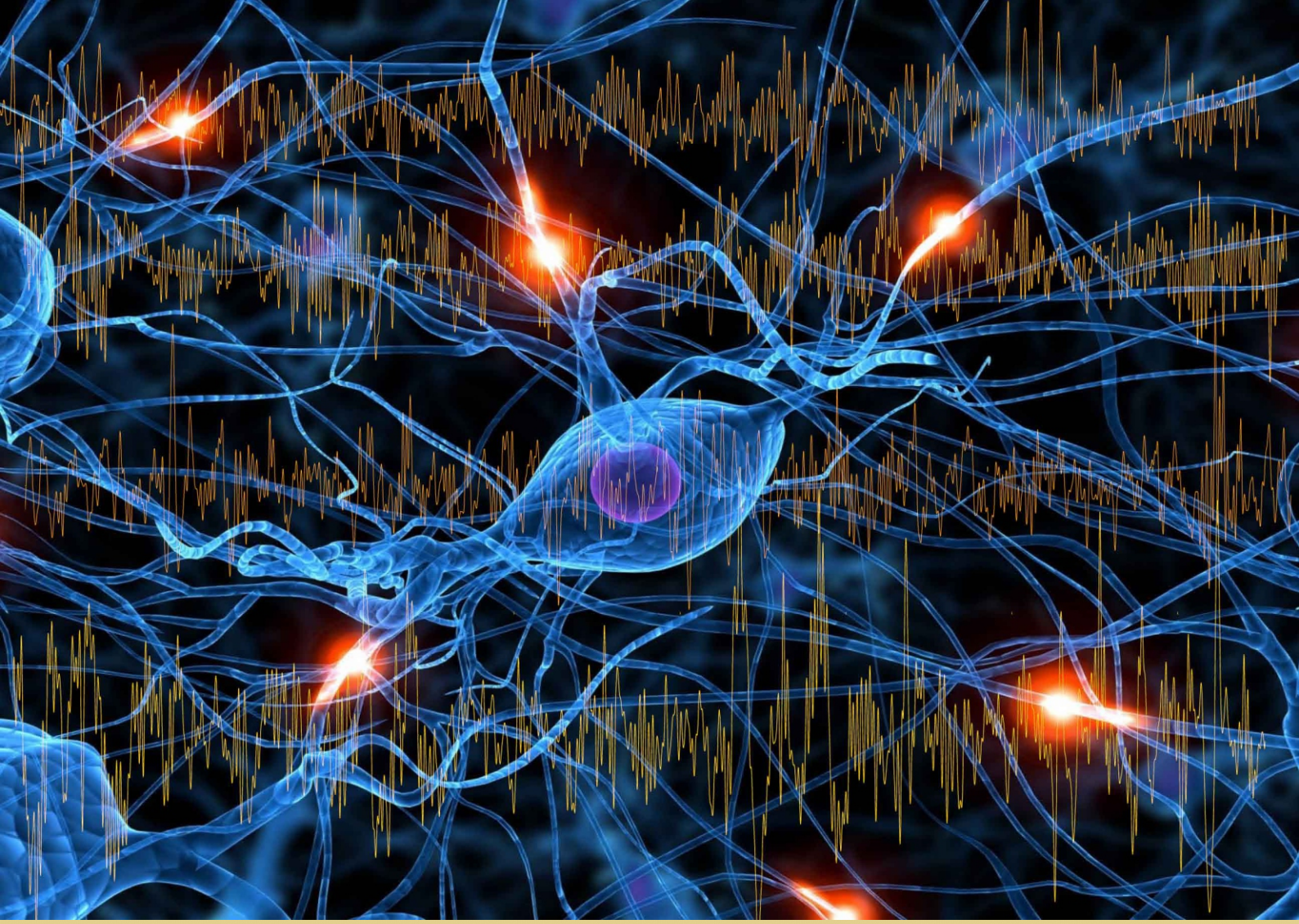


Fátima Machado

Automatic sleep staging based on classification methods

Thesis submitted to the University of Coimbra in compliance with the requisites for the degree of Master in Biomedical Engineering

October, 2015

· U C ·

Universidade de Coimbra

• U C •

FCTUC FACULDADE DE CIÊNCIAS
E TECNOLOGIA
UNIVERSIDADE DE COIMBRA

Fátima Machado

Automatic sleep staging based on classification methods

Dissertação apresentada à Universidade de Coimbra para cumprimento dos requisitos necessários à obtenção do grau de Mestre em Engenharia Biomédica

Thesis submitted to the University of Coimbra in compliance with the requisites for the degree of Master in Biomedical Engineering

Supervisors:
Professor César Teixeira

Coimbra, 2015

This thesis was developed in collaboration with:

Centro de Informática e Sistemas da Universidade de Coimbra

"All men dream: but not equally. Those who dream by night in the dusty recesses of their minds wake in the day to find that it was vanity: but the dreamers of the day are dangerous men, for they may act their dreams with open eyes, to make it possible."

Thomas Edward Lawrence

Abstract

During the sleep the brain generates different types of waves depending on the brain stage. To characterise these brain states, the two structures exist: the macrostructure and microstructure. The macrostructure is composed by five sleep stages (designated by N3, N2, N1, REM, W) whose classification is based on the present wave types. The microstructure is characterised by transitional states and the Cyclic Alternating Pattern (CAP) is an example of it. CAP is a periodic cerebral activity prevalent during NREM sleep-stage and composed by A-phases (A1, A2 or A3) and B-phases.

The visual scoring of macro- and microstructure are important elements for the diagnosis and prognosis of some diseases. Although this task of both is a time consuming process, which demands automatic scoring.

This thesis proposes different classifications methods (discriminate classifiers, k-NN and SVM) to detect automatically the sleep stages and A-phases. The classifiers are validated with a dataset that comprise 30 patients. For sleep stages the better model is SVM which obtained an accuracy off 72%, the sensitivities for the each sleep stage are 62%, 54%, 73%, 83% and 69% for W, N1, N2, N3 and REM stages. Regarding the CAP staging the best classifier method is also SVM with an accuracy of 71%, the sensitivities are 76%, 58%, 44% and 24% for B, A1, A2 and A3, respectively. The prediction of A-phases with the SVM yield the best results to date.

Resumo

Durante o sono o cérebro gera diferentes tipos de ondas, dependo do estado em que se encontra. Para caracterizar estes estados cerebrais existem duas estruturas: a macroestrutura e microestrutura. A macroestrutura é composta por cinco diferentes estados de sono (designados por N3, N2, N1, REM, W), cuja classificação é baseada de acordo com o tipo de onda gerado. A microestrutura é caracterizada por estados transicionais, sendo um exemplo o Padrão Cíclico Alternante (CAP). O CAP é uma actividade cerebral periódica prevalente durando o estado de sono NREM e composto por fases A (A1, A2 e A3) e fases B.

O estadiamento visual da macro e microestrutura são importante para o diagnóstico e prognóstico de algumas doenças. Contudo, esta tarefa é um processo bastante demorado para ambas as estruturas, o que gera uma necessidade de um estadiamento automático.

Esta tese propões diferentes métodos de classificação (classificadores discriminantes, k-NN e SVM) para detectar automaticamente os diferentes estados de sono e as fases A. Estes classificadores são validados com uma amostra composta por 30 pacientes. Para os estados de sono o melhor modelo é o SVM que obtém uma taxa de sucesso de 72% e de sensibilidades para cada estado de sono W, N1, N2, N3 e REM de 62%, 54%, 73%, 83% e 69%, respectivamente. Quanto ao estadiamento do CAP o melhor método de classificação é também o SVM com uma taxa de sucesso de 71% e sensibilidades de 76%, 58%, 44% e 24% para as fases B, A1, A2 e A3. As sensibilidades obtidas por este último método são muito acima das encontradas na literatura até à data.

Acknowledgements

I would like express my sincere gratitude to my advisor Prof. César Teixeira, for the continuous support during this thesis development, for his patience, motivation, enthusiasm, and immense knowledge. I am grateful to Clara Santos, which since we started work together was always available, and taught me about sleep concepts. I thank to Dr. Francisco Sales for give me access to some CHUC data from epileptic patients and to Nuno Vicente for the sleep stages scoring of some patients from CHUC.

To Liga Portuguesa Contra a Epilepsia I thank the financial support for this project.

I also would like also to thanks to the Laboratory for Advanced Computing at University of Coimbra (http://www.lca.uc.pt) for providing computing resources that have contributed to the research results reported within this thesis. Especially to professor Pedro Vieira Alberto and Pedro Almeida for giving access to the computational resources and helping with some problems.

To all my friends, which make part of my life I thanks for all the fun and the supporting. A more than a special thanks is due to Bruno, for all his incredible patience and friendship, for always helping when I need but in particularly during this thesis development.

Lastly, but the most important, to my all my family which have been supporting since I was born, to my grandparents, uncles and cousins. Especially to my parents and sisters which were always by my side, believing and supporting me and made possible being writing this thesis today.

Contents

List of Figures

List of Tables

Abbreviations

AASM	**A**merican **A**cademy of **S**leep **M**edicine
ANN	**A**rtificial **N**eural Network
ASDA	**A**merican **S**leep **D**isorders **A**ssociation
CAP	**C**yclic **A**lternating **P**attern
CV	**C**ross **V**alidation
DA	**D**iscriminant **A**nalysis
DTSTFT	**D**iscrete **T**ime **S**hort **T**ime **F**ourier **T**ransform
EEG	**E**lectro**e**ncephalo**g**ram
EMD	**E**mpirical **M**ode **D**ecomposition
EMG	**E**lectro**m**yo**g**ram
EOG	**E**lectro**o**culo**g**ram
FD	**F**ractal **D**imension
FT	**F**ourier **T**ransform
IMF	**I**ntrinsic **M**ode **F**unction
KMCFW	k-**M**eans **C**lustering **B**ased **F**eature **W**eighting
k-NN	**k**- **N**earest **N**eighbor
LDA	**L**inear **D**iscriminant **A**nalysis
LZC	**L**empel **Z**iv **C**omplexity
MMSD	**M**acro **M**icro **S**tructure **D**escriptor
MRMR	**M**inimum **R**edundancy **M**aximum **R**edundancy
NFLE	**N**octurnal **F**rontal **L**obe **E**pilepsy
NREM	**N**on **R**apid **E**yes **M**ovement
PCA	**P**rincipal **C**component **A**nalysis
PSG	**P**oly**S**omno**G**raphy

QDA	**Q**uadratic **D**iscriminant **A**nalysis
REM	**R**apid **E**yes **M**ovement
ROC	**R**eceiver **O**perating **C**haracteristic
SE	**SE**nsitivity
SEM	**S**low **E**ye **M**ovements
ShEnt	**Sh**anon **Ent**ropy
SFFT	**S**hort **F**ast **F**ourier **T**ransform
SL	**S**ynchronization **L**ikelihood
SP	**SP**ecificity
SVM	**S**upport **V**ector Machine
W	**W**akefulness
ZC	**Z**ero **C**rossing

Chapter 1

Introduction

1.1 Contextualisation

The quality of sleep has a great importance for a good cerebral activity. During sleep process the brain passes through different states that generate different types of brain waves. To organise the sleep two structures emerged: the macrostructure and microstructure. On the one hand, the sleep macrostructure describes the temporal organisation of sleep based on discrete levels that are related with the sleep deepness. On the other hand, sleep microstructure describes the transient and phasic events in the electrical activity of the brain [1–4].

Some macrostructure sleep-states of macrostructure are related to the neurotoxic waste remotion, the maintenance and creation of synapses (which has a creation impact on cognitive performance). Besides, many sleep disorders, that affect normal physical, mental, social and emotional functioning are detected analysing the brain stages during the sleep [1]. The microstructure can also provide complementary information of clinical importance. An important example of a microstructure phasic event is the Cyclic Alternating Pattern, whose rate has been found to be related with some pathologies and, in some cases, useful for the diagnosis and treatment of a disease. For instance, high rates are present in individuals with obstructive sleep apnea syndrome, the upper airway resistance syndrome and epilepsy [5, 6].

1.2 Motivation

About 150 million people worldwide are currently suffering from sleep problems. [1].

Sleep macrostructure and microstructure analysis are, nowadays, used as a diagnosis and prognostic tool. Besides, the visual scoring of sleep stages and micro-events is an exhaustive and time consuming process. In the last few years, the interest in automatic sleep staging as well as micro-events have been growing. As a consequence, a lot of models with different methodologies have been developed. Although, the performance of automatic staging it is not satisfactory enough to replace the visual scoring. In this thesis it is proposed a methodology based on some classification algorithms to both macro- and microstructure staging. At the end our methodology is compared with other algorithms present in literature.

1.3 Objectives

- Analysis of the existing algorithms for scoring the sleep stages and detect CAP sequences.
- Development of a bi-level algorithm for scoring sleep stages and CAP sequences.

Chapter 2

Background Concepts

In 1935, Alfred Loomis, discovered that sleep is not a continuous process, but has different depths that are related to brain waves [7, 8]. The brain waves in deep sleep are characterised by low frequencies (slow waves), whereas light sleep is more affiliated with high frequency waves (fast waves). Different studies along with different sleep classification methods, using different signals, have emerged since [9–11]. In 1968, a group of sleep researchers under the chairmanship of Rechtschaffen and Kales (R&K) developed the first standardised criteria for sleep staging [12] in healthy adult subjects, using polysomnography (PSG). The PSG uses the encephalogram (EEG), electromyogram (EMG), electrooculography (EOG) and electrocardiogram (ECG) to measure the brain, muscle, ocular and cardiac activity. These are the used metrics to divide the sleep into two main stages: Non-Rapid Eyes Movements (NREM) and Rapid Eye Movement (REM). Within NREM four possible stages were defined, namely: NREM1, NREM2 NREM3 and NREM4 from the lighter to the deepest sleep stage. These rules assign a sleep stage for each 30 seconds of the study. Nevertheless, within 30 seconds scored as stage X there can be a few seconds which are not characteristic of X but of the stage Y. These transitions within a stage are designated by micro-events and might provide important physiologic and pathologic information. One of this transitions are arousals, which basically are waves of awake stage within a sleep stage, but with a duration inferior to 15 seconds. The first definition and set of rules to detect an arousal was proposed by the American Sleep Disorders Association (ASDA) [13] in 1992. Different micro-events events were detected since 1968 which are not present in R&K manual, besides, it existed some disagreement about the four sleep stages within NREM due the resemblance between NREM3 and NREM4. Thus, a new

scoring system [14] came out in 2007, proposed by the American Academy of Sleep Medicine (AASM), presenting new sleep scoring rules and considering the micro-events (events whose duration is inferior to 30 seconds within a sleep stage). In this manual only three NREM stages are considered (due the similarity between NREM3 and NREM4 they are joined) designed by N1, N2 and N3. This thesis focus only in EEG which will be explained below, as well as its patterns and related sleep stages and micro-events.

2.1 EEG

The number of neurons in brain is of the order of 10^{11}. Each neuron can be connected to 100000 other neurons, which are organised in a highly interconnected network [15].

To measure the neural activity electrodes are placed over the scalp, the standard disposition is shown in Figure 2.1 along with the electrodes names, which will be used in this thesis. The signal acquired corresponds to the electrical difference between the electrode and a fixed reference, and is also designated by monopolar signal (channel).

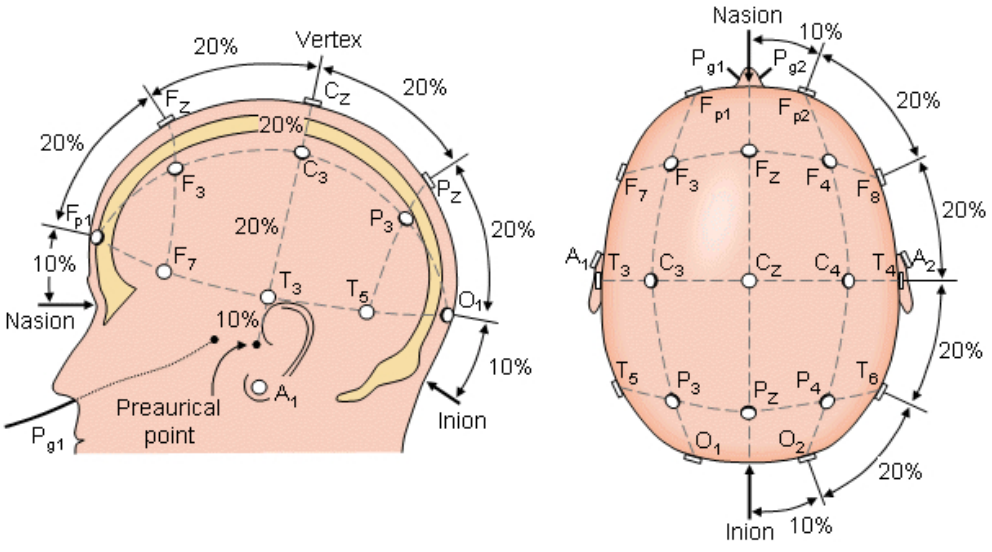


FIGURE 2.1: The international $10-20$ disposition of electrodes placed over the scalp (Adapted from: [15])

The difference between two monopolar channels is called bipolar signal (or channel). The difference is less susceptible to artefacts, since the bipolar processing

can remove mutual interferences on two adjacent electrodes, for instance, EOG or EMG artefacts. Besides, it provides a better spatial resolution compared to monopolar signals since it might reduce the volume conduction [16].

Yet, at any given instant, multiple neurons are activated, so each electrode measures a sum of the electrical currents that are being generated near it. The major drawback of scalp EEG is the attenuation effects of bones and tissues. This effect can be simulated by an low-pass filter with a determined cut-off frequency.

This measurement technique have been widely used for the diagnosis and treatment of several diseases, like epilepsy, since it gives the brain functioning state. It have been also used to study the different stages of brain during sleep.

Each sleep stage is characterised by a specific EEG pattern and some physiological changes like heart rate, muscle tone or breathing rate.

To score the macro and microstructure the recommended EEG monopolar channels are $F_4 - A_1$, $C_4 - A_1$ and $O_2 - A_1$, in the case of any of this channels is not available the alternatives derivations are $F_z - C_z$, $C_z - O_z$ and $C_4 - A_1$ (the configuration can be seen in Figure 2.1). Some sleep elements identify the sleep stages and micro events, and are used in the course of these thesis [1] It is also important to define that the brain waveforms can be subdivided into bandwidths known as delta ($\delta = 1 - 4$ Hz), theta ($\theta = 4 - 8$ Hz), alpha ($\alpha = 8 - 13$ Hz), sigma ($\sigma = 13 - 16$ Hz) and beta ($\beta = 16 - 35$ Hz), which along this thesis, will be designated by conventional frequency bands [17].

2.2 Sleep macrostructure

Sleep scoring is performed considering sequential epochs of 30 seconds (since the start of the study). For each epoch, a stage is assigned and in case of two or more stages coexist in the same epoch, the stage present in the majority prevails. According to the latest sleep-staging rules the possible sleep stages are REM, N1, N2, N3 and, if the subject is awake, the corresponding stage is assigned the letter W (Wakefulness).

An example of brain waves for each sleep stage is represented in Figure 2.2.

[1] to facilitate the reading they are described along the thesis but also in Appendix B.

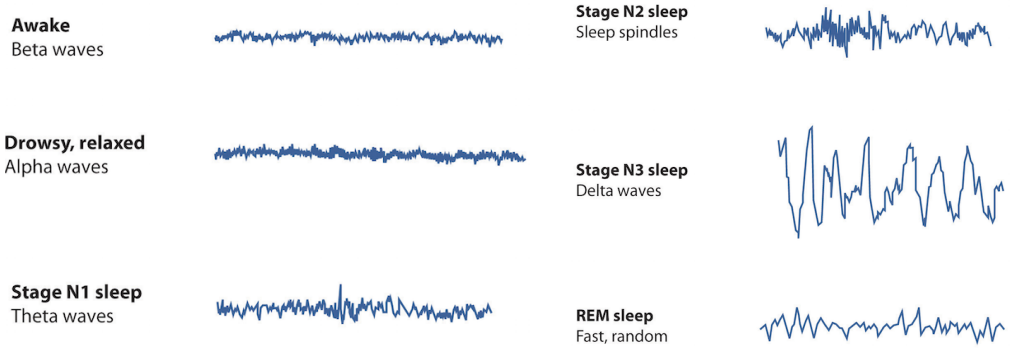


FIGURE 2.2: Examples of the characteristic brain waves for each sleep stage according to [14] scoring rules. (Adapted from: [18])

2.2.1 Wakefulness (W)

The physiology activity, in this stage, is the standard for an awakened person. Analysing the EEG, the W stage is identified when more than 50% of the epoch has alpha rhythms (trains of sinusoidal 8−13 Hz activity recorded over the occipital region with the eyes closed, attenuating with opening of the eyes) over occipital region. If there is no alpha rhythm but instead eye blinking, or if high muscle tone is detected, the epoch is also considered a W stage [14].

2.2.2 NREM stage 1 (N1)

Usually, in the absence of any pathology, when a person starts to sleep the first stage is N1. The alpha rhythm present in W state is attenuated and replaced by low amplitude and mixed frequency activity for more than 50% of the epoch. However, some individuals do not generate alpha rhythms. In this case, an epoch is assigned as N1 if most of frequencies present are in the range of $4-7$ Hz with slowing of background frequencies by more than 1 Hz, or if vertex sharp waves (sharply contoured waves with duration less than 0.5 seconds maximal over the central region and distinguishable from the background activity), or/and slow eye movement are present.

The normal duration of this stage is $1-7$ minutes in the initial cycle, constituting $2-5\%$ of total sleep. This stage is highly sensitivity to noise, which means that is easy to waken a person in this sleep stage.

Physiologic changes take place in this stage, the breathing becomes slow and even, heartbeat becomes regular, blood pressure falls, and blood flow to the brain reduces [14].

2.2.3 NREM stage 2 (N2)

In this stage the low frequencies gain relevance, although the delta activity should be present in less that 20% of the epoch. An epoch is considered a N2 if at least one K-complexes or sleep spindles unassociated with the arousals (i.e. a sudden frequency shifts toward faster rhythms) appear in it. The K-complex is a well-delineated negative sharp wave immediately followed by a positive component standing out from background EEG, with total duration > 0.5 seconds, usually maximal in amplitude when recorded using frontal derivations. The sleep spindle is a train of distinct waves with frequency $11 - 16$ Hz (most commonly $12 - 14$ Hz) with a duration > 0.5 seconds, usually with a maximal in amplitude using central derivations. Figure 2.3 shows these two sleep elements [14].

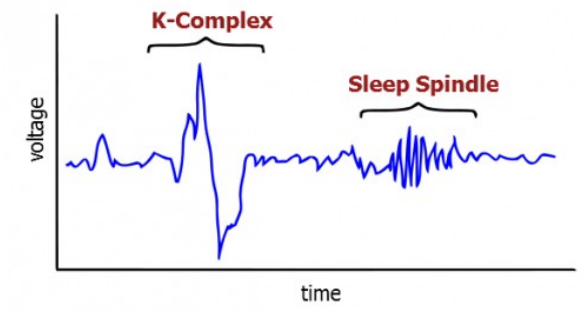


FIGURE 2.3: Example of a K-complex followed by a sleep spindle (Adapted from: [19])

After an epoch is scored as N2 the following epochs will also be classified as N2 unless two situations occur: when is a transition to other sleep stages (W, N2 or REM) or if an arousal not proceed by a K-complex or sleep-spindle appear, in the latter sleep stage changes to N1. This stage can occupy between $45 - 55\%$ of the total sleep episode. An individual in N2 requires more intense stimuli than in N1 to awaken [14].

2.2.4 NREM stage 3 (N3)

This is the most deep sleep stage, and is characterised by high-amplitude slow waves. For an epoch to be scored as N3, the slow wave activity must occupy more than 20% of the epoch. The sleep spindles can also be present in this stage, yet the same is not applied for eye movement which is not typically seen during this stage [14].

2.2.5 REM

In REM sleep the cerebral activity suddenly becomes much more active. REM sleep is distinguishable from NREM sleep by changes in physiological states, including rapid eye movements. In the absence of any pathology, the entrance in this stage leads to an increase of blood pressure, the breathing becomes irregular oxygen consumption increase and the muscle groups including submental muscles (muscles of chin and neck) have no activity.

The EEG waves observed have low amplitude and are irregular, i.e. they have mixed frequencies. In EMG the amplitude is low due to muscle paralysis, and in EOG the rapid eye movements could be detected. This last event is the one that most characterises this stage.

An epoch should be scored as a stage REM sleep, even in the absence of rapid eye movements, for epochs following one or more epochs of sleep stage, if the EEG continues to show low amplitude, mixed frequency activity without K-complexes or sleep spindles and the chin EMG tone remains low [14].

2.3 Sleep microsctruture

The term microstructure refers to EEG features below the time dimension of the conventional 20–30 seconds scoring epoch. The first microstructure defined and proved to be related with some pathological conditions was the arousal.

2.3.1 Arousal

The arousals are transient and generally do not result in behavioural awakening, reoccurring in some conditions as often as once per minute. They are defined as "an abrupt shift in EEG frequency, which may include theta, alpha and/or frequencies greater than 16 Hz but not spindle" [14]. This event must last at least three seconds preceded by ten seconds of stable sleep, the maximum duration of an arousal is 15 seconds.

The arousal origin can be extern or intern and can be associated with unstable sleep conditions [20]. Spontaneous arousals are an intrinsic component of physiological sleep and it can be a measure of the sleep quality [21]. Arousal and awakings are responsible for sleep fragmentation, and the sleep fragmentation is correlated with an increase in daytime sleepiness, in some extreme cases a night with high sleep fragmentation or sleep depravation cause similar effects in the daytime sleepiness [20, 22].

The arousals are associated with activation of neurovegetative functions (heart rate, blood pressure), and it was noticed that phasic slow wave activities during NREM (in particular K-complexes and delta bursts) also have an increase of neurovegetative functions which can trigger muscle activity [23]. However, in spite of their ascertained arousal-equivalent role, the rules proposed by ASSM [14] do not contemplate the K-complexes and delta bursts as arousal elements if it is not preceded by a frequency shift.

The K-complex is considered an elementary expression of arousal during sleep. It is a spontaneous marker of NREM sleep, specially of stage N2, but it also may be triggered by sensory stimulation. Neurophysiological evidence indicates that all types of K-complexes are accompanied by an increase of sympathetic activity, consequently they are considered as momentary arousals. However, it has been demonstrated that alpha arousals lead to significantly different increases in systolic and diastolic blood pressure than K-complexes, which means that all arousal events are different in terms of strength and duration [24].

The cellular behaviour underlying K-complexes are similar to the ones that characterise vertex waves. The only difference is that the two phenomena occur at different stages of the sleep process. The isolated appearance of vertex potentials during early sleep stages coincide with the onset of the slow oscillation, and with

the relatively lower synchronisation of the cortical network. As the network becomes more and more synchronised, and the slow oscillation spreads coherently over larger territories, vertex potentials become ampler and are recognised as K-complexes [25].

2.3.2 Cyclic alternating patterns

Cyclic alternating patterns (CAP) was first detected in patients in comatose [26] and it was used for diagnosis and prognostic of this disorder. A CAP sequence is composed by a succession of CAP cycles each with two elements: an A and B-phase. The A-phases in lighter stages of comas are closely related to hyperventilation and increase of pulse rate and can be associated with greater muscle activity, restless and cerebral spinal fluid. In contrast, automic and muscles activities are attenuated during the B-phases. Subsequent investigations discover that CAP is a physiologic component of NREM sleep stages, and do not occur, under normal conditions, in REM. Some pathological conditions generate CAP sequences in REM, therefore it can be used as a prognostic of such diseases.

CAP tends to appear associated to some dynamic events of sleep like sleep changes stages, falling asleep and arousal without waking [27].

2.3.2.1 Definition

In 2001, a group of researchers met and wrote down the CAP definition and the rules to score this event. CAP was defined as a periodic EEG activity of NREM sleep characterised by sequences of transient electrocortical events, that are distinct from background EEG activity and reoccur at up to 1 minute intervals [28].

The A-phase of CAP is a transient phenomena translated by an increase amplitude and/or frequency which is clearly distinguishable from background activity. This phase is related to a brain activation including cortical arousal and, for this reason, it is a potential trigger of somatomotor activities. B-phase is the background activity, thus is an EEG indicator of rebound deactivation, which induces somatomotor inhibition and becomes a potential limiting factor for the duration of any body movement during the NREM sleep periods. Each phase could have a

duration between 2 and 60 seconds. The mean duration, in young adults, of CAP sequences is two and half minutes, which contains in average six CAP cycles [29].

A B-phase finishes when an A-phase starts. If an A-phase is not preceded and succeeded for a A-phase in 60 seconds, it is called an isolated A-phase and does not belong to a CAP cycle. All CAP sequences have at least two consecutive CAP cycles, thus three or more consecutive A-phases are required.

An A-phase can be composed by high-voltage slow waves which are manifestations of EEG synchrony and/or low-amplitude fast rhythms are evidence the EEG desynchrony.

Depending on the percentage of each waves types present in the transition epoch, the A-phase is classified into different subtypes (details in Appendix B):

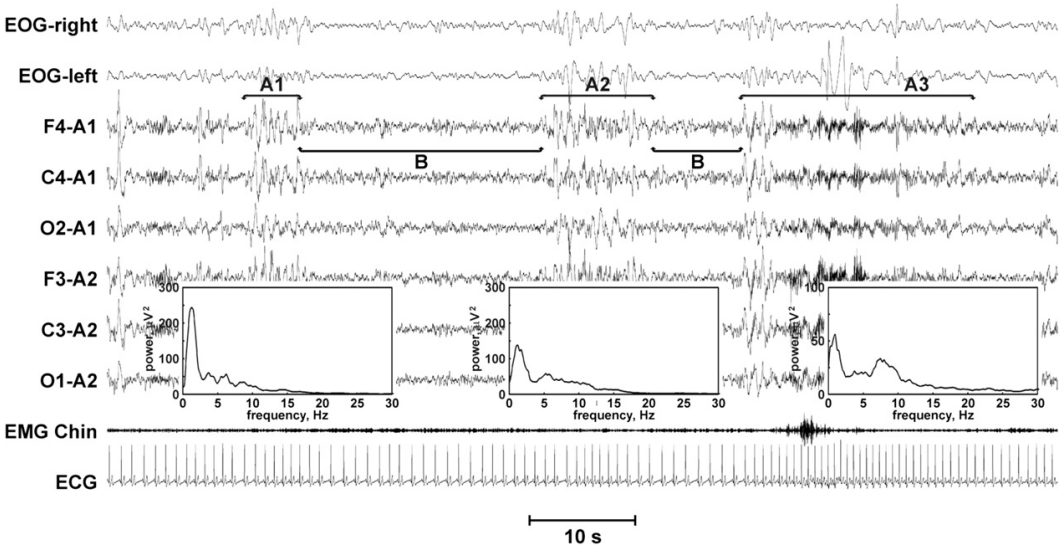


FIGURE 2.4: Example of the different subtypes of A-phases and their respectively spectrogram computed for the bipolar derivation F4-C4. (Adapted from: [29])

- **Subtype A1:** It is composed by high-voltage slow waves. This subtype can be classified by an increase in amplitude of at least 1/3 of the normal background activity. The synchronized EEG pattern must occupy more than 80% of the epoch classified as A1. Normally this event is originated in the frontal region.

  Micro events: Delta bursts, K-complex sequences

- **Subtype A2:** This subtype has elements from subtypes A1 and A3 so, it is composed by a mixture of fast and slow rhythms. The elements from the

subtype A1, which mean high-voltage slow waves, must occupy more than 50% of the length of an entire A-phase.

Micro events: polyphasic bursts

- **Subtype A3:** The rapid low voltage rhythms prevails in this subtypes, it exists an increase in frequency compared to the background. This event is generated in occipital region.

  Micro events: K-alpha, EEG arousals and polyphasic bursts.

  Note: A movement artefact within the CAP sequence is also classified as subtype A3.

Observing Figure 2.4 the maximum frequency for all subtypes are within the delta band, although A1 is the one which registered the high value and having the smooth spectral tail. In others subtypes the spectral tail is more thick, consequently the low and high frequencies have almost the same presence. The A3 have also a maximum, in alpha band, which almost reaches the one present in delta band, which is characteristic of an arousal.

2.3.2.2 Macrostructure and CAP relationship

The NREM portion of the sleep cycle starts with a slow descending branch sloping from the more superficial to the deeper NREM stages. It then continues with a central trough, that represents the deepest stages of the sleep cycle, and ends with a rapid reverse ascending branch, expressed by the more superficial NREM stages that precede REM sleep. Accordingly, the NREM sleep architecture delineates a continuous pattern of build-up (descending branch), maintenance (trough) and resolution (ascending branch) of EEG synchrony. A detailed investigation has ascertained that the spontaneous EEG fluctuations centered on the $20-40$ seconds periodicity of CAP are implicated in the subtle mechanisms that regulate the production and attenuation of slow-wave activities during sleep. In particular, there is evidence that the different components of CAP have a sculpturing effect on the profile of the sleep cycle. The regular EEG oscillations that proceed the transition from light sleep to deep stable sleep are basically expressed by the A1 subtypes. Within the sleep cycle, 90% of the A-phases detected in the descending branches and 92% of the A phases are detected in the troughs are subtypes A1,

while 64% of the A-phases identified in the ascending branches are subtypes A2 (45%) or A3 (19%) [30].

CAP is the EEG translation of unstable sleep which coordinates responses in disparate brain regions and follow the dynamic evolution of the sleep process such as falling asleep, stage shifts, NREM/REM transition and intra sleep awakenings, which are the crucial point for the motor events generation during sleep. The absence of CAP, for more that 60 seconds, is scored as non-CAP and reflects a condition of stable consolidated sleep in which body movement has less chance to appear, due to the overall multi-system stability [4].

These findings indicate that both slow and rapid EEG activating complexes are involved in the structural organisation of sleep.

2.3.2.3 Epilepsy, Sleep and CAP

In 2005 a task force of the International League Against Epilepsy (ILAE) formulated conceptual definitions of *epileptic seizure* and *epilepsy* [31]. It defined *epileptic seizure* as a transient occurrence of signs and/or symptoms due to abnormal excessive or synchronous neuronal activity in the brain, and *epilepsy* as a disorder of the brain characterised by an enduring predisposition to generate epileptic seizures, and by the neurobiological, cognitive, psychological, and social consequences of this condition. The definition of epilepsy requires the occurrence of at least one epileptic seizure.

The epileptic seizures occur because of the malfunctioning of the electrophysiological system of the brain where the normal neuronal network abruptly turns into a hyper-excitable state causing sudden excessive electrical discharge in a group of brain cells (i.e. neurons). Instead of controlled electrical discharges, there is an abrupt and huge surge of energy by the brain cells causing the epileptic seizures. Involvement of cerebral cortex neurons may leads to abnormalities of motor functions causing jerky spasms of muscles and joints [5].

The human knowledge regarding the functioning of the brain is still insufficient to understand the mechanisms behind epilepsy, but it is known that some biological factors, like the biological clocks, influences the development of seizures. The human body is regulated by several rhythms that exert control on behavioural and

physiological processes. Some of these rhythms which influence the seizures development are in some types epilepsies related with the sleep-wake process, sleep-cycle or CAP [32]. These three processes are related with some types of epilepsy, like the nocturnal frontal lobe epilepsy (NFLE) [33]. NFLE is primarily characterised by seizures occurring exclusively or predominantly during sleep. Patients often complain of nocturnal sleep discontinuity and some of them report either difficulty in waking, morning tiredness or excessive daytime sleepiness [6]. Nocturnal attacks are more numerous during non-REM sleep, while their occurrence during REM sleep is quite rare. These findings indicate that non-REM sleep is likely endowed with particular facilitatory properties. During NREM sleep, cerebral electrogenesis tends towards a synchronised functional activity that facilitates the neuronal discharges and the spreading mechanisms of the EEG paroxysmal abnormalities. Physiological desynchronisation of cerebral rhythms operating during REM sleep is responsible for an inhibitory action on the occurrence of epileptic discharges and for the more restricted spatial distribution [4, 33].

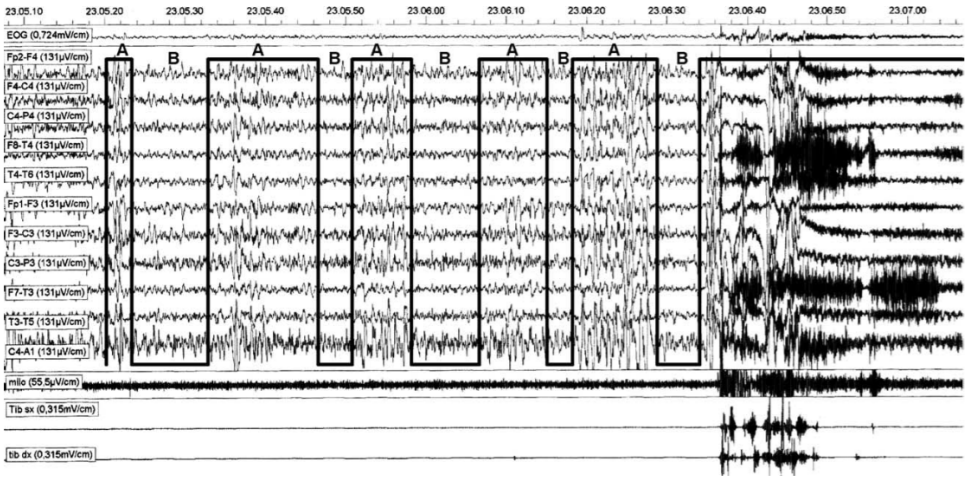


FIGURE 2.5: CAP before epilepsy (Adapted from: [4])

The number of spikes, in NREM, is significantly higher during CAP than during non-CAP. This enhancement is due to the activating properties of A-phase (mainly subtypes A1, around 82%), while B-phase exerts a powerful and prolonged inhibitory action [4]. A1-phase is the most synchronous of the A-phases, therefore it has characteristics which are propitious for the development of seizures.

Seizures cannot be regarded in isolation but require a process of changes in brain dynamics that starts long before its manifestation. Analysis of preictal synchronisations indicates that epileptic seizures do not occur in a behavioral vacuum, but

depends on the functioning of the brain before the seizure occurs. The preictal period may reflect a state of increased susceptibility for pathologic synchronisation, which acts as a route to the seizure. In patients with NFLE, major episodes lasting between ten and 60 seconds are preceded by a prolonged CAP sequence which reflects a condition of sustained arousal instability. An example of a seizure and the EEG that preceded it is shown in Figure 2.5, where it can be seen a high rate of A-phases before the seizure occurs. A powerful activating effect of CAP-related events has been described also for interictal epileptic discharges in A-phase (1.71 spikes/min) than in phase B (0.74 spikes/min). This study also reveals an intermediate activating effect of NCAP (1.18 spikes/min) between phase A- and phase B [32].

Chapter 3

Automatic sleep staging algorithms

All methodologies for automatic sleep staging share general steps, like feature extraction, feature selection, classification and performance evaluation, a briefly description of each is provided in this chapter.

At first the signals are filtered to remove undesired frequencies which are normally associated with muscular and ocular artefacts. The common practice is to filter the signal between $1 - 35$ Hz. A band-pass filter can be also applied to the EEG signal to generate the signal corresponding to one of the conventional frequency bands.

A sliding window is applied to the filtered signal(s), with a determined length and step, dividing the signal into portions with equal length. If the value of step is less than the window length, two or more consecutive windows will share a fraction of EEG, while if it is equal or higher than the window length the data contained in consecutive windows will be completely different (non-overlapping).

After, several transformations can be applied to each signal segment and the result will be a measures, designated as features. Features have a different type of signal information, that originally was not available, or, at least, not in an explicit way.

Finally, a classification method is used to assigned a class to each segment. Supervised classifiers assign a class based on a model built on prior knowledge about the relationship between features and the class labels. Classifiers developed without

any prior information about the relation between the features and the class label are unsupervised classifier.

Different classification algorithms have been widely used in different scientific areas, and they have a good performance. Some examples are artificial neural networks, support sector machines, discriminant analysis, K-nearest neighbors, among others. To evaluate the algorithm some performance measures like sensitivity, specificity and accuracy might be computed.

Different automatic classification algorithms have been emerging in the past years. Some of the most important ones will be presented in this Chapter.

3.1 Automatic macrostructure staging

Doroshenkov et. al [34] proposed an automated system for the classification of sleep stages based on Hidden Markov Model [35] using two bipolar EEG channels (Fpz-Cz and Pz-Oz). An amplitude feature is calculated for a non-overlapping window of 30 seconds for alpha, beta, theta and delta frequency-band signals. The highest accuracy is obtained for the REM stage with more than 86% of accuracy. However, N1 has a low accuracy of 5%. With this method N1 stage is incorrectly classified as N2 or REM, 48% and 47% of the times, respectively.

Artificial neural networks (ANN) [36] have been broadly used in automatic staging. Schaltenbrand et al [37] uses three polisomnogram signals: EEG (C4-A1), EOG and EMG to build a model for automatic sleep classification. In this study the signals are segmented into epochs of two seconds and frequency features are extracted. The classification system use a ANN trained with the back-propagation algorithm, which obtained a sensitivity superior to 80% in all stages except for W and N1 which have a sensitivity of 70% and 53% respectively. Oropesa et al [38] also uses ANN for sleep classification. From the channel C3 a wavelet decomposition [39] is performed to obtain the the conventional frequency-bands. Each signal is subdivided into epochs of 30 seconds and power and energy features are extracted. In total, 13 features are considered as the input to the neural network with three a layers and trained using the Levenberg-Marquardt algorithm. Although, they only used two patients in this method and, due the lower number of N3, the ANN is only trained to classify the W, REM N1 and N2. The overall

accuracy obtained for the training and test is 98% and 77% respectively. The N1 is the stage with more misclassified epochs, having an agreement around 20%.

Fraiwan et al [40] uses a single EEG bipolar channel (C3-A2 or Fpz-Cz) decomposed in the conventional frequency-bands using wavelets. The different features are computed for each 30-seconds epochs. The classification system used is the Linear Discriminant Analysis (LDA) [41] which obtains accuracies around 80% for all stages except for stage N1, that reaches an accuracy of 70% (mostly of the misclassification of this stage yields REM). Helland et al [42] also proposed a method based on LDA method but only for the stages W, N2, N3 and REM. For the monopolar channel C4-A1 the power for conventional frequency-bands and also the relative power between frequency bands are computed. The respiratory, EMG and ECG signals are also considered. For respiratory and EMG the mean, standard derivation, median, and root mean square standard deviation are computed. For ECG the heart rate variability parameters are computed. The features are extracted for signal windows of three minutes. With these set of features the accuracies obtained for each sleep stages are superior to 80%.

Güneş et al [43] proposed a model for sleep staging based on k-means clustering [44] based feature weighting (KMCFW) combined with k-nearest neighbors (k-NN) [45] or with a decision tree classifier [46]. At first, the signal is segmented in epochs of 30 seconds and for each the spectral features are extracted. From the 129 features collected four are selected based on statistical measures. A space transformation based on KMCFW is applied to the data. Afterwards a classification algorithm is applied, k-NN or a decision tree classifier (developed with the C4.5 algorithm) is applied. If the four features are not submitted to the KMCFW transformation the best result for the total accuracy with this method is 56%, for $k = 40$, whereas with the transformation the best accuracy is 82%, for $k = 30$. With KMCFW and k-NN most of stages have a sensitivity superior to 80%. The same do not applies to N1 and N3 whose sensitivities are 7% and 65%, respectively. The results for the decision tree are inferior to the ones obtained with k-NN.

Another method based on k-NN was proposed by Phan et al [47], which classifies 4-classes of sleep: W, N1/REM, N2 and N3. Due to the similarity between REM and N1 they are grouped. Statistical, frequency and amplitude related features are computed for each non-overlapping window of 30 seconds. The overall accuracy for this 4-class model is 94%, the stage better classified is the W with an accuracy of 98% and the poorly classified ones are the N1/REM with 76%.

Liang et al [48], for example uses a decision tree with EEG, EMG and EOG signals for sleep stage classification. Twelve frequency related features are extracted from EOG EMG and EEG signals. This method presents an overall accuracy of 87%, but N1 only has an accuracy of 35%.

Another model was proposed by Han et al [49] using fuzzy logic [50] to classify sleep stages. In this analysis a single monopolar channel, C3-A2, is used and the relative powers for each conventional frequency bands are computed. In total five features are extracted and considered as the inputs for a Mamdani-type fuzzy classifier in combination with a genetic algorithm. The accuracy of the method is 85% with an accuracy to detect N1 of 80% for all the four subjects studied. Although presenting a good result for N1, only four patients were used in this study. D. Álverez-Estévez et al [51] also used a fuzzy inference system of the Mamdani type to construct a classification model. In this models two centrals EEG derivations are used, C2-A2 and C4-A1 as well as EMG and EOG signals from both EOG electrodes. From the five signals amplitude and energy features are extracted with a moving window os 3-second and 1-second of length and step, respectively. Afterwards, the features extracted are combined resulting in five final features which are used in classification step in the fuzzy classification. The average accuracy, sensitivity and specificity for six patients are 95%, 95% and 95%, respectively.

Support Vector Machine [52] (SVM) has been widely used for sleep classification. For instance, Lajnef et al [53] uses a SVM method in a multichannel analysis. Their main methodology involves feature extraction using linear and non-linear time measures for all monopolar channels (C3, Cz, EOG1, EOG2 and EMG), and frequency features only for EEG channels. In total, 102 features are used related to amplitude, energy and frequency. The features are submitted to an statistical, *t-test*, to compare the mean of each feature across all pairs among the five stages, and he top 32 features are selected. They reach a mean accuracy of 92% and once again the stage with lower performance is in N1, with a 93%, 41% and 87% for specificity, sensitivity and accuracy. Also, Koley et al [54], used SVM for a automatic staging using only one EEG channel. In this method two types of features, amplitude and frequency, are used. A total of 21 features are computed for the C4-A1 channel. The reported results give a accuracy of 96% respectively. No disrepancy for the performance values between the sleep stages are reported by the authors.

Sotelo et al [55] proposed a method based on J-means clustering [56] where the features used are based on entropy. In this paper the set of features is also used for a neural network classification, and an adequate accuracy in N3 and N2, similar in W, is reported. Overall, the results using a neural network and J-means, achieves an accuracy of 80%. Also, the paper evaluates the accuracy of the method in patients from an another database, consequently the neural network shows a significantly drop of performance, while J-means maintains it. Therefore the unsupervised algorithms are promising methodologies for automatic scoring of sleep stages. The methods described in this this section are summarised in Table 3.2 along with the overall performance and individual sensitivities.

3.2 Automatic CAP scoring algorithms

Barcaro et al [57] proposed a simple methodology to detect A-phases. The Macro-Micro-Structure Descriptor (MMSD), described in more detail in section 4.2.3.1 is computed, for the conventional frequency bands. They developed a rule imposing two thresholds values, called: the existence and length thresholds. The segments of MMSD with amplitudes greater than the length threshold, and at least one of the values greater than the existence threshold, are considered as possible A-phases. An example is shown in Figure 3.1 where it can be noticed that some values above the length threshold are not considered as A-phase because the sequence of values above the the length threshold do not include any element above the existence threshold.

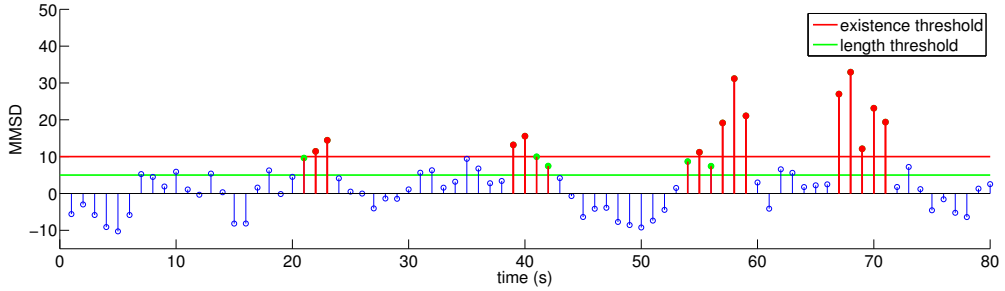


FIGURE 3.1: Example of the model used to classify the A-phases based on two threshold values with MMSD. The A-phases candidates are represented by the vertical red lines. Both coloured bullets, green and red points, represent the points above the length and existence threshold respectively within an A-phase candidate.

Barcaro et al proposed different combinations of the MMSD features to detected the A-phases subtypes. In a paper published in 1998 [57] all the bands are used for the detection when the rule mentioned before is satisfied for at least one of the bands. A different frequency-band combination was proposed in 2010 [58] here, an A-phase is only detected when the rule is satisfied in the delta or/and theta MMSD signals. Additionally, if the rule is also satisfied, at least, in one of the other MMSD signals (α, σ and β bands) the A-phase candidate is considered an A2/A3, otherwise it as considered an A1-phase. The reported results for correctness, specificity and sensitivity are 77% , 90% and 84%, respectively. Without discrimination between A-phases sub-types the performance values are 79%, 81% and 81% respectively. Lastly, in 2004 [59] if the rule is satisfied for the theta, sigma and beta bands it is considered an A2 or A3. If existence threshold are only crossed on MMSD in the delta band, the subtype is considered A1. If the time distance between two subtypes is less than 1.5 seconds they are joined. The reported correctness for the methodology is 83.5% for A-phase detection and 73.7% to distinguish between subtypes.

A proposed model explaining the origin of the EEG collected at the scalp level is outlined by Rosa et al [60]. It simulates activity based on excitatory and inhibitory neural populations interacting with each other and this can be modelled as feedback loops. A schematic representation of this is shown in Figure 3.2. The

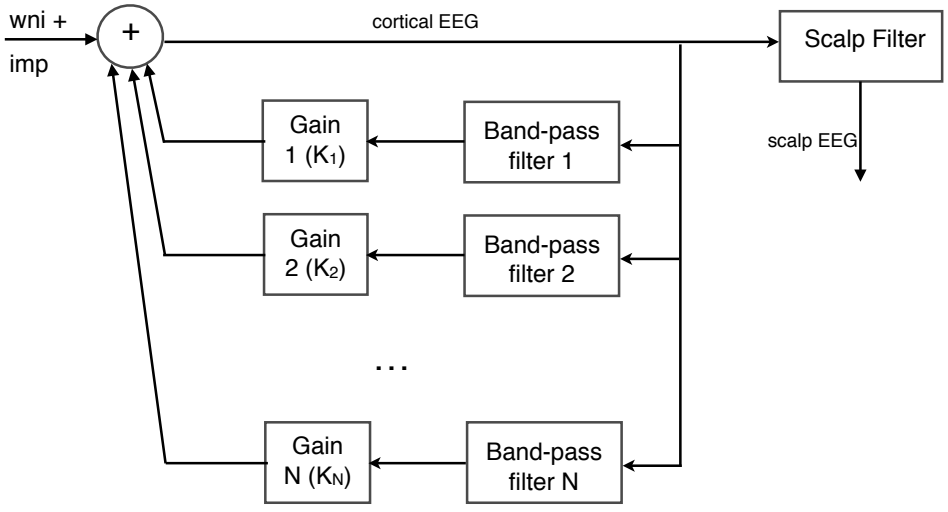


FIGURE 3.2: EEG Generation Model by feedback loops

model include several rhythms generators corresponding to the feedback loops. Each loop is composed by a bandpass filter with the same spectral characteristics

of the EEG rhythms and a loop gain modelling the global effect of the underlying excitatory and inhibitory neuronal populations. The input of the system is a combination of a white noise input (wni), modelling the tonic or background activity, and a random pulse input (imp), representing the afferent sensory stimuli, which corresponds to most of the phasic activity (the reared events). The input is processed by the different rhythms giving origin to cortical EEG[1]. The model can be inverted and from the EEG the reared impulses and the background activity can be estimated. This methodology was used to detect K-complexes [61] and vertex waves, where a true detection rate for all sleep stages of 89% is obtained. However the number of false positives is 49%. Also is was used to detect CAP sequences, by considering four rhythms generators which correspond to the frequency bands (δ, α, θ and σ) [62]. The results presented for this analysis are: a mean correctness of 90%, CAP detectability mean of 95% and a mean sensitivity of 90% [60].

A method based on wavelets and genetic algorithm tuning was proposed by R. Largo [63] to identify the A-phases independently of the subtype. Basically, the signal is decomposed in five different frequency bands, using the wavelets procedure, and for each one the MMSD is computed. The detection of the starts and end points of each A-phase is made, by combining and comparing the MMSD of each band a against threshold. The accuracy reported is 79%.

Stam and van Dijk [64] published a study which defines the synchronization likelihood (SL) between two channels and proposed a method to calculate this value. Synchrony was defined as a measure of the dynamical interdependencies between two channels. At first they proved that this measure is useful for epilepsy seizure prediction. Later, it was applied for A1 detection [65], where the signal is filtered between the frequencies $0.25 - 2.5$ Hz. Afterwards the SL for each possible combination of channels is computed. They concluded that the levels of SL in the frequency range of 0.25–2.5 Hz during sleep have significant fluctuations depending on the appearance of CAP. The subtype A1 appears to be related with a increase in SL, and the difference is more significant in N2. After, this feature was applied to distinguish between A1-phase and B-phase in sleep stage N2 with good results are obtained [66]. However in the other stages this feature did not provide a proper performance to differentiate between A1 and the background or to identify the others subtypes [65, 66].

[1] or intracranial EEG

In 2012, a method for detecting A-phase using SVM and other classifiers was implemented [67]. The model aimed the A-phase detection, without subtypes discrimination, and it used seven features. The features included the MMSD for the five frequency-bands, but instead of the absolute signal amplitude the amplitude squared is considered. The others two features are the *Hjorth activity*, which is the variance of the delta band for a window of three seconds and the EEG variance for each second. In addition to SVM other classifiers are considered: LDA, SVM, Adaboost [68] and ANN. The reported results are shown in Table 3.1

TABLE 3.1: Performance results for different classification methods for A phase detection

method	Sensitivity (%)	Specificity (%)	Accuracy (%)
Linear Discriminant	72.5 ± 10.9	86.6 ± 6.3	84.9 ± 4.9
SVM	70.1 ± 8.6	84.0 ± 11.1	81.9 ± 7.8
Adaboost	68.5 ± 6.7	79.3 ± 9.4	79.4 ± 5.5
Neural Network	72.9 ± 7.5	82.3 ± 7.1	81.5 ± 6.4

Analysing Table 3.1, the LDA method, which requires less computational resources than the other classification methods, obtains better results for for specificity and accuracy. However this values are only for A-phase detection and not take into account the different A phases subtypes.

TABLE 3.2: Summary of some methods for sleep automatic scoring

ref	year	patients	channels	method	# features	accuracy (%)	sensitivity (%)				
							W	N1	N2	N3	REM
[34]	2007	-	Fpz-Cz + Pz-Oz	Markov model	4	50	51	5	69	78	86
[37]	1996	60	C4-A1+EMG+EOG	ANN	17	76	69	53	89	82	84.92
[38]	1996	2	C3	wavelets+ANN	13	76	74	75	91	-	65
[40]	2010	32	C3-A2 or Fpz-Cz	wavelets+LDA	21	79	93	69	79	72	80
[42]	2010	10	C4-A1	LDA	10	91	83	-	97	94	90
[43]	2010	5	C4-A1	KMCFW+k-NN	4	65	80	7	89	65	81
[47]	2013	4	Fpz-Cz	k-NN	17	86	99	-	89	81	76
[48]	2010	20	C3-A2+EOG+EMG	decision tree	12	86	97	55	87	91	91
[49]	2010	4	C3-A2	GA+fuzzy logic	5	85	86	84		84	86
[51]	2009	6	C2-A1+C4-A1	Mandami type	5	96	-	-	-	-	-
[53]	2015	15	C3,Cz,EOG1,EOG2,EMG	SVM	32	74.8	90	41	70	76	97
[54]	2012	28	C4-A1	SVM	39	92	96	83	89	99	94
[55]	2014	39	Fpz-Cz+Pz-Cz	J-means	34	57	87	15	91	59	34

TABLE 3.3: Summary of some methods for A-phase scoring

ref	year	patients	channels	#features	method	correctness (%)	
						detectability	discrimination
[58]	2002	10	F4-C4	5	threshold	79	77
[59]	2004	10	F4-C4	7	threshold	84	74
[60]	1999	4	C4-A1	5	stochastic algorithm	95	90
[63]	2005	8	One EEG channel	5	wavelets + threshold	79	-
[67]	2012	8		7	LDA	85	-
					SVM	82	-
					Adaboost	79	-
					ANN	82	-

Chapter 4

Material and methods

Two analysis are performed in this thesis: algorithms development to predict A-phases and sleep stage, and a statistical analysis to compare the differences in macro and microstructure in patients with and without epilepsy. The latter aims to analyse the evolution of A-phases before and during a night with seizures, this analysis can be consulted in Annex A. In the previous Chapter some of the automatic methods present in literature are described. Yet, the studies developed, until now, are not sufficient to score the macro and microstructure without a technician revision. In this Chapter it is detailed a new methodology for sleep and A-phases automatic staging, as well as the data used in this thesis.

4.1 Materials

The dataset used in this thesis is available on an online database, called the *CAP Sleep Database* [69]. This database comprises several one-night polysomnographic recordings from different patients with different pathologies. Each polysomnographic record is provided in an European Data Format (EDF) file. Associated with the EDF file is a text file, where the manual macrostructure and A-phases scoring are listed. This database has been used in several studies [2, 58, 59, 63, 67].

Consequently, only patients with nocturnal frontal lobe epilepsy (NFLE), designated in the website as NFLE1-NFLE40, are considered for the development of the algorithms. Some patients, in this range are excluded from the study for two reasons. First, in some patients, the channels required are not present. Second, some

inconsistencies between the scoring and EEG were found on the recordings. The patients used to perform the automatic sleep-elements scoring are in Table 4.1.

Number	**File name**	**Sex**	**Age**	**freq sample**	**% A-phase**
1	NFLE2	F	41	512	13.72
2	NFLE3	M	29	512	10.76
3	NFLE5	F	22	512	15.68
4	NFLE6	F	32	128	11.50
5	NFLE8	M	14	256	12.51
6	NFLE9	F	21	512	13.94
7	NFLE10	M	18	128	14.50
8	NFLE11	M	31	128	12.85
9	NFLE12	F	67	512	12.06
10	NFLE14	M	35	512	15.22
11	NFLE15	F	29	512	13.55
12	NFLE16	F	30	512	19.77
13	NFLE17	M	25	512	16.67
14	NFLE18	M	25	512	11.74
15	NFLE20	F	26	128	13.95
16	NFLE21	M	27	512	18.90
17	NFLE22	F	42	512	21.22
18	NFLE24	M	39	512	12.33
19	NFLE26	M	38	128	12.85
20	NFLE28	F	28	512	11.87
21	NFLE30	F	26	512	10.27
22	NFLE31	M	33	128	12.45
23	NFLE32	F	34	512	14.27
24	NFLE34	M	26	512	13.72
25	NFLE35	M	44	512	14.55
26	NFLE36	F	18	512	12.68
27	NFLE37	M	16	512	12.42
28	NFLE38	M	31	512	23.67
29	NFLE39	M	24	512	11.23
30	NFLE40	F	60	512	22.80

TABLE 4.1: Patients used in statistics analysis and in algorithms development

The recordings were acquired at the Sleep Disorders Center of the Ospedale Maggiore of Parma, Italy, and included in the CAP-Sleep database. The polysomnographic data includes at least three EEG channels (F3 or F4, C3 or C4 and O1 or O2, referred to A1 or A2), two EOG channels, three electromyographic EMG signals, respiration signals and the ECG. For some patients, additional bipolar EEG monopolar channels were included: $Fp1 - F3$, $F3 - C3$, $C3 - P3$, $P3 - O1$ and/or $Fp2 - F4$, $F4 - C4$, $C4 - P4$, $P4 - O_2$.

The macrostructure scoring was annotated according to the R&K rules, while CAP was detected in agreement with Terzano reference atlas [28]. However, as mentioned in Chapter 2, the rules of R&K [12] were replaced by ASSM rules. Therefore, NREM stages 3 and 4 were joined together aiming to obtain a classification according to the actual rules.

4.2 Methods

4.2.1 Overview

The main methodological steps, performed in this thesis for monopolar channel C4-A1, are represented in Figure 4.1. The methodology used for automatic sleep staging and for automatic A phases classification is the same, although, the post-processing rules are different. In the next sections a closer look into the each one of the steps, will be presented.

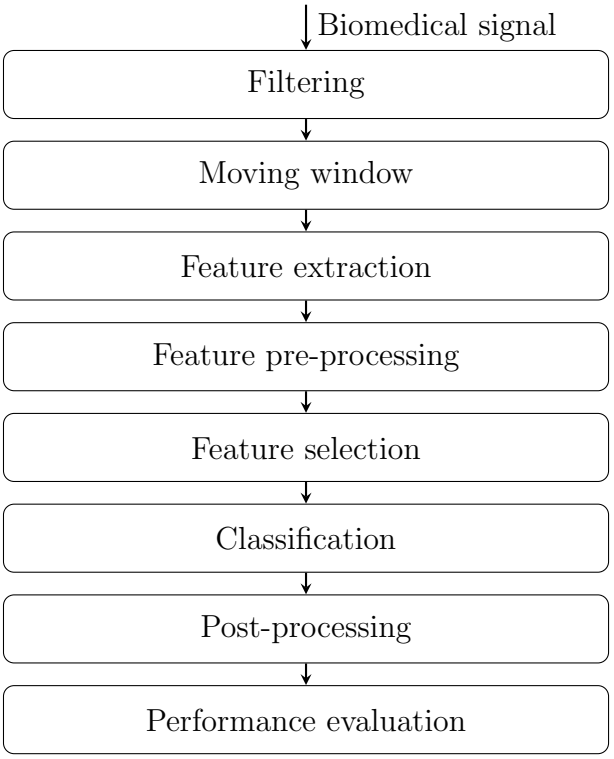


FIGURE 4.1: Methodological steps for automatic sleep staging and A phase staging

4.2.2 Filtering

Delta frequencies are prominent in N3, in N2 both delta and theta are present, N1 is characterised by alpha rhythms, REM sigma and alpha are observed whereas beta identify the awake stage. Regarding the A-phases, the subtype A1 occur within the delta band, A3 in alpha band and A2 in both (since it is a mixture of A1 and A3). For this reason, the signal in these frequency bands contains crucial information for automatic staging, of both sleep stages and A-Phases. Thus, the EEG signal is filtered, using a third-order *Butterworth* filter [70], to obtain the EEG signal filtered in the conventional frequency bands.

4.2.3 Feature extraction

The visual scoring of the macro and microstructure is based on the amplitude and frequency of the signal. Deep sleep stages are characterised by slow high-voltage waves while in lighter stages the signature is the existence of fast low-voltage waves. Therefore, the features extracted must be related with the amplitude, the frequency or both.

Some features extracted from the EEG, will be introduced, that will be used latter for classification.

4.2.3.1 Macro-Micro Structure Descriptor

MMSD is a adimensional and normalised measurement of how the mean amplitude, C, of the activity in a given frequency band φ differs, at a given instant t, from a signal portion, called its background. The MMSD is given by Equation 4.1.

$$\mathrm{MMSD}_{\varphi} = \frac{C_{\varphi,\tau_0}(t) - C_{\varphi,\tau}(t)}{C_{\varphi,\tau}(t)}, \tag{4.1}$$

Where $C_{\varphi,\tau}(t)$ represents the mean amplitude at time t over a certain time interval τ. If the τ is long enough it represents the background signal, while if it is too short it is related with instant activity. It can be proved that if the size of τ is approximately one minute, and φ is one of the conventional EEG frequency bands, the $C_{\varphi,\tau}(t)$ trend is generally related to the sleep macrostructure [57]. An interval

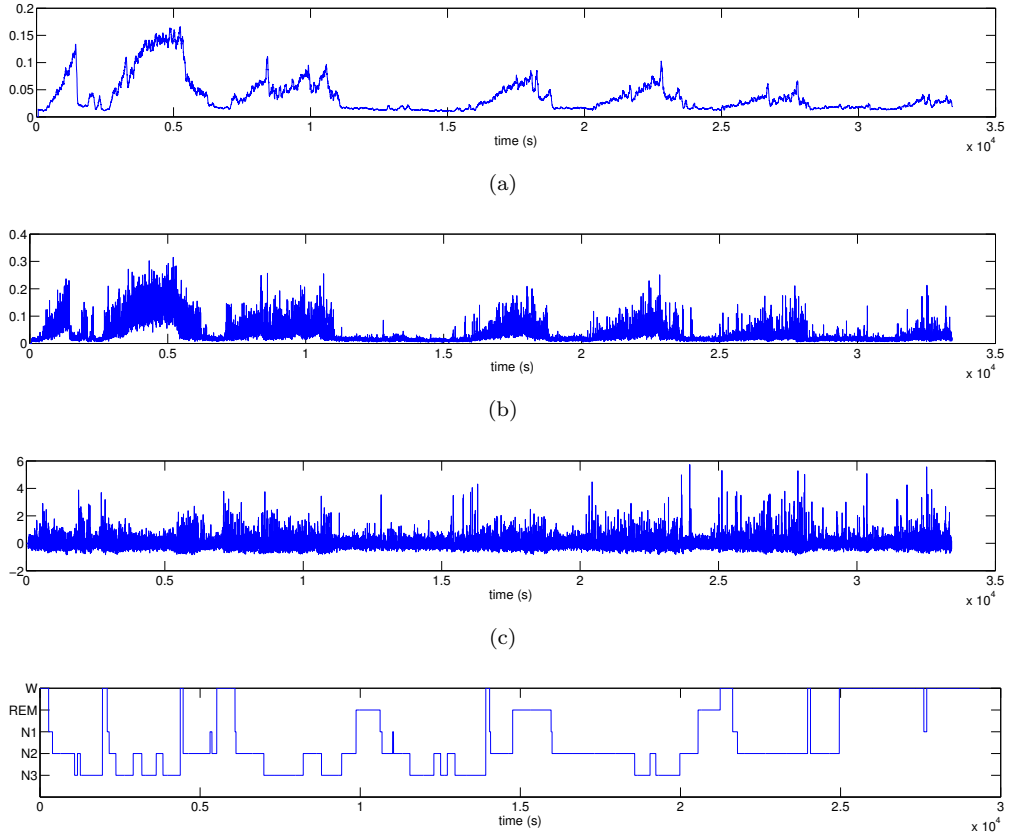


FIGURE 4.2: Evolution of the a) $C_{\delta,\tau}$, b) $C_{\delta,\tau 0}$ and c) MMSD_{δ} d) the corresponding hypnogram throughout a whole night of sleep for the patient NFLE2.

time (τ_0) is used to evaluate the instant activity, $C_{\varphi,\tau_0}(t)$, and it is related with sleep microstructure. It can be shown that the descriptors, MMSD_{φ}, are connected to the microstructure. Consequently, it is possible to identify epochs of transient increase in some frequency bands. This measure is used in some automatic CAP staging studies [57–59].

From the MMSD computation, three features can be obtained, namely: $C_{\tau,\varphi}$, $C_{\tau_0,\varphi}$ and MMSD and they are used in this thesis. Each one of these features are extracted from the signal correspondent to the conventional frequency bands, for a $\tau = 60$ and a $\tau_0 = 2$. A representation of the progress of these three features, for delta band, along with hypnogram [1], are presented in Figure 4.2.

[1] representation of the sleep stages with time

4.2.3.2 Teager Energy Operator

The Teager Energy Operator (TEO) can be defined as a measure of how much energy is required to generate a signal [71]. The TEO, Ψ, can also be viewed as an instantaneous measure of energy [72]. It is a non-linear energy-tracking operator is given, in its continuous form, by:

$$\Psi(x(t)) = \dot{x}(t)^2 - x(t)\ddot{x}(t), \tag{4.2}$$

where $\dot{x}$ is the first derivative of x, and $\ddot{x}$ represents its second derivative. The discrete form is given by:

$$\Psi[x[n]] = x[n]^2 - x[n-1] * x[n+1], \tag{4.3}$$

This operator has been successfully used in several signal processing applications. For example, it was used in speech processing [73–76], and feature extraction from EMG signals [77]. The TEO has been referred as adequate for the identification of some EEG elements which are crucial for the CAP scoring, like sleep-spindles (96.17% sensitivity and 95.54% specificity) [78], and K-complexes (with a rate of 7% false positives and 89% true positives) [79].

TEO is a remarkable feature to detect amplitude or frequencies changes relatively to the background. Consequently one may hypothesise that it is appropriate for A phase detection.

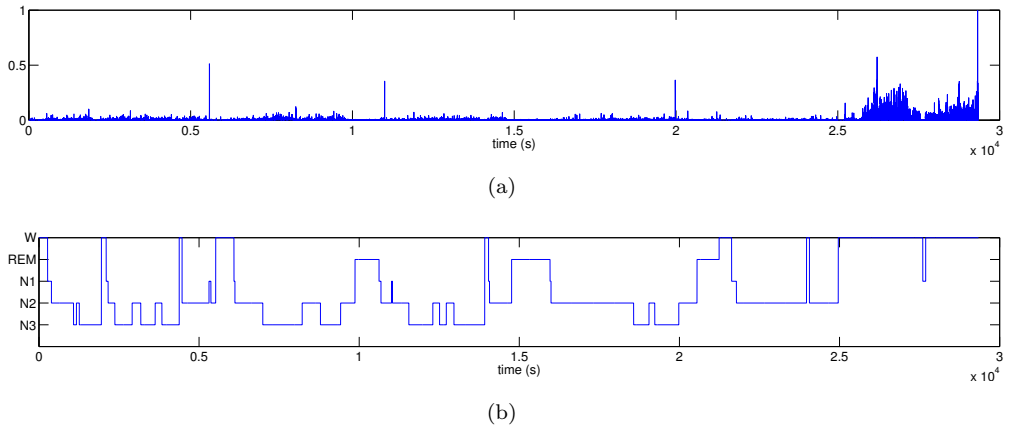


FIGURE 4.3: Evolution of a) normalised TEO feature in the delta band frequency band and b) the corresponding hypnogram over a whole night of sleep for patient NFLE2).

The TEO is computed for each signal sample, in all conventional bands, and for each second it is selected the maximum TEO value to represent it. An example of TEO progression, for delta band, is in Figure 4.3.

4.2.3.3 Zero-Crossing

Zero-Crossing (ZC) rate is a measure of the number of times in a given time interval that the amplitude of a signal passes through the value of zero. ZC is a fast, intuitive and low complex way to obtain information about the signal frequency in a short period of time [80, 81]. Basically, it is a measure of the central-frequency changes of a signal by counting the number of baseline crossings in a fixed time interval [82].

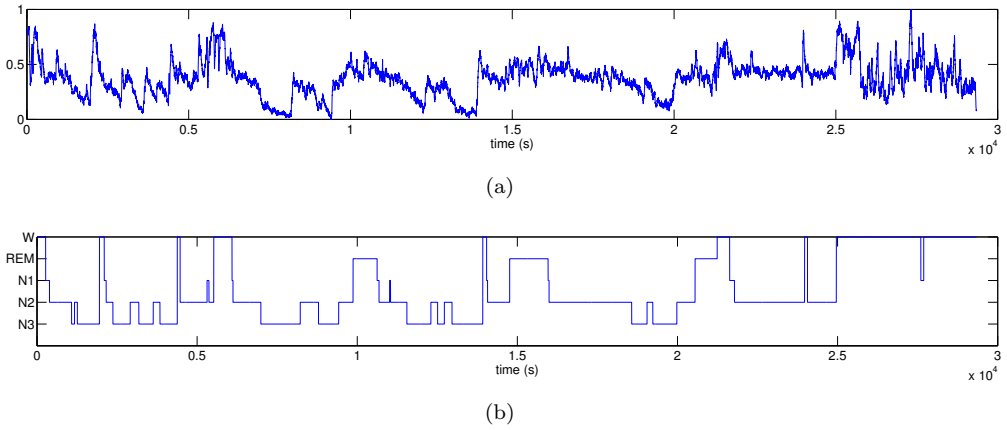


FIGURE 4.4: Evolution of a) ZC feature computed for EEG signal for a whole night of sleep, after a moving average window of 30 seconds was applied and b) the respective hypnogram, for patient NFLE2.

Theoretically, the ZC value will decrease as sleep becomes deeper. This feature was already used to map the sleep of newborns [82], along with another features. In the microstructure automatic scoring it was used in arousal detection [83]. Besides the sleep applications, one of the main areas of ZC applications in speech recognition [84–87].

The ZC rate is computed for moving window with 1-second length, for EEG and all conventional frequency-bands. The ZC on an EEG is shown in Figure 4.4, as well as the respective hypnogram.

4.2.3.4 Lempel-Ziv Complexity

Lempel-Ziv complexity (LZC) is a metric used to evaluate the randomness of finite sequences proposed by Lempel and Ziv, in 1976 [88]. To compute the LZC complexity, $c(n)$, a numerical sequence has to be transformed into a symbolic sequence. One popular approach is to convert the signal, $x(n)$, into a binary sequence, $P = s(n)$, comparing the signal with a threshold, T_d. The points whose value is greater than T_d are converted to 1, otherwise to 0. Usually the median is used as T_d, due of its robustness to outliers. Previous studies have shown that $0 - 1$ conversion is adequate to estimate the LZC in biomedical signals.

At first the $c(n) = 1$ and the P signal is scanned, from left to right, and very time that a new subsequence of consecutive characters is encountered is $c(n)$ increase one unit. The procedure, adapted from [89], is:

1. Being S and Q two sequences of P and SQ the concatenation of S and Q, while sequence $SQ\pi$ is derived from SQ after its last character is deleted (π denotes the denotes the operation of deleting the last character in the sequence). Let $v(SQ\pi)$ denote the vocabulary of all different subsequences of $SQ\pi$. At beginning: $c(n) = 1$, $S = s(1)$, $Q = s(2)$, this $SQ\pi = s(1)$;

2. In general, $S = s(1), ..., s(r)$ and $Q = s(r + 1)$, then $SQ\pi = s(1), ..., s(r)$ if Q belongs to $v(SQ\pi)$, the Q is a subsequence of $SQ\pi$, not a new sequence;

3. Renew Q to be $s(r+1), s(r+2)$ and verify if Q belongs or not to the existence vocabulary $v(SQ\pi)$;

4. Repeat the previous steps until Q does not belong to $v(SQ\pi)$. At this point, $Q = s(r + 1), ..., s(r + i)$ is not a subsequence of $SQ\pi = s(1), ..., s(r + i - 1)$, so increase $c(n)$ by one;

5. Thereafter, S is renewed to be $S = s(1), ..., s(r + i)$ and $Q = s(r + i + 1)$

This procedure is repeated until Q is the last character. Finally, for a sequence of two symbols (0 and 1), the LZC is given by:

$$\mathrm{LZC} = \frac{c(n)}{n} \log_2(n) \tag{4.4}$$

On the one hand, LZC parameter increases with the frequency, noise power present in quasi-periodic signals, signal bandwidth increase. On the other hand, LZC does not evidence any sensibility to amplitude modulation [89].

Coding, data compression, and generation of test signals were the first application areas of LZC [90]. Recently, it has been applied extensively in biomedical signal analysis as a metric to estimate the complexity of discrete-time physiologic signals. For instance, it was used in recognition of structural regularities, for complexity characterisation of DNA sequences [89].

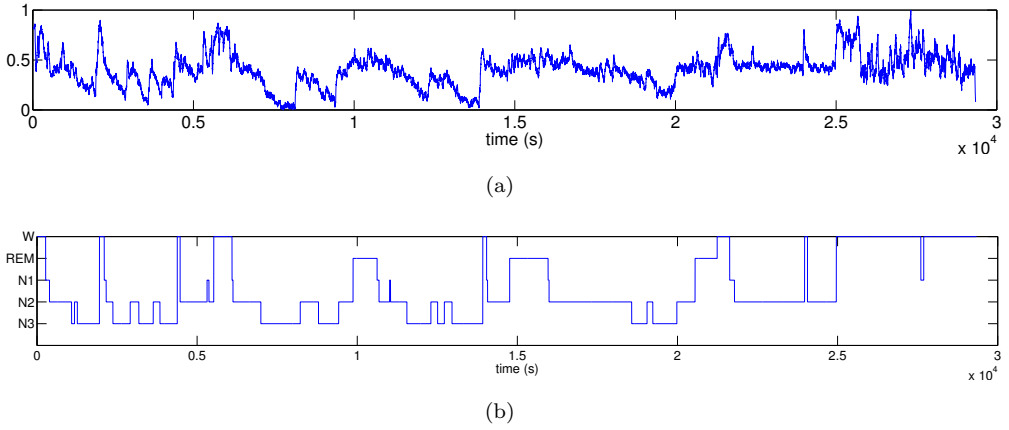


FIGURE 4.5: Evolution of a) LZC feature for a whole night of sleep, after a moving average window of 30 seconds was applied and b) the respective hypnogram, for patient NFLE2.

The LZC is also used to characterise the sleep [91]. The values of LZC increase about 30% at the transition from sleep to waking, while going back to sleep was associated with a comparable decrease. In addition, a LZC-based index (called the perturbational complexity index) was proposed and validated as a measure of consciousness in an extensive set of data obtained in patients recorded under anesthesia, in coma, persistent vegetative state, and during sleep [92]. Another study showed that activated brain states-waking and REM sleep are characterised by higher LZC compared with NREM sleep.

In this thesis the LZC rate is computed for a moving window with 1-second length. An example of the LZC as compared with the hypnogram is presented in Figure 4.5 for the EEG signal.

4.2.3.5 Discrete Time Short Time Fourier Transform

Discrete Fourier Transform (DFT) is the simplest form of time–frequency analysis for discrete stationary signals, and is given by:

$$X[k] = \sum_{n=0}^{N-1} x[n] e^{-j\frac{2\pi kn}{N}}, \tag{4.5}$$

where N is total the number of samples in a given signal segment, $x[n]$ is the value of the signal at instant n, k is the discrete frequency (0 to N-1 Hz).

The EEG is a non-stationary signal, with different portions of time having different frequencies. If the DFT is applied to the signal the information about frequencies present in a determined instant will be lost. Thus, the solution is to divide the signal into pieces through the application of a window, w, and apply the DFT [93], this technique is designated by discrete time short time fourier transform (DTSTFT).

$$X[n,k] = \sum_{m=-\infty}^{\infty} x[m] w[n-m] e^{-j\frac{2\pi km}{N}} \tag{4.6}$$

The window $w[n]$ is assumed to be non-zero only in a interval of length N_w and is referred to as the analysis window. The sequence $x[m]w[n-m]$ is called short section of $x[m]$ at time n [94].

The short-sections must be small enough to fulfil the requirement of stationarity. Besides, due the Heisenberg principle the time–frequency resolution of DTSTFT is directly determined by the segment size: the smaller the segment, the higher the time resolution and lower the frequency resolution whereas larger windows have less time resolution but greater frequency resolution.

This methodology has been widely used in the analysis of sleep EEG mainly due to its simplicity and straightforward implementation [95] in epileptic seizure prediction [96] and in automatic sleep staging [48].

In this thesis the spectrogram for each second is computed for a window of three seconds centred in the second of interest. From the spectrogram the frequency

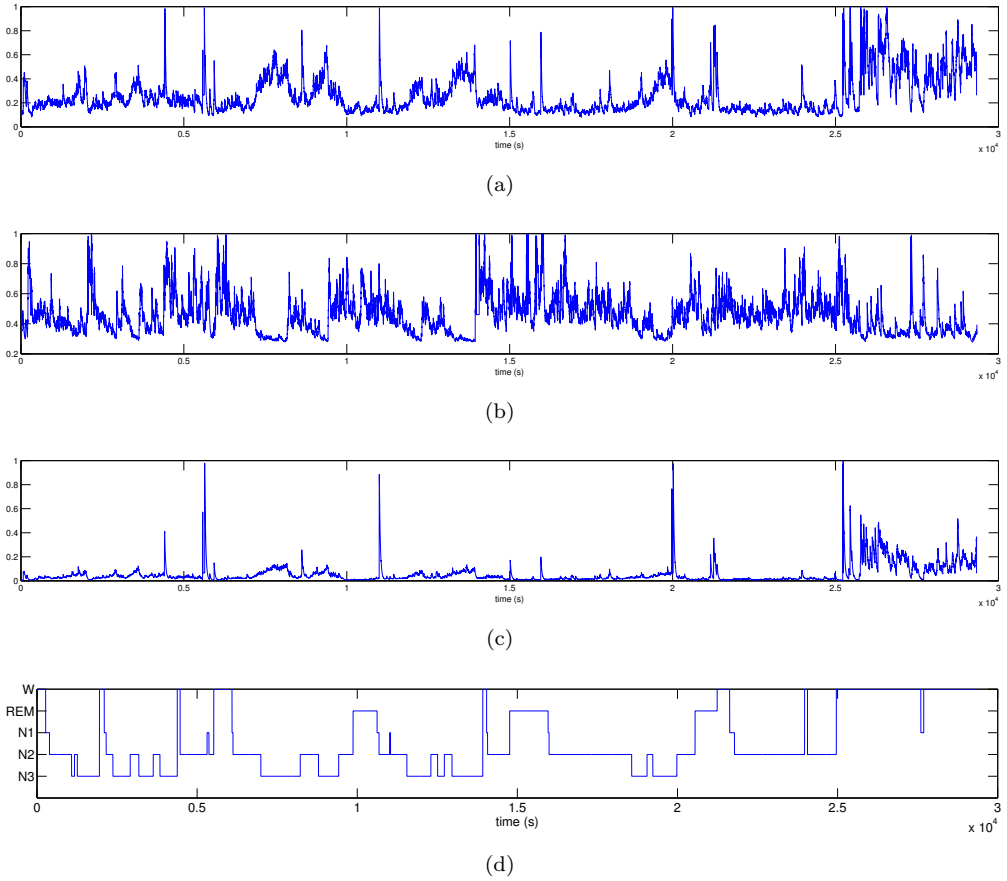


FIGURE 4.6: Evolution of frequency features a) f_{area} b) f_{max} c) f_{mean} and d) the respective hypnogram for a whole night of sleep, and after a moving average window of 30 seconds was applied, for patient NFLE2.

corresponding for the amplitude maximum, f_{max}, the mean amplitude of frequencies, f_{mean}, and also the area under the curve of spectrogram, f_{area}, are extracted. It can be seen in Figure 4.6 the evolution of each one of the three features.

4.2.3.6 Empirical Mode Decomposition

Empirical Mode Decomposition (EMD) is a decomposition technique to obtain a series of signals related to the characteristic oscillation of the signal. The signal is decomposed based on a Hilbert-Huang transformation [97], and the decompositions are designated by intrinsic mode functions (IMF) – each of them representing an embedded characteristic oscillation on a separated time-scale. The EMD application requires a continuous signal, with the number of maxima equal to the number of minima, and also a mean signal of zero.

After the restrictions above are satisfied the IMF's are computed as follow (adapted from [98]):

- For a given signal, $x(t)$, all the local minima and maxima of $x(t)$ are identified;
- The upper envelope $E_n(U)$ is computed by using a cubic spline to connect all the local maxima. Similarly, the lower envelope $E_n(L)$ is calculated from the local minima. The upper and lower envelopes should enclose all the signal samples;
- The mean $\overline{E}_n$ of the upper and lower envelopes are computed, and the signal, $x(t)$ is updated by subtracting the mean from it $x(t) = x(t) - \overline{E}_n$;

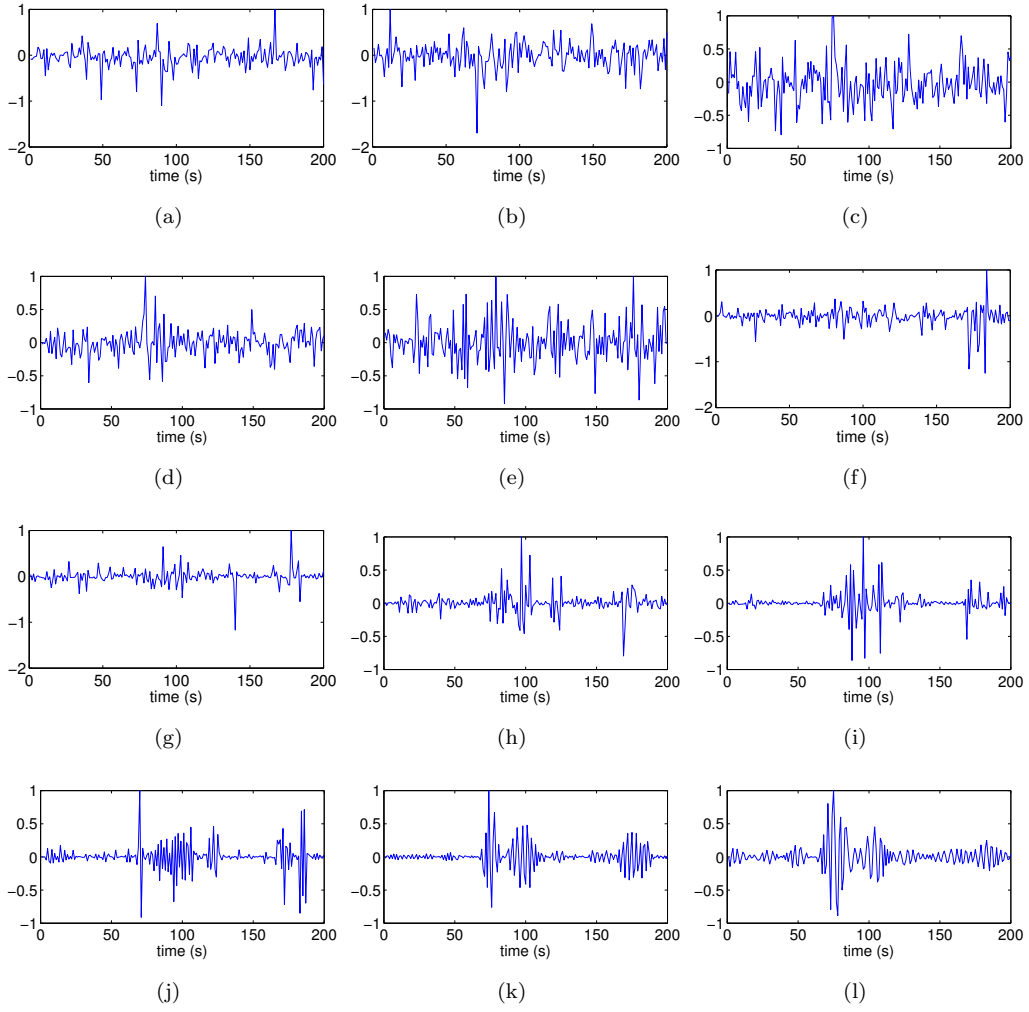


FIGURE 4.7: Twelve levels of decomposition a) IMF_1 b) IMF_2 c) IMF_3 d) IMF_4 e) IMF_5 f) IMF_6 g) IMF_7 h) IMF_8 i) IMF_9 j) IMF_{10} k) IMF_{11} k) IMF_{12} for the signal with EMD method, in 200 seconds of EEG for patient NFLE2.

- The $\overline{E}_n$ is the new potential IMF, although it must satisfy the restrictions above. If they are not the three previous steps are executed over $\overline{E}_n$ obtained until they are satisfied;
- The first IMF, $\text{IMF}_1(t)$, is subtracted from $x(t)$ to get the residue $r_1(t)$;
- The residue $r_1(t)$ is now taken as the starting point instead of $x(t)$, and the previously mentioned steps are repeated to find all the IMFs $\text{IMF}_i(t)$ so that the final residue r_n either becomes a constant, a monotonic function, or a function with a single maximum and minimum from which no further IMF can be extracted.

Therefore, at the end of the decomposition, the signal, $x(t)$, can be represented as the sum of IMFs and a residue r_n:

$$x(t) = \sum_{i=1}^{n} \text{IMF}_i(t) + r_n(t), \tag{4.7}$$

The IMFs with the lower indices correspond to high frequency oscillations, whereas those with higher indices correspond to the signal trend. Therefore it is possible to separate the trend from the original signal, and obtain a detrended signal.

The EMD has been used in EEG applications, for example, for classification of mental tasks [71, 99] and in automatic sleep staging using ECG [100].

In this thesis the EMD is computed for twelve decomposition levels for EEG, $n = 12$. In Figure 4.7 can be seen the evolution of a portion of the signal with 200 seconds in the different decomposition levels.

4.2.3.7 Shannon entropy

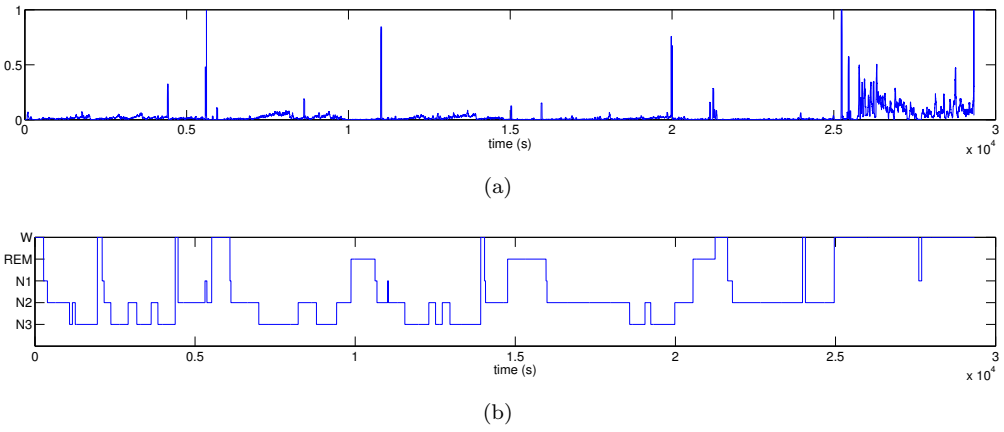
Shannon entropy (ShEnt) gives a value of relevance in the entire dataset based on a probability. The higher the probability is for an event to happen, less information exists in it. For instance, the following example exemplifies ShEnt with a real-world situation: "Alice arrives always on time at work, so the probability of this event is high. If today she is late, which is rare (low probability), something might happen, however if she arrived on time which is the usual there is no additional information with it".

Let $\mathbf{A} = (A, p)$ be a discrete probability space. Where $A = a_1, ..., a_n$ is a finite set, with the probability of each event p_i. The information gain $G(B|A)$ measures the gain obtained by the knowledge that the outcome belongs to a set $B \subset A$. Equation 4.8 defines the information gain

$$G(B|A) = \log_2 \left[\frac{1}{p(B)}\right] = -\log_2[p(B)], \tag{4.8}$$

where $p(B) = \sum_{i \in B} p_i$. The shannon entropy of A is:

$$H(A) = -\sum_{i=1}^{n} p_i \log_2(p_i). \tag{4.9}$$



(a)

(b)

FIGURE 4.8: Evolution of a) absolute ShEnt feature computed for EEG signal after a moving average window of 30 seconds was applied and b) the corresponding hypnogram for a whole night of sleep, for patient NFLE2

ShEnt has been widely used in EEG processing, for example to distinguished between normal and epileptic EEG [101]. It was proved that this feature is proportional to the sleep macro and microstructure[3] and it has been applied in the automatic staging of both [54, 55, 102, 103].

In this thesis, ShEnt is computed for each second. The feature evolution of this with time and the respective hypogram are shown in Figure 4.8

4.2.3.8 Fractal Dimension

Fractal Dimension (FD) quantifies how many times the same sequence appears in a signal. In other words a signal can be composed by basic blocks forming a pattern,

the FD is related to the number of these basic blocks. The algorithm proposed by Higuchi [104] is generally used for finding FD of EEG signals. The EEG signal is assumed as the time sequence $x(1), x(2), ..., x(n)$. From the EEG signal a new time series x_m^k is constructed as follow: $x(n)_m^k = x(m), x(m+k), ..., x(m+(N-m)/k)$, $m = 1, 2, ..., k$, where m and k are integers that indicate the initial time and the time interval, respectively. The length of the curve, $x(n)_m^k$, is defined as follows:

$$L_m(k) = \frac{\left(\sum_{i=1}^{\frac{N-m}{k}} |x(m+ik) - x(m+(i-1)k)|\right)\frac{N-1}{\left[\frac{N-m}{k}\right]k}}{k} \tag{4.10}$$

The length of the curve for time interval k, $< L(k) >$, is defined as the average values over k sets of $L_m(k)$. If $< L(k) > \propto k^{-D}$ then the curve is fractal with the dimension D.

FD have been broadly used in biomedical signal analysis as a measure of the complexity of signals. It can be used to study the dynamics of transitions between different brain stages.

FD is related with sleep macro and microstructure [3] and have been used in the automatic staging of sleep pattern [54, 103].

For each second, FD is computed and is represented in Figure 4.9.

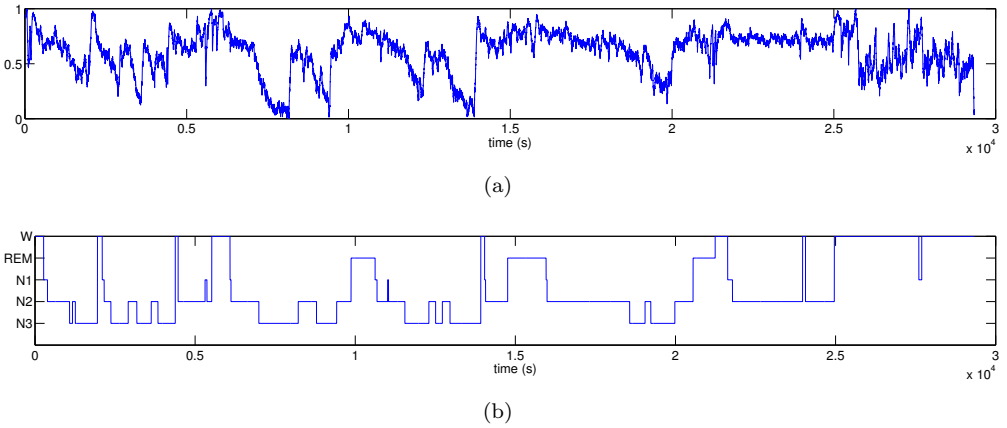


FIGURE 4.9: Evolution of a) FD feature after a moving average window of 30 seconds was applied and b) the respective hypnogram for a whole night of sleep, for patient NFLE2

4.2.3.9 Variance

The variance is the standard deviation of the a segment of data squared. This simple measure have been used in the automatic sleep staging [54] and A-phases detection [67]. In this thesis the variance is computed for each second. The evolution of variance with the time and the respective hypnogram is presented in Figure 4.10

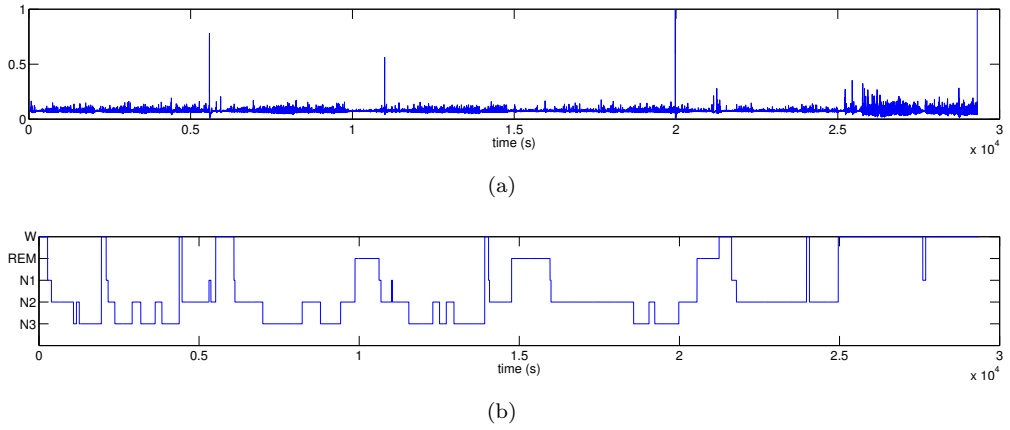


FIGURE 4.10: Evolution of a) variance feature after a moving average window of 30 seconds was applied and b) the respective hypnogram for a whole night of sleep, for patient NFLE2

4.2.4 Features Pre-processing

First, the features are smoothed using a moving average window filter [105], except for MMSD, TEO and EMD features. These features are excluded because they detect changes in amplitude and frequency, and if the smoothing technique is applied important information might be lost.

Secondly, the outliers are removed. Assuming the data follows a normal distribution, the points which are four standard deviations away from the mean are considered as outliers. Their value is replaced for the median of the feature. In literature, the usual value used is three standard deviation [106], although the A-phases are likely to be considered outliers. Therefore, the maximum distance allowed is extended.

Finally the values of each feature are normalised to be between 0-1.

4.2.5 Feature ranking and transformation

The selection or reduction of the features used for the classification task is essential for a good performance of the classifier, since it is desired that the features should be correlated with the class labels. Feature selection also contributes to reduce the computational cost associated with high dimensionality problems. So the features must be analysed to select the best set to use in the classification step. In this thesis the minimum redundancy maximum redundancy (MRMR) ranks the features and the principal component analysis (PCA) transform them.

MRMR: This algorithm tries to achieve two objectives: obtain the highest correlation between the selected features and class labels (maximum relevance) and reduce the redundancy between features (minimum redundancy). This technique used is described in detail in references: [107, 108].

PCA: In a few words, PCA implements transformation in the features with dimension d, to another space to maximize the data variance with a dimension n. This transformation combines features to obtain the new space which allows to reveal hidden information about the data, however the physical meaning of the axis is lost [109].

4.2.6 Classification

The different classification methods used in this thesis are presented next.

4.2.6.1 Discriminant Analysis

Discriminant Analysis (DA) [41] is generalisation of Fisher's linear discriminant used in a different areas like: statistics, pattern recognition and machine learning. In this thesis the DA is used as a classification method (machine learning). The basic principle is to find the best decision boundary, g, to separate the classes.

Considering a feature vector x of size $(1 \times d)$, and c classes ω_k $(k = 1, 2, ..., c)$, to find the best discrimination function, g, the type must be at first defined. It can be linear, $g(x) = wx + w_0$ and in this case the classifier is named by LDA, or quadratic, $g(x) = w_0 + \sum_{i=1}^{d} w_i x_i + \sum_{i=1}^{d} \sum_{j=1}^{d} x_i x_j w_{ij}$, and in this case is

designated by quadratic discriminant analysis (QDA) classifier where w is the weight vector and w_0 is the bias.

The next step is to minimise the following criteria, using n examples of training data:

$$J(w) = \frac{1}{n}\sum_{k=1}^{n}(t_k - g(x_k, w))^2 \tag{4.11}$$

where t is the vector with the labels of x.

In a multiclass problem with c classes discrimination procedure is to evaluate c discriminant functions, and classify x as ω_i is the discriminant function with the higher value. For DA the resulting classifiers are designated by linear machine and is represented in Figure 4.11, along with an example of space region divided by a linear machine for a $c = 3$.

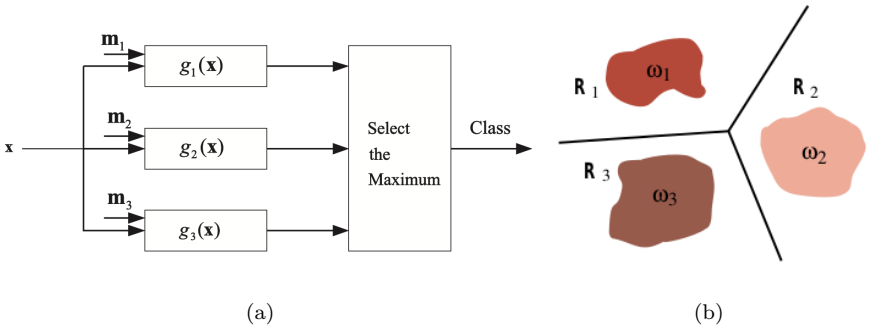


FIGURE 4.11: Example of a three class problem, a) the linear machine b) regions in space for a linear g

4.2.6.2 k-NN

k-NN is a non-parametric method, which means that there is no assumption about the underlying pattern distributions. To label the new data the k-NN calculates the probability density function (meaning the distribution must be well-behaved and smooth). The probability function $p(x)$ at point x is estimated by considering a sufficiently small region R with volume V, centred at x, as follows:

$$p(x)V \approx \int_R p(u)du \tag{4.12}$$

Suppose the training set has n classes and k of these are within the region R, based on Equation 4.12 the number of classes within R could be given by:

$$\hat{p}(x) \approx \frac{k/n}{V} \tag{4.13}$$

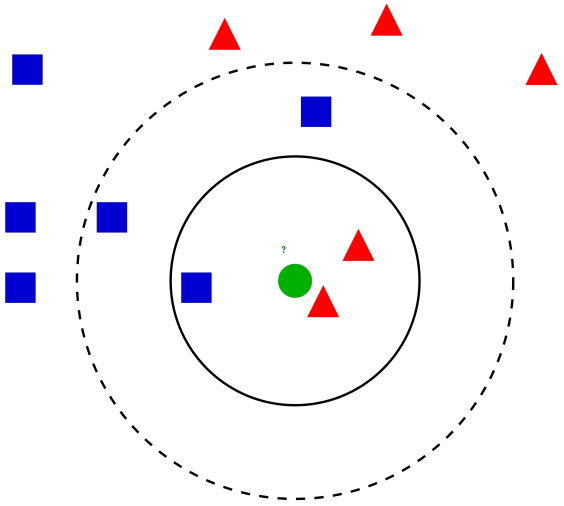


FIGURE 4.12: k-NN example, where the green point is the data to label and the red triangles and blue squares are the training data (adapted from: [110])

However, in the one hand if the number of samples is small compared to the volume an high variance will be obtained. On the other hand, if the volume is too big, a large number of samples will be included in it. So to overcame this problem one variable, V or k, must the fixed.

In k-NN the number of k samples is fixed, and the volume V is determined. At first, a smaller volume around x starts to increase until k samples are within the V, these k points are the nearest neighbors of x. Each k sample will belong to a certain class, and the label of x will be equal to the majority of the classes within the V. Observing the Figure 4.12 we can see that if $k = 3$ the green circle will be classified as a red triangle while if $k = 5$ the circle will be considered a blue square. So the choice of k is essential for a good classifier.

4.2.6.3 Support Vector Machine

Support Vector Machine (SVM) finds a decision hyperplane that maximises the separation between patterns belonging to two different classes.

Considering the the training data $\{x_i, y_i\}$, $i = 1, ..., n$, $y_i \in \{-1, 1\}$, $x_i \in R^d$, where and y_i are the class labels and x_i the features values. Supposing that a hyperplane exists which separates the two classes. The points x which lie on the hyperplane satisfy $wx + b = 0$, where w is the vector normal to the hyperplane and $|b|/\|w\|$ is the distance from the hyperplane to the origin (an example is shown in Figure 4.13). The distance of any data point x_i to the hyperplane is $|w'x_i + b|/\|w\|$. The vectors corresponding to the closest points are designed by support vectors and are distance from the hyperplane$1/\|w\|$. Meaning that all the other points satisfy:

$$t_i(w'x_i + b) \geq 1 \quad t \in -1; 1 \tag{4.14}$$

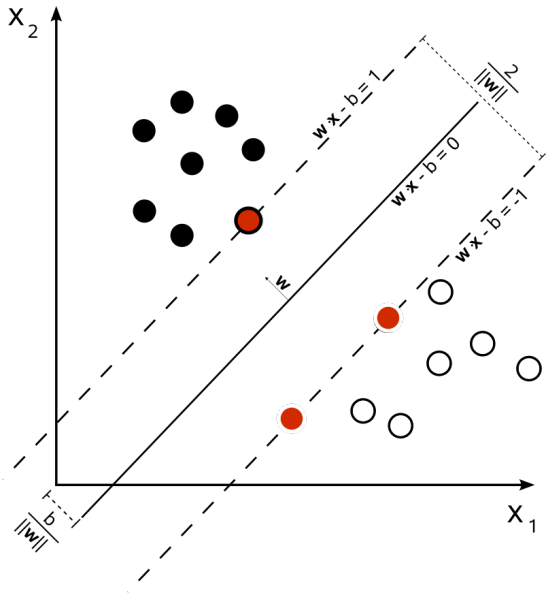


FIGURE 4.13: SVM for a linear separable two class problem ([111]). The support vectors are coloured by red.

Therefore, these vectors will define the margin of separation. The principle of SVM is to maximise of the separation margin $(2/\|w\|)$, which is translated by the minimisation criterion $\Psi(x)$, given by:

$$\Psi(w) = \frac{1}{2}\|w\|^2; \quad y_i(w'x_i + b) \geq 1, \text{with} i = 1, .., n \tag{4.15}$$

The equation above can be solved by the Lagrange multipliers method, α_i, where the saddle point of the Lagrange function is computed as:

$$J(w, b, \alpha) = \frac{1}{2}\|w\|^2 - \sum_{i=1}^{n} \alpha_i(y_i(w'x_i + b) - 1) \tag{4.16}$$

However, the example above is for a completely linear separable problem. For non-linearly separable classes a penalty part, $C\sum_{i=1}^{n}\xi$, is added to the minimisation criteria. The new equation system is:

$$\Psi(w) = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}\xi; \quad y_i(w'x_i + b) \geq 1 - \xi_i, i = 1, .., n \tag{4.17}$$

and the goal is to minimise $\Psi(w)$. The ξ defines the separation margins. If $\xi < 1$ the data points could be in the separation region between the hyperplane and the margin, otherwise misclassification of some points is allowed. C defines the influence of ξ in the minimisation criterion Ψ. If C is large the influence of ξ will be reduced, so there is a high tolerance to misclassification whereas if C is small the influence of ξ will be big, leading to points being misclassified and shorter margins.

In the case of a non-linear SVM the features space is projected into a higher dimensional space based on a Kernel projection where the linear principle still can be applied.

Considering the dataset $x_1, ..., x_l \subset \varphi$ (where φ is usually in the $\mathbb{R}^N$ space) is transformed into a new feature space through $K_{ij} := k(x_i, x_j)$, where k is a non-linear mapping of the actual features into a new space F where the features might be linear separable. This mapping is called *feature map associated with* k and is designated by $\Phi : \mathbb{R}^N \to F$ with the dot product being the projection of features into the new space, $k(x_i, x_j) = (\Phi(x_i)\Phi(x_j))$.

The function that describes the nonlinear decision boundary is given by Equation 4.18 where x_i are the training examples. The vectors for which $\alpha_i \neq 0$ are called support vectors [112].

$$f(x) = t_i\Big[\sum_{i=1}^{l} \alpha_i k(x_i, x_j) + b\Big] \tag{4.18}$$

The most popular feature space transformation is the by the radial basis function:

$$k(x_i, x_j) = e^{\gamma \|x - x_i\|^2}, \quad \gamma = \frac{1}{2\sigma^2} \tag{4.19}$$

The description above is for a two-class problem. However, the sleep stages and A-phases are multi-class problems. Due to various complexities, a direct solution of multi-class problems using a single SVM formulation is usually avoided. The better approach is to use a combination of several binary SVM classifiers to solve a given multiclass problem. Different methods exist but in this thesis the one used is one-against-all multi-class. Considering the training set with c classes: $\{x_i, y_i\}$, where i=1,...,n, and $y_i \in \{1, ..., c\}$. This method transforms the multi-class problem into a series of c binary subtasks that can be trained by the binary SVM. The ith classifier output function ρ_i is trained taking the examples from c_i as 1 and the examples from all other classes as -1. For a new example x this method assigns to x the class with the largest value of ρ_i [113, 114].

4.2.7 Post-processing

The sleep stages and A phases have different post-processing steps.

Macrostructure

For each second the features are computed and a sleep stage is assigned. Although, according to [14], the scoring is done considering epochs of 30 seconds. Hence, for each 30 seconds of the signal, the stage assigned to each one is the most present.

Microstructure

The length of A-phases must be within the interval $[2, 60]$ seconds, then, the A-phases with duration less than two seconds are removed as a possible A phase. Other approaches are tested, if A-phases is less than 2 seconds distant from another A-phase they are united. Although, the results are worse.

4.2.8 Performance evaluation

Cross validation (CV) is a popular strategy for the assessment of algorithms performance. It splits the data in several blocks and, in each iteration, two groups

are generated: the training and testing set. The training set is used to create the model, the test set is used to evaluate its performance. In this thesis it was used the Leave-one-out approach [115]. This is the most classical exhaustive CV procedure. Supposing that our data is composed n patients, the algorithm will be trained n times. In each iteration, a different patient is ´´left out" to be the test data and the remaining ones are training data. In macro and microstructure some stages or phases can be in majority. To avoid discrepancies in the train set, each sleep stage, in macro, and each phase, in micro, must have the same number of training points. The same is not applied for test patients where the whole night of sleep is evaluated.

The algorithm output for each patient is compared to the real staging and the confusion matrix and Table 4.2 is filled.

TABLE 4.2: Confusion matrix for a k-class problem, C_i is the real output while $\hat{C}_i$ is the one provided by the algorithm, where $i = 1, ..., k$

	C_1	...	C_k
$\hat{C}_1$	n_{11}	...	n_{k1}
$\vdots$	$\vdots$	$\ddots$	$\vdots$
$\hat{C}_k$	n_{1k}	...	n_{kk}

On the one hand, the diagonal terms, $n_{ii}(where, i \leq k)$, correspond to the instances where the algorithm's output, $\hat{C}$, was consistent with to the real class label, C. On the other hand, the values $n_{ij}(where 1 \leq i, and j \leq k)$ with $i \neq j$, are the number of instances classified by the algorithm as k when, in fact, the real class is j.

Considering the class i, four measurement can be taken:

- **True positive ($\mathbf{TP}_i$):** number of instances correctly classified as class i;
- **False positive ($\mathbf{FP}_i$):** number of instances classified as i when in fact they belong to other class;
- **True negative ($\mathbf{TN}_i$):** number of instances correctly not classified as i;
- **False negatives ($\mathbf{FN}_i$):** number of instances assigned to class i when in fact they belong to other classes;

The equations for this variables computation are given by:

$$n_{+,j} = \sum_{i=1}^{k} n_{ij} \qquad n_{i,+} = \sum_{j=1}^{k} n_{ij}$$
$$TP_i = n_{ii}; \quad FP_i = n_{i,+} - n_{i,i}; \quad FN_i = n_{+,j} - n_{i,i}; \quad TN_i = n - TP - FP - FN \tag{4.20}$$

To evaluate the algorithm's performance three measures are used: sensitivity (SE), specificity (SP) and accuracy (acc), which are given by:

$$SE = \frac{TP}{TP + FN}; \quad SP = \frac{TN}{TN + FP}; \quad acc = \frac{TN + TP}{TN + FP + FN + TN} \tag{4.21}$$

Regarding the number the A-phases percentage comparing to the B-phases (shown in Table 4.1), it can be seen that the B-phases are more present than the A-phases, sometimes seven times more. Therefore, when an algorithm correctly detects most of the B-phases (even through performance is not as good as A-phases), it usually has a higher accuracy. Since the main objective is to correctly detect the A-phases, this measurement, might not yeild a correct view of the overall performance of the algorithm. Thus, another performance measure is considered, which will be designated by a weighted accuracy, acc_w, that takes in account the number of a class instances for the accuracy computation. In other words, a miscomputation of B-phase instance will have lower impact on accuracy than one in A-phase instance.

The weighted accuracy is given by:

$$acc_w = \sum_{j=1}^{k} \frac{n_{jj}}{n_{+,j}}; \tag{4.22}$$

The simple accuracy is also computed to better compare the results with the literature. The weight accuracy is only computed for microstructure algorithms.

4.3 Summary

To summarise, for each patient, the following steps are implemented:

1. Filter the signal, to generate: EEG filtered (EEG filtered between $1 - 35$ Hz), EEG_δ, EEG_θ, EEG_α, EEG_σ, EEG_β (this last 5 are the signals corresponding to the $\delta, \theta, \alpha, \sigma$ and β frequency bands);

2. Extract features from the signals (55 features in total);
3. Apply a moving average filter for all features except for MMSD, TEO and EMD;
4. Remove the outliers (all points whose value is above mean(feature)+4×STD(feature);
5. Normalise the features values to range between 0 and 1;

At this point each patient has 55 features, for each second, for a whole night recording. Different algorithms, with different conditions, will be validate to predict the sleep stages and A-phases for each patient. A patient is 'left out', while the others are used for train the algorithm. The patient 'left out' is used to test the algorithm. This is performed for all patients.

The algorithms used are: linear algorithm (threshold), LDA, k-NN, SVM. For the linear algorithm the threshold value ranged from 0.0001 to 1 spaced by 0.0005. The discrimination functions used for DA are linear and quadratic. For KNN the k values are: 3, 5 ,9,11,15 and 25. Lastly, for SVM it is performed a grid search for $c = 2^{-5}, 2^{-3}, ..., 2^{15}$ and $\gamma = 2^{-15}, 2^{-13}, ..., 2^{5}$ values;

For each patient (*test_patient*) and for each algorithm with a specific condition the following steps are executed:

1. Create the matrix train (*matrix_train*): for all patients (p except for *test_patient*):
   (a) Discover the class with less instances (i) for patient p;
   (b) Select randomly i instances from each class in patient p;
   (c) Add the instances chosen, with the all features, to the matrix train;
2. Select the x best features based on MRMR;
3. Construct the training set, $train\_MRMR$, with x best features chosen by MRMR, x_MRMR;
4. Train the classifier with $train\_MRMR$;
5. Construct the *test_set* with all instances of the test patient and with x_MRMR features;
6. Predict the instances for *test_set* with the algorithm;
7. Compute the performance values;
8. Apply the PCA to *matrix_train*, and extract the axis transformation;

9. Create the $train\_set\_PCA$ with the best x projections of PCA;
10. Train the classifier with $train\_set\_PCA$;
11. Apply the axis transformations to the features of test patient, and select the same projections, $test\_PCA$;
12. Predicted the classes instances, the test patient with $test\_PCA$ and the classifier previous trained;
13. Compute the performance values;

Chapter 5

Results

5.1 Feature Ranking and Tranformation

MRMR

Aiming to facilitate the analysis for each feature a numerical ID is assigned and they are represented in Table 5.1. Some features are computed for the filtered EEG while others are computed for the signal corresponding to one of the conventional frequency bands. To distinguish both situation the nomenclature used is *feature name*$_X$ where X is eeg in first case and one of the conventional bands ($\delta, \theta, \alpha, \sigma$ and β) in the second case. The variance is represented by v, due the used of σ to represented one of the conventional frequency bands.

Table 5.1: Feature ID and the corresponding name

ID	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**
name	$C_{\tau_0,\delta}$	$C_{\tau_0,\theta}$	$C_{\tau_0,\alpha}$	$C_{\tau_0,\sigma}$	$C_{\tau_0,\beta}$	C_δ	C_θ	C_α	C_σ	C_β	φ_δ
ID	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**
name	φ_θ	φ_α	φ_σ	φ_β	TEO_δ	TEO_θ	TEO_α	TEO_σ	TEO_β	ZC_{eeg}	ZC_δ
ID	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**	**31**	**32**	**33**
name	ZC_θ	ZC_α	ZC_σ	ZC_β	LZ_{eeg}	LZ_δ	LZ_θ	LZ_α	LZ_σ	LZ_β	f_{areas}
ID	**34**	**35**	**36**	**37**	**38**	**39**	**40**	**41**	**42**	**43**	**44**
name	f_{mean}	f_{max}	IMF_1	IMF_2	IMF_3	IMF_4	IMF_5	IMF_6	IMF_7	IMF_8	IMF_9
ID	**45**	**46**	**47**	**48**	**49**	**50**	**51**	**52**	**53**	**54**	**55**
name	IMF_{10}	IMF_{11}	IMF_{12}	ShEnt	FD	v_{eeg}	v_δ	v_θ	v_α	v_σ	v_β

Since MRMR is a supervised technique (one of the terms analyse the correlation between features and the output), a feature rank is generated for macro and microstructure which might be equal or not. Therefore, for each patient two ranks exist, one for macro and microstructure. An extended version of the results are displayed in Annex C in Tables C.1 and C.2.

In both ranking sequences, whether applied for sleep stages or A-phases, there are no patient with the same feature ranking sequence. The features obtained from the EEG are not correlated with the output in the same way in all patients, being an evidence of the interindividual variability.

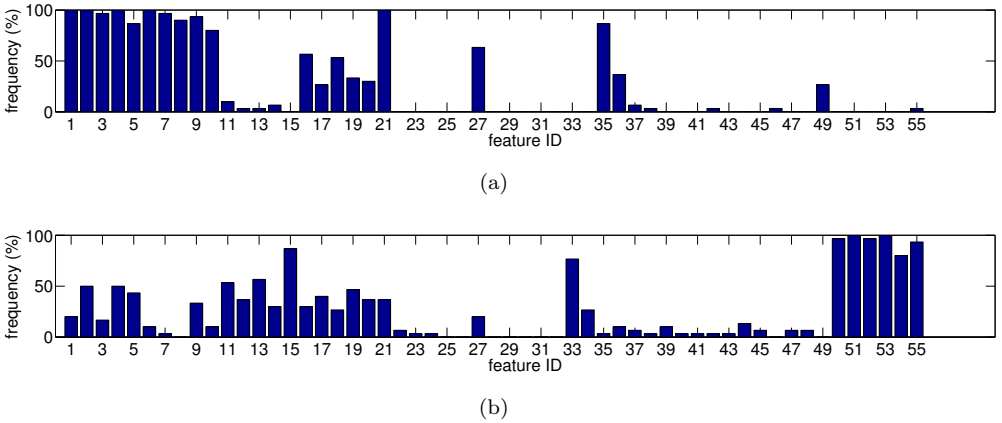


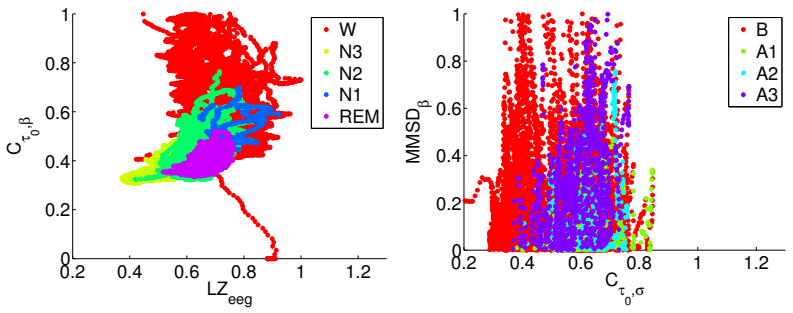
FIGURE 5.1: Frequency of occurrence of each feature in the top 15 of all patients using the MRMR with a) macrostructure and b) microstrcuture

Analysing the MRMR results for macrostructure the set of features since the rank 46 is the same for all individuals, and they are $23-26$, $28-32$, 49. Considering the *top 15* ranked features for all patients, their frequency can be analysed in Figure 5.1(a). The most frequent features are $1-11$, $16-21$, 35 and 27. Thus, these features are the most informative and less redundant among all features. Moreover, some features with frequency equal to zero correspond to the ones below the rank 46 (mentioned above).

Regarding the microstructure and contrary to the macro results, the last ranks do not converge for the same features. Consequently, there are no agreement between the patients, about the features that have a smaller amount of information to predict the A-phases. In Figure 5.1(b) can be observed the features frequency in *top 15*, the most present are [50-55], [9-21], 33 and 15.

Comparing top ranked features for both outputs, macrostructure have more features with frequency equal to 100 and 0 than the microstructure. Hence, the macrostructure results show better consistency. Additionally, the macro and microstructure do not share the same best features, in spite of this, some features does not appear on *top 15* in both (25,20-32).

In Figure 5.2(a) it is represented the distribution of instances belonging to the different sleep stages classes for the two high-ranked features, LZ_{eeg} and $C_{\tau_0,\beta}$. Firstly, in can be seen a greater superposition of the classes. The W stage is superposed with almost all sleep stages. N1 is almost covered by the REM stage. This last observation is in



(a) (b)

FIGURE 5.2: Representation of the distribution of the values of two best features for a) macrostructure and b) microstructure classes for patient NFLE2

accordance with the literature, since in some algorithms these two stages are merged due to their similarity.

In Figure 5.2(b) is represented the distribution of A- and B-phases for the two best features, MMSD_β and $C_{\tau_0,\sigma}$. It can be seen a resemblance with 5.2(a) where the W stage is behind all the other sleep stages. In this case the same applies to B-phase shows in Figure 5.2(b). Although, apparently, the microstructure problem is less separable between the classes than in the macrostructure problem.

PCA

For each patient, the PCA method is applied, and afterwards the eigenvalues are analysed. This technique does not depend on the output, therefore the transformation is the same for macro and microstructure. The evolution of the mean value and the standard deviation, of eigenvalues, with the number of components, is shown in Figure 5.3. As expected the mean eigenvalue decreases with the number of components.

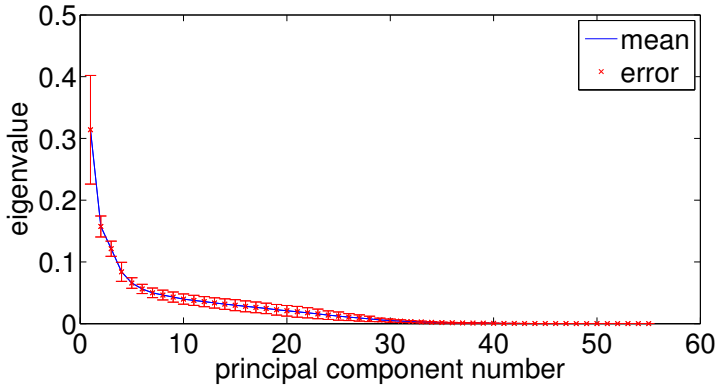
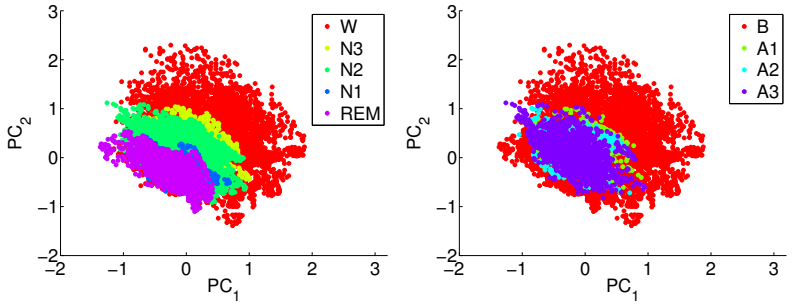


FIGURE 5.3: Mean of eigenvalues for 30 patients for the different principal components

In Figure 5.4(a) it is represented the distribution of sleep stages in the new PCA space, for the top two principal components. The classes seems to overlap more than using the simple features (Figure 5.2(a)). Considering the PCA for microstructure, shown in Figure 5.4(b), again the separation between the classes is clearly reduced compared to using the original features, presented in Figure 5.2(b).



(a) (b)

FIGURE 5.4: Representation of the distribution of the values of the two principal components with the a) macrostructure and b) microstructure classes

Concluding, the results shows that for two dimensions, the classes seem to be more separable using the features than using the PCA transformation. Although, this cannot be extrapolated for other dimensions, using more principal components might improve the separability.

5.2 Classification

5.2.1 Macrostructure

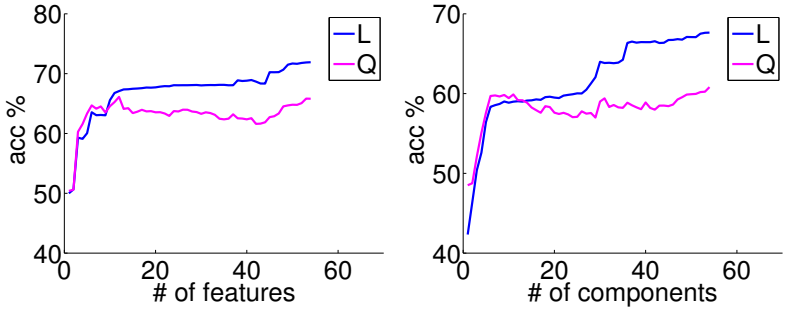
Different supervised models were build to predict the classes of macrostructure. The results obtained for each model are presented in the next sections.

5.2.1.1 Discriminant analysis

The accuracy obtained with LDA and QDA, with both feature selection and transformation techniques, are presented in Figure 5.5. QDA reaches a clear maximum in performance, whereas the performance of LDA always monotonically increases with the dimensionality. The QDA shows better performance than LDA for lower dimensions, although when more features are considered (more than 13) the results obtained for LDA are better. For a LDA the maximum is 72%, for features, and using PCA is 68%

(for 55 components). Therefore, to achieve the best accuracy with this method all the features and components are needed.

More details about the specificity and sensitivity for each stage are given. In Appendix D, in Figure D.1. Figure 5.6 is shows the sensitivity (SE) and specificity (SP) evolution for each sleep-class with the dimensionality. Considering MRMR, the SE shown in Figure 5.6(a) is quite different between the stages, being N1 the one with low SE (47% for 55 features). N2 and N3 are the stages more correctly classified, with a SE of 70% approximately. Regarding SP, Figure 5.6(b) shown the value for all stages are between 90-95%, except for N2 with 85%. The W (N2) stages have the higher (lower) value of SP, therefore, they are the stages less (most) confused by the algorithm with another stages. Regarding PCA, the SE values (shown in Figure 5.6(c)) for all stages are close to each other and above 50%. The REM stage is the one with higher sensitivity, around 70%, for more than 40 principal components. The SP values are more disperse than SE values, as presented in Figure 5.6(d). The stage with less (most) confused with other stages are N3 (N2), since they are the ones with high (low) SP. Regarding the feature selection methods, the best SE for PCA is obtained for REM stage whereas with MRMR is the N3. The high value of SP in MRMR is reached for W stage while in PCA is for N3 stage.



(a) (b)

FIGURE 5.5: Performance comparison between different DA types (Q: QDA and L: LDA) for different selection methods a) MRMR b) PCA

The conclusions taken from Figures 5.6 can be complemented looking at Table 5.2, where it can be seen that, in general, when a stage is not detected correctly by the algorithm, significant confusion occurs with other stage. Some misclassified stages are common to both feature selection techniques and are the N3, REM and W are confused with the stages N2, N1and N1, respectively. For N1 and N2 the situation differs using features and principal components. Most of the N2 are considered by the algorithm as N3, using features, while using PCA the misclassification with N3 is almost non-existant but greater with N1. Combining LDA with PCA around 31% and 9% of N1 are classified as

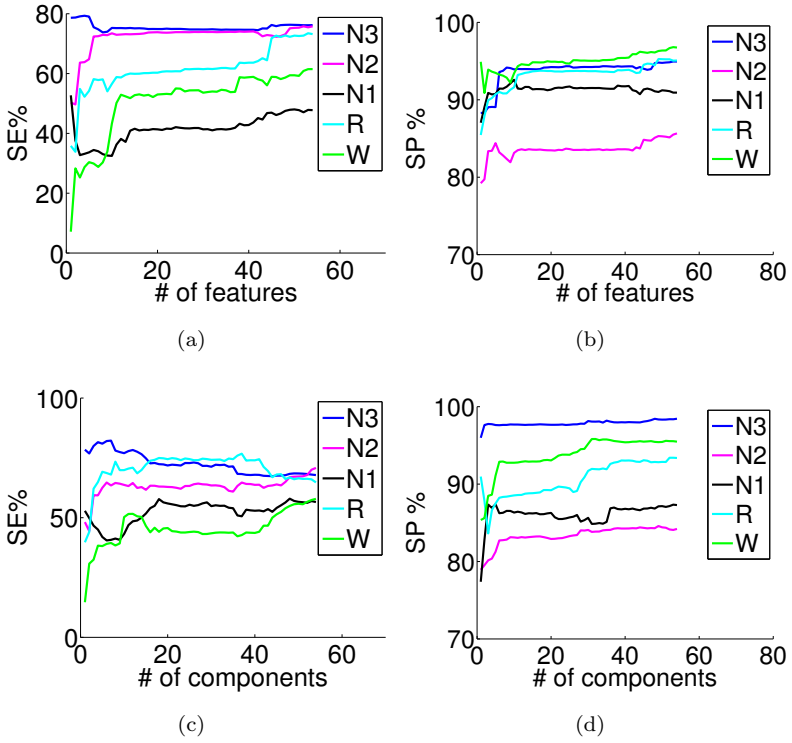


FIGURE 5.6: Evolution of SE (a and c) and SP (b and d) for two LDA models using 55 features (a and b) and principal components (c and d)

W and REM. While using features the the N1 misclassification increase 26%, 15%, 11% and are confused with W, REM and N2, respectively.

The SE is larger for all sleep stages using features than using principal components with LDA, except in N1. Besides, with PCA transformation the N1 and REM are never confused with N3 and in feature selection this misclassifications happens. It is important to notice that with this method the performance is different from patient to patient, because sometimes the standard deviation have the same order of magnitude than the SE. The distribution of accuracies both models are present in Figure 5.7. Using a LDA model with dimension 55 (all features) the high and lower accuracy obtained is 83% and 40% respectively. Using a model build based on the 55 principal components the high and lower accuracy obtained is 87% and 40%, respectively.

5.2.1.2 k-NN

The computational cost of this method greatly increases with the dimensionality. Therefore each k-NN was only computed for a maximum of 40 features or principal components. The k-NN classification method is performed for six values of k=[3 5 9 11 15 25]. The accuracy results are present in Figure 5.8 for all k's, the SE and SP are detailed in

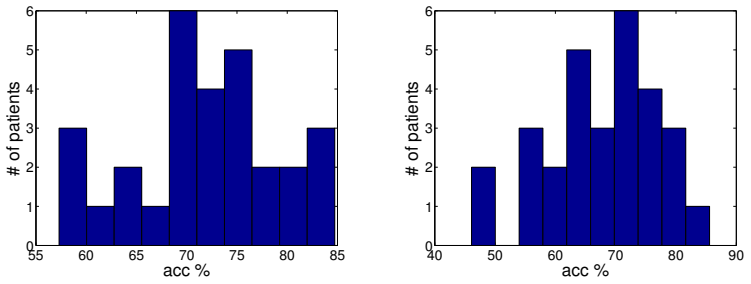
TABLE 5.2: Mean of confusion matrix for the LDA model with 55 dimensions using a) features and b) principal components

		real (%)				
		N3	**N2**	**N1**	**REM**	**W**
outcome	**N3**	**77 ± 16**	8 ± 10	0.5 ± 1.5	0.18 ± 0.45	6 ± 7
	N2	23 ± 16	**75 ± 12**	11 ± 16	7 ± 10	5 ± 7
	N1	0.08 ± 0.18	5 ± 5	**47± 19**	17 ± 22	25 ± 20
	REM	0.3 ± 0.7	7 ± 7	15 ± 18	**74 ± 24**	4 ± 5
	W	0.2± 0.6	3 ± 4	26 ± 20	2.1 ± 2.5	**62 ± 25**

(a)

		real (%)				
		N3	**N2**	**N1**	**REM**	**W**
outcome	**N3**	**63 ± 16**	1.8 ± 2.0	0 ± 0	0 ± 0	1.04 ± 1.3
	N2	35 ± 15	**70 ± 10**	1.8 ± 2.6	2.4 ± 3.5	0.62 ± 1.37
	N1	0.29 ± 0.50	13 ± 6	**58 ± 19**	16 ± 14	30 ± 27
	REM	0.7 ± 1.2	11 ± 4	9 ± 8	**78 ± 16**	4 ± 5
	W	0.7 ± 1.4	4.7 ± 3.1	31 ± 22	4 ± 5	**64 ± 29**

(b)

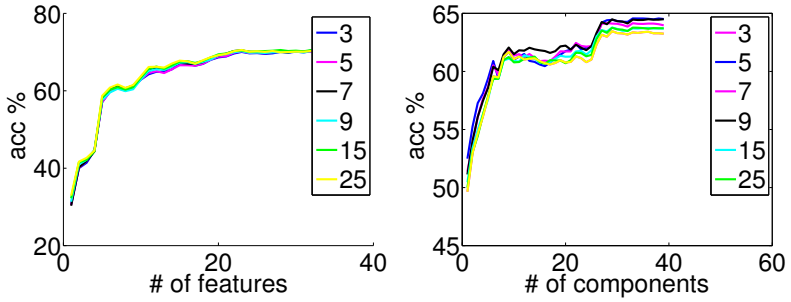


(a) (b)

FIGURE 5.7: Accuracy distribution for the different patients for LDA constructed with a) 55 features b) 55 principal components

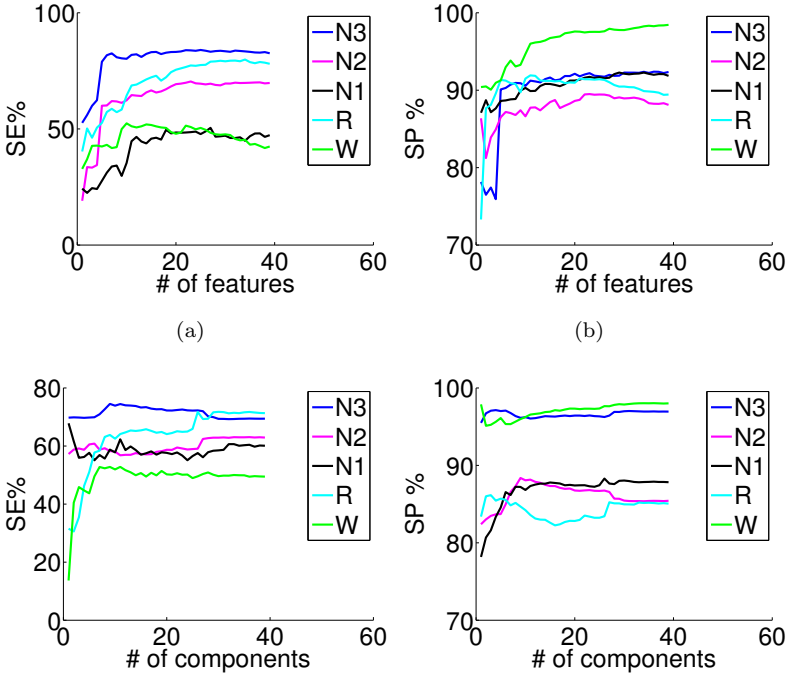
Annex D in Figure D.2 for all sleep stages. Analysing the accuracy, both MRMR and PCA contribute with monotonically increasing performance with the number of input dimensions. The performance of k-NN using a small feature dimension, d, is better use d principal components than features. For instance, the accuracy of k-NN for $d = 2$ principal components is 50% while using the two best features is 30%. Besides, for high dimensions the features chosen by MRMR have a better performance than principal components of PCA. In both cases not much difference between the performance are observed for different k values. For PCA the best mean accuracy is 65%, and is achieved with $k = 3$ and for a dimension of 34, although, since eight principal components are used the accuracy value is above the 60%. For MRMR is with $k = 25$ and the best accuracy value is 71% for a dimension equal to 24.

The progress of SE and SP using a k-NN classifier, with $k = 25$, relative to dimensions is shown in Figure 5.9, for the different feature selection methods (PCA and MRMR). N3



(a) (b)

FIGURE 5.8: Accuracy evolution with the dimensionality for different k values using k-NN classification method with the dimensions chosen by a) MRMR and b) PCA



(a) (b)

(c) (d)

FIGURE 5.9: Replace caption by: Evolution of SE (a and c) and SP (b and d) for two k-NN models for $k = 25$ with 24 features (a and b) and $k = 3$ with 34 principal components (c and d)

is the stage which most benefits with the k-NN approach. In both approaches the SE tends to be equal or greater than 50%. In spite of some differences in PCA and MRMR, regarding the disposition of performance for sleep stages, the N3 and W registered a higher and lower SE values, respectively, for both feature selection techniques. Regarding the SP for MRMR the values are between 85% and 95%, with W having the best specificity, while N2 and N1 having the lower one. For PCA there is a contrast between W and N3 which have a SP around 95% and W, REM and N1 whose values are close

to 85%. Concluding the best model for using MRMR is k-NN with $k = 25$ and a dimensionality of 24, while using PCA, $k = 3$ with 34 components yields the best results.

TABLE 5.3: Mean of confusion matrix for MRMR+k-NN, $k = 25$ and $d = 24$ (a) and PCA+k-NN, $k = 3$ and $d = 34$ (b), for sleep staging prediction

		real (%)				
		N3	**N2**	**N1**	**REM**	**W**
outcome	**N3**	**84 ± 11**	14± 10	1 ± 4	0.12 ± 0.23	5 ± 6
	N2	15 ± 11	**71 ± 11**	10 ± 12	6 ± 6	5 ± 5
	N1	0.15 ± 0.29	6 ± 4	**54 ± 14**	17 ± 17	30 ± 19
	REM	0.14 ± 0.48	8 ± 6	17 ± 13	**69 ± 24**	9 ± 8
	W	0.25 ± 0.54	1.8 ± 1.6	17 ± 13	7 ± 12	**52 ± 23**

(a)

		real (%)				
		N3	**N2**	**N1**	**REM**	**W**
outcome	**N3**	**68 ± 13**	7 ± 7	0.6 ± 1.5	0.08 ± 0.22	2.4 ± 2.7
	N2	30 ± 12	**66 ± 9**	6 ± 5	2.7 ± 3.0	6 ± 5
	N1	0.12 ± 0.30	5.8 ± 2.6	**53 ± 10**	24 ± 15	31 ± 16
	REM	0.9 ± 1.8	19.5 ±9	22 ± 14	**70 ± 17**	8 ± 9
	W	0.09 ± 0.21	1.4 ± 1.3	18 ± 13	3 ± 3	**52 ± 22**

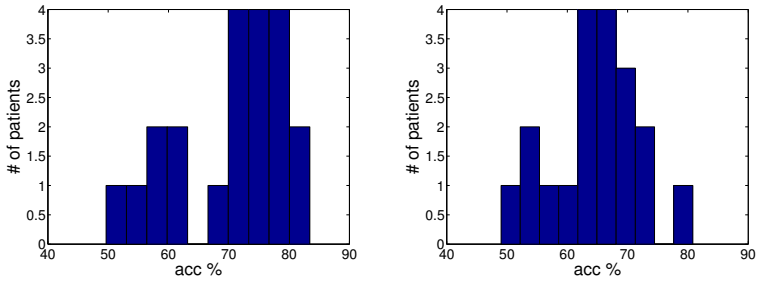
(b)

Table 5.3 presents the confusion matrix, using all patients, with the two best models. Comparing this table with Table 5.2 and regarding the LDA and k-NN with MRMR feature selection method, the big changes that are the real N3 classified as N2 drops to half, and the misclassification of N2 as N3 increase twofold. Relative to PCA when a stage is misclassified using PCA and LDA it was confused N1 and REM, 11% and 13% of the times, while in k-NN+PCA is 6% and 20%, respectively. The misclassification with N1 decrease whereas in REM it increases. Regarding N1, 30% of times is confused with the W stage and almost none with others.

The accuracies distributions are represented in Figure 5.10. The distribution in 5.10(b) are more concentraded than in 5.10(a), therefore the model built with PCA is more consistent. The high and lower accuracy for a model build with the 24 best features chosen by MRMR are 83% and 50%, respectively, while using the 34 best components 81% and 49%, respectively.

5.2.1.3 SVM

The SVM classifiers training is computationally costly, because it requires an appropriate parameter search. Thus some restrictions were considered, as describe next. Considering the previous results obtained with DA and k-NN the accuracy starts to stabilise, in some cases, when 15 features are attained. Therefore, the SVM models are build for

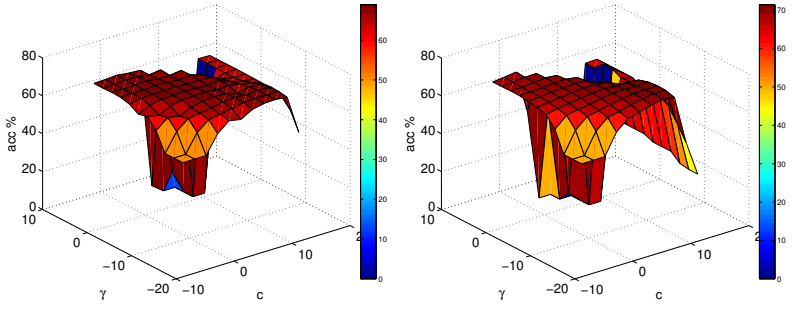


(a) (b)

FIGURE 5.10: Accuracy distribution for the different patients a) $d = 24$ chosen by MRMR and $k = 25$ and b) $d = 34$, principal components, and $k = 3$.

two dimensions, one with $d = 15$ and other with $d = 30$. With LDA and k-NN the MRMR achieves an better accuracy in comparison to PCA. Therefore, the SVM models are computed for 15 and 30 best features using MRMR. The accuracy, for the grid search, is shown in Figure 5.11, and SE and SP for each sleep stage class are shown in Annex D in Figures **??**. For some c and γ the SVM model does not converge, the points are represented with an accuracy equal to zero. Additionally, points that do not converge, have a representation of the accuracy proportional to the hottest of the colour. With $d = 15$ the better accuracy average obtained is 70%, for $c = 2^{-3}$ and $\gamma = 2$. For this c and γ the best accuracy is 85% and the lowest is 46%. Figure 5.12(a) shows the distribution of average accuracies obtained for each patient. For $d = 30$ the better average value for all patients is 72% for $c = 2^{-3}$ and $\gamma = 2^{-1}$. The higher accuracy obtained is 88% and the lower is 50%, shown in Figure 5.12(b). Each patient obtained a maximum for different c's and γ's.

Considering each patient individually the maximum accuracy ranges from $76 - 89\%$, although these good values are for different c's and γ's.



(a) (b)

FIGURE 5.11: Accuracy for SVM classifier with a) $d = 15$ and b) $d = 30$, for a grid search where $c = 2^{-5}, 2^{-3}, ..., 2^{15}$ and $\gamma = 2^{-15}, 2^{-13}, ..., 2^{5}$

The confusion matrix for the c and γ which obtained the highest accuracies, are presented in Table 5.4. The previous observations, about the misclassification of the stages observed in LDA and k-NN using the MRMR are the same, for both dimensions. The number of features is shown in the Table 5.4, which has a much more higher computational cost, despite the performance do not change significantly. The performance differences for N3, REM, N2, N3 increase 8%, 6%, 3%, 1% from $d = 15$ to $d = 30$, respectively. In W stage the sensitivity is the same. It is also notable that the differences in the accuracies obtained differ less from patient to patient in the SVM with higher dimensionality.

TABLE 5.4: Mean of confusion matrix for MRMR and SVM for $c = 2^{-3}$ and $\gamma = 2$, for 15 dimensions (a) and $c = 2^{-3}$ and $\gamma = 2^{-1}$, for 30 dimensions (b) for sleep staging prediction

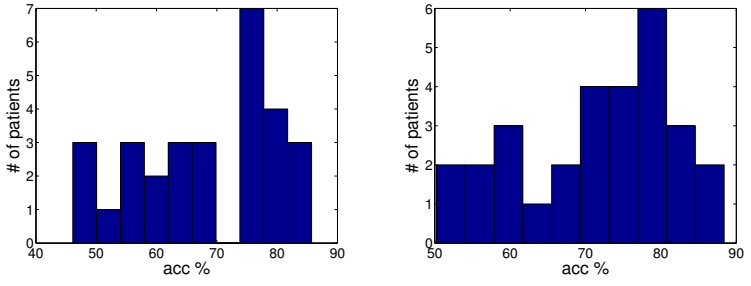
		real (%)				
		N3	**N2**	**N1**	**REM**	**W**
outcome	**N3**	**82 ± 14**	12 ± 11	0.4 ± 1.9	0.20 ± 0.70	3 ± 6
	N2	17 ± 13	**70 ± 16**	11 ± 20	7 ± 8	5 ± 6
	N1	0.07 ± 0.19	7.32 ± 8.14	**48 ± 22**	21 ± 26	22 ± 16
	REM	0.2 ± 0.7	8 ± 14	17 ± 22	**65 ± 32**	8 ± 16
	W	0.8 ± 1.3	3 ± 3	24 ± 20	6 ± 145	**62 ± 22**

(a)

		real (%)				
		N3	**N2**	**N1**	**REM**	**W**
outcome	**N3**	**83± 12**	11 ± 11	0.2 ± 0.9	1.4 ± 0.5	3 ± 5
	N2	16 ± 12	**73 ± 14**	10 ± 17	6 ± 6	6 ± 8
	N1	0.05 ± 0.17	6 ± 6	**54 ± 18**	19 ± 23	23 ± 18
	REM	0.2 ± 0.8	7 ± 13	16 ± 17	**69 ± 29**	6 ± 12
	W	0.7 ± 1.5	2.9 ± 3.3	20 ± 19	5 ± 14	**62 ± 24**

(b)

Finally, from the histograms in Figure 5.12, it is seen that the distribution is more homogenous in an high dimensional SVM, more than 50% of the patients have an accuracy superior to 70%, and the number of patients with low accuracy is low.



(a) (b)

FIGURE 5.12: Accuracy distribution for the different patients a) $d = 15$ ($c = 2^{-3}$ and $\gamma = 2$) and b) $d = 30$ ($c = 2^{-3}$ and $\gamma = 2^{-1}$)

5.2.2 Microtructure

Different models are build for DA and k-NN in macrostructure, the metric used to choose the best model of a determined classification method is accuracy. Although there, due the A-phases being in lower percentage relative to the B-phases, background, the metric used is weighted accuracy, more detail is specified in section 4.2.8.

5.2.2.1 Discriminant Analysis

The LDA and QDA are used to classify the B and A-Phases. The SE and SP results are presented for each stage and the different methods, in Figure D.7. The accuracy and weighted accuracy are shown in Figure 5.13. For MRMR the performance in both accuracies measures monotonically increases with dimensionality and stabilise between $30 - 40$ features. Using features to build a model the linear boundary is the best shape (better weighted accuracy). For 30 features in LDA, the mean weighted and simple accuracy is 45% and 61%, respectively. Using the principal components, the maximum mean weighted accuracy, 49%, is achieved using 12 components with a quadratic discrimination function. For QDA with 12 principal components, the mean accuracy is 68%. In Figures 5.14(a) and 5.14(b) it is shown the sensitivity and specificity of B and A-phases with MRMR with LDA. SE and SP stabilise as well as the accuracy after 30 features. The performance for B-phases and A1-phases monotonically increases, while for A2 and A3 phases the performances reach a maximum at 10 and 3 features, respectively and, after that, they start to decrease. For B-phases specificity increases and maintains (having a slight decrease at 18 features), while for A1-phases declines with an increase in features the A2 and A3 specificity enhances with the number of features.

In Table 5.5 is represented the confusion matrix for the best combination of features. B-phases are misclassified by the algorithm almost the same number of times as the other classes, around 10%. A1 is confused around 22% of the times with the A2 subtype, and it is less confused as a A3 (8%). Most of the times that A2 is misclassified as A1. The A3 is classified as an A2 and B-phase 30%, of the times, approximately, while with A1 is only 20%. Comparing to the macrostructure classification, here all the A-phases are much more misclassified. Considering the PCA with LDA for A-phases subtypes the disposition is the same. Although here the B-phases are classified as A2 or A3 only 6%, with A1 14% of the times. Despite the same disposition it is notorious that the misclassification of A2 as an A1 increases 10% using the PCA.

An accuracy distribution, for the different patients, with the different methods is shown is Figure 5.15. It can be seen that the both accuracies for QDA + PCA have a gaussian

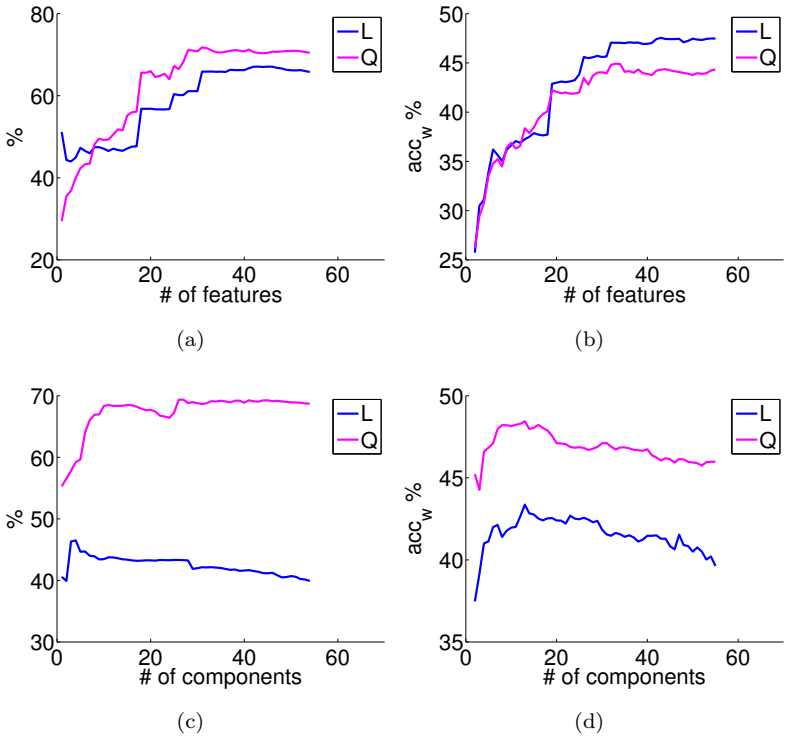


FIGURE 5.13: Performance evaluation, of accuracy simple (a and c) and weight (b and d) for the A phase using and LDA with different number of features chosen by MRMR (a and b) and LDA with different number of components (c and d)

shape, while in the LDA + MRMR they are more equally distributed. Considering the simple accuracy, the maximum obtained for MRMR + LDA was 81% while for PCA + QDA was 75%. The minima registered are 43% and 60%, respectively. Due the fact the fact that LDA method registered both best and worst values, the results are most disperse than the ones obtained for QDA.

Detection

Considering no discrimination between the A-phases subtypes. If the features are used to build a model, the maximum weighted of 72% is achieved for 43 dimensions. The SP, SE and accuracy are 65%, 78% and 67% respectively, for LDA. Using the principal components, the maximum is achieved for 15 dimensions with 76%. The SE, SP and accuracy are 79%, 74% and 73%, respectively for QDA.

5.2.2.2 k-NN

The k values analysed to classify the macro- and microstructure are the same. The quantified results using both accuracies are shown in Figure 5.16. The weighted accuracy

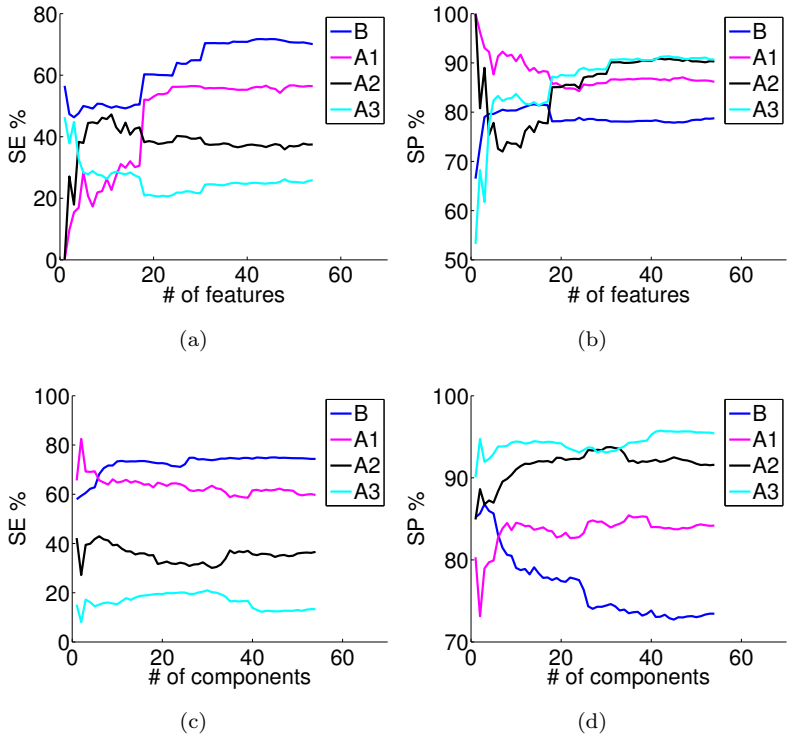


FIGURE 5.14: Sensitivity (a and b), and specificity (c and d), using MRMR+LDA (a and c), and PCA+ QDA (b and d)

is the criterion to choose the best k value, the k that achieved higher weighted accuracy is 25. Considering a k-NN model with 30 features the mean for weighted and simple accuracy is 46% and 70%, respectively. If it is build based on principal components it is 47% and 61%, respectively. In Appendix D can be consulted the SE and SP for each class of CAP in Figure D.10.

The SE and SP for each A-phase subtypes and background are represented in Figure 5.17. Considering the disposition of SE in both feature selection method, the A3 and A2 are the ones with lower SE and the disposition is the same in both methods, independently of the dimension (except in four dimensions). In PCA the A1 is always greater than B-phase for a dimension inferior to 19, although the situation inverts for higher dimensions than 19. Analysing SP obtained with MRMR the classes have approximately the same value from 2 until 19 dimensions. For more than 19 dimensions the SP value of A-phases and to decrease for B-phases decrease. Using PCA the major changes occurs at the beginning and for dimensions superior to 20 the value for each class is practically the same.

The confusion matrix for a k-NN classification problem, with $k = 25$ is presented in Table 5.6. Analysing the table the B and A3 classes have a better agreement with features than with principal components, while for A1 and A2 the situation is in reverse.

TABLE 5.5: mean of confusion matrix for MRMR+LDA linear for a) 30 dimensions and b) PCA+QDA with 12 dimensions, for A-phase prediction

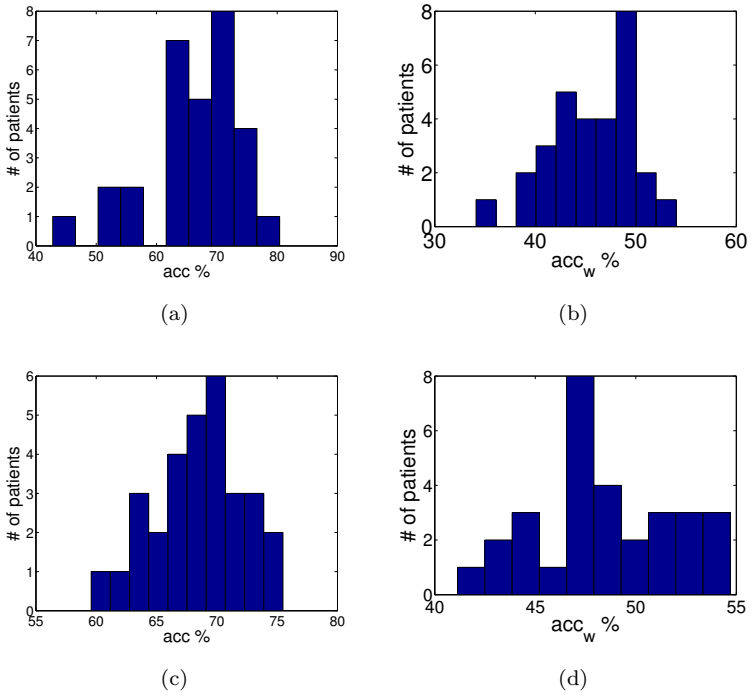
		real (%)			
		B-phase	**A1**	**A2**	**A3**
outcome	**B-phase**	**71 ± 10**	14 ± 10	22± 15	30 ± 12
	A1	12 ± 5	**55 ± 19**	30 ± 18	20 ± 11
	A2	10 ± 8 ± 5	22 ± 14	**37 ± 16**	26 ± 12
	A3	10 ± 6	8 ± 10	11 ± 11	**25± 9**

(a)

		real (%)			
		B-phase	**A1**	**A2**	**A3**
outcome	**B-phase**	**73 ± 5**	13 ± 8	18 ± 9	30 ± 8
	A1	14 ± 4	**66 ± 13**	40 ± 10	25 ± 7
	A2	6.3 ± 2.0	18 ± 9	**37 ± 8**	27 ± 7
	A3	6.1± 2.1	2.3 ± 3	5 ± 4	**18 ± 6**

(b)

The misclassification of B-phase as A1 and A2 with MRMR is practically the same 4.1% and 4.5%, respectively. Although, in PCA, if is quite different being 10.7% and 4.5%, and the sensitivity of A1 in PCA is twofold than in MRMR. The majority of the misclassified



(a) (b) (c) (d)

FIGURE 5.15: Accuracy simple (a and c) and weight (b and d) distribution for MRMR and LDA (a and b) and PCA and QDA (c and d)

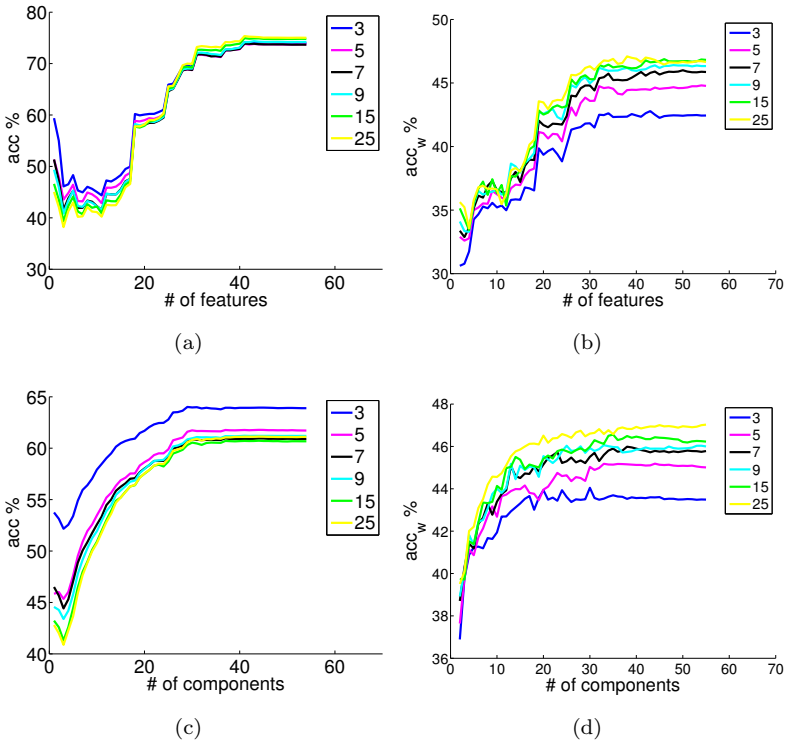


FIGURE 5.16: Performance evaluation, of accuracy simple (a and c) and weight (b and d) for the A phase using and k-NN classification method with features chosen by MRMR (a and b) and PCA (c and d)

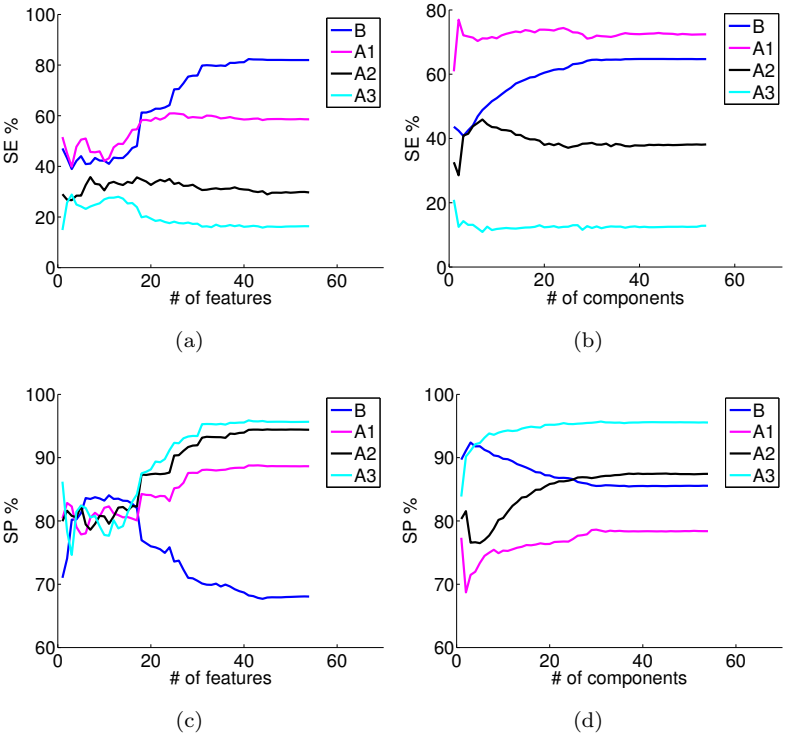


FIGURE 5.17: Sensitivity (a and b), and specificity (c and d), using $k = 25$ for MRMR (a and c), and PCA (b and d)

TABLE 5.6: Mean of confusion matrix for kNN, $k = 25$, with 30 best features chosen by a) MRMR and the b) 30 principal components for A-phase staging prediction

		real (%)			
		B-phase	**A1**	**A2**	**A3**
outcome	**B-phase**	**81 ± 4**	23 ± 9	30 ± 12	40 ± 9
	A1	10.32 ± 3	**59 ± 12**	34 ± 14	21 ± 8
	A2	4.1 ± 1.5	16 ± 6	**31 ± 9**	22 ± 6
	A3	4.5 ± 1.1	2.2 ± 1.8	4.9 ± 3.9	**16 ± 5**

(a)

		real (%)			
		B-phase	**A1**	**A2**	**A3**
outcome	**B-phase**	**65 ± 4**	7.1 ± 2.7	10 ± 4	25 ± 8
	A1	20.15 ± 3	**73 ± 9**	49 ± 10	32 ± 7
	A2	10.7 ± 1.8	19 ± 7	**39 ± 8**	31 ± 7
	A3	4.5 ± 1.4	1.02 ± 0.9	2.4 ± 1.9	**12.7 ± 4.7**

(b)

A1 are B-phases when MRMR is applied and A2 when PCA is applied. All the non-detected A2 as such in PCA are considered A1 (49%) and using MRMR are classified as A1 and B-phases (around 30% for each). For A3 in MRMR, around 20% are considered by the algorithm as A1 and A2, and 40% are considered as B-phases.

The distributions of simple and weighted accuracy are shown in Figure 5.18. Considering the accuracy the high and low values of sensitivity are given with a k-NN build with 30 features is 76% and 63%, respectively for a k-NN model build with 30 principal components is 68% and 52%, respectively. Regarding the weighted accuracy the maximum and minimum values obtained for accuracy with 30 best features is 53% and 39%, respectively and with 30 principal components is 53% and 40%, respectively.

Detection

Using the k-NN model, with $k = 25$ and 55 features, the better performance values are 75% and 80% for simple and weighted accuracy, respectively, and for sensitivity and specificity is 82% and 68%, respectively. Regarding k-NN, $k = 25$, and PCA principal components for the same model, the maximum of accuracy weight is achieve for 54 components with the value 75%. The SE, SP and accuracy simple is 86%, 65% and 68%, respectively.

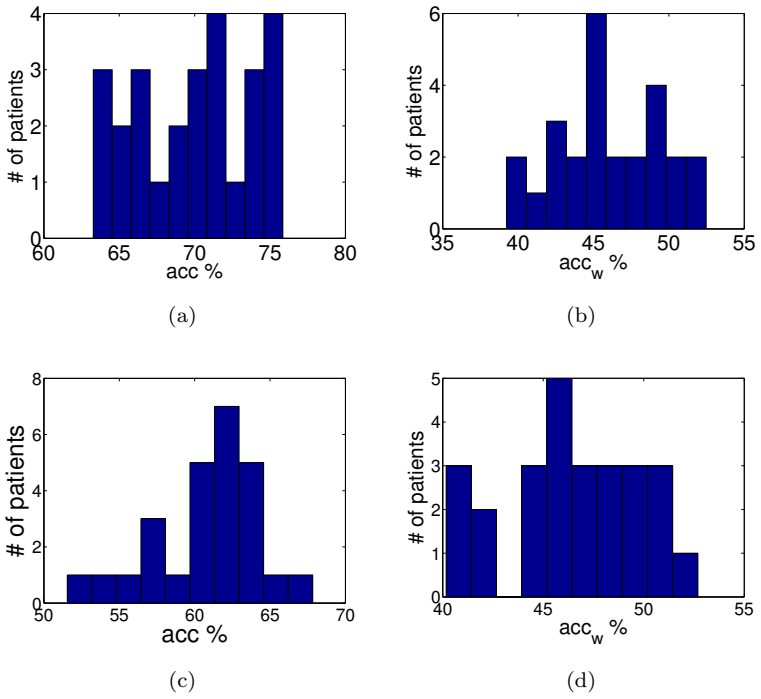
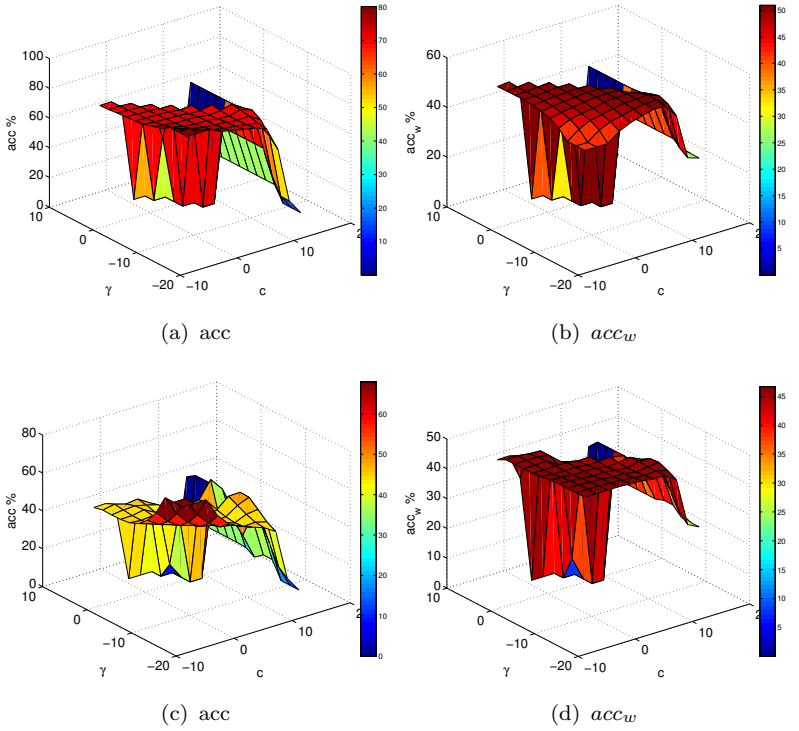


FIGURE 5.18: Accuracy simple (a and c) and weight (b and d) distribution for MRMR (a and b) and PCA(c and d) for a k-NN with a $k = 25$ for 30 dimensions

5.2.2.3 SVM

The SVM is also used for microstructure classification. From the DA and k-NN it can be seen that the PCA reaches stability at around 30 dimensions, while for the features, more dimensions are needed. Thus, two grid searches are performed, one for the 30 principal components and other for 40 features. The accuracy results, for weighted and simple accuracy are shown in Figure 5.19. SE and SP for each A-phase and B-phase class can be seen on the Annex D, in Figures **??** and **??**. The criterion to find the best c and γ is the weighted accuracy, the best values for c and γ are different when the model is build with features or if designed with principal components. For a SVM build with features the better mean weighted accuracy is obtained for the values, $c = 2^{-1}$ and $\gamma = 2^{-1}$, and the mean weighted and simple accuracy are 51% and 71%, respectively. While for a best model build with principal components is for the a $c = 2^{-5}$ and $\gamma = 2^{-9}$, for a weighted and simple accuracy of 47% and 56%, respectively. A confusion matrix for the best c and γ is shown in Table 5.19

Table 5.7, shows that for SVM with $c = 2^{-1}$ and $\gamma = 2^{-1}$, it can be seen that the B-phases are confused around 10% of the times with the A-phase subtypes. Approximately 21% of the A1 phases are confused with A2 but only 3.6% with A3. A2 is more confused with A1, although 20% of these phases are considered by the algorithm as B-phases.



(a) acc (b) acc_w

(c) acc (d) acc_w

FIGURE 5.19: Accuracy (a and c) and accuracy weight (b and d) for each class of CAP events using SVM classifier build with 40 features (chosen by MRMR) (a and b) and 30 principal components (c and d) for a grid search with $c = 2^{-5}, 2^{-3}, ..., 2^{15}$ and $\gamma = 2^{-15}, 2^{-13}, ..., 2^{5}$

A3 is the one with lower sensitivity and the misclassification is quite high, around 30% are confused with A2 or B-phase and 15% with A1. The A3 and A1 are the most distinct types, thus it was expectable that they were less confused. Regarding the results obtained with $c = 2^{-5}$ and $\gamma = 2^{-9}$, the A1 and A2 are more misclassified with only one class, A2 and A1 respectively. Besides, the classes are more correctly classified using features than using PCA, except the A2.

The simple and weighted accuracy for a SVM with 40 features are shown in Figure 5.20(a) and 5.20(b). The maximum and minimum, for the simple accuracy, obtained with this approach is 77% and 63%, respectively. Regarding the weighted accuracy the maximum and minimum is 58% and 42%, respectively. For a SVM based on 30 principal components the distributions are shown in Figure 5.20(c) and 5.20(d). The maximum and minimum simple accuracy for the $c = 2^{-5}$ and $\gamma = 2^{-9}$ are 65% and 52%, respectively. Concerning the weighted accuracy the maximum and minimum are 55% and 38%, respectively.

For this (c, γ) pair a simple and weighted accuracy of 71% and 51%, respectively, is obtained. The detection without discrimination yields 88.3% and the A-phase is not

TABLE 5.7: Mean of confusion matrix for MRMR and SVM for a) $c = 2^{-1}$ and $\gamma = 2^{-1}$ with features and b) for $c = 2^{-5}$ and $\gamma = 2^{-9}$ with 30 principal components, for CAP prediction

		real (%)			
		B-phase	**A1**	**A2**	**A3**
outcome	**B-phase**	**76 ± 5**	15 ± 8	21 ± 12	27 ± 11
	A1	11 ± 36	**59 ± 14**	27 ± 13	17 ± 10
	A2	11 ± 6 ± 3	22 ± 12	**44 ± 15**	32 ± 11
	A3	7.2 ± 2.3	5 ± 4	8 ± 8	**24 ± 10**

(a)

		real (%)			
		B-phase	**A1**	**A2**	**A3**
outcome	**B-phase**	**61 ± 3.1**	5.0 ± 2.4	9 ± 4	26 ± 9
	A1	17 ± 4	**60 ± 17**	35 ± 12	25 ± 8
	A2	16 ± 4	32 ± 15	**53 ± 10**	38 ± 9
	A3	6.3 ± 2.0	2.0 ± 2.5	3.1 ± 2.5	**13 ± 5**

(b)

totally found, but only 76.5%, with discrimination. This values drops to 50.2% for discrimination but only in 42.6% of its duration.

Detection

Considering now a two-class problem where the objective is to separate the A-phases from B-phases, for a SVM with 40 features. A new range of accuracies are obtained. On the one hand, the high weighted accuracy for SVM is achieved with $c = 2^{-1}$ and $\gamma = 2^{-1}$ and is 78%, for this pair of (c, γ) the simple accuracy, SP and SE is 76%, 79% and 77%, respectively. On the other hand, the higher accuracy simple is 84%, for the $c = 2^3$ and $\gamma = 2^1$, and in this case the accuracy simple, SP and SE is 50%, 90% and 70%, respectively.

5.3 Literature algorithms

Some algorithms, described in Chapter 3.2, are implemented and aiming to develop trustworthy comparisons.

5.3.1 MMSD

One of the methodologies, with different versions, were proposed by Barcaro et al. The principle is simple, an A-phases is detected based on a simple model of two thresholds

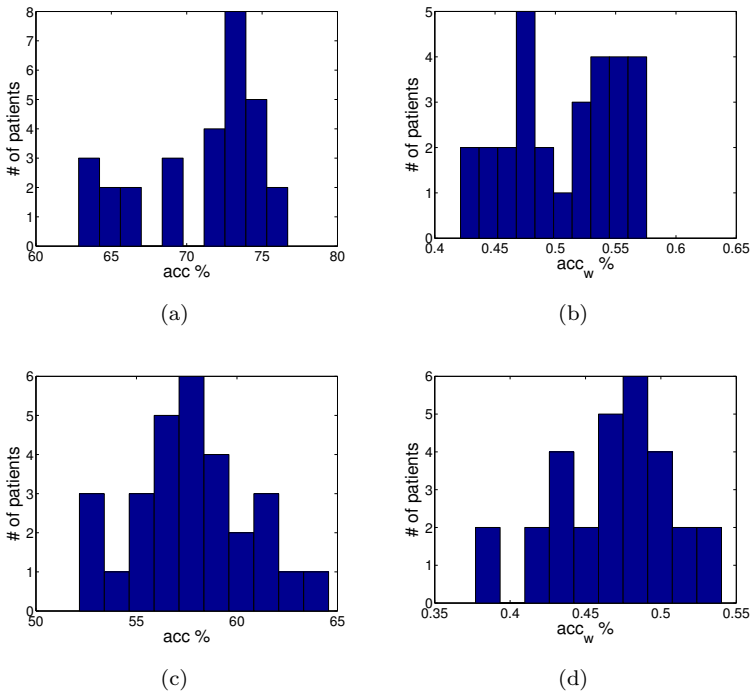


FIGURE 5.20: Distribution of (a and c) accuracy simple and accuracy weight (b and d) distribution for $c = 2^{-1}$ and $\gamma = 2^{-1}$ with 40 features (a and b) and for $c = 2^{-5}$ and $\gamma = 2^{-9}$ with 30 principal components (c and d)

applied to a MMSD feature (section 4.2.3.1). In total, five MMSD features are extracted from the signal (one for each conventional band) and are combined, in different ways, to detect the A-phases.

The algorithm is implemented, using the reported parameters, $\tau = 64$, $\tau_0 = 2$ and a length and existence threshold of 0 and 1, respectively.

The method reported in [57], only detected A-phases, without any distinction between the subtypes. The A-phase is detected using the MMSD_δ, MMSD_θ, MMSD_α and MMSD_σ signals. The specificity, sensitivity, simple and weighted accuracy are 81.2%, 58%, 77.8%, 71.3, respectively. The sensitivity of this method is quite poor, nevertheless the specificity and accuracies are reasonable. The weighted accuracy drops almost 10% relatively to simple accuracy due to the misclassification of A-phases that have a greater impact than B-phase misclassification.

The second version of the algorithm distinguishes the considered two subtypes of A-phase (the A2 and A3 are considered together). The results obtained are in Table 5.8 and the mean of confusion matrix for all epileptic patient is in Table 5.9. The simple and weighted accuracies are 79.4% and 48.7%, respectively. This method classify correctly most of the B-phases, although only 50% of A1 and 7% of A2 and A3-phases. Analysing Table 5.9, around 50% of the A2 and A3-phases are detected but considered as A1-phase.

Moreover, when the method classify an epoch as A3-phase it is rarely wrong, therefore it has a high specificity.

TABLE 5.8: Results of sensitivity and specificity of A-phase subtype with the method proposed in [58]

	B	**A1**	**A2 & A3**
SE (%)	89.0 ± 1.8	51 ± 8	6.5 ± 3.0
SP (%)	49 ± 6	86.5 ± 2.7	99.50 ± 0.63

TABLE 5.9: Confusion matrix with the method [58]

		real (%)		
		B	**A1**	**A2&A3**
outcome	**B**	89.1 ± 1.8	48 ± 9	52 ± 5
	A1	10.6 ± 1.7	51 ± 8	41 ± 5
	A2&A3	0.42 ± 0.25	1.7 ± 1.8	6.5 ± 3.0

Considering the same algorithm but without any discrimination between the A-phases subtypes, the performance values are 89.40%, 51.59%, 83.3%, 69.23% for specificity, sensitivity, simple and weighted accuracy, respectively.

Another algorithm, with different MMSD signal combinations, is used. The results are shown in Table 5.10, and the respective confusion matrix is on Table 5.11, and the simple and weighted accuracy for this method is 71.9% and 48.9%, respectively.

TABLE 5.10: Results of sensitivity and specificity of A-phase subtype with the method proposed in [59]

	B	**A1**	**A2 & A3**
SE (%)	79.3 ± 2.7	57 ± 8	10 ± 5
SP (%)	67 ± 6	76.4 ± 2.7	99.80 ± 0.63

TABLE 5.11: Mean confusion matrix with the method [59]

		real (%)		
		B	**A1**	**A2&A3**
outcome	**B**	79.3 ± 2.7	41 ± 9	28 ± 7
	A1	19.6 ± 2.4	57 ± 8	61 ± 5
	A2&A3	1.05 ± 0.57	1.8 ± 1.9	10 ± 5

This version of the method, has a better A-phases classification. The mean of sensitivity for A1 and A2&A3 increased around 8% and 6%. Although, this increase in sensitivity is accompanied by a decrease in specificity around 10% for A1 and 1% for A3. Regarding the confusion matrix (in Table 5.11) and comparing to the previous method, it can be seen that more 20% of A2 and A3 are classified by the algorithm as A1-phase. If

the subtypes are not considered around 58% and 70% of A1 and A2&A3 are detected, respectively. The algorithm classifies less A1 and A2 phases as a B-phase which translates in a high specificity for the B-phase, although the B-phases correctly classified drops 10%. For detection this method obtained 80.17 ± 2.49, 58.00 ± 8.60, 77.50 ± 2.43, 73.08 ± 3.17 for specificity, sensitivity, accuracy simple and accuracy weight, respectively.

The length threshold values are varied between -1 and 1.5 in steps of 0.05 with the existence thresholds one unit above, to verify if the threshold values proposed in the Barcaro et al methods (which are 0 and 1), are the appropriate to the problem. The ROC curve is shown in Figure 5.21, with black bullets representing the best threshold values for B, A1, A2/A3 phases, and they are different for each classes. Therefore, a mean is performed between B and A1 threshold best points. The A2/A3 is not considered because do not achieve great values. For the length threshold the best values are -0.05 and 0.01 for A1 and B-phase, respectively. Thus, the best length threshold is 0.

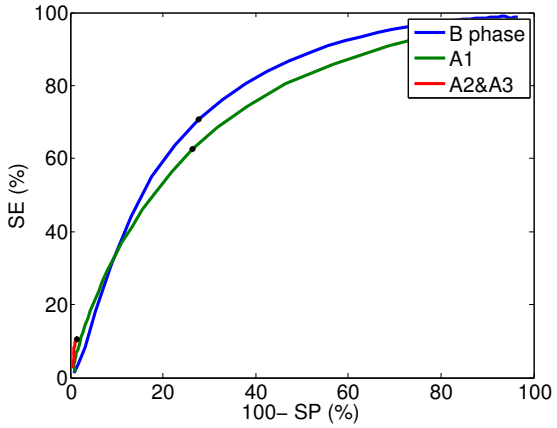


FIGURE 5.21: ROC curve for the methodology proposed in reference [59], for different length and existence threshold values

Comparing the different MMSD features combination, the best performance is achieved with the algorithm proposed in reference [59]. In spite of the reasonable values found for weighted accuracy without discrimination, it is concluded that some subtypes are more correlated with some diseases. For instance, the A2 is greater for epileptic individual than in the control group, while A1 is practically the same. Therefore, is important discriminate the A-phases types. Although, in any version of this model the weighted accuracy is below 50%. The accuracy reported in reference [59] is 84% that compares to 72% obtained with the same method. The difference between the results might be due the number of patients. While in reference [59] uses only 10 patients, here 30 patients are used making this study more reliable due to the statistical increase.

5.3.2 Stochastic Algorithm

As previously explained (see section 3.2), this method simulated the EEG brain with a feedback loop model. The loops are the EEG rhythms, in the conventional frequency bands, each one associated with a gain proportional to the presence of the rhythm in the signal. The expression proposed, to compute the gains (K), assumes that the signal segments, used for estimation procedures, are extracted from a stationary and from a gaussian regular processes [62]. Therefore the KS-test is applied to each epoch, for the EEG and all conventional bands. Two scenarios are tested, epochs with one and two second duration. The results are present in Table 5.12. The Table 5.12 shows that the

TABLE 5.12: Percentage of epochs that rejected the null hypothesis of Kolmogorov–Smirnov, do nor follow a normal distribution, at the 5% significance level for different sleep stages

mean	**sEEG**		**delta**		**theta**		**alpha**		**sigma**	
	1sec	2sec	1sec	2sec	1sec	2sec	1sec	2sec	1sec	2sec
all	57	55	82	83	40	41	23	26	22	27
N3	64	63	83	84	40	42	23	25	24	31
N2	57	55	82	83	40	42	23	26	24	30
N1	45	44	81	83	39	40	24	27	20	23
REM	53	52	80	82	39	41	23	25	19	21
W	41	40	84	85	39	41	28	30	20	23

TABLE 5.13: Percentage of times that a A-phase is included epoch that does not follow a gaussian distribution, for all subtypes and, in particular, for each subtype

mean	**sEEG**		**delta**		**theta**		**alpha**		**sigma**	
	1sec	2sec	1sec	2sec	1sec	2sec	1sec	2sec	1sec	2sec
all	66	66	85	86	41	44	24	28	23	30
A1	77	77	85	87	42	44	23	26	25	32
A2	69	69	84	86	42	45	24	28	23	30
A3	53	54	84	86	40	42	25	29	21	26

rejection of null-hypothesis is greater for deeper sleep stages. With Table 5.13 it can conclude that the majority of A-phases are included in a segment that does not follow a gaussian distribution. Also, the rejection of null-hypothesis is proportional to the size of epochs. Another interesting conclusion is that the epochs corresponding to the lower frequencies have higher null-hypothesis rejection. This is in concordance with the fact that deeper stages have a greater rejection, since they are composed by low frequencies.

The K-values are computed according to the literature [62, 116], for the instances that follow a normal distribution. This feature revealed to be highly discriminative, although it could not be computed for all time instances.

The model applies a filter to transform the scalp EEG in cortical EEG to reverse the attenuation effects from tissues and bones. An exhaustive search in the literature was

performed, aiming, to construct the filter, but not enough information was found. This step was therefore skipped, and might have lead to the high null-hypothesis rejection obtained.

Chapter 6

Discussion

Analysing the features ranking for macrostructure, the eight lower ranked features are equal for all patients. This observation lead the conclusion that these eight features were irrelevant, and the other 47 features were important to the classification process. Although, the accuracy evolution with the number of features for macrostructure classifiers (DA in Figure 5.5(a) and k-NN in Figure 5.8(a)) stabilise for more than twelve features. Therefore, it can concluded that, for macrostructure, only the first 12 best ranked features provide informative and non-redundant information. Contrarily to the macrostructure, the rank sequences in microstructure do not have any resemblance between the patients from a given rank. Thus, it was supposed that all of them were important for the classification task. Besides, since more than 30 features are used to build a classifier, the accuracy does not change significantly (either in LDA, shown in Figure 5.13(a), or in k-NN, shown in Figure 5.16(a)). Consequently, the more informative features for microstructure are the first 30 ranked features.

Considering now the performance evolution with the number of components (PCA) for macrosctructure (for LDA shown in Figure 5.5(b), and k-NN shown in Figure 5.8(b)), the accuracy performance increases until approximately 12 features. After these number of features, the accuracy values are approximately constant until $27 - 28$ features and starts to increase again until 30 features, maintaining afterwards. For microstructure staging (in DA shown in Figure 5.13(c), and in k-NN shown in Figure 5.16(c)), the accuracy function monotonically increases until 30 features, and afterwards the accuracy values maintain. This stabilisation of accuracy for a dimension superior to 30 can be understood by looking at the Figure 5.3, that shows that from 30 components onwards, the information added for a component is reduced or almost zero.

TABLE 6.1: Comparison of the best classifier model proposed in this thesis with the ones present in literature to predict sleep stages.

Method		**sensitivity (%)**					**acc (%)**
		W	**N1**	**N2**	**N3**	**REM**	
DA	thesis	62	47	75	77	74	72
	L. Fraiwan [40]	93	69	79	82	80	79
	V. Helland [42]	83	-	97	94	90	91
k-NN	thesis	52	54	72	84	52	71
	S. Güneş [43]	80	7	89	81	76	65
	H. Phan [47]	99	-	89	81	77	86
SVM	thesis	62	54	73	83	69	72
	T. Lajnef [53]	90	41	70	76	97	75
	B. Koley [54]	96	83	89	99	94	92

The models with the higher accuracy for each classification technique (DA, k-NN and SVM) are considered and compared with some methods presented in the literature. A summary of the comparisons is presented in Table 6.1. Regarding the LDA, two classifiers present in literature, L. Fraiwan [40] and V. Helland [42], are compared with the best model for LDA proposed in this thesis. The LDA classifier proposed by L. Fraiwan have the higher accuracy of 91%, however their model only considers four sleep stages (REM stage and N1 stage are joined) and are only validated with 10 patients. The LDA proposed by V. Helland are applied to 32 patients and considers all the sleep stages, which is more comparable with the LDA of this thesis. V. Helland method has higher sensitivities for the all stages than the ones reported in this thesis.

Two k-NN classifiers present in literature, S. Güneş [43] and H. Phan [47], reported the accuracies of 65% and 86%, respectively. The k-NN proposed in this thesis have a higher and lower accuracy comparing to the previous two methods, respectively. Besides, the H. Phan method do not considers N1 stage and is only performed in four patients, in spite of having a greater accuracy the results are less reliable. Comparing the k-NN of the thesis with S. Güneş method, the first one has a greater N1 sensitivity relative to the second (54% versus 7%), the N3 stage in this thesis its 3% above of the S. Güneş method. Although, for the other sleep stages, N2, W and REM the accuracy is higher in the literature method.

Finally, considering the two SVM classifiers proposed in literature, T. Lajnef [53] and B. Koley [54], present accuracies of 74.8% and 92%, respectively. In comparison, the SVM proposed in this thesis has higher sensitivities for N3, N2 and N1 stages than the T. Lajnef SVM classifier, but for REM and W the literature method has the better performance. Relative to B. Koley SVM method it has always a better performance for all stages.

All the methods present in literature report N1 as the sleep stage with lower sensitivity, the proposed classifiers in this thesis also follow this tendency. The stage with more discrepancies between our classification methods and the literature is W stage. Some hypnograms, used in this thesis, were analysed by a technician expert in sleep staging to verify if the W epochs are correctly classified. It was noted that there exist some A3 in the database which occupy more than 50% of the epochs, and are complete epochs (30 seconds), and the epochs with these A3 are considered one of the NREM stages. Besides an A3 is an arousal and, according to the ASSM rules, when an arousal occupies more than 50% of an epoch this epoch should be classified as W stage. In their database the minimum duration of an W stage is 60 seconds, when exist a transition for W stage with less than 60 seconds they considered that is an A3 subtype in an NREM stage. The NREM stage assign to the W stage is the equal to the previous epoch. Consequently, there exist stages classified as NREM which have characteristics of W stage. This have a great impact in algorithm performance. Once the algorithm 'learn' with the training cases exist epochs belonging to the W stage that are wrongly are labeled as NREM. Consequently, the algorithm will learn in a wrong way since it learn based on wrong examples. This might be an explanation for lower performance obtained. Other explanation might be due the use of a moving window with a length of 1 or 3 seconds (depending on the feature) while in literature the length used is 30 seconds. Besides, observing the Table 3.2 it can be seen that some of the high accuracies are obtained when N1 is joined or not considered or a short number of patients are used to validate the model. Concluding, in general the approach presented in this thesis have an acceptable performance since the number of patients used is 30, considered all sleep stages and in the dataset some instances are mislabeled.

The majority of the models for automatic staging of A-phases are only for detection without discrimination among different subtypes. In this thesis it is evaluated the proposed approach considering both detection and discrimination performance.

The detection of the methods proposed in this thesis and the ones present in literature are shown in Table 6.2. The method proposed by Rosa et al [60], is not implemented in this thesis because it does not fulfil the requirement of normality. Besides, the method is only tested in four patients, which makes their validation less reliable. In general, the accuracy reported in literature is higher than the ones for the methods reported in this thesis. This is due to the fact that they reported the standard accuracy that not accounts for class imbalance. Therefore they give more importance to the B-phases, that are in majority, than to the A-phases. Comparing their values of SE and SP with the one obtained with the different methods proposed in thesis. The SP presented in the literature has higher SE values. In models proposed in this thesis, the data imbalance is taken into account, which leads to a higher SE than SP, and consequently to a lower

TABLE 6.2: Comparison of the best classifier model proposed in this thesis with the ones present in literature to detect A-phases.

Method	**paper**	**SE (%)**	**SP (%)**	**acc (%)**
DA	thesis	79	74	76
	Mariani, S. [67]	72.5	86.6	84.9
k-NN	thesis	82	69	75
SVM	thesis	79	76	77
	S Mariani [67]	70	84	82
MMSD	Barcaro et al [57]	58	81	78
	Barcaro et al [58]	52	89	83
	Barcaro et al [59]	58	80	78

accuracy comparative to literature methods. The highest accuracy mean for a SVM with 40 features is 84%, although the SE and SP are 50% and 90%, respectively. This results are comparable to the literature, although the algorithm goal is not to find the best accuracy but to have the best trade off SE and SP, given more importance to SE since they correspond to rare events.

Only two methods [58] and [59] that discriminate between A-phases are found in literature, proposed Barcaro et al. In these methods the subtypes A2 and A3 are joined and classified as one, A2&A3. Due to its simplicity they are implemented in this thesis and their results, as well the results for the other methods, are described in Table 6.3. Observing this table it can be seen that the sensitivities of A2&A3 are quite low, in the two versions proposed by Barcaro et al (below the 10%). Although, the accuracy of this method is high. As explained before, this value is due a high sensitivity of the B-phases, which occupies 70% of the all data. Then, due to the a high sensitivity of B-phases, the accuracy will be high as well, even if the sensitivity of A-phases is low. The methods proposed in this thesis for A-phases discrimination are built to detect the three A-phases subtypes, which is an innovation. All the A-phases sensitivities are higher than the literature methods. The accuracies relative to the A-phases, in descending order, are: A1, A2, and A3 in all methods. Which is interesting since the frequencies in A1 to A3 increase. Most of the A3 are arousals which are events with characteristics of W sleep stage. Therefore, the model was trained with B-phases similar to A3 phases, which lead to a lot of A3 phases being classified as B-phase. This can be seen observing the confusion Tables 5.5, 5.6 and5.7. The characteristic waves of the A1-subtype are only typical of it (an K-complex) which explained the higher results for this subtype, compared to the A2 and A3. Therefore, to identify better an A3, it might be useful to have information about the context (sleep stage).

According to the ASSM rules an A-phase should have a minimum duration of 2 seconds, and maximum of 60 seconds. Therefore it is possible for a subtype A3 to have a duration

TABLE 6.3: Comparison of the best classifier model proposed in this thesis with the ones present in literature to predict the A-phases subtypes

Method	**paper**	**sensitivity (%)**				**acc (%)**
		B	**A1**	**A2**	**A3**	
LDA	thesis	73	66	37	18	68
k-NN	thesis	81	59	31	16	70
SVM	thesis	76	58	44	24	71
MMSD	Barcaro et al [58]	89	51	6.5		79.4
	Barcaro et al [59]	79.3	57	10		71.9

of 40 seconds, and this event corresponds to an EEG arousal. Although, an arousal with 40 seconds is against the rules to scored the macrostructure in both manuals (R&K and ASSM). As mentioned before, the signal is divided into epochs of 30 seconds and the signal is analysed. An epoch is classified based on the majority of the signal characteristics. The previous example has a duration superior to 15 seconds, so the EEG arousal occupies 50% of the epoch, so in this case, the epoch should be classified as a Wake state than as an NREM stage with a A3 phase. Therefore, it is a disagreement between the two scoring rules which should be revised.

Chapter 7

Conclusion and Future Work

This thesis presents a machine learning multi-feature approach to classify both sleep macro and microstructure. Concerning macrostructure, the proposed approach presents an improved performance when the classifier is of the SVM type. The better performance achieved was an accuracy of 72% which is not always better than the results published in the bibliography. However, it is believed that the hypnogram available may have problems. Future steps will encompass the review of the hypnograms aiming to improve the macrostructure classification.

Concerning microstructure algorithms are new comparing to the microstructure the results are satisfactory and from an overall point of view better than the performance obtained with the algorithms in the literature. The best performance for microstructure is with SVM with a accuracy of 71%, and a sensitivity of 76%, 58%, 44%, and 24% for B, A1, A2 and A3-phases, respectively. The lower sensitivity for A3 might be resolved taken in count the sleep stage. Therefore a future work is used the output of a microstructure algorithm to predict microstructrure. The CAP sequences should be studied as a precursor of epileptic seizures.

Future steps will encompass the consideration of the macro-structure classification as an additional information for micro classification. This information could be important to discriminate A3 phases, given that it can be confused with the awake state. Other future step is the analysis of the benefits of micro-structure staging as a disease biomarker, more particularly as a precursor of epileptic seizures.

Appendix A

Statistical Analysis

One of the objectives of this thesis is to evaluate the utility of A-phases to predict epileptic seizures. It was already proved that some pathologies, like insomnias, bruxism, epilepsy, etc., are related with higher rates of CAP. CAP, in normal conditions, only appear in NREM sleep. Therefore, the percentage of CAP also depends on the sleep stages distribution through the night.

This Chapter focus to compare the percentage of each sleep stage and A-phases subtypes in a night of sleep in epileptic patients comparing to a control group.

A.1 Materials

Besides the dataset presented in Table 4.1 it is also used a control group (without epilepsy) designated by N1-N16 is also used. To analyse the A-phases evolution with seizures patients from Centro Hospitalar e Universitário de Coimbra (CHUC) are used, Table A.1.

The N1-N16 have the same properties than NFLE1-NFLE40 described in section 4.1.

EPI1-EPI2 were recorded at CHUC including, at least, three EEG channels (F3 or F4, C3 or C4 and O1 or O2, referred to A1 or A2) but without the EMG and EOG. The staging was performed by a technician from Centro de Sono do Hospital dos Covões. The macrostructure was scored according to the ASSM rules and the A-phases according to the Terzano reference atlas [28].

The technician was not trained by specialists but instead learned based on the examples provided in the CAP sleep database [69]. To validate the scores, the patient NFLE12

Number	File name	Sex	Age
1	N1	F	37
2	N2	M	34
3	N3	F	35
4	N4	F	25
5	N5	F	35
6	N6	M	31
7	N7	M	31
8	N8	F	42
9	N9	M	31
10	N10	M	23
11	N11	F	28
12	N12	M	29
13	N13	F	24
14	N14	F	35
15	N15	M	34

patient	Sex	Age	night	# of seizures
EPI1	M	55	night 2	2
			night 3	0
EPI2	M	18	night 2	3
			night 4	0

TABLE A.1: Patients used in statistics analysis, the control patients in left side and the patients to study the evolutions of A-phases in right side

was used to compare the output obtained by Covões technician to the one provided in database. The comparison is shown in Table A.2. In [2] it is presented an agreement between four scores in 11 patients relative to A-phases, a summary of the results is presented in Table A.3. One of the parameters in Table A.3 is the Kendall's coefficient of concordance [117] which is a nonparametric test to measure the concordance among the raters. It can be seen that agreement is greater in A1, and lower in A3, in duration as well as in total number of A-phases. Analysing the Table A.2 it can seen that the differences between the scores are in concordance with the ones found in literature.

TABLE A.2: Comparison between two scores of NFLE12 patient

	italian scorer	covões scorer
Total no. A1	227	254
Total no. A2	122	133
Total no. A3	205	243
A mean duration (s)	6.97	7.52
A1 mean duration (s)	4.93	5.59
A2 mean duration (s)	6.27	7.34
A3 mean duration (s)	9.64	9.63

The total agreement, without discrimination between the subtypes is 78.58%, and for A1, A2, and A3 is respectively 74.38%, 64.71%, 73.94% considering that the italian score is correct. These values, only considering the A-phases, are the ratio between A-phase scored at Covões which are in concordance the italian database and the total number of A-phases of the italian scorer. The global agreement, non A-phase and subtypes of

TABLE A.3: Analysis of the inter-rater variability of A-phases parameters between the four different scorers and the Kendall's coefficient of concordance (W) ((Adapted from: [2])

CAP Parameter	Scorer 1		Scorer 2		Scorer 3		Scorer 4		W
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	
Total no. A1	262.5	81.16	277.0	89.58	378.6	90.59	358.6	114.98	0.844
Total no. A2	111.8	57.99	42.8	18.98	30.5	20.32	41.3	22.04	0.713
Total no. A3	41.9	24.71	34.5	18.97	13.5	11.27	27.7	15.73	0.527
A mean duration (s)	8.4	0.86	10.5	2.37	7.0	0.90	8.8	1.29	0.646
A1 mean duration (s)	6.7	0.95	8.8	2.07	6.4	0.75	7.8	1.17	0.535
A2 mean duration (s)	10.2	1.79	13.9	4.63	11.9	3.52	11.8	2.17	0.772
A3 mean duration (s)	14.6	2.00	20.5	5.43	14.0	5.04	17.2	4.86	0.519

A-phases is 91.33%. Therefore, even though the Covões scorer lack of training with experts, the results are accordance to the literature, thus are reliable.

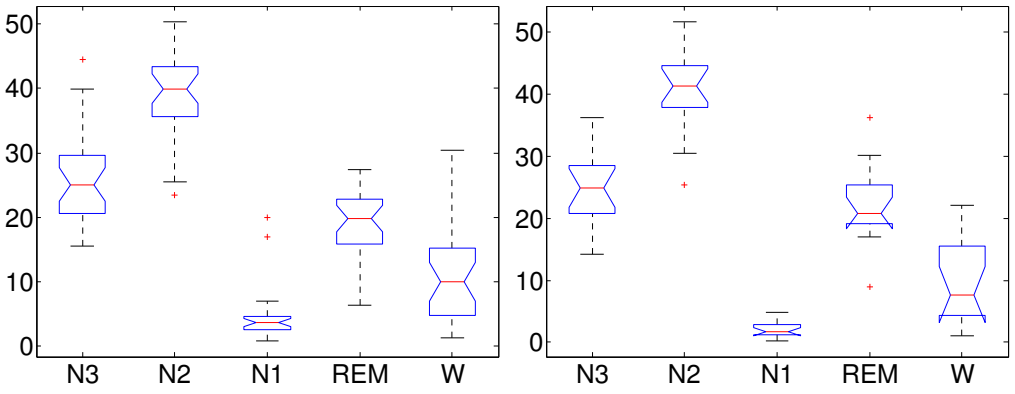
A.2 CAP presence analysis

This section presents a comparison between a normal and epileptic group relative to A-phases and sleep stages distribution, using different statistical tests. The control patients are also available on database, and were considered in this part. The output of a statistical test is the p-value which ranges from 0 to 1, and p is proportional to the certainly of null-hypothesis. Is considered that exist statistical difference, null-hypothesis rejection, when the p-value is less than 0.05 or 5% [118].

Regarding the sleep stages, their percentage through the night, for all patients in a determined group, is represented in Figure A.1. At first glance, it appears that no different exist between the two groups, comparing the sleep stages. Besides, a statistical test must applied to verify this conclusion. The choice of the proper test depend on the data distribution, if it has or not a normal distribution. Thus, at first, the Kolmogorov–Smirnov (KS) test [119, 120], whose null-hypothesis is that data come from a standard normal distribution, was applied to distribution of each sleep stage, in both groups. When data follows a standard distribution the appropriate test is one-way ANOVA [121, 122] while if does not the a non-parametric test must be applied, Kruskal-Wallis [123]. The null hypothesis, in the these two tests, is the same: The groups come from the same population.

The null hypothesis was never rejected, thus, the proper test to verify if the sleep stages distributions is affected by NFLE is the ANOVA test. The results, of p-values, are present in Table A.4. Considering the significance level imposed, the null hypothesis was never rejected, then the sleep stages are have, approximately, the same presence in individuals with and without epilepsy. The stage with similar appearance in both groups

is N2, contrarily to N3 which is the most different. It can also be seen, in Figure A.1, that N2 stage prevailed over the others sleep stages while N1 is the one with less presence.



(a) control group (b) epileptic group

FIGURE A.1: Sleep Stages percentage for the both groups study

TABLE A.4: p values for ANOVA test for the sleep stages with the two groups

	N3	**N2**	**N1**	**REM**
p value	0.26	0.97	0.71	0.64

Considering the A-phases, it was computed, for each patient, the ratios between number of seconds occupied by an A-phase(s) in stage x and number of seconds of stage x, where x are one of the sleep-stages N1, N2, N3, NREM. As previously, the KS-test must be applied to decide which test should be used to verify if there are differences between the epileptic and control group, ANOVA or Kruskal-Wallis. For all ratios, except NREM stage with A1 and A2 subtypes, the distribution of patient's ratio follows a standard distribution. Therefore, for this patients the ANOVA test was applied, and for the other two cases (NREM stage with A1 and A2 subtypes) the non-parametric test, Kruskal-Wallis, was performed. The results for the mean of the ratio in each group and the p-value, for the test that verify if the groups come from the same population, are present in Table A.5.

TABLE A.5: Mean of percentage of A phases in one or more NREM sleep stages for epileptic population ($\bar{x}_{epi}$) and a control population ($\bar{x}_{control}$), p-value of the ANOVA test comparing both groups, epileptic and control.

	all subtypes			**subtype A1**			**subtype A2**			**subtype A3**		
	$\bar{x}_{epi}$	$\bar{x}_{control}$	p	$\bar{x}_{epi}$	$\bar{x}_{control}$	p	$\bar{x}_{epi}$	$\bar{x}_{control}$	p	$\bar{x}_{epi}$	$\bar{x}_{control}$	p
NREM	14.50	10.30	1.70×10^{-4}	5.48	4.81	0.28	3.16	2.11	6.17×10^{-3}	5.86	3.38	1.76×10^{-4}
N1	20.25	23.35	0.47	0.59	0.75	0.65	0.29	2.50	0.48	19.36	20.09	0.87
N2	19.89	12.98	8.94×10^{-5}	3.90	3.34	0.38	5.55	3.58	0.007	10.45	6.06	2.70×10^{-4}
N3	22.45	18.33	0.023	15.53	14.10	0.39	3.64	2.32	0.04	3.28	1.89	5.38×10^{-3}

The generic case, for all A-phases subtypes in NREM stage, have p-value below 0.05. Consequently, the A-phases occupied more space in NREM stage in epileptic group than

in control group, and this difference is significantly. Regarding the A-phases subtypes in NREM, the A1 have the same presence in both groups, on contrary, the A2 and A3 are increased in this disease. Analysing the p-values for N1 stage and for A1 subtype, they are always above the significance level, thus, N1 stage and A1-phase are not affected by NFLE. For A2 and A3 in N2 and N3, the null-hypothesis is rejected, therefore the NFLE interfere with this subtypes and results, always, in an increase of the A-phases presence. Whether the lower is p-value more different is the presence of an A-phase subtype between the groups. Therefore, the greater difference between the groups is in stage N2 considering all A-phases.

Besides the differences, some conclusions are common for both groups. In N1 the subtype A3 appearance prevails over A1 and A2. Doing deeper, in N2 the disposition remains, however, the A1 and A2 are more present while A3 subtype drops, at least, 50%. In N3, on the contrary, the majority of the A-phases subtypes are A1. This disposition was expectable, since the A1 is a subtype characterised by a decrease in frequency relative to the background while A3 is by an increase. Therefore, the A1 must appear in the stage composed by high-voltage slow waves, N3, whereas A3 must be present in the stage with low-voltage fast waves.

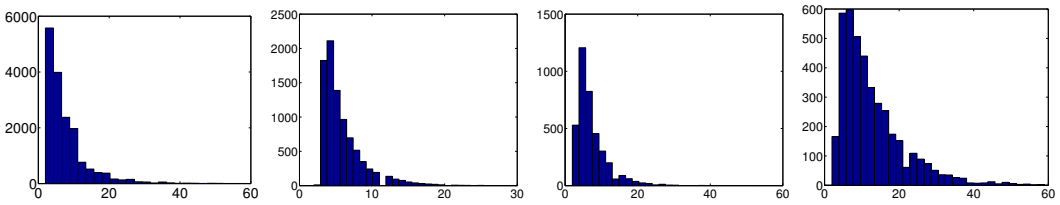
Respecting to A-duration, it was compared the mean duration of A-phases and the number of A-phases in patients with epilepsy. The increase of the number of seconds of A-phases in a determined stage can be due the increase of the number of A-phases or to the duration of A-phases. Once more, to analyse if the epilepsy alters the duration or number of A-phases, the ANOVA or Kruskal-Wallis must be applied. The KS-test was performed for the two groups, and the null-hypothesis was never rejected. Thus, ANOVA was applied for all cases, a resume is shown on Table A.6. For A-phases duration, the the null hypothesis was never rejected, consequently, the NFLE does not interfere, significantly, with the the A-phases duration. Looking at A-phases number it can be seen that the null.hypothesis is rejected for all A-phases, A2 and A3 subtypes. Therefore, the differences observed in A.5 do not results from A-phases more longer but, instead, in the A-phases number.

TABLE A.6: Mean duration and A-phases number in epileptic and control group

	all subtypes			**subtype A1**			**subtype A2**			**subtype A3**		
	$\bar{x}_{epi}$	$\bar{x}_{control}$	p	$\bar{x}_{epi}$	$\bar{x}_{control}$	p	$\bar{x}_{epi}$	$\bar{x}_{control}$	p	$\bar{x}_{epi}$	$\bar{x}_{control}$	p
duration	7.79	7.72	0.98	5.77	5.88	0.57	7.23	8.10	0.20	13.69	13.75	0.37
number	557	402	1.0×10^{-4}	294	246	0.14	128	80	2.6×10^{-3}	135	76	4.0×10^{-4}

Considering only the A-phases of epileptic patients, their the distribution, overall, and concerning the A-subtypes are shown in Figure A.2. In total, 16704 A-phases exist in all dataset, used in this thesis, being 8805, 3840 and 4059 A1, A2 and A3, respectively. The minimum duration, two seconds, of an A-phase is independently on the subtype.

However, for maximum duration, the situation is different, A1, A2 and A3 have 29, 57 and 59 seconds, respectively. The number of A1 is, at least, twice than the number of A2 and A3. Although, the maximum duration of A2 and A3 is two times greater than the observed for A1. Observing the Figure A.2, the A1 curve is more narrow than the A3 duration curve, thus, most of A3 have a major duration. This explains why in spite of the A1 have double number of phases, in NREM the number of seconds occupied by A3, 5.86%, is greater than the A1, 5.81%, Table A.5.



(a) A (b) A1 (c) A2 (d) A3

FIGURE A.2: Distribution of the A-phases duration in patients with epilepsy

Statiscal example from another disease with CAP [124], paroxysmal dystonia where the patients do not show a change in macro but in micro yes.

A.2.0.1 CAP predict seizure

The percentage of the sleep stages in the patients from the *CAPsleep database* and *CHUC* are in Table A.7. The W stage are much greater in patients from *CHUC*, and the percentage of N3 is below the average of the patients from the CAP sleep database. This results might be associated with the hospital conditions, which in CHUC worse than in italian hospital.

TABLE A.7: Percentage of each sleep stage, since the moment that patients started to sleep until they awaken

	N3 (%)	**N2 (%)**	**N1 (%)**	**REM (%)**	**W (%)**
database	27.31±7.40	40.70 ± 6.85	4.81 ± 4.24	19.92 ± 5.27	7.27 ± 6.59
EPI1 with crisis	14.01	31.01	17.58	9.86	27.55
EPI2 with crisis	4.69	37.64	13.09	10.10	34.49
EPI2 without crisis	19.36	36.67	20.64	0	23.33

For patient 285, in the night with a epileptic crisis, comparing with the normal night, the W stage have more presence whereas N3 have less appearance. Relatively to A phases, A3 is higher in the night with crisis while the values for the others subtypes do not differ much. The range of A1 of CAPsleep database include the values found for the CHUC patients. Although, for A2, in patient 169 the values is quit high while 256 have a low

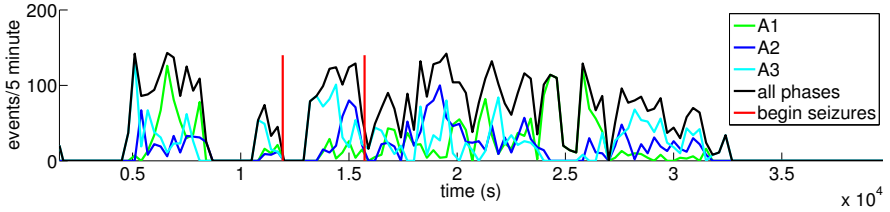
TABLE A.8: Percentage of each A-phase subtype in NREM stage, since the moment that patients started to sleep until they awaken

	A1 (%)	**A2 (%)**	**A3 (%)**
database	8.12 ± 3.04	4.39 ± 1.60	8.23 ± 2.96
EPI1 with crisis	10.79	9.60	13.04
EPI2 with seizure	9.28	0.8260	9.81
EPI2 without seizure	8.46	0.70	5.28

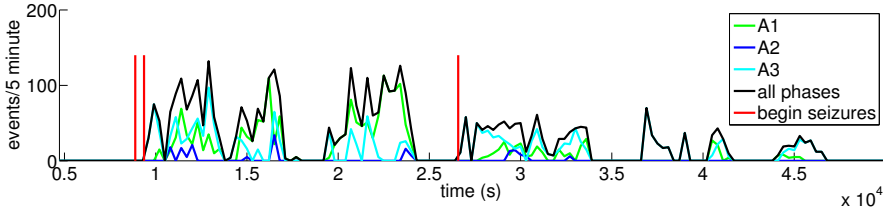
values, for both nights. Although, this discrepancies could be related with the patient characteristics.

In Figure A.3 is represented the rate, per 5 minutes, of the A-phases and, in patients from CHUC with seizures, the are marked. The only remarkable thing that can be concluded is that the percentage of A tends to diminish along the night. Although, the same, is not observed, with such evidence, in patients from CAP sleep database.

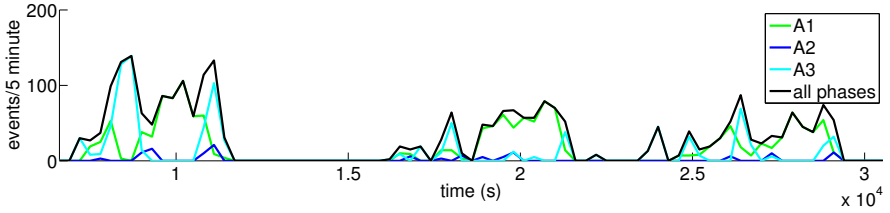
Finally, no relationship was found between the A-phases appearance with the begin of the seizure. However, in literature is reported that the appearance of A1-subtype is prominent before the seizure whereas the A2 and A3 are likely after.



(a)



(b)



(c)

FIGURE A.3: Distribution of the A-phases rate (five per minute) for patient EP1 (a) and EP2 (b and c) in a night with (a and b) and without seizures (c)

Appendix B

Sleep Elements

- **Alpha rhythm:** Trains of sinusoidal $8-13$ Hz activity recorded over the occipital region with eye closed, atenuating with eye opening.
- **Eye blinks:** Conjugate vertical eye movements at a frequency of $0.5 - 2$Hz present in wakefulness with the eyes open or closed.
- **Rapid eyes movement (REM):** Conjugate, irregular, sharply peaked eye movement with an initial deflection usually lasting < 500 ms. While rapid eye movement is a characteristic of REM stage, they may also be seen in wakefulness with the eyes opened when subjects scan the environment.
- **Slow eye movement (SEM):** Conjugate, reasonably regular, sinusoidal eye movement with an initial deflection usually lasting > 500 ms
- **Low Amplitude, mixed frequency activity:** Low amplitude, predominantly $4 - 7$ activity.
- **Vertex sharp waves (V waves):** Sharply contoured waves with duration less than 0.5 seconds maximal over the central region and distinguishable from the background activity
- **Sleep onset:** The start of the first epoch scored as any stage other than stage W.
- **K-complex:** A well-delineated negative sharp wave immediately followed by a positive component standing out from background EEG, with total duration greater than 0.5 seconds, usually maximal in amplitude when recorded using frontal derivations. For an arousal to be associated with a K-complex, it must start no more than one second after termination of the K-complex.

- **Sleep spindle:** A train of distinct waves with frequency $11 - 16$ Hz (most commonly $12 - 14$ Hz) with a duration greater than 0.5 seconds, usually with a maximal in amplitude using central derivations.

- **Slow wave activity:** Waves of frequency $0.5-2$ Hz and peak amplitude $> 75\mu V$, measured over the frontal regions.

- **Delta bursts:** Sequence of at least two waves in the frequency bandwidth ranging from 0.5 to 4 Hz and with an amplitude 1/3 higher, or more, than the background activity. Delta bursts are most prominent in the fronto–temporal regions. They may appear in consolidated stage N2, and become increasingly common in stages N3 os sleep. Compared to the background EEG delta rhythm of stages N3, delta bursts tend to be lower in frequency.

- **Vertex sharp transients:** Transients are EEG potentials of 50–200 ms duration and variable amplitude (up to 250 mV) expressed maximally on derivations at the central vertex areas. Sequences of vertex sharp transients, composed of two or more repetitive potentials lasting 2 s or more, often appear at the transition from stage N1 to stage N2 of sleep.

- **K-complex sequences:** Series of two, or more, consecutive K- complexes. Each K-complex presents a bi-/triphasic pattern consisting of an initial rapid negative component followed by a slower positive wave. The K-complex may be mixed with or followed by a sleep spindle. The duration of a single K-complex ranges from 0.5 to 2 s; therefore, a K-complex sequence duration is generally two seconds. K-complex sequences can appear in sleep stages N2 and N3.

- **Polyphasic bursts:** Clusters of high-voltage delta waves, intermixed with theta, alpha or beta rhythms. Polyphasic bursts can include two or more delta peaks and occur in sleep stages N2 and N3. However, polyphasic bursts most commonly appear in stage N2, especially before REM sleep onset.

- **K-alpha:** K-complex followed immediately by an alpha burst.

- **Intermitten alpha:** Alpha EEG rhythm (8 - 13 Hz), usually prominent in tracings from posterior derivations is commonly recorded from occipital areas. At sleep onset, the alpha rhythm field tends to spread anteriorly, then, in sleep stage N1, it fragments into intermittent sequences and, finally, as sleep progresses, it disappears. Alpha EEG activity may also increase in amplitude and decrease in frequency just before it vanishes. In addition to occurring at sleep onset, intermittent alpha may appear when stage 1 reemerges and during REM sleep.

- **EEG arousals:** Sudden frequency shifts towards faster rhythms (theta, alpha, beta, but not spindles) that shortly interrupt sleep continuity for 3 s.

Appendix C

Feature Selection Results

	patient																												
rank	**2**	**3**	**5**	**6**	**8**	**9**	**10**	**11**	**12**	**14**	**15**	**16**	**17**	**18**	**20**	**21**	**22**	**24**	**26**	**28**	**30**	**31**	**32**	**34**	**35**	**36**	**37**	**38**	**39**	**40**
1	27	9	4	9	21	9	9	21	27	4	21	9	4	2	4	21	27	4	9	9	21	4	21	9	5	21	6	21	10	21
2	10	21	6	1	6	21	6	48	8	6	4	21	21	21	3	27	4	6	2	27	3	21	3	34	3	9	4	4	6	5
3	9	34	34	7	8	4	10	4	21	3	27	27	7	4	21	2	7	2	10	4	4	1	5	4	21	4	7	10	4	2
4	5	4	9	4	10	1	7	5	6	1	18	4	6	7	9	7	6	21	8	6	2	9	6	1	8	6	9	9	21	6
5	4	8	1	6	9	7	1	1	3	2	9	5	9	34	8	4	21	9	1	2	5	3	10	2	10	8	21	2	5	10
6	6	1	5	2	1	6	4	6	5	8	1	7	2	9	7	10	1	1	4	1	6	17	2	6	1	1	1	6	3	4
7	21	7	3	8	3	19	21	10	1	21	10	8	1	1	5	1	2	7	5	5	8	10	4	10	6	2	2	5	1	1
8	1	37	21	19	7	2	3	35	7	9	2	18	34	3	2	6	9	48	7	21	7	2	1	21	34	3	8	8	9	7
9	34	3	2	21	5	35	2	34	10	48	6	10	27	6	18	5	5	3	6	7	9	8	27	3	27	34	34	7	8	34
10	3	10	10	3	4	3	5	20	9	5	8	3	5	8	1	3	10	16	3	16	1	6	7	7	2	7	48	1	2	9
11	7	6	8	20	27	18	8	19	2	18	7	2	8	17	6	8	48	8	21	10	10	5	34	8	9	27	3	34	34	27
12	12	27	7	17	2	16	16	16	18	17	5	34	3	19	10	17	8	20	20	3	17	27	8	16	16	16	35	3	27	18
13	2	2	19	5	16	48	48	45	4	7	3	6	16	35	19	20	16	35	16	20	18	7	16	5	7	5	16	27	7	3
14	8	36	16	34	34	17	27	2	34	10	16	1	10	20	48	19	18	19	27	18	34	35	9	18	4	11	13	14	35	8
15	16	5	20	18	17	34	35	3	20	35	34	19	35	18	36	34	34	18	34	34	19	34	18	35	11	54	14	11	18	41
16	11	18	11	16	35	8	17	9	19	16	35	20	17	16	17	9	35	17	19	19	16	16	35	17	18	12	17	36	48	11
17	36	35	15	12	48	20	18	18	16	27	19	13	36	14	20	16	3	34	11	17	48	48	48	48	13	51	18	45	17	37
18	51	17	17	15	40	11	34	47	11	19	48	41	19	48	35	18	20	11	36	36	35	20	12	14	20	13	19	43	16	20
19	37	16	14	14	14	14	19	7	12	34	20	12	48	54	37	41	11	12	40	41	11	18	17	19	55	10	27	40	37	17
20	41	12	54	13	19	33	13	12	17	12	17	35	41	12	11	40	13	50	17	40	13	19	40	11	19	15	15	20	36	12
21	13	19	35	10	51	54	20	8	13	20	11	36	13	52	16	36	55	14	39	42	20	11	19	53	12	35	11	44	19	19
22	40	13	12	37	33	15	11	17	42	53	13	40	40	51	41	14	19	36	15	54	27	12	51	36	36	53	36	54	41	42
23	42	33	18	55	18	51	14	39	55	11	12	37	18	27	40	51	17	5	12	37	14	51	11	12	17	14	53	51	42	36
24	15	40	48	11	50	12	54	11	41	13	36	16	37	11	38	12	14	51	46	39	53	13	14	41	37	44	12	41	20	39
25	52	41	13	54	54	13	51	13	48	14	50	17	11	50	39	37	15	15	18	51	54	46	50	20	15	19	54	13	39	13
26	14	43	55	40	39	5	12	46	35	15	53	39	39	53	14	55	53	10	35	14	41	14	52	54	41	33	33	42	43	38
27	33	42	53	36	41	50	50	50	14	51	51	11	51	43	42	50	12	37	43	8	39	33	13	39	48	45	50	47	13	40
28	50	39	36	50	13	10	33	53	51	33	55	48	14	33	54	11	39	27	51	43	40	50	43	51	14	50	5	12	44	14
29	45	51	40	48	42	55	15	33	54	54	40	43	20	5	44	52	51	54	48	50	12	15	53	40	51	18	10	52	14	51
30	39	48	33	52	15	27	53	15	15	40	14	15	42	55	12	15	33	33	41	12	50	55	20	37	38	55	51	15	53	15
31	55	50	27	35	53	45	52	44	46	50	54	55	15	15	53	42	54	52	54	15	33	45	42	13	35	17	20	55	45	52
32	46	44	50	51	52	52	55	51	52	44	41	42	54	10	27	54	50	55	50	11	51	47	45	42	40	52	52	38	50	47
33	47	20	52	53	38	41	36	14	43	52	33	51	12	13	34	46	46	53	55	55	36	54	54	33	39	20	41	39	47	50
34	48	14	51	46	55	46	40	55	47	55	42	14	50	44	43	13	45	44	33	47	15	52	55	50	44	39	42	50	12	55
35	38	11	39	39	12	36	43	54	53	42	37	38	45	41	45	35	42	42	45	13	52	36	44	55	47	38	40	46	33	45
36	35	54	42	33	20	47	41	43	50	41	47	50	33	42	51	47	43	39	44	48	42	53	39	38	43	36	39	53	46	46
37	54	47	38	43	11	53	44	36	44	36	43	33	38	22	52	43	44	41	47	35	38	44	33	43	53	48	37	37	15	43
38	17	45	45	45	46	42	47	37	33	43	39	54	52	39	13	53	47	47	37	52	55	40	15	15	33	22	43	35	51	35
39	44	53	44	38	44	40	39	38	40	45	15	52	43	37	33	38	41	46	53	33	44	37	41	47	42	41	38	48	55	54
40	43	52	37	41	36	38	45	42	45	47	52	47	47	40	46	39	52	38	13	53	45	38	37	52	54	40	46	17	11	44
41	53	46	46	47	47	43	42	52	37	37	38	46	46	23	50	45	36	13	42	44	47	39	36	44	50	46	55	19	54	16
42	20	15	43	44	43	37	46	27	36	46	44	22	55	24	15	33	37	40	38	46	37	42	38	45	46	43	44	16	40	48
43	19	55	47	42	37	44	37	40	39	39	46	44	44	25	55	48	38	43	14	38	46	43	46	46	52	37	45	33	38	33
44	18	38	41	22	45	39	38	41	38	38	45	53	53	26	47	44	40	45	52	45	43	41	47	22	45	42	47	18	52	53
45	22	22	22	23	22	22	22	22	22	22	22	45	22	46	22	22	22	22	22	22	22	22	22	23	22	47	22	22	22	22
46	23	23	23	24	23	23	23	23	23	23	23	23	23	28	23	23	23	23	23	23	23	23	23	24	23	23	23	23	23	23
47	24	24	24	25	24	24	24	24	24	24	24	24	24	29	24	24	24	24	24	24	24	24	24	25	24	24	24	24	24	24
48	25	25	25	26	25	25	25	25	25	25	25	25	25	30	25	25	25	25	25	25	25	25	25	26	25	25	25	25	25	25
49	26	26	26	27	26	26	26	26	26	26	26	26	26	31	26	26	26	26	26	26	26	26	26	27	26	26	26	26	26	26
50	28	28	28	28	28	28	28	28	28	28	28	28	28	32	28	28	28	28	28	28	28	28	28	28	28	28	28	28	28	28
51	29	29	29	29	29	29	29	29	29	29	29	29	29	49	29	29	29	29	29	29	29	29	29	29	29	29	29	29	29	29
52	30	30	30	30	30	30	30	30	30	30	30	30	30	38	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30
53	31	31	31	31	31	31	31	31	31	31	31	31	31	47	31	31	31	31	31	31	31	31	31	31	31	31	31	31	31	31
54	32	32	32	32	32	32	32	32	32	32	32	32	32	36	32	32	32	32	32	32	32	32	32	32	32	32	32	32	32	32
55	49	49	49	49	49	49	49	49	49	49	49	49	49	45	49	49	49	49	49	49	49	49	49	49	49	49	49	49	49	49

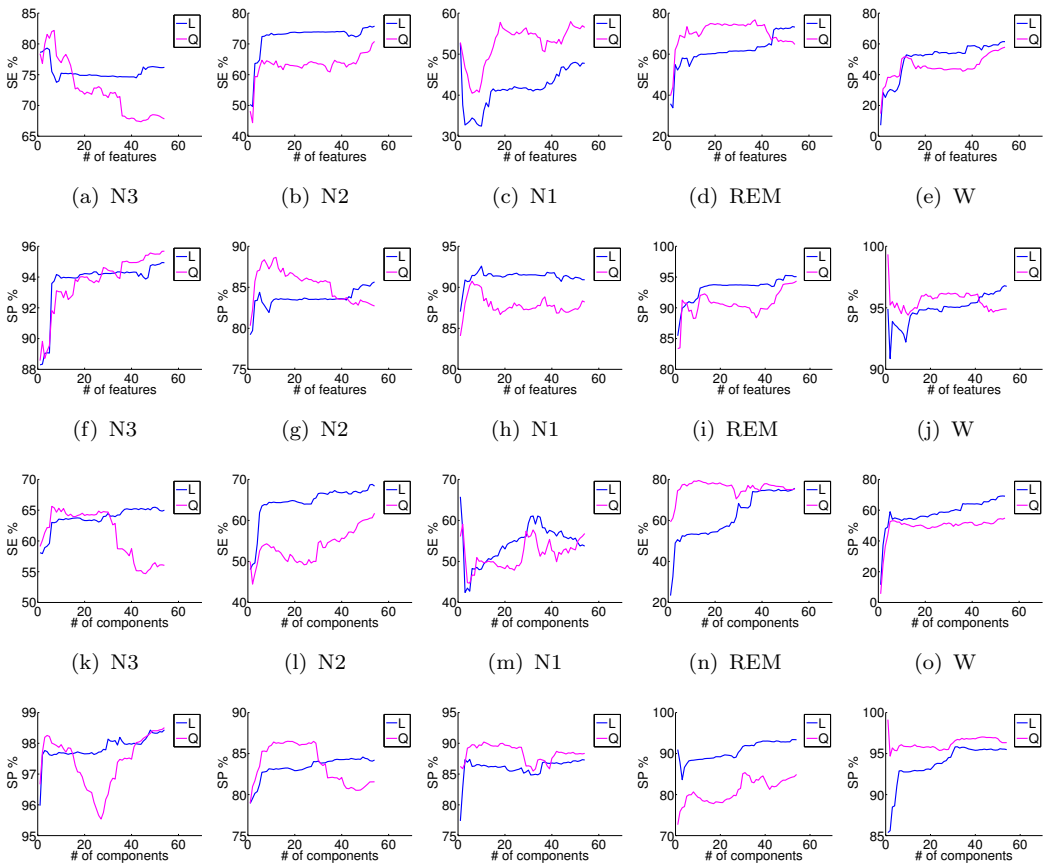
TABLE C.1: MRMR feature selection for each patient to predict sleep stages

	patient																													
rank	**2**	**3**	**5**	**6**	**8**	**9**	**10**	**11**	**12**	**14**	**15**	**16**	**17**	**18**	**20**	**21**	**22**	**24**	**26**	**28**	**30**	**31**	**32**	**34**	**35**	**36**	**37**	**38**	**39**	**40**
1	9	2	4	2	3	2	21	1	2	4	21	3	4	2	2	21	1	4	3	9	1	5	21	2	19	2	21	2	21	5
2	15	11	51	51	12	50	15	50	55	50	52	50	52	55	5	51	52	50	50	55	55	50	50	50	34	4	52	5	52	51
3	53	55	52	52	52	53	52	13	52	11	51	20	50	50	50	19	55	52	53	50	51	52	53	11	50	50	50	51	19	4
4	51	51	34	15	51	52	50	52	15	52	3	33	21	21	55	50	51	11	14	15	50	33	14	52	54	52	15	13	50	52
5	54	15	50	50	33	33	33	51	51	6	33	52	12	53	52	52	4	51	16	52	53	55	51	33	14	15	53	4	55	50
6	5	20	53	53	15	12	19	5	53	51	50	15	53	15	33	17	15	1	10	17	5	11	3	51	52	55	51	55	11	1
7	34	52	18	55	53	51	17	19	5	15	4	51	33	51	51	55	50	20	51	11	52	53	52	15	4	51	55	14	51	53
8	35	34	55	12	50	17	51	53	54	14	53	21	51	17	15	53	13	33	13	51	19	51	5	55	51	33	33	52	53	55
9	44	53	13	33	55	55	55	55	19	55	16	12	18	52	53	27	17	53	55	27	11	20	54	34	27	11	4	50	17	19
10	13	19	15	20	10	20	53	18	50	18	15	53	15	13	11	14	53	19	4	53	54	36	27	54	53	53	12	53	54	54
11	23	50	33	16	17	4	5	4	9	54	17	9	19	33	16	4	54	55	52	33	15	34	11	13	15	21	44	15	5	11
12	27	33	54	54	13	15	14	15	16	20	11	55	2	18	19	20	33	9	12	5	16	18	19	20	55	13	37	1	13	21
13	37	13	17	39	54	34	54	54	14	53	55	11	55	54	54	16	6	15	54	12	40	13	2	53	17	10	7	54	33	15
14	24	54	38	48	44	18	12	16	22	42	27	45	39	20	17	54	19	2	2	39	2	45	33	47	11	9	13	9	15	9
15	22	9	36	5	47	13	2	33	6	9	13	34	54	16	43	11	20	54	18	36	33	44	13	41	33	48	54	12	12	14
16	3	45	39	4	21	27	20	12	11	12	20	46	13	27	20	2	44	44	11	46	36	15	55	10	5	39	47	33	27	44
17	25	14	11	42	22	11	41	20	46	40	54	43	43	4	14	18	39	14	27	45	14	48	43	22	39	12	43	11	20	10
18	50	22	45	37	23	45	34	34	23	5	12	47	41	44	21	15	2	34	17	54	4	54	1	23	38	35	42	43	39	33
19	43	43	2	11	24	39	18	3	24	45	37	19	3	42	27	12	12	27	15	41	20	22	15	24	13	43	14	21	9	38
20	41	23	35	22	25	48	42	36	25	33	43	13	34	23	12	33	22	42	19	44	47	9	18	25	35	37	34	47	40	3
21	26	24	37	23	26	22	22	11	33	17	9	27	11	24	22	5	42	13	48	2	13	23	9	7	46	42	22	19	14	36
22	28	25	47	24	2	23	27	21	26	47	36	5	36	35	23	3	23	40	43	43	43	24	38	26	37	20	23	42	41	17
23	47	26	22	35	11	8	23	14	34	13	5	23	9	25	37	34	16	22	1	37	17	25	12	12	40	5	46	27	1	34
24	16	28	23	45	28	24	4	17	4	22	35	24	42	26	24	1	24	18	22	13	44	26	40	27	16	40	24	3	36	43
25	29	21	24	25	29	25	11	41	28	19	42	54	44	28	18	13	25	12	23	42	22	28	45	28	43	23	19	36	38	42
26	30	29	12	26	20	26	24	43	29	43	22	8	38	29	44	9	21	46	9	40	23	21	7	42	47	3	25	10	22	45
27	31	30	25	18	30	7	25	7	30	23	23	25	1	1	10	46	26	23	44	18	45	29	46	29	9	16	26	39	37	37
28	32	31	26	38	40	28	26	45	31	36	47	26	5	14	25	40	28	21	24	6	10	30	4	46	22	24	2	22	23	27
29	21	5	42	27	31	47	28	39	32	21	24	28	47	10	9	22	29	24	34	47	24	31	47	38	23	25	28	23	24	12
30	49	32	28	28	32	29	29	27	49	24	1	42	45	3	26	7	9	25	33	22	37	32	35	19	24	26	29	20	2	22
31	33	44	29	29	49	30	36	42	38	25	39	44	46	30	28	23	30	26	37	19	25	49	37	30	25	14	11	34	25	7
32	18	49	30	1	14	1	30	48	27	41	25	39	7	31	29	24	45	28	25	23	6	40	42	31	26	28	30	24	26	23
33	36	37	48	30	36	31	31	22	45	26	2	29	22	32	30	25	40	29	26	24	26	38	6	21	28	29	31	25	28	24
34	48	46	16	31	46	32	3	23	48	2	26	30	20	49	31	26	27	30	21	25	39	41	44	32	18	7	41	26	42	25
35	19	42	31	32	37	14	32	9	39	44	28	31	23	22	32	8	31	7	28	20	28	47	41	49	29	30	32	28	29	20
36	46	38	32	49	38	3	49	24	36	28	29	32	48	43	1	28	32	31	29	14	29	12	10	43	30	31	49	38	16	26
37	20	12	14	47	42	49	45	46	44	29	30	49	24	6	49	6	49	32	8	3	12	37	22	45	21	17	20	29	30	28
38	42	47	44	41	45	38	6	25	13	30	31	38	25	47	47	35	41	49	30	26	30	39	39	14	31	45	5	30	31	29
39	14	4	19	10	4	46	40	6	37	31	40	4	26	41	34	29	48	36	31	28	31	42	23	44	32	54	3	31	44	18
40	38	41	49	17	39	54	47	26	17	27	32	22	14	40	13	30	47	10	32	21	27	43	24	37	48	32	45	7	18	48
41	17	48	40	13	43	10	37	28	40	32	49	14	28	36	40	10	11	17	49	29	32	27	25	17	49	49	40	32	43	30
42	55	39	5	36	41	44	13	29	47	49	44	2	37	11	45	31	43	45	45	30	49	14	26	39	44	47	10	49	32	31
43	39	36	41	44	35	40	44	30	43	38	38	36	17	19	46	32	36	43	39	31	21	46	28	40	2	46	39	18	49	32
44	8	40	27	40	34	37	43	31	3	39	14	40	29	46	38	49	34	47	41	32	41	19	16	4	36	41	38	46	4	49
45	45	10	43	34	18	16	46	32	35	37	6	37	30	37	42	45	46	39	38	49	38	2	8	36	41	36	9	37	6	35
46	12	3	46	43	8	41	39	49	42	34	45	48	8	38	35	38	38	41	40	4	42	17	29	5	45	18	18	48	45	39
47	52	27	20	14	19	36	38	40	12	46	41	16	31	7	39	37	37	5	42	35	46	4	30	16	20	1	27	45	46	41
48	1	16	1	46	1	42	9	38	41	1	46	41	32	45	41	41	5	37	46	38	34	16	34	48	12	22	36	16	47	13
49	40	7	7	7	48	43	16	35	20	3	34	7	49	39	4	39	35	38	20	48	35	10	31	18	42	44	6	8	34	2
50	4	35	8	19	7	9	48	2	18	35	19	17	40	5	36	42	14	48	47	1	9	1	32	9	1	38	17	41	7	46
51	11	8	21	21	5	5	10	44	10	7	7	10	6	12	3	47	8	35	36	8	48	35	20	35	8	19	8	17	48	16
52	7	1	10	3	9	21	7	37	8	8	48	6	10	48	48	43	7	16	7	34	8	8	49	3	10	34	16	35	3	47
53	10	18	3	9	27	6	1	10	1	48	10	35	27	9	7	44	18	3	35	10	18	6	36	6	3	8	48	6	35	6
54	2	6	6	6	16	19	8	8	21	10	8	18	35	34	6	48	3	8	6	16	7	7	48	1	7	27	1	40	8	40
55	6	17	9	8	6	35	35	47	7	16	18	1	16	8	8	36	10	6	5	7	3	3	17	8	6	6	35	44	10	8

TABLE C.2: MRMR feature selection for each patient to predict A phases

Appendix D

Results



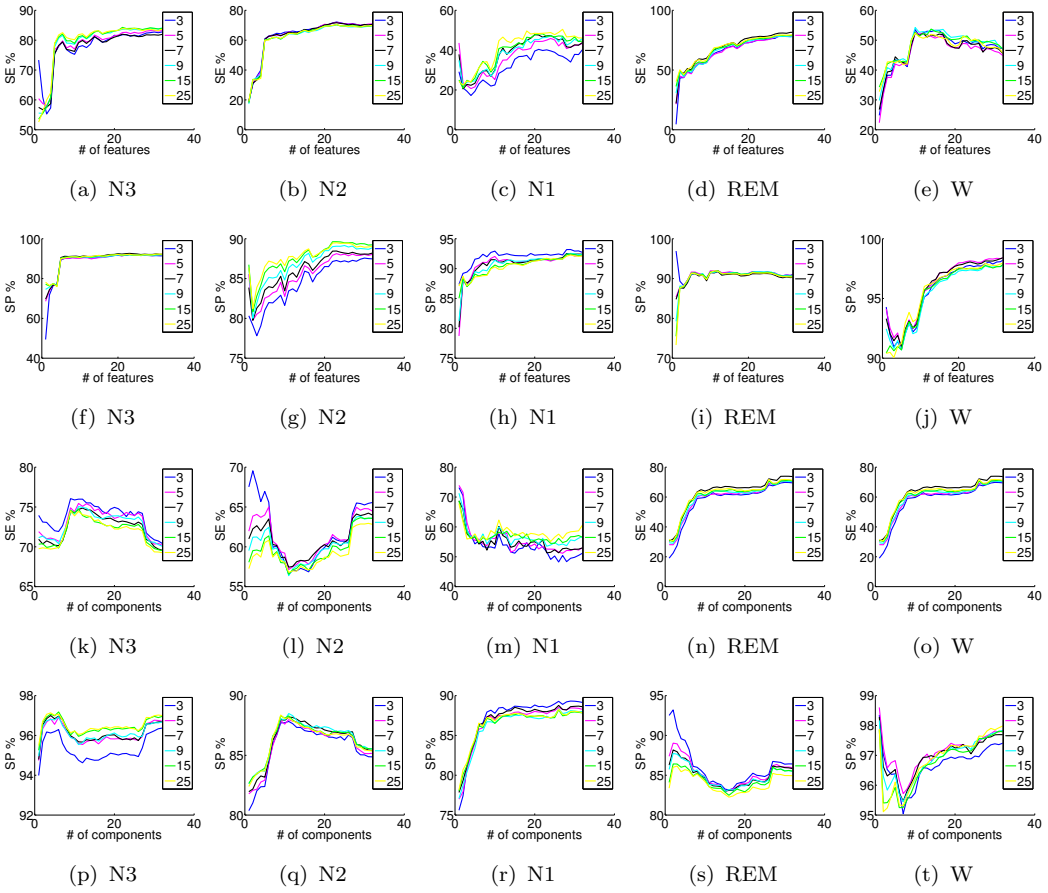
(a) N3 (b) N2 (c) N1 (d) REM (e) W

(f) N3 (g) N2 (h) N1 (i) REM (j) W

(k) N3 (l) N2 (m) N1 (n) REM (o) W

(p) N3 (q) N2 (r) N1 (s) REM (t) W

FIGURE D.1: Sensitivity (a-e and k-o) and specificity (f-j and p-t) evolution for the different sleep stages with the dimensionality increase using a-j) MRMR and k-t) PCA as feature selection for different LDA classification method (L: Linear; Q: Quadratic)



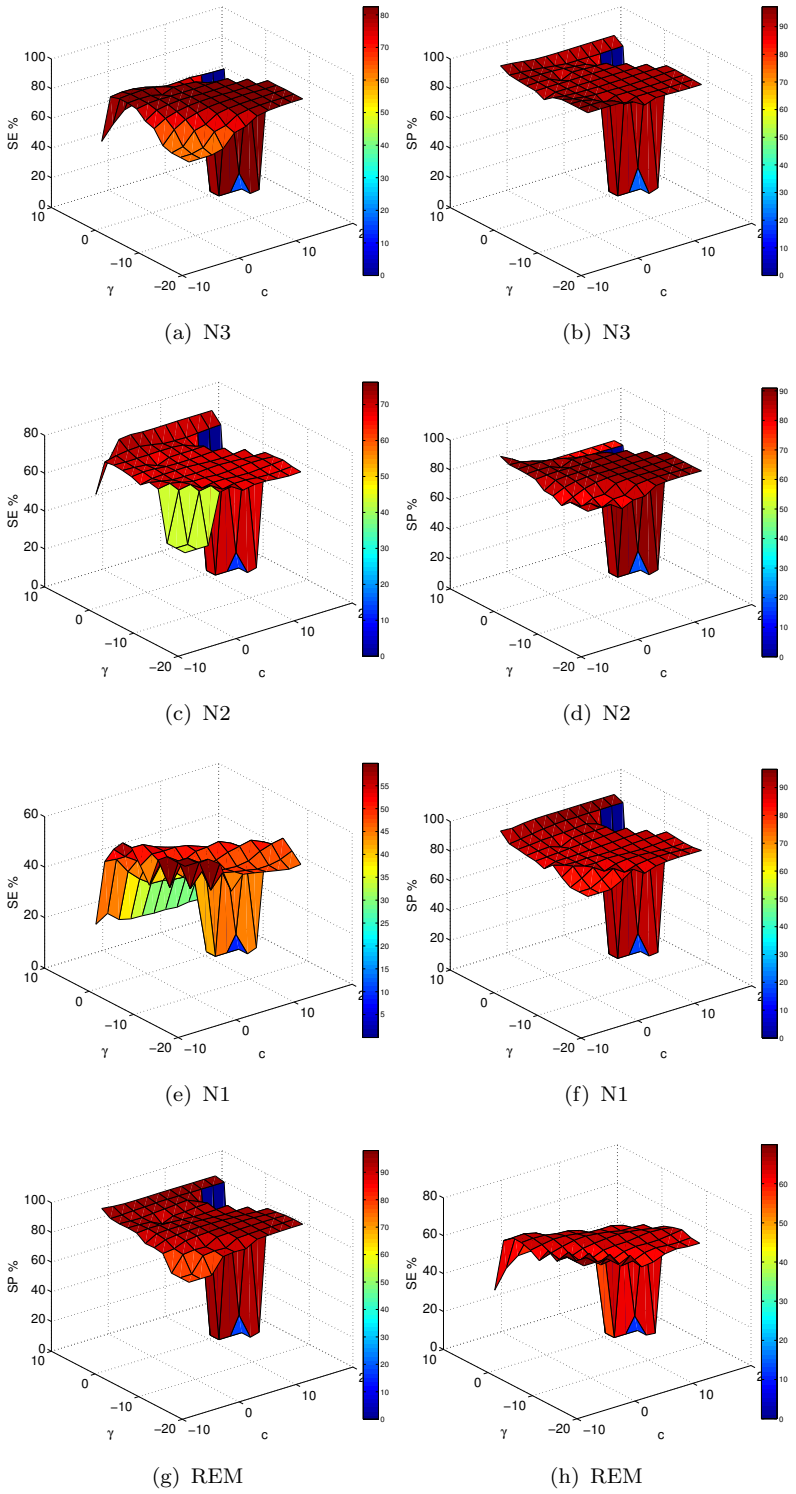
(a) N3 (b) N2 (c) N1 (d) REM (e) W

(f) N3 (g) N2 (h) N1 (i) REM (j) W

(k) N3 (l) N2 (m) N1 (n) REM (o) W

(p) N3 (q) N2 (r) N1 (s) REM (t) W

FIGURE D.2: Sensitivity (a-e and k-o) and specificity (f-j and p-t) evolution for the different sleep stages with the dimensionality increase using a-j) MRMR and k-t) PCA as feature selection and k-NN classification algorithm for different k's values (3, 5, 7, 9, 15 and 25)



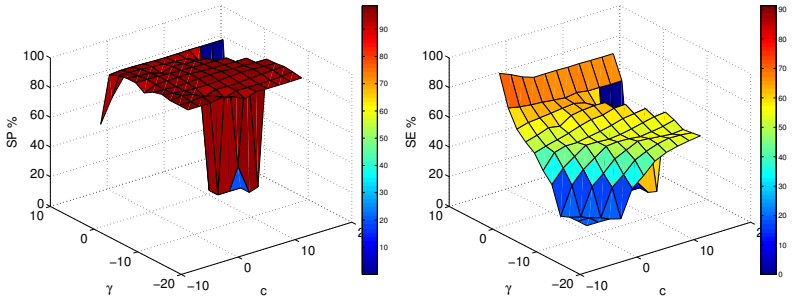
(a) N3 (b) N3

(c) N2 (d) N2

(e) N1 (f) N1

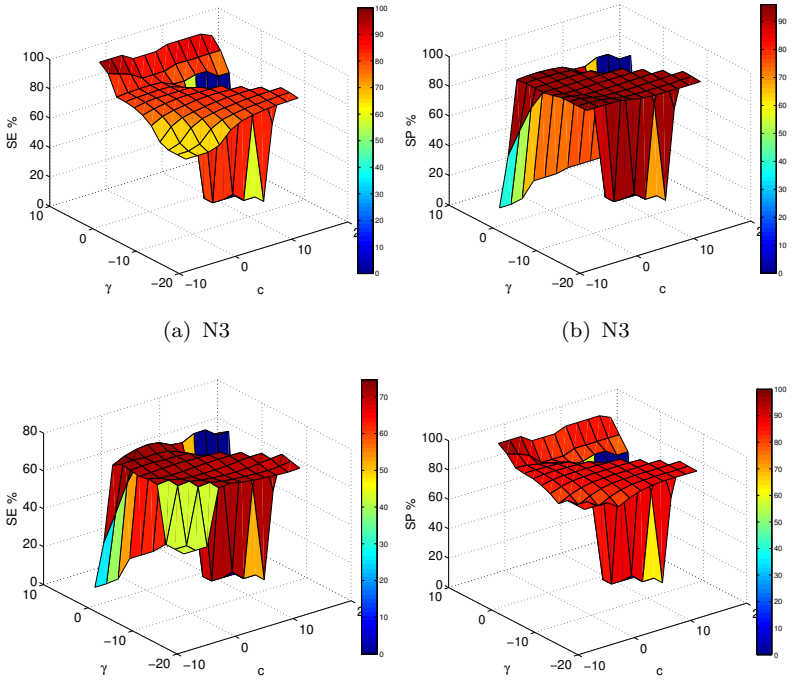
(g) REM (h) REM

FIGURE D.3: Sensitivity (a, c, e and g) and specificity (b, d, f and h) using a SVM classification algorithm, for a grid search where $c = 2^{-5}, 2^{-3}, ..., 2^{15}$ and $\gamma = 2^{-15}, 2^{-13}, ..., 2^{5}$, using 15 best features for a-b) N3, c-d) N2, e-f) N1 and g-h) REM sleep stages



(a) W (b) W

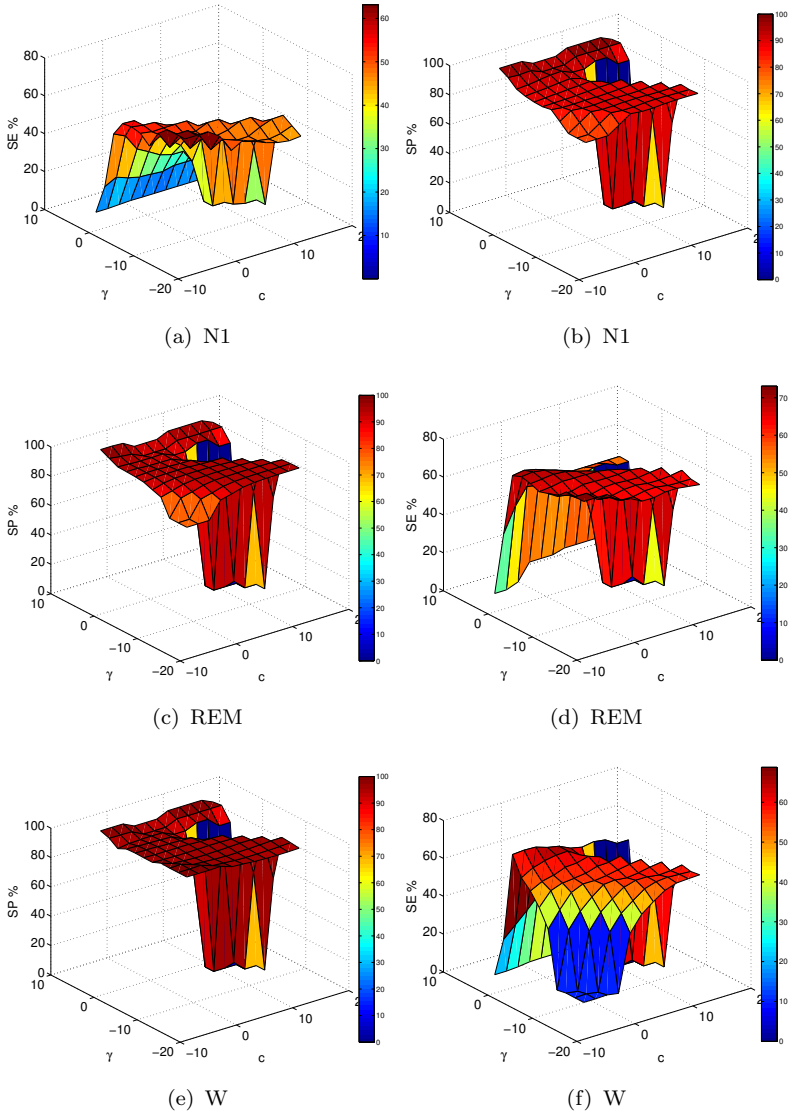
FIGURE D.4: Sensitivity (a) and specificity (b) using a SVM classification algorithm, for a grid search where $c = 2^{-5}, 2^{-3}, ..., 2^{15}$ and $\gamma = 2^{-15}, 2^{-13}, ..., 2^{5}$, using 15 best features for W sleep stage



(a) N3 (b) N3

(c) N2 (d) N2

FIGURE D.5: Sensitivity (a, c and e) and specificity (b, d and f) using a SVM classification algorithm for a grid search where $c = 2^{-5}, 2^{-3}, ..., 2^{15}$ and $\gamma = 2^{-15}, 2^{-13}, ..., 2^{5}$ using 30 best features for a-b) N3 and c-d) N2 sleep stages

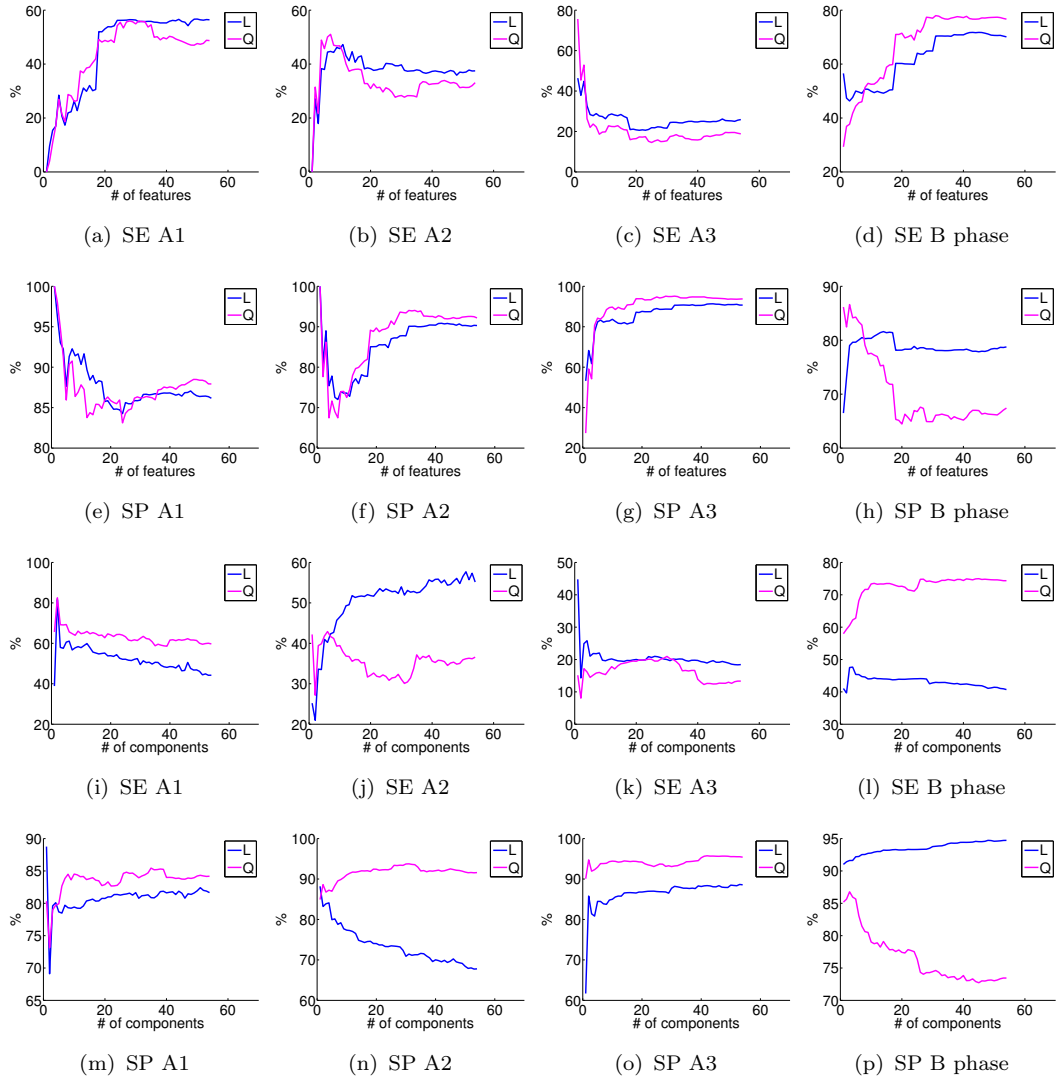


(a) N1 (b) N1

(c) REM (d) REM

(e) W (f) W

FIGURE D.6: Sensitivity (a, c and e) and specificity (b, d and g) using a SVM classification algorithm for a grid search where $c = 2^{-5}, 2^{-3}, ..., 2^{15}$ and $\gamma = 2^{-15}, 2^{-13}, ..., 2^{5}$ using 30 best features for a-b) N1, c-d) REM and e-f) W sleep stages



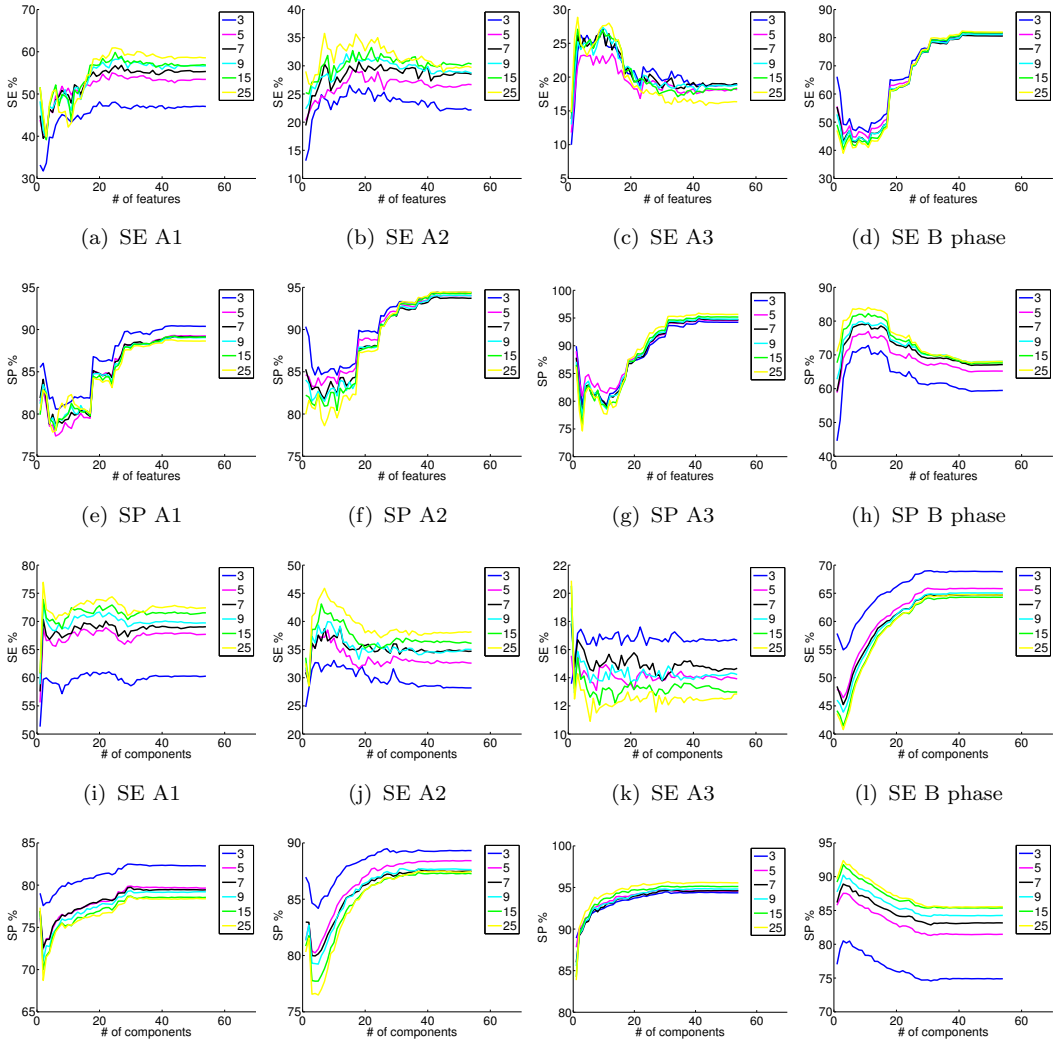
(a) SE A1 (b) SE A2 (c) SE A3 (d) SE B phase

(e) SP A1 (f) SP A2 (g) SP A3 (h) SP B phase

(i) SE A1 (j) SE A2 (k) SE A3 (l) SE B phase

(m) SP A1 (n) SP A2 (o) SP A3 (p) SP B phase

FIGURE D.7: Sensitivity (a-e and k-o) and specificity (f-j and p-t) evolution for A- and B-phases with the dimensionality increase using a-j) MRMR and k-t) PCA as feature selection for different LDA classification method (L: Linear; Q: Quadratic)



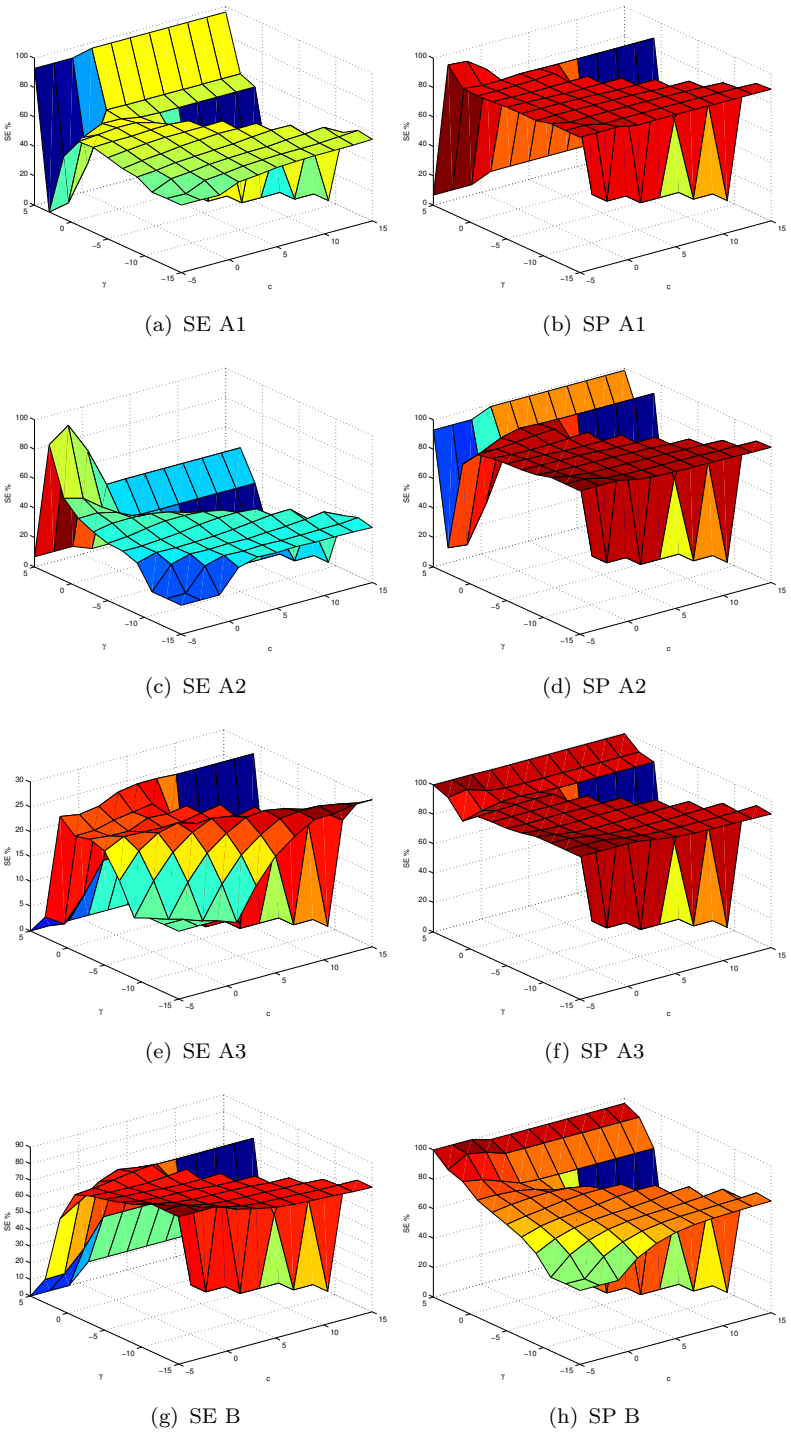
(a) SE A1 (b) SE A2 (c) SE A3 (d) SE B phase

(e) SP A1 (f) SP A2 (g) SP A3 (h) SP B phase

(i) SE A1 (j) SE A2 (k) SE A3 (l) SE B phase

(m) SP A1 (n) SP A2 (o) SP A3 (p) SP B phase

FIGURE D.8: Sensitivity (a-e and k-o) and specificity (f-j and p-t) evolution for the A- and B-phases with the dimensionality increase using a-j) MRMR and k-t) PCA as feature selection and k-NN classification algorithm for different k's values (3, 5, 7, 9, 15 and 25)



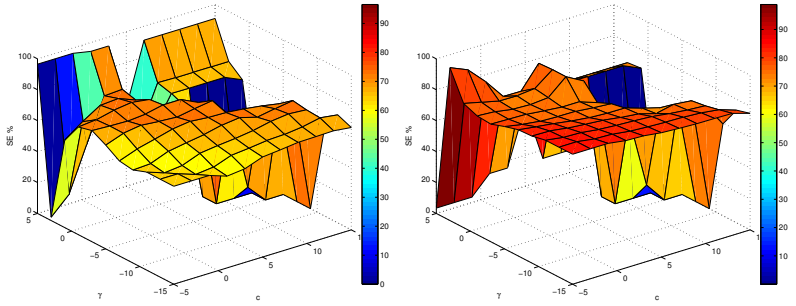
(a) SE A1 (b) SP A1

(c) SE A2 (d) SP A2

(e) SE A3 (f) SP A3

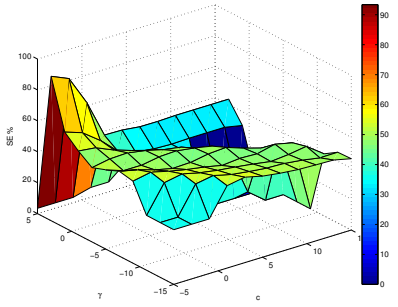
(g) SE B (h) SP B

FIGURE D.9: Sensitivity (a, c, e and g) and specificity (b, d, f and h) using a SVM classification algorithm for a grid search where $c = 2^{-5}, 2^{-3}, ..., 2^{15}$ and $\gamma = 2^{-15}, 2^{-13}, ..., 2^{5}$ using 40 best features for A- and B-phases

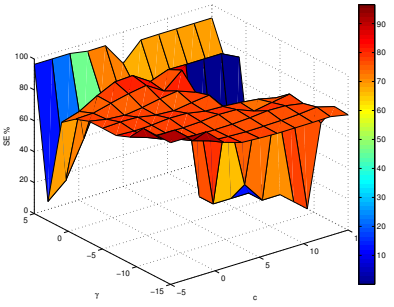


(a) SE A1

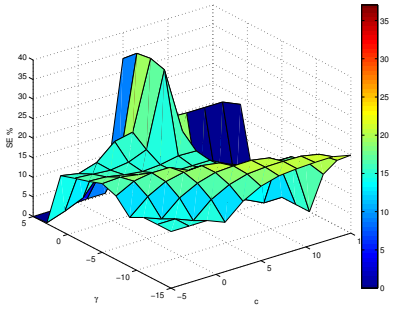
(b) SP A1



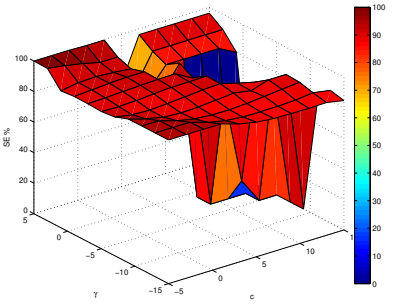
(c) SE A2



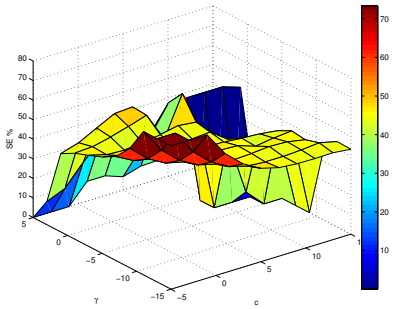
(d) SP A2



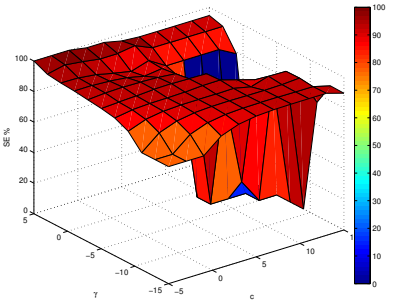
(e) SE A3



(f) SP A3



(g) SE B



(h) SP B

FIGURE D.10: Sensitivity (a, c, e and g) and specificity (b, d, f and h) using a SVM classification algorithm for a grid search where $c = 2^{-5}, 2^{-3}, ..., 2^{15}$ and $\gamma = 2^{-15}, 2^{-13}, ..., 2^{5}$ using 30 principal components for A- and B-phases

Bibliography

[1] P. Achermann, *The Two-Process Model of Sleep Regulation Revisited*, Aviation, Space, and Environmental Medicine **75** no. 3, (2004-03-01T00:00:00) A37–A43.

[2] F. R, B. O, M. S, S. A, S. K, and T. M., *Inter-rater reliability of sleep cyclic alternating pattern (CAP) scoring and validation of a new computer-assisted CAP scoring method*, Clinical Neurophysiology **116** no. 3, (2005) 696 – 707.

[3] I. Chouvarda, M. Mendez, V. Rosso, A. Bianchi, L. Parrino, A. Grassi, M. Terzano, N. Maglaveras, and S. Cerutti, *Predicting EEG complexity from sleep macro and microstructure*, Physiological Measurement **32** no. 8, (2011) 1083.

[4] L. Parrino, P. Halasz, C. T. Alberto, and M. Terzano, *CAP, epilepsy and motor events during sleep: the unifying role of arousal*, Sleep Medicine Reviews **10** no. 4, (2006) 267 – 285.

[5] U. Acharya, S. Sree, G. Swapna, R. Martis, and J. Suri, *Automated EEG analysis of epilepsy: A review*, Knowledge-Based Systems **45** no. 0, (2013) 147 – 165.

[6] P. Ryvlin, S. Rheims, and G. Risse, *Nocturnal Frontal Lobe Epilepsy*, Epilepsia **47** (2006) 83–86.

[7] A. Loomis, E. Harvey, and G. Hobart, *Potential rhythms of the cerebral cortex during sleep.*, Science (1935).

[8] A. Loomis, E. Harvey, and G. Hobart, *Further observations on the potential rhythms of the cerebral cortex during sleep.*, Science (1935).

[9] A. Loomis, E. Harvey, and G. Hobart, *Cerebral states during sleep, as studied by human brain potentials.*, Journal of experimental psychology **21** no. 2, (1937) 127.

[10] H. Schulz, *Rethinking sleep analysis*, Journal of Clinical Sleep Medicine **4** no. 2, (2008) 99–103.

[11] E. Aserinsky and N. Kleitman, *Regularly Occurring Periods of Eye Motility, and Concomitant Phenomena, During Sleep*, Science **118** no. 3062, (1953) 273–274.

[12] A. Rechtschaffen and A. Kales, *A manual of standardized terminology, techniques and scoring for sleep stages of human subjects.*, Los Angeles, CA: University of California; 1968 (1968) 1–57.

[13] M. Bonnet, D. Carley, M. Carskadon, P. Easton, C. Guilleminault, R. Harper, B. Hayes, M. Hirshkowitz, P. Ktonas, S. Keenan, M. Pressman, T. Roehrs, J. Smith, J. Walsh, S. Weber, and P. Westbrook, *EEG arousals: scoring rules and examples: a preliminary report from the Sleep Disorders Atlas Task Force of the American Sleep Disorders Association.*, Tech. Rep. 2, American Sleep Disorders Association, 1992.

[14] C. Iber, S. Ancoli-Israel, J. Chesson, A., and S. Quan, *The AASM Manual for the Scoring of Sleep and Associated Events: Rules Terminology and Technical Specifications 1st ed.*, tech. rep., American Academy of Sleep Medicine, 2007.

[15] J. Malmivuo and R. Plonsey, *Bioelectromagnetism: principles and applications of bioelectric and biomagnetic fields*. New York, Oxford University Press, 1995.

[16] P. Nunez, R. Srinivasan, A. Westdorp, R. Wijesinghe, D. Tucker, R. Silberstein, and P. Cadusch, *EEG coherency: I: statistics, reference electrode, volume conduction, Laplacians, cortical imaging, and interpretation at multiple scales*, Electroencephalography and Clinical Neurophysiology **103** no. 5, (1997) 499 – 515.

[17] W. Tatum, *Ellen R. Grass Lecture: Extraordinary EEG*, The Neurodiagnostic Journal **54** no. 1, (2014) 3–21.

[18] 8, 2015. `http://2012books.lardbucket.org/books/beginning-psychology/section_09/0f909be2c2ff6ecccb6d269f8a6c2d26.jpg`.

[19] *A fictional EEG showing a sleep spindle and K-complex in stage 2 sleep*, 2008. `https://en.wikipedia.org/wiki/Sleep_spindle#/media/File:Stage2sleep.svg`. [Online; accessed on 23-September-2015].

[20] F. De Carli, L. Nobili, P. Gelcich, and F. Ferrillo, *A method for the automatic detection of arousals during sleep.*, Sleep **22** no. 5, (1999) 561–572.

[21] M. Boselli, L. Parrino, A. Smerieri, and M. Terzano, *Effect of age on EEG arousals in normal sleep.*, Sleep **21** no. 4, (1998) 351–357.

[22] L. De Gennaro, M. Ferrara, and M. Bertini, *EEG arousals in normal sleep: variations induced by total and selective slow-wave sleep deprivation.*, Sleep **24** no. 6, (2001) 673–679.

[23] J. Sassin and L. Johnson, *Body motility during sleep and its relation to the K-complex*, Experimental Neurology **22** no. 1, (1968) 133 – 144.

[24] M. Terzano and L. Parrino, *Origin and Significance of the Cyclic Alternating Pattern (CAP)*, Sleep Medicine Reviews **4** no. 1, (2000) 101 – 123.

[25] P. Halász, *Hierarchy of micro-arousals and the microstructure of sleep*, Neurophysiologie Clinique/Clinical Neurophysiology **28** no. 6, (1998) 461 – 475.

[26] M. Terzano, P. Gatti, G. Manzoni, E. Formentini, and D. Mancia, *Is the EEG Cyclic Alternating Pattern a True Autonomous Entity*, European Neurology **21** no. 5, (1982) 324–334.

[27] M. Terzano, D. Mancia, M. Salati, G. Costani, A. Decembrino, and L. Parrino, *The cyclic alternating pattern as a physiologic component of normal NREM sleep.*, Sleep **8** no. 2, (1985) 137–145.

[28] M. Terzano, L. Hollberg, L. Parrino, A. Sherieri, and R. Chervin, *Atlas, rules, and recording techniques for the scoring of cyclic alternating pattern (CAP) in human sleep*, Sleep Medicine **2** (2001) 537–553.

[29] L. Parrino, R. Ferri, O. Bruni, and M. Terzano, *Cyclic alternating pattern (CAP): The marker of sleep instability*, Sleep Medicine Reviews **16** (2012) 27–45.

[30] M. Terzano, L. Parrino, M. Boselli, A. Smerieri, and M. Spaggiari, *CAP components and EEG synchronization in the first 3 sleep cycles*, Clinical Neurophysiology **111** (2000) 283–290.

[31] R. Fisher, C. Acevedo, A. Arzimanoglou, A. Bogacz, J. Cross, C. E. Elger, J. Engel, L. Forsgren, J. French, M. Glynn, D. Hesdorffer, I. Lee, G. Mathern, S. Moshé, E. Perucca, I. Scheffer, T. Tomson, M. Watanabe, and S. Wiebe, *ILAE Official Report: A practical clinical definition of epilepsy*, Epilepsia **55** no. 4, (2014) 475–482, `http://dx.doi.org/10.1111/epi.12550`.

[32] L. Parrino, A. Smerieri, M. Spaggiari, and M. Terzano, *Cyclic alternating pattern (CAP) and epilepsy during sleep: how a physiological rhythm modulates a pathological event*, Clinical Neurophysiology **111, Supplement 2** no. 0, (2000) S39 – S46, Sleep and Epilepsy Supplement.

[33] A. Mirzoev, E. Bercovici, L. Stewart, M. Cortez, O. Snead, and v. . Desrocher, M.“, *Circadian profiles of focal epileptic seizures: A need for reappraisal*,.

[34] L. Doroshenkov, V. Konyshev, and S. Selishchev, *Classification of human sleep stages based on EEG processing using hidden Markov models*, Biomedical Engineering **41** no. 1, (2007) 25–28.

[35] S. Eddy, *Hidden markov models*, Current opinion in structural biology **6** no. 3, (1996) 361–365.

[36] B. Yegnanarayana, *Artificial neural networks*. PHI Learning Pvt. Ltd., 2009.

[37] N. Schaltenbrand, R. Lengelle, M. Toussaint, R. Luthringer, G. Carelli, A. Jacqmin, E. Lainey, A. Muzet, and J. Macher, *Sleep stage scoring using the neural network model: comparison between visual and automatic analysis in normal subjects and patients.*, Sleep **19** no. 1, (1996) 26–35.

[38] E. Oropesa, H. Cycon, and M. Jobert, *Sleep Stage Classification using Wavelet Transform and Neural Network*, 1999.

[39] P. Addison, *The illustrated wavelet transform handbook: introductory theory and applications in science, engineering, medicine and finance*. CRC press, 2002.

[40] L. Fraiwan, K. Lweesy, N. Khasawneh, M. Fraiwan, H. Wenz, and H. Dickhaus, *Classification of Sleep Stages Using Multi-wavelet Time Frequency Entropy and LDA*, Methods of Information in Medicine **49** no. 3, (2010) 230–237.

[41] A. Izenman, *Linear discriminant analysis*, pp. , 237–280. Springer, 2008.

[42] V. Helland, A. Gapelyuk, A. Suhrbier, M. Riedl, T. Penzel, J. Kurths, and N. Wessel, *Investigation of an automatic sleep stage classification by means of multiscorer hypnogram*, Methods of Information in Medicine **49** no. 5, (2010) 467–472.

[43] G *Efficient sleep stage recognition system based on EEG signal using k-means clustering based feature weighting*, Expert Systems with Applications **37** no. 12, (2010) 7922 – 7928.

[44] A. Likas, N. Vlassis, and J. Verbeek, *The global k-means clustering algorithm*, Pattern recognition **36** no. 2, (2003) 451–461.

[45] N. Altman, *An introduction to kernel and nearest-neighbor nonparametric regression*, The American Statistician **46** no. 3, (1992) 175–185.

[46] S. Safavian and D. Landgrebe, *A survey of decision tree classifier methodology*, IEEE transactions on systems, man, and cybernetics **21** no. 3, (1991) 660–674.

[47] H. Phan, Q. Do, T. Do, and D. Vu, *Metric learning for automatic sleep stage classification*, pp. , 5025–5028. July, 2013.

[48] S. Liang, C. Kuo, Y. Hu, and Y. Cheng, *A rule-based automatic sleep staging method*, Journal of Neuroscience Methods **205** no. 1, (2012) 169 – 176.

[49] H. Jo, J. Park, C. Lee, S. An, and S. Yoo, *Genetic fuzzy classifier for sleep stage identification*, Computers in Biology and Medicine **40** no. 7, (2010) 629 – 634.

[50] T. J. Ross, *Fuzzy logic with engineering applications*. John Wiley & Sons, 2009.

[51] D. Estevez, J. Pastoriza, and V. Bonillo, *A continuous evaluation of the awake sleep state using fuzzy reasoning*, pp. , 5539–5542. Sept, 2009.

[52] C. Cortes and V. Vapnik, *Support-vector networks*, Machine learning **20** no. 3, (1995) 273–297.

[53] T. Lajnef, S. Chaibi, P. Ruby, P. Aguera, J. Eichenlaub, M. Samet, A. Kachouri, and K. Jerbi, *Learning machines and sleeping brains: Automatic sleep stage classification using decision-tree multi-class support vector machines*, Journal of Neuroscience Methods **250** (2015) 94 – 105, Cutting-edge EEG Methods.

[54] B. Koley and D. Dey, *An Ensemble System for Automatic Sleep Stage Classification Using Single Channel EEG Signal*, Comput. Biol. Med. **42** no. 12, (2012) 1186–1195.

[55] J. Sotelo, A. Forero, A. Rodríguez, D. Frau, E. Roldán, and D. Peluffo, *Automatic Sleep Stages Classification Using EEG Entropy Features and Unsupervised Pattern Analysis Techniques*, Entropy **16** no. 12, (2014) 6573.

[56] N. Belacel, P. Hansen, and N. Mladenovic, *Fuzzy J-means: a new heuristic for fuzzy clustering*, Pattern Recognition **35** no. 10, (2002) 2193–2200.

[57] U. Barcaro, C. Navona, S. Belloli, E. Bonanni, C. Gneri, and L. Murri, *A simple method for the quantitative description of sleep microstructure*, Electroencephalography and Clinical Neurophysiology **106** (1998) 429–432.

[58] C. Navona, U. Barcaro, E. Bonanni, F. Martino, M. Maestri, and L. Murri, *An automatic method for the recognition and classification of the A-phases of the cyclic alternating pattern*, Clinical Neurophysiology **113** no. 11, (2002) 1826 – 1831.

[59] B. U., E. Bonanni, M. Maestri, L. Murri, L. Parrino, and M. Terzano, *A general automatic method for the analysis of NREM sleep microstructure*, Sleep Medicine **5** no. 6, (2004) 567 – 576.

[60] A. Rosa, L. Parrino, and M. Terzano, *Automatic detection of cyclic alternating pattern (CAP) sequences in sleep: preliminary results*, Clinical Neurophysiology **110** no. 4, (1999) 585 – 592.

[61] A. Rosa, B. Kemp, T. Paiva, F. Lopes, and H. Kamphuisen, *A model-based detector of vertex waves and K complexes in sleep electroencephalogram*, Electroencephalography and Clinical Neurophysiology **78** no. 1, (1991) 71 – 79.

[62] N. Martins and A. Rosa, *Selective Linear Prediction for Rhythmic Activity Modelling*, Power1–8.

[63] R. Largo, C. Munteanu, and A. Rosa, *CAP event detection by wavelets and GA tuning*, pp. , 44–48. Sept, 2005.

[64] C. Stam and B. Dijk, *Synchronization likelihood: an unbiased measure of generalized synchronization in multivariate data sets*, Physica D: Nonlinear Phenomena **163** no. 3–4, (2002) 236 – 251.

[65] R. Ferri, F. Rundo, O. Bruni, M. Terzano, and C. Stam, *Dynamics of the EEG slow-wave synchronization during sleep*, Clinical Neurophysiology **116** no. 12, (2005) 2783 – 2795.

[66] R. Ferri, F. Rundo, O. Bruni, M. Terzano, and C. Stam, *Regional scalp EEG slow-wave synchronization during sleep cyclic alternating pattern A1 subtypes*, Neuroscience Letters **404** no. 3, (2006) 352 – 357.

[67] S. Mariani, E. Manfredini, V. Rosso, A. Grassi, M. Mendez, A. Alba, M. Matteucci, L. Parrino, M. Terzano, S. Cerutti, and A. Bianchi, *Efficient automatic classifiers for the detection of A phases of the cyclic alternating pattern in sleep*, Medical & Biological Engineering & Computing **50** no. 4, (2012) 359–372.

[68] Y. Freund and R. E. Schapire, *A desicion-theoretic generalization of on-line learning and an application to boosting*, pp. , 23–37, Springer. 1995.

[69] A. Goldberger, L. Amaral, L. Glass, J. Hausdorff, P. Ivanov, R. Mark, J. Mietus, G. Moody, C. Peng, and H. Stanley, *PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals*, Circulation **101** no. 23, (2000 (June 13)) e215–e220.

[70] S. Butterworth, *On the theory of filter amplifiers*, Wireless Engineer **7** no. 6, (1930) 536–541.

[71] M. Kaleem, L. Sugavaneswaran, A. Guergachi, and S. Krishnan, *Application of Empirical Mode Decomposition and Teager energy operator to EEG signals for mental task classification*, pp. , 4590–4593. Aug, 2010.

[72] E. Kvedalen, *Signal processing using the Teager Energy Operator and other nonlinear operators*, tech. rep., University of Oslo - Department of Informatics, May, 2003.

[73] P. Maragos, T. Quatieri, and J. Kaiser, *Speech nonlinearities, modulations, and energy operators*, pp. , 421–424 vol.1. Apr, 1991.

[74] M. Bahoura and J. Rouat, *Wavelet speech enhancement based on the Teager energy operator*, Signal Processing Letters, IEEE **8** no. 1, (2001) 10–12.

[75] F. Jabloun, A. Cetin, and E. Erzin, *Teager energy based feature parameters for speech recognition in car noise*, Signal Processing Letters, IEEE **6** no. 10, (1999) 259–261.

[76] H. Patil and S. Viswanath, *Effectiveness of Teager energy operator for epoch detection from speech signals*, International Journal of Speech Technology **14** no. 4, (2011) 321–337.

[77] R. Lauer and L. Prosser, *Use of the Teager-Kaiser Energy Operator for Muscle Activity Detection in Children*, Annals of biomedical engineering **37** no. 8, (2009) 1584–1593.

[78] F. Duman, A. Erdamar, O. Erogul, Z. Telatar, and S. Yetkin, *Efficient sleep spindle detection algorithm with decision tree*, Expert Systems with Applications **36** no. 6, (2009) 9980 – 9985.

[79] A. Erdamar, F. Duman, and S. Yetkin, *A wavelet and teager energy operator based method for automatic detection of K-Complex in sleep EEG*, Expert Systems with Applications **39** no. 1, (2012) 1284 – 1290.

[80] H. Cai, *Fast frequency measurement algorithm based on zero crossing method*, in *ICCET 2010 - 2010 International Conference on Computer Engineering and Technology, Proceedings*, pp. , 606–608. 2010.

[81] B. Milenko and R. Željko, *Frequency measurement of distorted signals using Fourier and zero crossing techniques*, Electric Power Systems Research **78** no. 8, (2008) 1407 – 1415.

[82] M. Carrozzi, A. Accardo, and F. Bouquet, *Analysis of sleep-stage characteristics in full-term newborns by means of spectral and fractal parameters.*, Sleep **27** no. 7, (2004) 1384–1393.

[83] M. Drinnan, A. Murray, J. White, A. Smithson, C. Griffiths, and G. Gibson, *Automated recognition of EEG changes accompanying arousal in respiratory sleep disorders*, Sleep **19** no. 4, (1996) 296–303.

[84] V. Lipovac, *Zero-crossing-based linear prediction for speech recognition*, Electronics Letters **25** no. 2, (1989) 90–92.

[85] Y. Aye, *Speech Recognition Using Zero-Crossing Features*, pp. , 689–692. Feb, 2009.

[86] A. Saxena and A. Singh, *A Microprocessor based Speech Recognizer for Isolated Hindi Digits*, Design**0** .

[87] K. Elleithy Advanced Techniques in Computing Sciences and Software Engineering.

[88] A. Lempel and J. Ziv, *On the Complexity of Finite Sequences*, Information Theory, IEEE Transactions on **22** no. 1, (1976) 75–81.

[89] M. Aboy, R. Hornero, D. Abasolo, and D. Alvarez, *Interpretation of the Lempel-Ziv Complexity Measure in the Context of Biomedical Signal Analysis*, Biomedical Engineering, IEEE Transactions on **53** no. 11, (2006) 2282–2288.

[90] S. M. Pincus and A. L. Goldberger, *Physiological time-series analysis: what does regularity quantify?*, American Journal of Physiology - Heart and Circulatory Physiology **266** no. 4, (1994) H1643–H1656.

[91] D. Abásolo, S. Simons, R. Morgado da Silva, G. Tononi, and V. Vyazovskiy, *Lempel-Ziv complexity of cortical activity during sleep and waking in rats*, Journal of Neurophysiology **113** no. 7, (2015) 2742–2752.

[92] A. Casali, O. Gosseries, M. Rosanova, M. Boly, S. Sarasso, K. Casali, S. Casarotto, M. Bruno, S. Laureys, G. Tononi, and M. Massimini, *A Theoretically Based Index of Consciousness Independent of Sensory Processing and Behavior*, Science Translational Medicine **5** no. 198, (2013) 198ra105–198ra105.

[93] R. Bracewell, *The Fourier Transform and Its Applications*. Electrical engineering series. McGraw Hill, 2000.

[94] J. Allen, *Short-term spectral analysis, and modification by discrete Fourier transform*, IEEE Transactions on Acoustics Speech and Signal Processing **25** no. 3, (1977) 235–238.

[95] E. Derya Übeyli, *Analysis of EEG signals by combining eigenvector methods and multiclass support vector machines*, Computers in Biology and Medicine **38** no. 1, (2015) 14–22.

[96] K. Sivasankari and K. Thanushkodi, *An Improved EEG Signal Classification Using Neural Network with the Consequence of ICA and STFT*, Journal of Electrical Engineering and Technology **3** no. 3, (2014).

[97] N. E. Huang, S. Long, and Z. Shen, *The mechanism for frequency downshift in nonlinear wave evolution*, Advances in applied mechanics **32** (1996) 59–117C.

[98] T. Rutkowski, D. Mandic, A. Cichocki, and A. Przybyszewski, *EMD approach to multichannel EEG data—the amplitude and phase components clustering analysis*, Journal of Circuits, Systems, and Computers **19** no. 01, (2010) 215–229.

[99] P. Diez, V. Mut, E. Laciar, A. Torres, and E. Avila, *Application of the empirical mode decomposition to the extraction of features from EEG signals for mental task classification*, pp. , 2579–2582. Sept, 2009.

[100] F. Ebrahimi, S. Setarehdan, J. Ayala-Moyeda, and H. Nazeran, *Automatic sleep staging using empirical mode decomposition, discrete wavelet transform, time-domain, and nonlinear dynamics features of heart rate variability signals*, Computer Methods and Programs in Biomedicine **112** no. 1, (2013) 47 – 57.

[101] N. Kannathal, M. Choo, U. Acharya, and P. Sadasivan, *Entropies for detection of epilepsy in EEG*, Computer Methods and Programs in Biomedicine **80** no. 3, (2005) 187 – 194.

[102] I. Chouvarda, M. Mendez, V. Rosso, A. Bianchi, L. Parrino, A. Grassi, M. Terzano, N. Maglaveras, and S. Cerutti, *CAP sleep in insomnia: New methodological aspects for sleep microstructure analysis*, pp. , 1495–1498. Aug, 2011.

[103] I. Chouvarda, M. Mendez, A. Alba, A. Bianchi, A. Grassi, E. Arce-Santana, V. Rosso, M. Terzano, and L. Parrino, *Nonlinear analysis of the change points between A and B phases during the Cyclic Alternating Pattern under normal sleep*, pp. , 1049–1052. Aug, 2012.

[104] T. Higuchi, *Approach to an irregular time series on the basis of the fractal theory*, Physica D: Nonlinear Phenomena **31** no. 2, (1988) 277–283.

[105] A. Savitzky and M. J. Golay, *Smoothing and differentiation of data by simplified least squares procedures.*, Analytical chemistry **36** no. 8, (1964) 1627–1639.

[106] D. Ruan, G. Chen, and E. Kerre, *Intelligent data mining: techniques and applications*, vol. 5. Springer Science & Business Media, 2005.

[107] H. Peng, F. Long, and C. Ding, *Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy*, Pattern Analysis and Machine Intelligence, IEEE Transactions on **27** no. 8, (2005) 1226–1238.

[108] C. Ding and H. Peng, *Minimum redundancy feature selection from microarray gene expression data*, Journal of bioinformatics and computational biology **3** no. 02, (2005) 185–205.

[109] J. Shlens, *A Tutorial on Principal Component Analysis*, CoRR **abs/1404.1100** (2014).

[110] A. Ajanki, *Example of k-nearest neighbour classificationnb*, 2007. `https://commons.wikimedia.org/wiki/File:KnnClassification.svg`. [Online; accessed on 3-September-2015].

[111] Cyc, *Graphic showing the maximum separating hyperplane and the margin*, 2007. `https://commons.wikimedia.org/wiki/File:Svm_max_sep_hyperplane_with_margin.png`. [Online; accessed on 3-September-2015].

[112] B. Scholkopf, S. Mika, C. Burges, P. Knirsch, K. Muller, G. Ratsch, and A. Smola, *Input space versus feature space in kernel-based methods*, Neural Networks, IEEE Transactions on **10** no. 5, (1999) 1000–1017.

[113] K. Duan and S. Keerthi, *Which is the best multiclass SVM method? An empirical study*, pp. , 278–285. Springer, 2005.

[114] V. Franc and V. Hlavác, *Statistical pattern recognition toolbox for Matlab*, Prague, Czech: Center for Machine Perception, Czech Technical University (2004).

[115] R. Picard and R. Cook, *Cross-Validation of Regression Models*, Journal of the American Statistical Association **79** no. 387, (1984) pp. 575–583.

[116] A. Rosa and L. Allen, *Fuzzy classification of microstructural dynamics of human sleep*, in *Systems, Man, and Cybernetics, 1996., IEEE International Conference on*, pp. , 1108–1113 vol.2. Oct, 1996.

[117] J. Cohen et al., *A coefficient of agreement for nominal scales*, Educational and psychological measurement **20** no. 1, (1960) 37–46.

[118] N. J. Salkind, *Encyclopedia of measurement and statistics*. Sage Publications, 2006.

[119] J. Hodges, J.L., *The significance probability of the smirnov two-sample test*, Arkiv för Matematik **3** no. 5, (1958) 469–486, `http://dx.doi.org/10.1007/BF02589501`.

[120] A. Kolmogoroff, *Confidence Limits for an Unknown Distribution Function*, Ann. Math. Statist. **12** no. 4, (1941) 461–463.

[121] R. C. Littell, W. W. Stroup, and R. J. Freund, *SAS for linear models*. SAS Institute, 2002.

[122] P. Lewicki and T. Hill, *Statistics: methods and applications*, Tulsa, OK. Statsoft (2006).

[123] L. Moore, *The basic practice of statistics*, Technometrics **38** no. 4, (1996) 404–405.

[124] M. Terzano, M. Monge-Straws, F. Mikol, M. Spaggiari, and L. Parrino, *Cyclic alternating pattern as a provocative factor in nocturnal paroxysmal dystonia*, Epilepsia **38** no. 9, (1997) 1015–1025.