the overall variance within all the visits. Hence, even if their average value is possibly exhibiting a trending behavior across the observation period, they are both not reliably associable to their median value and they are not showing a significative change taking into account their variability.
On the other hand, the other features mentioned in Table 3.3 and represented in the successive figures, show a lower variance within sessions of the same visit (from the boxplots) and a statistically significant trending behavior, reflected in the regression plots.
Another observation that can be made from the figures reported below is the capacity of the regression through Mann Kendall's slope to avoid the influence of outliers. It is clear that, also looking at the PSDs evolution reported in Figure 3.7, visit 2 can probably be considered as an outlier from all the representations made, since it has higher spectral power and features values completely different from the ones obtained for the other visits. It is noticeable how the slope used for the regression, being the median of the slopes computed on all pairs of subsequent values, is not taking it into account; a common regression applied on features values would otherwise have shifted the slope towards visit 2's values, since their absolute value is most of the times higher than the ones related to the other visits.

Grouping the obtained results for frequency bands they can be resumed as:

- θ: increase relative power and spectral entropy (F3);
- β: decrease in spectral range in registered in Cz, while in C1 and F3 increase in relative power and spectral entropy;
- γ: increase in flatness (F2).

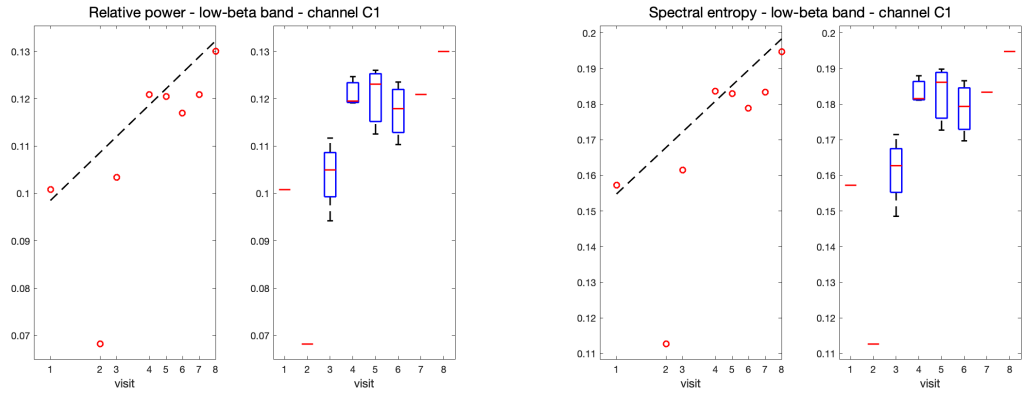


(A) (B)



(C)

FIGURE 3.17: Features selected - channel Cz - Patient 11



(A) (B)

FIGURE 3.18: Features selected - channel C1 - Patient 11

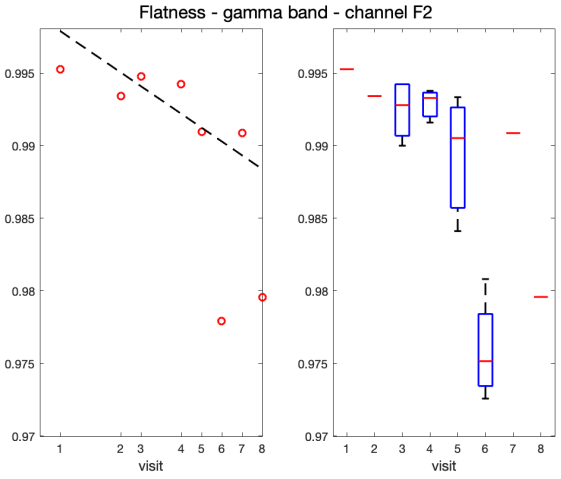
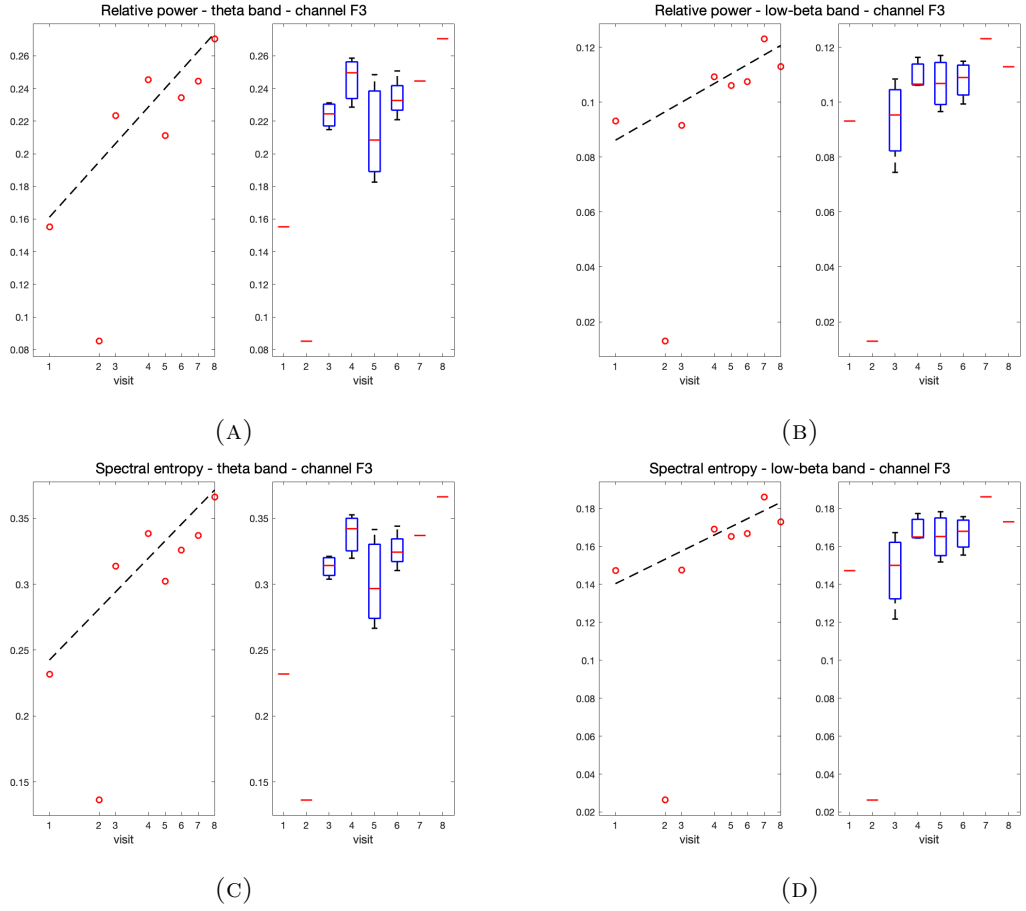


FIGURE 3.19: Features selected - channel F2 - Patient 11



(A) (B)

(C) (D)

FIGURE 3.19: Features selected - channel F3 - Patient 11

Chapter 4

Discussion

The discussion will focus on the evidence obtained from the analysis, related to the first questions mentioned in Chapter 1:
is EEG changing in its features during the evolution of the disease? If yes, is it evolving in the same features that differentiate patients from healthy subjects?
All the stages of the pipeline used are going to be discussed in the following sections, focusing on their usefulness and their limits and reliability.

4.1 Data preprocessing

EEG preprocessing is the first and most delicate step to face in data analysis. It is of crucial importance because the interpretation of EEG is not unique and neither exist a standard procedure to approach raw data. Most of the time the way of preprocessing data is application specific, but it has to be paid attention not to manipulate data to the specific purpose of the analysis, since in this way the analysis would be balanced and not reliable.
In this study, excluding the basic steps of filtering and downsampling, three preprocessing stages need to be discussed: indipendent components removal, channel interpolation and power spectral density normalization.

The application of independent component analysis to EEG signals is widely used in the literature and, as explained in Section 2.3.1, EEG signal is particularly suitable to this decomposition into independent sources.
Rejecting artifacts by manual removal of EEG epochs can be problematic since it needs

the capacity to visually evaluate the quality and "cleanness" of EEG signal, which usually requires extensive experience in EEG evaluation and interpretation. Furthermore, ALS patients' brain signals, as happens with other neurological disorders, are different from the ones that can be observed in healthy patients, so evaluating them in the same way is probably imprecise.
In this study the choice of this procedure to reject artifacts was adopted to avoid making subjective evaluations on EEG epochs, with the overall objective of approaching data analysis in a blind and unbalanced way.
The main heuristics used in component removal for EEG artifacts rejection are [27]:

- Eye movements should project mainly to frontal sites with a lowpass time course.
- Eye blinks should project to frontal sites and have large punctate activations.
- Temporal muscle activity should project to temporal sites with a spectral peak above 20 Hz.

With these premises components removal can be used as an artifact rejection tool without an excessive subjective intervention on data manipulation. It has been shown in Section 3.1.1 to be effective for removing both eyes and general noise artifacts.

Channel interpolation is another tool applied to EEG data analysis, usually used to reconstruct one or more channels have to be rejected for their noise content, overcoming their absence.
In this study channel interpolation was applied in a different context; here the available channel sets were not always the same between different recordings, and to perform a uniform longitudinal comparison it was needed to operate always with the same electrode. Channel Cz was the one present in almost all the recordings, and when it was not available the neighbor electrodes were acquired. For this reason, it has been considered admissible to derive it through interpolation when it was missing.
In addition to that, being Cz channel located on the central region of the scalp, it is expected to reflect mainly the activity of the sensorimotor area (the correspondent underlying region), which is actually the one thought to be highly involved by the neuronal degeneration occurring with ALS disease. That is another reason why the principal analysis was carried on this channel.
For Patient 11 a larger set of electrodes was interpolated, and the results of this interpolation have been shown in Section 3.1.2. In this case it has to be paid more attention on data processing, considering that the signals are obtained from mathematical operations on the neighbor channels, and therefore completely dependent on them.

EEG signals, as mentioned in Section 1.2, has not a high spatial resolution and so single channels' activity cannot be considered independent within electrodes; anyway there exist synchronization measures that analyse delay and correlation between channels relying on their phase difference, and this aspect cannot be recreated through interpolation.
That is the reason why the features analysed here are single channel features in time and frequency domains, and the usage of interpolated channels to extract them is admissible.

The need of normalizing EEG signals comes from the inner difference that can be observed in single patient recordings acquired in different circumstances, since they are affected from condition-specific offsets that are possibly due to patient state or acquisition equipment. Applying normalization implies the deletion of the absolute amplitude features of the signals, so after EEG normalization only relative amplitudes in their spatial configuration are retained.
Also in this case the choice of the normalization is highly application specific. In this study normalization was applied first zscoring signal in time domain across sessions of the same visit, and secondly to single-channel power spectral densities dividing them by their median. The choice of using the median as normalization factor for PSDs was made since it was observed that its adoption allowed to compare PSDs within days and visits; furthermore, the usage of median instead of mean is genrally more robust against outliers.

4.2 Motivation for features

In this study the EEG features that have been analyzed are taken from both the state of the art of the comparison between ALS patients and controls and the one of the analysis and detection of disorders of consciousness.
The goal of this choice of features is to investigate whether the markers able to discriminate patient groups from healthy subjects increase their discrimination power with the progression of the disease. Furthermore, analysing the DOC's EEG biomarkers is a way to assess whether ALS patients go progressively into a state similar to one of limited consciousness in the transition from LIS to CLIS. Comparing the EEG features that usually mark unconsciousness states is possible to confirm that those patients' brain activity is more similar to the one that is usually registered for healthy patients.
In the literature this topic is addressed in several reviews, and Figure 4.1 shows the primary features in EEG resting state analysis for DOC.

The features investigated in this study covers approximatively all the single-channel features mentioned there (see 2.4 and 2.5). Other features were evaluated but not reported since they have been found to be not significant or redundant with the ones already selected for the analysis.

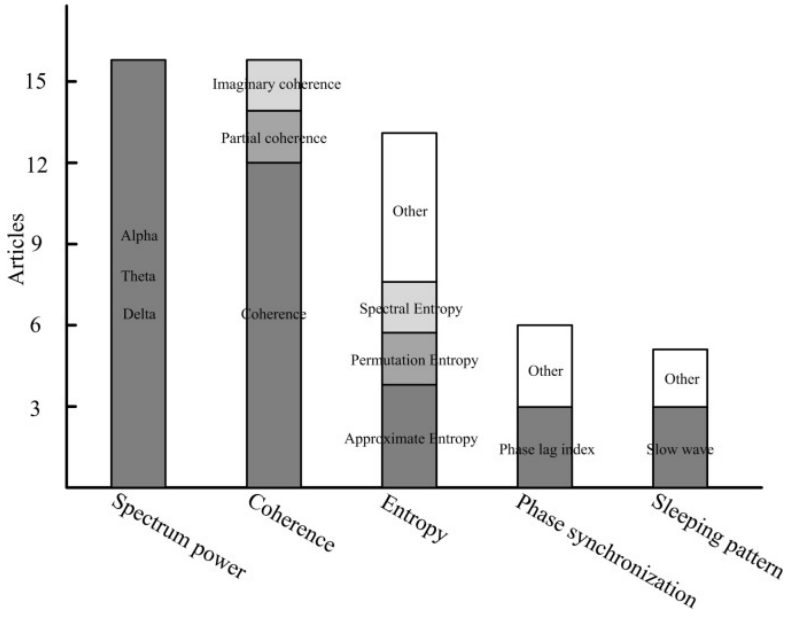


FIGURE 4.1: Primary EEG resting state features in DOC [20].

Common characteristics found in DOC patients are [20] [19]:

- increase of low spectral power (δ and θ)
- decrease of high spectral power (α)
- disconnection between frontal and other regions (decreased coherence)
- low complexity both in time domain (Kolmogorov complexity) and in frequency domain (spectral and permutation entropy)

Much attention is paied to α frequency band since it is usually the dominant frequency band in healthy patients, directly connected to attention and vigilance.
Referring to the state of the art in EEG resting state analysis for ALS syndrome, the primary features are the same as the ones reported in Figure 4.1, but with different results. EEG features resulted as the most effective in the separation of patients and controls are mainly:

- decrease in spectral power in medium-high frequency bands (α,β,γ)
- shift of α peak towards lower frequencies

- increased co-modulation parallel to a decreased synchrony

In Figure 4.2 are shown the results obtained by Maruyama et al. in their study [23] comparing EEG resting state recordings of CLIS patients vs controls. Notably, in this study performed in collaboration with the Institute of Medical Psychology and Behavioral Neurobiology of the University of Tübingen, one of the patients corresponds to Patient 6 and the acquisition protocol and analysis are comparable with the ones adopted for this work. PSD's absolute value is not comparable with the analysis performed here due to the normalization applied.
From subfigure C of Figure 4.2 is clear how they found a separation between patients and controls, occurring at several frequency bands.

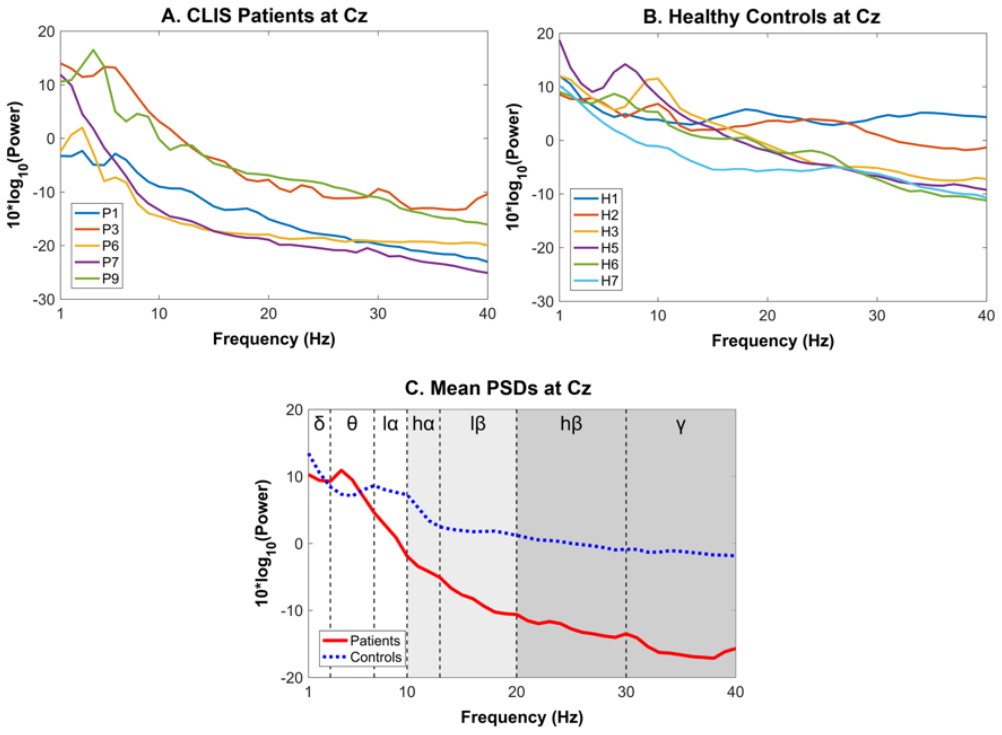


FIGURE 4.2: PSD comparison between CLIS patients' group and healthy participants, taken from [23]. The frequency bands reported in the third subfigure refers to: delta (1 to 3 Hz), theta (4 to 7 Hz), low alpha (8 to 10 Hz), high alpha (11 to 13 Hz), low beta (14 to 20 Hz), high beta (21 to 30 Hz), gamma (31 to 40 Hz)

Are these features able to trace the degeneration occurring in ALS syndrome in the same way?
Drawing this conclusion from the results obtained in this study, reported in Section 3.4 for the three patients analysed, is not entirely possible.
It has not been found consistency between the features selected for Patient 4, as the ones changing over the period of observation, and the ones found for Patient 11 (Patient 6 is not considered here since the features extracted from his dataset resulted not changing

within the observed period).
For Patient 4 it has been encountered a general power decrease in alpha and low beta frequency bands with a parallel increase in gamma band, while for lower frequencies as the ones in delta and theta band it has been stated an increase and a decrease in flatness respectively.
For Patient 11 the outcoming features are different between the observed channels and can be generalized as an increase of relative power in theta (F3) and beta bands (C1 and F3), while an increase in flatness has been found in gamma band (F2).
In this case of multichannel analysis it has to be considered that some of the electrodes positions have been interpolated for some sessions and that frontal channels, even though eye movements' artifacts should have been removed through independent component removal, can possibly be affected by some residual artifacts.
Commonly to all the three patients, none of the complexity indexes used as features in time domain emerged from this feature analysis.

These results bring to different observations and reflections.
First of all, as we can see from Figure 4.2, where CLIS patients group and healthy participants group are compared, PSDs vary widely within groups where subjects should have similar features. Due to that, comparing single-patients results with the ones obtained in group studies can easily take to misleading conclusions if we don't consider this issue. Averaging PSDs within groups is the usual practice for this kind of studies, but that aspect make single-patient PSDs not directly comparable with the ones obtained by a group average.
This inter-group diversity is particularly noticeable also in the control group, and this should prevent us to compare patients' PSDs to those related to healthy subjects.
Observing this, is difficult to expect that the features resulting from group analysis can be expressed in the same way at a single patient level.

What is more interesting after this evaluation is thus to focus on each patient singularly, considering the results on the basis of their pathophysiological condition characterizing the time range were EEG data has been acquired.
Table 4.1 gives an overview of each patient's general clinical information, reporting the ALS form diagnosed with the respective date, to be related to the time range of the available data.

Referring to Patient 4, she was diagnosed with juvenile ALS between 2012 and 2013 and, relying on the available clinical information, she had control on her eye movement

Patient	**Diagnosis**	**Diagnosis date**	**Transition date**	**Data's time range**
Patient 4	juvenile	2012	2015	2016-2017-2019
Patient 6	bulbar	2009	2012	2017-2019
Patient 11	non bulbar	2015	2019	2018-2019

TABLE 4.1: The table reports the diagnosis made for each patient, with the respective date (year), the time when the LIS to CLIS transition occurred and the time range relative to the datasets analysed.

until January 2015. Hence, we can assume that the transition to CLIS occurred 2 years after the arise of the first symptoms. The available dataset correspond to 2016, 2017 and 2019, so they cannot be related to the transition from LIS to CLIS but cover a wide period when that patient is in a completely locked-in state.
From this analysis, power spectral density computed from EEG recordings in these three different years seems to go through significant changes, but these cannot be generalised as a characterization of the neurodegeneration occurring during CLIS state, since they are obtained from single EEG recordings (one for each year). The only conclusion that can be hypothesised is that during the CLIS state, with an observation window of several years, some changes in EEG are possibly encountered, especially in case of severe complications of patient's health condition as happened in this case with a collapsed lung.
In this specific case what seems more interesting is the increase in gamma band power, that is generally found reduced in patients vs controls studies, but other researches [40] found out a possible non linear relationship between high frequencies power and the neurodegeneration occurring in ALS.
Another aspect that has to be discussed is the particular clinical phenotype related to Patient 4, that is juvenile ALS. The symptoms started when the patient was 23 years old, placing her clearly in this specific form of the syndrome. Juvenile ALS is different from the other forms in which this disease can manifest, and is thought to be one with longest life expectation and slower symptoms degeneration.
That aspect, together with the information about the current utilization of BCI for communication, could reinforce the hypothesis of a CLIS state that is still evolving with totally preserved consciousness.

Patient 6 is in a condition totally different from the one associable to Patient 4.
Patient 6 was diagnosed with bulbar ALS in 2009 (when he was 30 years old), and he can be considered in CLIS since 2012. Also in this case the observation period (2017-2019) is completely included in the CLIS stage of the disease, with a time gap of around five years from the presumed occurrence of the transition.

Knowing this information, what can be observed is that data refers to a period in time when the patient is completely stable in its CLIS condition, while for Patient 4 the first EEG recording was acquired just one year after the definitive transition. This point, together with the shorter observation period (less than two years for Patient 6 vs 3 years for Patient 4), is probably connected and coherent with the absence of significant changes in EEG time and frequency contents.
Furthermore, bulbar ALS is one of the most aggressive phenotypes and it is usually associated with a faster degeneration.
Examining all these evidence together, the final hypothesis that can be drawn is that, in case of faster degeneration, the brain activity represented by EEG recordings ends up into a stable state after CLIS stage is reached. Alternatively, it can be at least assumed that period of observation is too short to find some significant changes in EEG.

Patient 11's clinical situation is again different from the two conditions previously examined. He was diagnosed with non-bulbar ALS in 2015 and he started using BCI for communication using his eyes just one year after, until 2019 when he completely lost his capability to fixate his gaze going into CLIS. In this case, the transition from LIS to CLIS was slower and data analysed in this study are acquired within a period of one year, coincident with the transition (last visit's date is March 2019).
Differently from the case of Patient 6, even if the duration of the observation period was similar to the one of Patient 11's data, the analysis performed found out a significant difference in almost all the features examined. Moreover, it has been identified the presence of some EEG features changing monotonously within the transition period.
As previously marked in Chapters 2 and 3, the reliability of the findings on the features is not completely trustable for many reasons, from the different numbers of visits and the different numbers of sessions to which each visit refers to, to the actual significance of the statistical tests on a low number of observations. However, even if any conclusion can be drawn from the results on features values, it can be assumed that in this case EEG is changing in some of the analyzed features and those changes are possibly reflecting the neural degeneration happening during the transition period.
Patient 11's particular course of disease and his rapid adoption of BCI for communication, can exemplify and prove what is theorized by Kubler and Birbaumer in their study about the extinction of goal-directed thinking [22]. There they hypothesize that a brain that is trained to BCI usage before the transition to a CLIS can preserve its capability to associate his behaviour with the external feedback received, and thus keep the possibility to use BCI even in a completely locked-in state. In addition, for this patient it can be observed a slower degeneration in terms of time delay between the onset date and the transition to CLIS, and even if his recent definitive transition to CLIS prevent us to

conclude it, is possible that keeping the brain "trained" and used to imagery thinking with reward has contributed to a delay in this late stage presentation.

Putting together all the findings it is hypothesized that:

- the period including the transition from LIS to CLIS is characterized by a modification in the EEG signal, visible from the changes in its features. Those changes are possibly connected to specific features' monotonous evolution.
- CLIS condition can be considered stable regarding changes in EEG signal after the transition. The period that occurs between the transition and the reaching of a stable condition is probably related to the diagnosis and to the velocity of the degeneration.

4.3 Pipeline validation

In this work a great importance has been given to the construction of a solid pipeline to preprocess and process single patients data, with the purpose of performing a longitudinal analysis on the acquired EEG resting state recordings.
The importance of this kind of studies is given by their wide field of application in medicine, since they can represent support to medical decisions and help in diagnosis and in monitoring diseases within their evolution.
This work specifically focuses on ALS syndrome, but the proposed pipeline could be used to study the implications to brain signals due to other neuropathies going through a neuro-degenerative process as well, as Alzheimer or Parkinson's diseases.
Indeed, excluding the choice of features that is partially referring to ALS studies' state of the art, the preprocessing and processing stages where designed in an unbalanced way that is not specifically suited to this syndrome.
Moreover, having few observations to analyse, the biggest effort was made with the objective of operating in a blind way to the processing of data, since drawing a precise conclusion from the results would have been premature.

For this reason, even if applying statistical tests to a low number of observations possibly brings to misleading results, it was tried to build a statistical procedure to assess the significance of the obtained results, through ANOVA and Mann Kendall trend test in parallel.
Their aim is to answer the questions: are the changes registered in EEG features related

to different observations isignificant? If yes, does these changes occur as monotonous trends (are increasing or decreasing in a monotonous way)?
It is possible to answer to these questions with ANOVA and Mann Kendall test respectively, and the extent to which their answer is significant can be assessed through their respective p-values, giving a measure of how far is the specific condition under examination from a random occurrence.
Another aspect that justifies the parallel use of two statistical tests is the possibility to use them to check the coherence and reliability of their results; as an example, a positive result related to a trending behavior is meaningless without a concomitant positive result for a significant change in features values through different visits. Even if in this specific case obtaining coherent results between statistical tests was not to be taken for granted, the ones reported in Sections 3.4.2 and 3.4.3 show how this procedure can possibly be applied even to a small set of samples.
All the considerations that rely on statistical results in this study have to be considered together with the limits of the dataset, but anyway it is assumable that the statistical pipeline for assessing the results would be reliable in presence of a higher number of samples.

4.4 Limits and future improvements

In the previous chapters, it has been marked several times that the composition of the available dataset for the single-patient analysis was limiting the reliability of the results, considering both the overall number of EEG resting state recordings and the different numbers of sessions that were unified into visits. An ideal situation would have been to have an equal number of EEG acquisitions for each visit, with visits equally spaced and covering a wide period in time. Anyway, dealing with patients, similar conditions are rarely reproducible due to their real-time needs and the limited possibility of acquiring a large dataset from the same patient during a wide time span.
The different electrode placement between different recordings is also limiting the reliability of the results and the possibilities of computing more complex features, but even in this case having a large and uniform set of EEG channels among single patient recordings and within different patients is an ideal condition, reproducible just in large population experiments where data are acquired exactly to the study purpose. This was not the case for this work, since the available data were acquired in the previous years without this specific purpose but concomitantly to other BCI experiments and analysis on patients.
The condition of these patients itself is already limiting data acquisition, since they are lying in bed and need frequent interventions of caretakers to fulfill their basic needs.

Quantitative comparison of features values within patients is impracticable not because of limits in the dataset, but due to physiological inconsistency of this kind of comparison. Quantitative evaluations are performed in group studies where a features' group average is applied, but EEG is so different within patients that this comparison would have no meaning at a single patient level.

A possible improvement for this study would be the enlargement of the dataset available for each patient with further acquisitions, together with the increase of the number of patients under analysis, acquiring EEG resting states systematically for all of them. The usage of a uniform protocol for these studies could allow also to compare data registered from patients in different countries and institutions, enlarging the dataset for the analysis.

Having the possibility to compare single-patient data evolution with other patients presumably in a similar clinical condition could possibly reinforce the hypothesis made here, concerning the correlation of EEG evolution with phenotypes and overall course of disease.

In addition to that, in the literature many connectivity metrics are found effective in the analysis of ALS patients' EEG, resulting in being able to discriminate control groups from patients [17] [16]. Probably connectivity metrics would be effective in monitoring the progression of the disease too, and the use of EEG/MEG acquisition systems together with TMS studies are found to be a promising way of assessing neurodegeneration in longitudinal studies [41].

The principal motivation for that is that there is clear evidence that neural degeneration is related to progressive changes in brain networking, and those changes can be identified in functional terms through altered patterns of brain connectivity and neuronal transmission.

This kind of analysis needs a high-density electrode placement on the scalp, but these methods are still more accessible than other imaging techniques as fMRI, with the same potential of localising precisely sources of information in the brain.

In the future would be interesting to have the possibility to investigate those metrics longitudinally for ALS patients, acquiring data with a fixed set of electrodes that cover a wider area on the scalp.

Chapter 5

Conclusion

This research aimed to perform a longitudinal analysis on EEG resting state recordings, acquired from three ALS patients with different backgrounds and clinical conditions. This choice was made since it is known that, in particular in the case of neurodegenerative disorders, a great inter-patients difference is met over many aspects. The progression of the disease occurs with different timings and implications between different patients, depending on the specific disease phenotype that characterizes them but also on other subjective factors, as the current health conditions of the patient, the environment in which he is inserted and its motivation despite the critical and difficult situation.
Assuming that EEG can reflect the degeneration happening in patients' brains, is thus reasonable to conclude that an investigation on a single patient or few patients cannot allow to characterize a general condition or the implications due to the evolution of ALS disease.
Analysing three patients data allowed to compare EEG evolution over different courses of ALS syndrome: Patient 4 is a case of juvenile ALS, analysed in three different years of CLIS; Patient 6 is a case of bulbar ALS, analysed over a period of almost two years in which he is in CLIS; Patient 11 is a case of non-bulbar ALS, analysed over one year period when the transition from LIS to CLIS occurred.

The use of EEG for its capability to reflect brain states is well known, but being acquired externally from the brain is at the same time a good point for its wide applicability and a drawback for its limited spatial precision. Showing that it is still able to reveal a change happening at a neural level is useful for many medical applications, since in many cases using more invasive ways of investigating brain activity is not possible for patients affected from constrictive diseases.

The objective has been, first of all, to evaluate if a change in EEG signal related to neural degeneration can be observed; the answer to this point is positive since the results, even if data was limited in the number of observations for all the patients, demonstrated that the degeneration occurring at a neuronal level seems to be expressed through a significant change in the spectral content of the signal. Assessing which features are specifically tracking the evolution of the disease is more challenging and to draw reliable conclusions more EEG samples would be necessary.
The potential of using EEG for analysing the completely locked-in state, as late stage of ALS disease or other neuropathies, could have important implications for its comprehension and to assess its inner difference from those classes of states characterized by a reduced consciousness. The transition occurring between LIS and CLIS is another interesting aspect to investigate since, to the actual state of the art, the analysis made focus just on some physiological markers, and the transition is defined basing mailnly on the loss of control on eye movements. In this direction, what has been found of particular interest are the encountered changes in EEG recordings of Patient 11, since they correspond to the transition from LIS to CLIS occurred in the patient in the period in which they were acquired, parallel to an absence of significant changes found over a period of comparable length of stable CLIS condition for Patient 6, further in time from the transition.
These findings lead to hypothesize that the transition between LIS and CLIS is characterized by a dynamic evolution in the EEG signal, while CLIS state, as the last stage encountered in the degeneration of ALS, seems to be more stable. This stability, anyway, is possibly relative to the short time range of observation and related to the distance in time from the moment when the transition occurred in the patient. Indeed, from the analysis done for Patient 4 in CLIS during the whole period of observation of four years but closer in time to the transition, came out that major changes in EEG's frequency content are still happening.

After these observations other questions conceivably come out.
Does EEG evolve in the direction of those states of limited consciousness?
This point is very complex in its implications, and answering that would be premature. Performing some group analysis comparing late-CLIS patients to MCS patients and VS patients, as done in the literature for discriminating those last two from controls, could be useful to identify major differences between those groups, that for the moment can just be hypothesised by the knowledge of preserved consciousness for late stage ALS patients. From the analysis made in this study no changes related to those features characterizing DOC have been registered.
Are CLIS patients all similar from the consciousness point of view?

Probably not, and we can see that in a preliminary way from Patient 6 stability compared to the potentially more dynamic state of EEG for Patient 4. Patients have all different backgrounds and are surrounded by different environmental conditions that are likely to influence the overall progression in their own course of disease.

In addition to that, the application of the proposed pipeline for a longitudinal analysis of single patient data on three patients demonstrated its solidity within totally different datasets. The overall preprocessing and processing procedure was build with the aim of possibly applying it in the future to a larger and more consistent dataset, with good chances of being reliable and able to assess changes in specific EEG features also from a statistical point of view.

This study could answer some preparatory questions that prepared the ground for other future investigations. After stating EEG's capability to reflect the neural degeneration occurring in neurodegenerative disorders, which features can really act as markers for those changes occuring in patients' brains? Are there some features common to many patients and ALS phenotypes?
These questions could not be evaluated here due to the limited number of observations, but the results obtained here are suggesting that it would be worth to go deeper into this analysis acquiring a larger EEG dataset to explore those aspects.

Appendix A

Appendix

A.1 Welch's PSD parameters

Figure A.1 shows the effects of different choices of parameters in PSD computation with Welch's method. This method implies the computation of power spectral densities over windows of fixed length and then the average of the resulting PSDs to obtain the final spectrum.
It is particularly useful and commonly adopted with nonstationary signals as EEG.
As we can see from the figure, choosing a too short window length (2 s) brings to a smooth PSD, due to the higher number of segments that are averaged. On the other hand, a too long window returns a signal that is noisy and with larger confidence bounds.
For these reasons, applying a trade off between PSD's complexity and the precision needed, the choice of parameters applied to this analysis is of a window of length 5 s with an overlapping period of 2 s.

A.2 Features values variability

Figures A.2 and A.3 represent the variability of features values computed on channel Cz for Patient 6 and Patient 11respectively.
X axis refers to visits' indexes, y axis refers to features values. Each box corresponds to a visit and shows in red the median value of features computed day wise, while the upper and lower bound of each box represent the 25th and 75th percentiles respectively.

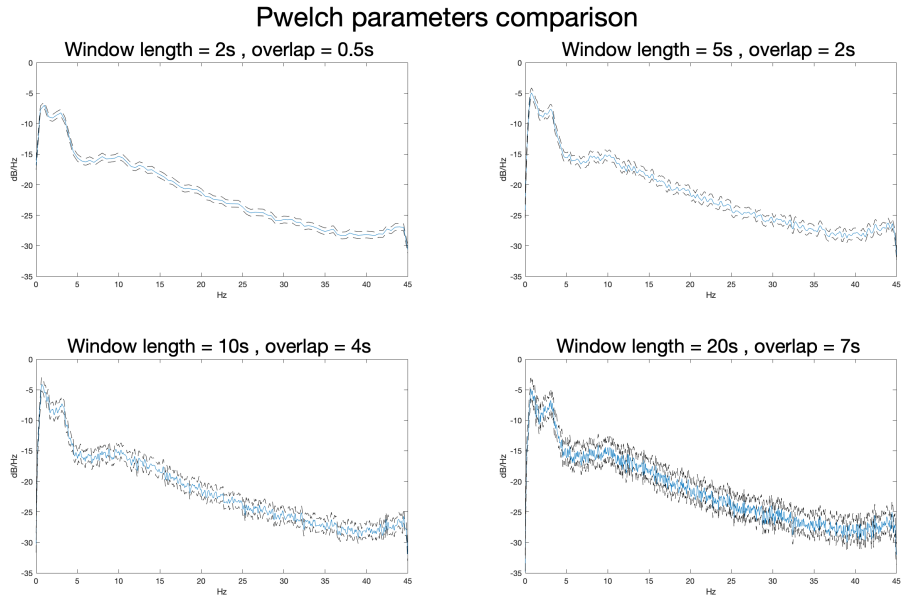


FIGURE A.1: In the figure is shown a comparison between different choices of window length and overlap period used for the computation of PSD through Welch's method. Data used for this example are taken from Patient 4, visit 2.
In blue is represented the final PSD correspondent to the specified choice of parameters, and in dashed-black are displayed the confidences bounds, with coverage probability to 0.99

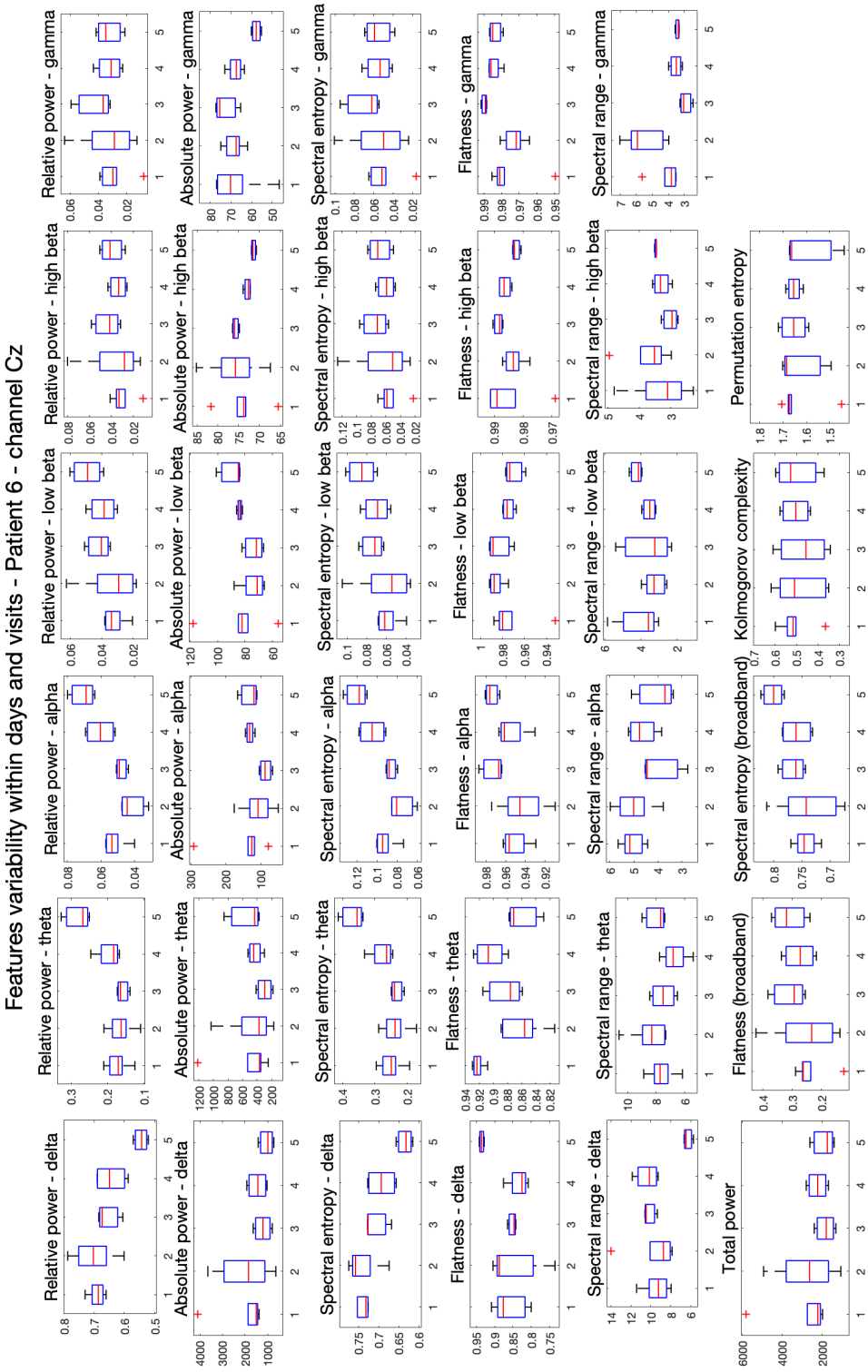


FIGURE A.2: Variability of features values computed on channel Cz for Patient 6, each subfigure represents a different feature with the respective frequency band of interest.

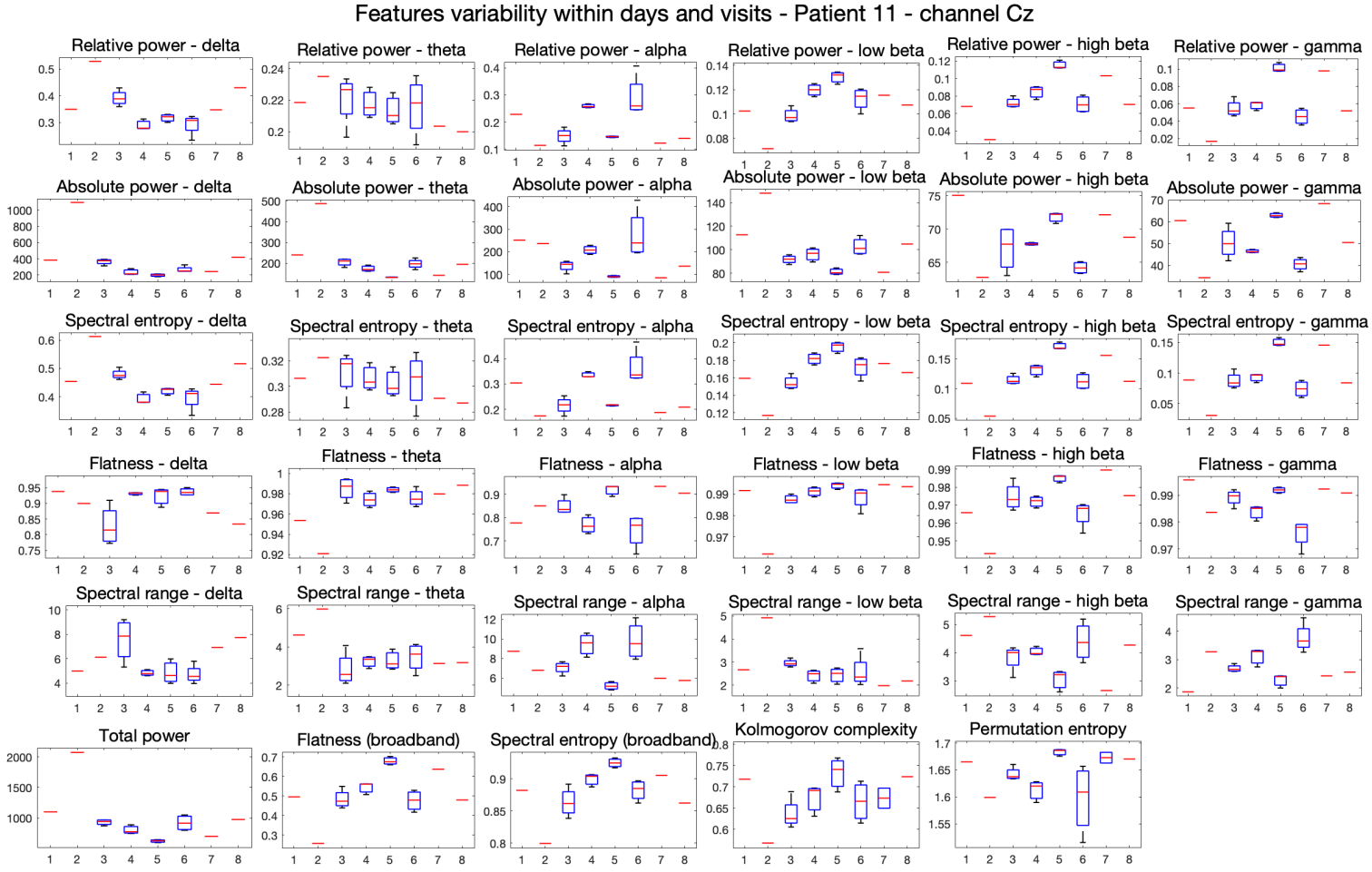


FIGURE A.3: Variability of features values computed on channel Cz for Patient 11, each subfigure represents a different feature with the respective frequency band of interest.

Bibliography

[1] Lewis P. Rowland and Neil A. Shneider. Amyotrophic lateral sclerosis. *New England Journal of Medicine*, 344(22):1688–1700, 2001.

[2] Giancarlo Logroscino, Bryan J Traynor, Orla Hardiman, Adriano Chiò, Douglas Mitchell, Robert J Swingler, Andrea Millul, Emma Benn, Ettore Beghi, et al. Incidence of amyotrophic lateral sclerosis in europe. *Journal of Neurology, Neurosurgery & Psychiatry*, 81(4):385–390, 2010.

[3] Amyotrophic lateral sclerosis. *The Lancet*, 377(9769):942–955, 2011.

[4] PH Gordon, B Cheng, IB Katz, M Pinto, AP Hays, H Mitsumoto, and LP Rowland. The natural history of primary lateral sclerosis. *Neurology*, 66(5):647–653, 2006.

[5] G. Bauer, F. Gerstenbrand, and E. Rumpl. Varieties of the locked-in syndrome. *Journal of Neurology*, 221(2):77–91, Aug 1979.

[6] El escorial world federation of neurology criteria for the diagnosis of amyotrophic lateral sclerosis. *Journal of the Neurological Sciences*, 124:96 – 107, 1994.

[7] Martin R Turner, Matthew C Kiernan, P Nigel Leigh, and Kevin Talbot. Biomarkers in amyotrophic lateral sclerosis. *The Lancet Neurology*, 8(1):94–109, 2009.

[8] Marc Riviere, Vincent Meininger, Phillipe Zeisser, and Theodore Munsat. An analysis of extended survival in patients with amyotrophic lateral sclerosis treated with riluzole. *Archives of neurology*, 55(4):526–528, 1998.

[9] Mitchell JD Lyon M Miller, RG, Moore D Moore, DH, and M Lyon. Riluzole for amyotrophic lateral sclerosis (als)/motor neuron disease (mnd). *Cochrane Database of Systematic Reviews*, (2), 2002.

[10] Analysis of electroencephalography (eeg) signals and its categorization–a study. *Procedia Engineering*, 38:2525–2536, 2012.

[11] Resting state eyes-closed cortical rhythms in patients with locked-in-syndrome: An eeg study. *Clinical Neurophysiology*, 121(11):1816–1824, 2010.

[12] R Mai, D Facchetti, A Micheli, and M Poloni. Quantitative electroencephalography in amyotrophic lateral sclerosis. *Electroencephalography and clinical neurophysiology*, 106(4):383–386, 1998.

[13] Case series: Slowing alpha rhythm in late-stage als patients. *Clinical Neurophysiology*, 129(2):406–408, 2018.

[14] V. Jayaram, N. Widmann, C. Förster, T. Fomina, M. Hohmann, J. M. vom Hagen, M. Synofzik, B. Schölkopf, L. Schöls, and M. Grosse-Wentrup. Brain-computer interfacing in amyotrophic lateral sclerosis: Implications of a resting-state eeg analysis. pages 6979–6982, Aug 2015.

[15] Yalda Shahriari, Theresa M Vaughan, Lynn McCane, Brendan Z Allison, Jonathan R Wolpaw, and Dean J Krusienski. An exploration of bci performance variations in people with amyotrophic lateral sclerosis using longitudinal eeg data. *Journal of neural engineering*, 2019.

[16] Bahman Nasseroleslami, Stefan Dukic, Michael Broderick, Kieran Mohr, Christina Schuster, Brighid Gavin, Russell McLaughlin, Mark Heverin, Alice Vajda, Parameswaran M Iyer, et al. Characteristic increases in eeg connectivity correlate with changes of structural mri in amyotrophic lateral sclerosis. *Cerebral Cortex*, 29 (1):27–41, 2017.

[17] Stefan Dukic, Roisin McMackin, Teresa Buxo, Antonio Fasano, Rangariroyashe Chipika, Marta Pinto-Grau, Emmet Costello, Christina Schuster, Michaela Hammond, Mark Heverin, Amina Coffey, Michael Broderick, Parameswaran M. Iyer, Kieran Mohr, Brighid Gavin, Niall Pender, Peter Bede, Muthuraman Muthuraman, Edmund C. Lalor, Orla Hardiman, and Bahman Nasseroleslami. Patterned functional network disruption in amyotrophic lateral sclerosis. *Human Brain Mapping*.

[18] A Ramos Murguialday, J Hill, M Bensch, S Martens, S Halder, Femke Nijboer, Bernhard Schoelkopf, N Birbaumer, and A Gharabaghi. Transition from the locked in to the completely locked-in state: a physiological analysis. *Clinical Neurophysiology*, 122(5):925–933, 2011.

[19] Jacobo Diego Sitt, Jean-Remi King, Imen El Karoui, Benjamin Rohaut, Frederic Faugeras, Alexandre Gramfort, Laurent Cohen, Mariano Sigman, Stanislas Dehaene, and Lionel Naccache. Large scale screening of neural signatures of consciousness in patients in a vegetative or minimally conscious state. *Brain*, 137(8): 2258–2270, 2014.

[20] Yang Bai, Xiaoyu Xia, and Xiaoli Li. A review of resting-state electroencephalography analysis in disorders of consciousness. *Frontiers in Neurology*, 8:471, 09 2017.

[21] Yvonne Höller, Aljoscha Thomschewski, Jürgen Bergmann, Martin Kronbichler, Julia S Crone, Elisabeth V Schmid, Kevin Butz, Peter Höller, Raffaele Nardone, and Eugen Trinka. Connectivity biomarkers can differentiate patients with different levels of consciousness. *Clinical Neurophysiology*, 125(8):1545–1555, 2014.

[22] Andrea Kübler and Niels Birbaumer. Brain–computer interfaces and communication in paralysis: Extinction of goal directed thinking in completely paralysed patients? *Clinical neurophysiology*, 119(11):2658–2666, 2008.

[23] Y. Maruyama, N. Yoshimura, A. Rana, A. Malekshahie, A. Tonin, A. J. Gonzalez, N. Birbaumer, and U. Chaudhary. Eeg in the completely locked-in state. 2019.

[24] MATLAB. *update 3 (R2018b)*. The MathWorks Inc., Natick, Massachusetts, 2018.

[25] Arnaud Delorme and Scott Makeig. Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis. *Journal of neuroscience methods*, 134(1):9–21, 2004.

[26] Importing channel location with eeglab. URL `https://sccn.ucsd.edu/wiki/A03:_Importing_Channel_Locations#Importing_channel_location`.

[27] Tzyy-Ping Jung, Scott Makeig, Colin Humphries, Te-Won Lee, Martin J Mckeown, Vicente Iragui, and Terrence J Sejnowski. Removing electroencephalographic artifacts by blind source separation. *Psychophysiology*, 37(2):163–178, 2000.

[28] Scott Makeig, Anthony J Bell, Tzyy-Ping Jung, and Terrence J Sejnowski. Independent component analysis of electroencephalographic data. In *Advances in neural information processing systems*, pages 145–151, 1996.

[29] Christoph Bandt and Bernd Pompe. Permutation entropy: a natural complexity measure for time series. *Physical review letters*, 88(17):174102, 2002.

[30] F Kaspar and HG Schuster. Easily calculable measure for the complexity of spatiotemporal patterns. *Physical Review A*, 36(2):842, 1987.

[31] Jacob Ziv and Abraham Lempel. A universal algorithm for sequential data compression. *IEEE Transactions on information theory*, 23(3):337–343, 1977.

[32] Lian-Yi Zhang and Chong-Xun Zheng. Analysis of kolmogorov complexity in spontaneous eeg signal and it's application to assessment of mental fatigue. pages 2192–2194, 2008.

[33] Sebastian Berger, Gerhard Schneider, Eberhard Kochs, and Denis Jordan. Permutation entropy: Too complex a measure for eeg time series? *Entropy*, 19(12):692, 2017.

[34] Thomas Burns and Ramesh Rajan. Combining complexity measures of eeg data: multiplying measures reveal previously hidden information. *F1000Research*, 4, 2015.

[35] Mike X Cohen. Data analysis lecturelets by mike x cohen.

[36] Aihua Zhang, Bin Yang, and Ling Huang. Feature extraction of eeg signals using power spectral entropy. 2:435–439, 2008.

[37] Henry B Mann. Nonparametric tests against trend. *Econometrica: Journal of the Econometric Society*, pages 245–259, 1945.

[38] Steven A. Dressing Donald W. Meals, Jean Spooner and Jon B. Harcum. Statistical analysis for monotonic trends, tech notes 6. 2011. URL `https://www.epa.gov/polluted-runoff-nonpoint-source-pollution/nonpoint-source-monitoringtechnical-notes`.

[39] Jan Terje Kvaløy, Bo Henry Lindqvist, and Hakon Malmedal. A statistical test for monotonic and non-monotonic trend in repairable systems. In *Proc. European Conference on Safety and Reliability–ESREL*, pages 1563–1570. Citeseer, 2001.

[40] CS Herrmann and T Demiralp. Human eeg gamma oscillations in neuropsychiatric disorders. *Clinical neurophysiology*, 116(12):2719–2733, 2005.

[41] Rubika Balendra, Ashley Jones, Naheed Jivraj, I Nick Steen, Carolyn A Young, Pamela J Shaw, Martin R Turner, P Nigel Leigh, Ammar Al-Chalabi, UK-MND LiCALS Study Group, et al. Use of clinical staging in amyotrophic lateral sclerosis for phase 3 clinical trials. *J Neurol Neurosurg Psychiatry*, 86(1):45–49, 2015.