Differences in sleep patterns among healthy sleepers and patients after stroke

Bui Thi Mai Phuong

Bratislava 2014

Comenius University Bratislava
Faculty of Mathematics, Physics and Informatics
Department of Applied Mathematics and Statistics

Bachelor thesis

Bui Thi Mai Phuong

Differences in sleep patterns among healthy sleepers and patients after stroke

Supervisor: Ing. Mgr. Roman Rosipal, PhD.

Study programme: Economic and Financial Mathematics

Branch of study: 1114 Applied Mathematics

Bratislava 2014

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky
Katedra aplikovanej matematiky a štatistiky

Bakalárska práca

Bui Thi Mai Phuong

Rozdiely v štruktúre spánku u pacientov po mozgovej príhode a zdravých subjektov

Vedúci práce: Ing. Mgr. Roman Rosipal, PhD.

Študijný program: Ekonomická a finančná matematika

Študijný odbor: 1114 Aplikovaná matematika

Bratislava 2014

Comenius University in Bratislava
Faculty of Mathematics, Physics and Informatics

THESIS ASSIGNMENT

Name and Surname:	Phuong Bui Thi Mai
Study programme:	Economic and Financial Mathematics (Single degree study, bachelor I. deg., full time form)
Field of Study:	9.1.9. Applied Mathematics
Type of Thesis:	Bachelor´s thesis
Language of Thesis:	English
Secondary language:	Slovak

Title: Differences in sleep patterns among healthy sleepers and patients after stroke

Aim: With the aim of improving our understanding of a challenging question how a typical, good quality, structure of sleep should look like, a novel probabilistic sleep model has been developed. The model was extensively tested and validated on a wide cohort of healthy sleepers. The aim of this bachelor study is to apply the model to patients with specific cerebral lesions and to carry out a thorough statistical comparison of the sleep patterns and extracted sleep biomarkers with healthy sleepers.
By self-study and experimental work a student in his bachelor work should meet the following objectives i) to understand basic principles of electrophysiological recordings and signal processing ii) to understand principles and being able to apply the developed probabilistic sleep model to new set of clinical data of patients after stroke iii) to carry out a comprehensive statistical comparison of extracted sleep patterns and selected sleep biomarkers with the aim to identify pathological aspects of the sleep process in subjects after stroke.

Supervisor:	Ing. Mgr. Roman Rosipal, PhD.
Department:	FMFI.KAMŠ - Department of Applied Mathematics and Statistics
Head of department:	prof. RNDr. Daniel Ševčovič, CSc.
Assigned:	14.11.2013
Approved:	14.11.2013

doc. RNDr. Margaréta Halická, CSc.
Guarantor of Study Programme

.....
Student

.....
Supervisor

Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Phuong Bui Thi Mai
Študijný program: ekonomická a finančná matematika (Jednoodborové štúdium, bakalársky I. st., denná forma)
Študijný odbor: 9.1.9. aplikovaná matematika
Typ záverečnej práce: bakalárska
Jazyk záverečnej práce: anglický
Sekundárny jazyk: slovenský

Názov: Rozdiely v štruktúre spánku u pacientov po mozgovej príhode a zdravých subjektov / *Differences in sleep patterns among healthy sleepers and patients after stroke*

Cieľ: S cieľom lepšie porozumieť otázke typickej štruktúry kvalitného spánku bol navrhnutý nový pravdepodobnostný model spánku.Tento model bol otestovaný narozsiahlej vzorke zdravých subjektov. Cieľom bakalárskej práce je aplikovať model na súbor dát pacientov so špecifickými mozgovými léziami a vykonať dôkladné štatistické porovnanie získaných spánkových biomarkerov so zdravými subjektmi.
Vo svojej bakalárskej práci by mal študent samoštúdiom a experimentálnou prácou dosiahnuť nasledujúce ciele i)porozumieť základným princípom elektrofyziologických záznamov a spracovania signal ii) porozumieť princípom navrhnutého pravdepodobnostného modelu spánku a byť schopný aplikovať model na nový súbor klinických dát pacientov pomozgovej príhode iii) vykonať dôkladné štatistické porovnanie vybraných spánkových biomarkerov s cieľom identifikovať patologické aspekty spánkového procesu u pacientov po mozgovej príhode.

Vedúci: Ing. Mgr. Roman Rosipal, PhD.
Katedra: FMFI.KAMŠ - Katedra aplikovanej matematiky a štatistiky
Vedúci katedry: prof. RNDr. Daniel Ševčovič, CSc.

Dátum zadania: 14.11.2013

Dátum schválenia: 14.11.2013 doc. RNDr. Margaréta Halická, CSc.
garant študijného programu

.....
študent

.....
vedúci práce

Acknowledgement

Many thanks to Roman Rosipal for practical advice, a highly responsible attitude and active support far beyond his role as a supervisor. Our collaboration has been a most instructive and fun experience.

Poďakovanie

Veľká vďaka Romanovi Rosipalovi za praktické rady, zvlášť zodpovedný prístup a aktívnu podporu presahujúcu jeho školiteľské povinnosti. Naša spolupráca bola nanajvýš obohacujúcou skúsenosťou a navyše to bol kopec srandy.

Abstrakt

Bui Thi Mai Phuong: Rozdiely v štruktúre spánku u pacientov po mozgovej príhode a zdravých subjektov [Bakalárska práca], Univerzita Komenského v Bratislave, Fakulta matematiky, fyziky a informatiky, Katedra aplikovanej matematiky a štatistiky; školiteľ: Ing. Mgr. Roman Rosipal, PhD., Bratislava, 2014.

Spánková deprivácia, spôsobená či už ochorením alebo životným štýlom, akútna či chronická, predstavuje značné riziko pre denný kognitívny výkon a môže vyústiť v excesívnu somnolenciu, narušenú pozornosť alebo zníženú úroveň motorických schopností. Ischemická cievna mozgová príhoda je dobre známe akútne ochorenie, ktoré zanecháva postihnutých pacientov náchylných voči poruchám spánku, ktoré často vedú k vyššie spomenutému zhoršeniu pozornosti a kognície. V tejto práci sme analyzovali a porovnávali štruktúru spánku zdravých subjektov a pacientov po mozgovej príhode.

Za účelom obísť dobre známe nedostatky štandardizovanej klasifikácie spánku do niekoľkých diskrétnych stupňov sme použili nedávno navrhnutý pravepodobnostný model spánku, ktorý reprezentuje spánkový proces ako súbor pravdepodobnostných kriviek. Túto novú reprezentáciu spánku sme analyzovali dvomi spôsobmi: i) z kriviek boli vypočítané fyziologicky plauzibilné spánkové parametre. Následne boli aplikované metódy modifikovaného bootstrapu a opakovanej analýzy rozptylu za účelom identifikovať spánkové parametre rozlišujúce zdravých subjektov od subjektov po mozgovej príhode, ii) na pravdepodobnostné krivky sme sa pozerali ako na istú formu funkcionálnych dát. Mikroštruktúru a časovú dynamiku spánku sme preskúmali za použitia funkcionálnej analýzy hlavných komponentov a klasterizácie.

Identifikovali sme viacero spánkových parametrov s významne odlišnými hodnotami pre zdravých subjektov a subjektov po mozgovej príhode. Výsledky získané metódou bootstrap a analýzou rozptylu boli konzistentné. Napriek tomu, že naša štúdia ohľadne klasterizácie funkcionálnych dát predstavuje len úvodné pokusy o separáciu daných dvoch skupín subjektov, boli sme schopní identifikovať viacero fyziologicky izolovaných spánkových štruktúr a tiež sme identifikovali štruktúru spánkových mikrostavov ako potenciálny faktor pre diskrimináciu medzi zdravými subjektmi a subjektmi po mozgovej príhode.

Kľúčové slová

štruktúra spánku, mozgová príhoda, pravdepodobnostný model spánku, podmienený bootstrap, funkcionálna analýza hlavných komponentov

Abstract

Bui Thi Mai Phuong: Differences in sleep patterns among healthy sleepers and patients after stroke [Bachelor thesis], Comenius University Bratislava, Faculty of Mathematics, Physics and Informatics, Department of Applied Mathematics and Statistics; supervisor: Ing. Mgr. Roman Rosipal, PhD., Bratislava, 2014.

Sleep deprivation, whether from disorder or lifestyle, whether acute or chronic, poses a significant risk to day-time cognitive performance and may result in excessive somnolence, impaired attention or decreased level of motor abilities. Ischemic stroke resulting in cerebral lesions is a well-known acute disorder that leaves affected patients strongly vulnerable to sleep disturbances that often lead to the above-mentioned cognitive and attentional impairments. In this thesis, we analysed and compared sleep patterns of healthy sleepers and patients after stroke.

To overcome the well-known limits of the standardised classification of sleep into several discrete stages, we employed the recently proposed probabilistic sleep model that represents the sleep process as a continuum in terms of a set of probability curves. We analysed this novel sleep representation in two ways: i) physiologically plausible sleep parameters were computed from the curves. Then a modified bootstrap method and repeated ANOVA were applied to identify sleep parameters differentiating healthy and stroke subjects, ii) probability curves were considered to represent a form of functional data. The microstructure along with time dynamics of sleep were explored using functional principal components analysis and clustering.

We identified several sleep parameters with significantly different values for healthy subjects and stroke patients. The two statistical approaches, the bootstrap and ANOVA testing, produced consistent results. Although our functional data clustering study represents a preliminary attempt to separate the two groups of subjects, we were able to identify several physiologically separate sleep patterns and we also identified sleep microstate patterns as a potential factor of discrimination between healthy subjects and stroke patients.

Keywords

sleep patterns, stroke, probability sleep model, subset-conditional bootstrap, functional principal component analysis

Contents

List of abbreviations

ANOVA	Analysis of Variance
CI	Confidence Interval
EDF	Empirical Distribution Function
EEG	Electroencephalography, Electroencephalogram
EM	Expectation-Maximisation
fPCA	Functional Principal Component Analysis
GMM	Gaussian Mixture Model
MLE	Maximum Likelihood Estimate
NREM	Non-REM
PSG	Polysomnography, Polysomnogram
PSM	Probabilistic Sleep Model
REM	Rapid Eye Movement
RK	Rechtschaffen and Kales
SWS	Slow Wave Sleep

Introduction

Sleep is a time-dynamic process coming in diverse forms. Polysomnography (PSG) is the standard diagnostic procedure for capturing the structure of sleep. It consists of a battery of diagnostic tests monitoring the subject's cerebral and cardial activity, eye movement, muscle tension, breath rate, and other physiological and neurophysiological variables, throughout the whole night. The key component of a PSG is the recording of brain electrical activity on the scalp, the electroencephalography (EEG).

A hypnogram, a form of PSG, represents the graph of clinically well-defined sleep stages as a function of time. It is usually obtained by visual scoring of EEG, electrooculogram (EOG) and electromyogram (EMG) recordings. Cycles of rapid eye movement sleep (REM) and non-rem (NREM) stages are major components of the sleep process. According to the scoring protocol of the American Academy of Sleep Medicine (AASM) [4], NREM sleep can be further classified into stage 1,2, and 3, which includes previously considered stage 4 of the Rechtschaffen and Kales scoring protocol [11] into stage 3. This stage is called slow wave sleep (SWS) and represents the deepest sleep.

Strokes are caused by an interruption of blood flow to the brain. If the condition persists for more than a few seconds, it results in the loss of consciousness and leads to brain ischemia, consequently resulting in irreversible brain disorders and brain tissue damage. Sleep disorders may be direct functional consequences of a stroke. For instance, obstructive sleep apnea is the major sleep disorder associated with an ischemic stroke [6]. On the other hand, sleep disruption, partial or full sleep deprivation is ubiquitous following hemispheric strokes. Characterising the sleep structure of brain damaged patients is a challenging task, since the condition of patients often does not allow any measurement to take place. Targeted experiments with a proper monitoring and evaluation of sleep are scarce, studies in this area are therefore limited. We highlight but a recent study by Jiang et al. [5] who found polysomnographic abnormalities in patients with a specific cognitive impairment. The results hint at the potential of this line of research.

This bachelor thesis captures a part of the project "Effects of sleep disturbances on day-time neurocognitive performance in patients with stroke", overseen by the Slovak Academy of Sciences. The project aims to answer the following three questions: a) Are there differences in sleep patterns between healthy sleepers and stroke patients? b) How does sleep affect the subject's day-time neurocognitive performance? c) Is the sleep-performance relationship different in healthy subjects and patients with stroke? Here, we focus on and give a discussion of the first problem. In particular, we carry out a statistical analysis and evaluation of sleep EEG recordings, which may be viewed as a set of extensive time series, each being a first-instance representation of one subject's sleep.

The thesis is organised as follows. In Section 1, we provide technical details about the dataset in question and about the preprocessing steps taken. Next, we note that the raw EEG time series is in itself cumbersome to work with and on that account, we seek a more workable transcription of sleep. This leads us to the probabilistic sleep model (PSM) proposed by Lewandowski et al. [8], to which we give an overview in section 2. The PSM produces a new sleep representation in the form of posterior probability curves representing sleep microstructure defined by a finite set of sleep microstates. In the subsequent sections, two ways of analysing sleep posterior curves are presented. First, we reduce the curves to scalar indicators and face a one-dimensional multi-sample problem. Second, we treat the posteriors as a form of functional data and tackle the problem using tools of functional data analysis [10]. The methodological background, as well as results of applying these procedures to our dataset are given in Sections 3 and 4.

1 Data specification and preprocessing

In total, 99 EEG recordings were analysed, 34 of which correspond to stroke subjects. These had been collected in accordance with the objectives of the project, not later than 7, and in most cases within 5 days after the occurrence of stroke. A sample of recordings of healthy sleepers taken from the SIESTA database [7] was used as the control group. The database consists of full PSG recordings of two consecutive nights for 175 healthy adults of varying ages. However, as one would expect, stroke patients are mostly elderly people. As age is one of the major factors of sleep architecture, only healthy subjects over 60 years of age were included in our study to avoid age related bias in the analysis. The first-night recording was taken, since adaptation to the clinical environment may influence subsequent measurements. All subjects in the control group had a Pittsburgh Sleep Quality Index [2] of at most 5.

A brief quantitative characterisation of the dataset is provided in Table 1. The RK hypnograms used for sleep marker computation had been obtained by manual scoring for stroke subjects and via an automatic RK classifier, Somnolyzer24×7 [1], in the case of healthy subjects.

Though it would be interesting to look at other PSG components as well, we left it as a future research question and used EEG data only. Primarily, the C3 channel was considered. Noise-infected segments were replaced by corresponding C4 segments, if artifact-free, otherwise the problematic time intervals were ignored. All data were downsampled to 100Hz.

Table 1: Summary of selected RK sleep parameters of healthy subjects and stroke patients considered in the study.

	Healthy	Stroke
Men : Women	31 : 34	20 : 14
Age	73.0 ± 7.5	68.1 ± 11.5
Time in bed, tib [h]	7.97 ± 0.46	7.09 ± 1.50
Total sleep time, tst [h]	5.97 ± 1.07	6.57 ± 1.55
Sleep latency [min]	22.0 ± 19.9	9.35 ± 7.19
Sleep latency to REM [h]	2.42 ± 1.38	2.60 ± 1.74
Number of awakenings	25.3 ± 12.4	7.58 ± 7.21
Wake within sleep [min]	89.6 ± 46.4	19.0 ± 20.3
Sleep efficiency = tst / tib [%]	74.9 ± 12.4	92.6 ± 7.7
Proportion of S1 in sleep [%]	12.7 ± 7.8	34.5 ± 21.0
Proportion of S2 in sleep [%]	58.5 ± 10.9	28.8 ± 13.1
Proportion of SWS in sleep [%]	12.5 ± 10.4	26.3 ± 14.4
Proportion of REM in sleep [%]	16.3 ± 7.5	10.5 ± 7.8
Number of REM cycles	2.98 ± 1.14	5.04 ± 3.07

mean ± standard deviation of selected parameters

2 Probabilistic sleep model

The PSM is a novel sleep model proposed and validated by Lewandowski et al. [8]. For the complete description of PSM, we refer the reader to [8], and an interesting application of the PSM is addressed in [12]. Below we provide some background and motivation behind the model.

Based on the specific PSG patterns produced, various forms of sleep may be identified. The traditional Rechtschaffen and Kales (RK) classification [11] distinguishes between phases involving REM and phases characterised by the absence of this sort of movement, NREM. Phases of NREM sleep are further classified according to depth as either S1, S2, S3 or S4, in order from lighter to deeper. S3 and S4 are considered as SWS. The classical representation of sleep, the hypnogram, is then derived by dividing a PSG recording into 30s epochs and assigning each epoch to one of the five or six classes mentioned above. The scoring is performed by an attending specialist according to the RK rules.

The RK model has been criticised for dependence on inherently subjective judgements involved in sleep stage classification. Also, it is at best a crude description of sleep, both in terms of time resolution and the simplistic classification into a small number of fixed stages. The PSM aims to address these issues by introducing two main changes. First, the time window length is set to 3s, constituting a compromise between a more fine-grained partition and the ability to capture low-frequency EEG waves with reasonable reliability. Second, and more importantly, it explains sleep in terms of a larger number of sleep states, which we denote microstates, and transitions between them. Instead of a class label, the PSM assigns each time window a vector of non-negative values summing to one. The components of the vector correspond to the microstates, and may be interpreted as weights or probabilities that the subject was in the respective microstate at the time. Thus, the PSM is able to capture a much wider range of sleep forms.

Plotting probabilities that apply to a given microstate against time produces a posterior curve. A complete PSM transcription of sleep comprises a set of posterior

curves, one for each microstate. Alternatively, microstate posteriors can be combined to produce RK posteriors, which may be conceived of as a continuous counterpart of the indicator function implied by an RK hypnogram.

2.1 Construction

The conversion of a polysomnogram to the PSM format is not as straightforward as the conversion to a hypnogram. One way of obtaining the probabilities would be to implement a suitable modification of the RK manual as a computer algorithm. However, this would be quite demanding in terms of artificial intelligence and machine learning methods. Also, the manual only defines the few basic stages and there is no obvious answer how to determine the probabilities.

Instead of the RK classification rules, the PSM uses PSG recordings that have already been RK scored, to learn the typical characteristics of the basic sleep stages. Additionally, it is supplied with the information on specific patterns visible in the EEG, sleep spindles, which are predominantly associated with the S2 stage. We used an automatic detector that recognises four spindle presence levels: 0,1,2 and 3, in order of increasing probability of spindle presence within a given time interval. Spindle labels output by the detector are fed into the PSM with the aim of improving the accuracy of sleep microstates identification.

Before the PSM can be used for generating posteriors from PSGs, it must be trained on a sample of RK and spindle labelled recordings. It follows that the behaviour of a PSM, and consequently all posteriors derived from it, depend on the training dataset. In particular, the training dataset and the posteriors generated by the PSM are not statistically independent. This issue will be discussed in later sections. Below we give an overview of the training procedure. However, it is important to note that the PSM can be trained without the RK and/or spindle labels. This allows the construction of microstate sleep structure driven solely by EEG data. The RK-free version of the PSM is discussed in Section 2.3.

In: A training dataset consisting of m EEG recordings with RK and spindle labels.

:: Partition each EEG time series into disjoint 3s epochs. Obtain

$$\left\{(x, rk, sp)^{(i)} | i = 1, \ldots, n\right\}$$

where n is the total number of epochs in all m recordings. At this stage, the time ordering of epochs is ignored. Provided that the EEG is sampled at 100Hz, x is a vector of 300 data points corresponding to the i-th epoch. $rk \in \{W, S1, S2, SWS, REM\}$ is the respective RK class and $sp \in \{0, 1, 2, 3\}$ marks the spindle level. Note that due to the specific time resolutions of the RK and PSM models, we consider consecutive epochs within a 30s long RK window as having the same RK label.

:: Fit an autoregressive model of order 10 to each epoch. Symbolically,

$$x \xrightarrow{AR(10)} y$$

where y is a 10-dimensional vector of AR coefficients that characterises the frequency spectrum of a given EEG segment. It is taken to be the new representation of the signal.

:: Fit a Gaussian mixture model (GMM) to the observations.

Here, a parametric model is introduced to explain the distribution of the data

$$\left\{(y, rk, sp)^{(i)} | i = 1, \ldots, n\right\}$$

Each observation is assumed to have been generated by one out of k microstates. Each microstate occurs with a fixed probability π_j and is assumed to generate AR vectors from a Gaussian distribution. Independently from the AR vector, an RK class and a spindle state are determined following discrete distributions ρ_j, ψ_j that vary from microstate to microstate. Formally, the assumed joint density is

given by

$$f(y, rk, sp) = \sum_{j=1}^{k} \pi_j \mathcal{N}\left(y|\mu^{(j)}, \Sigma^{(j)}\right) \rho_j(rk)\psi_j(sp)$$

where k we may set and $\pi_j, \mu^{(j)}, \Sigma^{(j)}, \rho_j, \psi_j$ are unknown parameters whose values are estimated from data during the training step using the maximum likelihood (ML) approach. In the absence of a closed-form expression for the MLE of parameters, estimates are derived using the expectation-maximisation (EM) algorithm, as is common practice in a GMM context.

A similar GMM fitting procedure for semi-supervised learning, but with only one class variable, was considered by Miller and Uyar [9]. The one-variable version was termed separator model by Seeger [13].

Out: The fitted Gaussian mixture model.

Note that the optimal values of parameters have a natural interpretation. Variables $\mu^{(j)}, \Sigma^{(j)}$ represent centers and covariance matrices of the identified clusters, π_j indicates how common a microstate is, ρ_j specifies a level of similarity between a microstate and the traditional RK stages, and ψ_j sets out the level of spindle activity typical for the microstate.

2.2 PSM in action

A successfully trained PSM can be used for converting an EEG into a set of posterior probabilities. Below we outline the mechanisms behind the process.

In: An EEG signal with spindle presence levels.

:: Partition the signal into 3s time frames and fit an AR(10) model to each. Thus derive

$$\left\{(y, sp)^{(i)}\right\}_{i=1}^{n}$$

where y is the AR(10) vector, sp is the spindle class corresponding to the i-th epoch and n is the number of epochs for the given recording.

:: for $i = 1 : n$

Evaluate the responsibility of each microstate for the i-th data point. The responsibility $\gamma_j^{(i)}$ represents the probability that microstate j generated $(y, sp)^{(i)}$, conditonal on the value of $(y, sp)^{(i)}$:

$$\gamma_j^{(i)} := \frac{\pi_j \mathcal{N}\left(y^{(i)} | \mu^{(j)}, \Sigma^{(j)}\right) \psi_j\left(sp^{(i)}\right)}{\sum_{j=1}^{k} \pi_j \mathcal{N}\left(y^{(i)} | \mu^{(j)}, \Sigma^{(j)}\right) \psi_j\left(sp^{(i)}\right)}$$

end

:: The posterior curve for microstate j is given by

$$\left\{\gamma_j^{(i)}\right\}_{i=1}^{n}$$

In order to acquire a sleep representation in terms of RK stages, microstate posteriors are combined using the fitted similarity values, ρ_j:

$$\gamma_{rk}^{(i)} = \sum_{j=1}^{k} \gamma_j^{(i)} \rho_j(rk)$$

Out: Microstate posterior curves and RK posterior curves.

If the recording input into the model has not previously been RK-scored, it may in some contexts be useful to recover a discrete, pseudo-RK hypnogram. We can obtain this by assigning each epoch the RK state that has the highest probability.

2.3 RK-free PSM

As noted earlier, we can also consider a variant of the PSM that does not require the training dataset to be RK- or spindle-scored. Such a PSM can be constructed by making straightforward changes to the algorithm set out in section 2.1. The assumed

density of AR coefficients takes the form

$$f(y) = \sum_{j=1}^{k} \pi_j \mathcal{N}\left(y|\mu^{(j)}, \Sigma^{(j)}\right)$$

and is equivalent to the density of the basic GMM. When this PSM is used for data conversion, RK labels and/or spindle labels are ignored. The microstate posteriors are calculated as

$$\gamma_j^{(i)} := \frac{\pi_j \mathcal{N}\left(y^{(i)}|\mu^{(j)}, \Sigma^{(j)}\right)}{\sum_{j=1}^{k} \pi_j \mathcal{N}\left(y^{(i)}|\mu^{(j)}, \Sigma^{(j)}\right)}$$

The model is completely RK-free. It is not given any information on the RK stages and it does not output RK posteriors. We may see it as a pure model, in the sense that it aims to define sleep microstates based solely on the data, unaffected by any a priori RK based scoring.

3 Sleep markers

In this section we apply a reductionist approach to the problem of comparing the posterior curves of healthy subjects with those of stroke patients. The idea was to represent each curve by a summary statistic, thus transforming the dataset into two one-dimensional samples. Potentially, these could be easily evaluated by standard tests. Since a scalar cannot possibly contain all the information stored in a posterior curve, a number of different summary statistics were considered for each curve. The choice of the statistics reflected aspects of the posteriors that we wished to explore. Given the specific applied nature of the problem, most of the statistics were defined following a physiological interpretation and were therefore called sleep markers or features. For instance, the area under the wake curve gave a relative measure of wake present in sleep.

Subsequently, features were considered one by one and the difference between the two samples was assessed under the respective transformation. Hence, the scalarisation of the posteriors enabled us to analyse various features of sleep individually and to identify the relevant ones. Thus we could answer the question of what the differences in sleep patterns are, rather than whether there are any.

The line of action might appear obvious. For example, given a particular sleep marker, one might apply the two-sample t-test for a difference in means. Tracing back, however, the features to be compared are determined by posteriors, which in turn are a product of a fitted PSM. Therefore, the final comparison of features depends on the availability of a training dataset for our PSM. With limited data, it is necessary to exclude some PSGs from the comparison process and use them for model fitting, where the recordings to be set aside are chosen as a random subset of all recordings. The specific selection of training and testing subjects is likely to have a marked effect on the results, though. With the aim of mitigating these random fluctuations, we propose a scheme for repeated model fitting and testing based on the bootstrap method.

3.1 The standard bootstrap

Suppose $x_1, x_2, \ldots, x_n$ are observations of a scalar random variable $\mathbf{x}$ distributed according to an unknown law F, and we wish to find a 95% confidence interval (CI) for some parameter $\theta(F)$, without assuming a parametric model for F. If we could generate another, say, 999 n-samples from the same probability distribution, then a straightforward way to get an approximation of the CI is to calculate an estimate of $\theta(F)$ from each sample and take the 25th lowest and the 25th highest estimate value to be the lower and upper bound. A bootstrap solution stems from the following idea. To circumvent the difficulty posed by the unknown true distribution F, samples are generated from its empirical counterpart instead. The empirical distribution function (EDF) is given by

$$\hat{F}(x) = \frac{1}{n} \sum_{i=1}^{n} H(x - x_i)$$

where H is the Heaviside step function defined as

$$H(x) = \begin{cases} 0 & \text{for} \quad x < 0 \\ 1 & \text{for} \quad x \geq 0 \end{cases}$$

The EDF puts an equal probability on each of the observed outcomes of $\mathbf{x}$. Estimating the CI from sample quantiles is also an instance of substituting the true distribution by its empirical version. Here the distribution of the θ-estimate is sought and its EDF is generated rather than observed.

The described method works well under the implicit assumption that the EDF approximates the true distribution well enough. Furthermore, sampling from the EDF is equivalent to sampling with replacement from the set of observations, $\{x_1, x_2, \ldots, x_n\}$, so the technique is easy to implement. The number of samples to generate may be arbitrarily chosen, but in general, extra samples improve the accuracy of the estimate. Finally, note that the CI derived by taking the 2.5% and 97.5% quantile, is one out of many versions of the bootstrap CI. For others, see for example [3].

3.2 Subset-conditional bootstrap

We consider a modification of the bootstrap that incorporates repeated random sampling of a basis subset on which the rest of the data is to be conditioned. The standard bootstrap is then nested in this procedure. The method is suited for problems in which the data under consideration require some form of processing in order to become more tractable, but the processing itself depends on additional data input. The algorithm is detailed below.

In: A sample of high-dimensional data, $X = \left\{x^{(1)}, x^{(2)}, \ldots, x^{(n)}\right\}$.

:: for $i = 1 : B$

:: Generate a random m-subset $X_T \subset X$ and fit a model $M(\,\cdot\,|X_T)$.

Formally, we treat the model as a function that accepts $x^{(i)}$ as its argument and outputs its scalar representation, $M : X \to \mathbb{R}$. The output depends in a non-trivial way on the training dataset X_T.

:: Apply the model to the data not used in model fitting. Obtain

$$X_E = \left\{M(x^{(i)}|X_T) \,|\, x^{(i)} \in X \setminus X_T\right\}$$

This is a set of numbers.

:: Treat X_E as a sample from which $\theta(F)$ is to be estimated using the standard bootstrap: draw b samples with replacement from X_E and from each derive an estimate of $\theta(F)$. Denote the estimates $\theta_1^{(i)}, \theta_2^{(i)}, \ldots, \theta_b^{(i)}$.

end

:: From the $B \times b$ estimates of $\theta(F)$,

$$\left\{\theta_j^{(i)} |\; i = 1, \ldots, B;\; j = 1, \ldots, b\right\}$$

compute the 2.5% and 97.5% quantile. These constitute an approximate lower and upper bound of the desired CI.

Out: 95% confidence interval for $\theta(F)$.

This procedure inherits the rationale from the basic bootstrap method and adapts it for a more complex situation. Here, the original sample provides only indirect observations of the measure of interest. This ambiguity produces variation in the employed EDFs. Note that the method can be used to test hypotheses of the form $H_0 : \theta(F) = c$, where c is a constant. We reject the hypothesis if the computed CI does not contain c.

We can also derive a two-sample variant of the subset-conditional bootstrap for testing hypotheses defined as $H_0 : \mathbb{E}(\mathbf{x}) = \mathbb{E}(\mathbf{y})$. We define $\theta = \mathbb{E}(\mathbf{x}) - \mathbb{E}(\mathbf{y})$ and find a CI for the value of the parameter. Subsequently, the direction of inequality can be inferred from the position of the interval. However, the generalisation to three or more samples is not so straightforward. The main obstacle is represented by the problem of defining θ as a function of the expected values in a way that would permit subsequent extraction of information about the individual pairwise comparisons. This problem is a serious one and alternative techniques must be employed.

3.3 Two-sample study

Our aim was to apply the introduced methods to identify aspects of sleep in which the two groups of subjects differ. Therefore, we analysed a series of sleep markers one at a time. For each marker, we quantified the dissimilarity between the two samples and evaluated its statistical significance.

Suppose we are interested in a feature whose distribution in the healthy population we model by a scalar random variable $\mathbf{h}$ and its distribution among stroke patients by $\mathbf{s}$. We proceed by finding a CI for the quantity $\theta = \mathbb{E}(\mathbf{h}) - \mathbb{E}(\mathbf{s})$. This may be accomplished using a two-sample version of the subset-conditional bootstrap developed in the preceding section, where one of the samples consisted of 65 EEG recordings of healthy sleepers and the other sample was constituted by 29 sleep transcriptions of stroke patients. The number of outer iterations was set to $B = 1000$. In each iteration, an RK-based PSM composed of 20 Gaussians was fitted to a random subset of 36 control subjects, so that the number of healthy and stroke subjects in the comparison

stage was equal. Stroke patients were omitted from the fitting process on the grounds that the RK scoring rules are not properly defined for individuals with cerebral damage. Next, RK posterior representations were computed for each group of subjects and the selected sleep marker was extracted from each set of posterior curves. The inner loop was set to $b = 100$ iterations. In each run, two bootstrap samples were drawn, one from the set of features characterising the healthy population, and one representing the stroke group: $\{h_1^*, h_2^*, \ldots, h_{29}^*\}$ and $\{s_1^*, s_2^*, \ldots, s_{29}^*\}$ respectively. The bootstrap estimate of θ was defined as the difference in means

$$\frac{1}{29}\sum_{i=1}^{29} h_i^* - \frac{1}{29}\sum_{i=1}^{29} s_i^*$$

Ultimately, the 100,000 θ-estimates were used to compute a 90%, 95%, 99% and 99.9% CI for θ and to resolve the hypothesis of no difference, $H_0 : \theta = 0$.

Table 2 lists the sleep markers for which a significant difference was found between the two groups. Code explanations and details on the computation of individual features are given in Table 3. Most sleep markers, in addition to providing global information, are computed for each quarter of the night separately.

The results indicate that stroke patients have a shorter total sleep period than healthy individuals, though this result may be strongly influenced by the clinical protocols of the two studies (on average, we can observe higher values of the time in bed sleep parameter in healthy subjects, see Table 1). Findings concerning sleep latency imply that stroke patients fall asleep faster, which could make up for a higher proportion of S2 in the 1st quarter of the night. With regard to sleep dynamics, stroke patients exhibit fewer state changes including abrupt ones from SWS to wake. Also, higher entropy implies a more uniform distribution of both SWS and S2 sleep. The number of wake periods is lower, which offers an explanation for higher levels of S2 and SWS. It can be observed that once in S2 stage, a stroke subject is more likely to stay there than a healthy subject. There is frequent alternation between wake and S2 in the 1st phase of sleep. Overall, stroke patients' sleep appears shorter, less interspersed with wake and more regular, with a weaker tendency to change from one state to another.

Table 2: Sleep features that were found to be significantly different between healthy and stroke subjects. Feature definitions are given in Table 3. For each parameter, the group exhibiting higher values of the parameter is indicated. The significance of global, as well as quarter-night comparisons is stated. Unavailable quarter-night statistics are shown as '-'.

		Quarter of the night				
feature	healthy $\lessgtr$ stroke	Q1	Q2	Q3	Q4	Q1-4
tsp	healthy >	-	-	-	-	**
rauc-s2	< stroke		.			.
rauc1-s2	< stroke	*				
rauc2-s2	< stroke	*				
rauc2-sws	< stroke					.
sl-s1	healthy >	-	-	-	-	***
sl-s2	healthy >	-	-	-	-	***
fw	healthy >			*	*	*
sc	healthy >	-	-	-	-	**
rent-s1	healthy >				.	
rent-s2	< stroke		.			.
rent-sws	< stroke		*	*	.	.
rsc-sws-w	healthy >		*			*
rsc-s2-s2	< stroke		.			.
rsc-s2-w	< stroke	.				
rsc-w-s2	< stroke	.				

significance codes: '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' '

Table 3: Description of the used sleep markers and their abbreviations. The prefix r- indicates features normalised by time – these do not represent absolute values.

In the following, X and Y stand for any of the states W, S1, S2, SWS or REM. The last six features are computed from the RK-like discretisations of posteriors.	
-q1, ...-q4	Suffix indicating quarter of the night.
tsp	Total sleep period. Time from the first appearance of any sleep stage until final awakening.
rauc-x	Mean value of the X posterior curve.
rauc1-x	Mean value of the first derivative of the X posterior.
rauc2-x	Mean value of the second derivative of the X posterior.
rent-x	Entropy computed from the X curve after normalisation.
path-length	Normalised length of the path described by a point in 5D whose coordinates are given by the posteriors.
sl-x	Sleep latency. Time until the first appearance of X.
wtsp	Number of wake epochs within the sleep period.
fw	Frequency of wake. Number of coherent periods of wake.
rfw	fw divided by the length of the time frame.
sc	Number of state changes.
rsc-x-y	Number of state changes from X to Y divided by the length of the time frame.

3.4 Three-sample study

Here we address the question of whether different types of cerebral impairment have different effects on sleep architecture. We divided the sample of patients after stroke into two groups according to stroke severity. Type 1 stroke was defined as clinically insignificant stroke in the non-dominant hemisphere or in the brainstem. Subjects with type 1 stroke should not have a severe motoric or speech impairment. Type 2 stroke constituted the complement to type 1. In this study, we compared 14 type 1 stroke subjects, 20 type 2 stroke subjects and 65 healthy subjects.

We adopted the same approach of one-by-one feature analysis as above and a similar multiple-model framework was used in combination with analysis of variance (ANOVA). Again, 1000 iterations were run. Each iteration involved training a PSM on a random subset of 45 healthy subjects. The number was chosen to ensure approximately equal sample sizes in the following steps. Then, we converted the EEGs to posterior representations and obtained three samples of sleep markers of sizes 14 (stroke patients type 1), 20 (stroke patients type 2) and 20 (healthy). Each sample was normality-checked using the Kolmogorov-Smirnov test. The assumption about homogeneity of variances across samples was verified using Bartlett's test and ANOVA was performed. Finally, Tukey-corrected pairwise 95% confidence intervals for differences in means were computed.

In contrast to the two-sample case, this procedure does not produce a definite answer. Instead, we obtain results from 1000 tests, possibly conflicting. In some iterations, the samples may not even be normally distributed or homoskedastic. To our knowledge, there are no statistically rigorous means of aggregating the output, we therefore have to apply a heuristic. A summary of results is presented in Table 4.

Healthy subjects were found to sleep significantly longer than subjects with stroke. Again, note that this finding can be influenced by different clinical protocols used in the two studies. The assumptions of normality and homoskedasticity are violated here, nevertheless, the outcome of the test is consistent with our earlier findings. We also confirmed results regarding increased SWS and S2 in stroke patients in the 1st phase

of sleep, as well as globally. The association with less wakefulness during sleep, both in terms of frequency and length of wake periods, remains valid. Likewise, uniformity of sleep in the form of higher SWS entropy was detected.

The distinction between the two stroke groups is less clear-cut, and concerns chiefly the 1st quarter of sleep. Subjects with type 2 stroke appear to have more wake and S1, which is consistent with findings concerning the number of wake periods. This group's entropy of wake and S1 in the 1st quarter is higher as well. The S2-wake oscillation pattern mentioned in the preceding section seems to be characteristic for this group and is reflected by both a higher number of relevant state changes compared to the smaller stroke group, and longer path length compared to healthy subjects. The other 14 stroke patients exhibit a form of alternations in the early sleep stage as well, specifically between S2 and SWS. This agrees with our earlier observation regarding this group's lower levels of wake and S1.

In the first instance, the ANOVA approach produced results similar to those output by the bootstrap. ANOVA's drawback lies in the specification of a parametric structure which led to assumption testing that could not be readily dealt with in the iterative context. The test is likely to perform poorly given strongly skewed data. However, if the samples under consideration are approximately normal, ANOVA outperforms the bootstrap. Also, it has allowed a more fine-grained discrimination between subjects.

Table 4: Sleep features that were found to be significantly different across the groups of healthy subjects (h), type 1 stroke patients (s14) and type 2 stroke patients (s20). Feature definitions are given in Table 3. For each parameter, significant pairwise differences are indicated, along with test assumptions verification and ANOVA results.

	Tukey pairwise	sample normality				
feature	comparisons	h	s20	s14	Bartlett	ANOVA
tsp	h > s20	!			*	*
rauc2-w-q1	s20 > s14					.
rauc-s1-q1	s20 > s14					.
rauc1-s1-q1	s20 > s14					.
rauc2-s1-q1	s20 > s14					.
rauc1-s2	h < s20					.
rauc1-s2-q1	h < s20					*
rauc2-s2	h < s20					.
rauc2-s2-q1	h < s20					*
rauc-sws-q3	h < s20					.
wtsp	h > s20					.
wtsp-q2	h > s14					.
fw	h > s14					.
fw-q1	s20 > s14					.
rfw-q1	s20 > s14					*
rent-w-q1	s20 > s14					.
rent-s1-q1	s20 > s14					*
rent-s2-q3	s20 > s14		!		*	*
rent-sws	h < s20					.
rent-sws-q2	h < s20				.	.
rent-sws-q4	h < s20					.
path-length-q1	h < s20					.
rsc-w-s2	s20 > s14					.
rsc-w-s2-q1	h < s20 > s14					*
rsc-s2-w-q1	h < s20 > s14					*
rsc-s2-sws-q1	h < s14					.
rsc-sws-s2-q1	h < s14					.
rsc-w-w-q3	s20 < s14		!		.	.

Significance codes: '*' 0.05 '.' 0.1 ' '

Indicated significance relates to the mean P-value across iterations.

Normality rejected in more than 5% of cases is denoted by '!'.

4 Principal components

Though correct and capable of yielding reasonable results, the approach developed in the preceding sections cannot capture the variation of sleep in time. Some sleep markers can be modified to operate on one quarter of the night only, but the role of time in this process is nevertheless minor. In order to describe sleep dynamics, we abandoned the notion of sleep markers and sought a form of reduction that would retain time-related properties of posterior curves.

One such technique, and a common first stop in continuous-curve-data mining, is the functional principal component analysis (fPCA) [10]. As the name suggests, the method represents an alternative to the standard statistical PCA technique of extracting directions of maximal variance, designed for functional data such as posterior curves. In fPCA the original posterior curves are represented in a new basis providing a lower-dimensional representation. Computationally, the process involves finding dominant eigenvalues and eigenfunctions of the variance-covariance function estimated from the posterior curves. For a detailed treatment of fPCA and related methods, see [10]. We made use of the PACE implementation [14].

4.1 Beyond Rechtschaffen and Kales

In addition to incorporating time, it was observed by Rosipal et al. [12] that microstate posteriors themselves may contain more information about sleep than RK-based posterior representations. Moreover, it was interesting to build a model independent of any a priori RK classification and see what we could learn purely from the EEG data. Before we could adopt this approach, two problems needed to be addressed.

First, one could argue that the interpretation of microstates is problematic. While RK stages have a well established physiological meaning, all that characterises a microstate is the corresponding Gaussian in $\mathbb{R}^{10}$. But there is a plausible way of assigning meaning to a microstate via the centre of the Gaussian, a typical representative of the microstate. From the vector of AR coefficients, the microstate-characteristic frequency spectrum can be recovered and easily read [8, 12] .

Second, the design is incompatible with the repeated model-fitting framework. As follows from the construction of the PSM, the microstates found by the EM algorithm depend on the training dataset. Different datasets produce different models and no determinate matching between the microstates of two PSMs is possible. As opposed to RK posteriors, the meaning of microstate posteriors may vary from model to model. We chose to overcome the problem by dropping the multiple-model paradigm and using a single PSM in our analysis.

4.2 Empirical study

In the following, we present the methodology and results of applying functional data analysis techniques to our dataset. We trained an RK-free PSM using all 99 subjects as the training set and computed posteriors for each subject. On the one hand, by fitting the model to all available data, the subjects were treated symmetrically and we avoided the uncertainty connected with either random or criterion-based training set selection. On the other hand, converting recordings that had already been used for model fitting violates the inter-subject statistical independence of the resulting posteriors. Nevertheless, this led to a faithful and rich representation of sleep. Effectively, we sacrificed independence for greater explanatory power, which is justifiable in an exploratory context.

The model was set to comprise 15 Gaussian mixture components. The number was chosen empirically as a compromise between reasonable adaptive power and stability of the model. Concerns over the model's instability arise from the PSM's definition via maximum likelihood, which includes the EM algorithm that tends to converge to local extrema. This was not considered a problem in our previous attempts, since a large number of models were fitted and individual inaccuracies were likely to smooth out.

Next, PSM-based posteriors were derived and processed using positive spline smoothing. Each out of 99 available polysomnograms was thus represented by 15 smoothed microstate posterior curves. For each microstate, fPCA was performed on the respective 99 curves leading to a reduction to lower-dimensional vectors of scores.

Seeking structure in the data, the vectors were subsequently clustered using the 2-means algorithm. The procedure was performed repeatedly, with cluster centres initialised as two randomly selected data points in each run. Although the algorithm did not always converge to the same state, the results generally differed by fewer than three subjects. In Figures 1, 2 and 3, we consider the clustering given by the 2-means method with initial cluster centres chosen as the mean vector across stroke patients and the mean across the healthy control group. The fraction of variance explained by fPCA was set to 85%, but varying the number of principal components had little effect on the resulting clusters. The assumption underlying this procedure is that if there is no significant difference in sleep architecture of the two populations, then the clusters will be roughly independent of the healthy-stroke partition.

Results for each microstate are summarised in Figures 1, 2 and 3. Taking into account the different lengths of recordings, the value of a representative posterior in a given time point is computed as the mean across available posterior values in that time point. The physiological interpretation of individual microstates can be inferred from their frequency spectra as well as from posteriors. As a rule of thumb, SWS occurs later in the night and is associated with lower-frequency waves.

It can be observed that microstates 2 and 10 are more representative of patients with stroke. We find microstate 2 very interesting, as the cluster posteriors follow the two groups fairly closely from about the middle of the night onwards. In our future investigations, the microstate will be explored in greater depth, for example by analysing the second half of the night separately.

Microstates 3 and 12 capture the specificity of sleep of a small subset of healthy subjects. As can be noted from high probability values, the microstates explain a considerable component of these subjects' sleep. In the next step, a hierarchical clustering method should be used to separate the specific subjects. Alternatively, we could increase the number of clusters.

Microstate 4 describes an aspect of healthy sleep as well, but corresponds to a rather different physiological state and involves a greater portion of the population. Due to the fact that one of the clusters includes virtually all stroke patients, there is a fairly

Microstate	1		2		3		4		5	
Healthy	14	51	15	50	58	7	49	16	21	44
Stroke	9	25	15	19	34	0	30	4	13	21

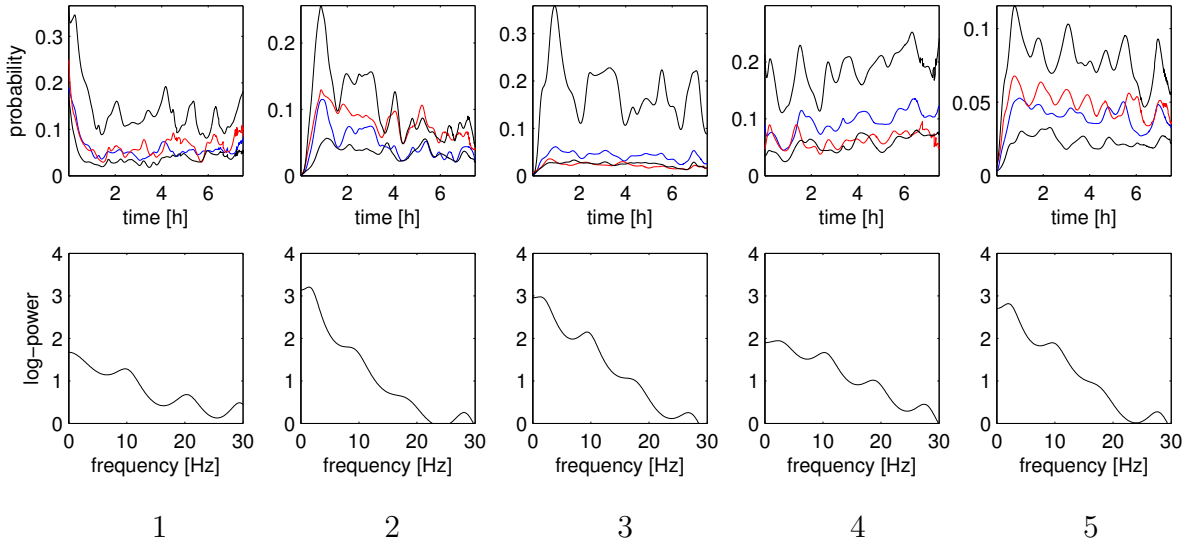

Figure 1: Microstates 1 to 5. From top to bottom: Two-way table of counts, with rows corresponding to clinical classification and colums representing clusters. Mean posterior curves with blue, red and black representing healthy subjects, stroke subjects and the two clusters respectively. Microstate frequency spectrum.

close fit between this cluster and stroke patients' posterior curves.

Based on its spectral and temporal characteristics, microstate 9 could be seen as a close-to-wake state, highly typical of healthy individuals and virtually unachievable by patients after stroke. Especially values of posteriors early in the night differ markedly between the two groups. Microstate 15 is also a quasi-wake state related mainly to healthy subjects, but seems to capture differences between the groups later in the night instead. Apart from those already mentioned, microstates 1 and 13 seem to be closely linked to wake. In neither of these four microstates does the stroke mean posterior dominate the posterior of healthy subjects. This is in line with our earlier findings regarding less frequent occurrence of wake in stroke patients. As expected, stroke patients are more dominant in connection with microstates 2, 8 and 10, which correspond to SWS.

Microstate	6		7		8		9		10	
Healthy	16	49	24	41	18	47	48	17	23	42
Stroke	13	21	15	19	12	22	33	1	22	12

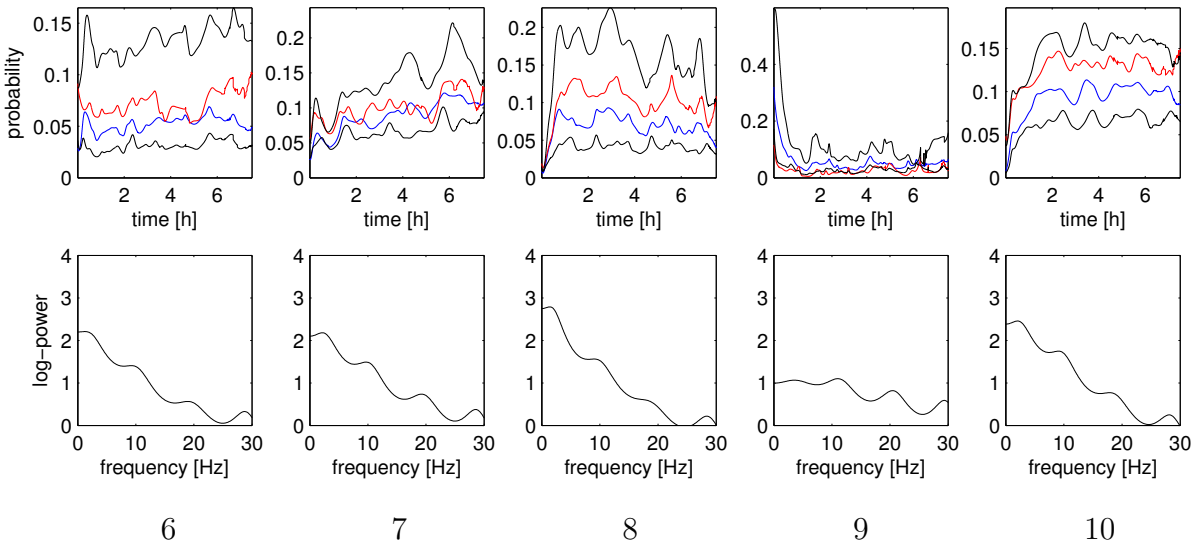

Figure 2: Microstates 6 to 10. From top to bottom: Two-way table of counts, with rows corresponding to clinical classification and colums representing clusters. Mean posterior curves with blue, red and black representing healthy subjects, stroke subjects and the two clusters respectively. Microstate frequency spectrum.

Microstate	11		12		13		14		15	
Healthy	21	44	59	6	49	16	25	40	51	14
Stroke	11	23	34	0	27	7	13	21	31	3

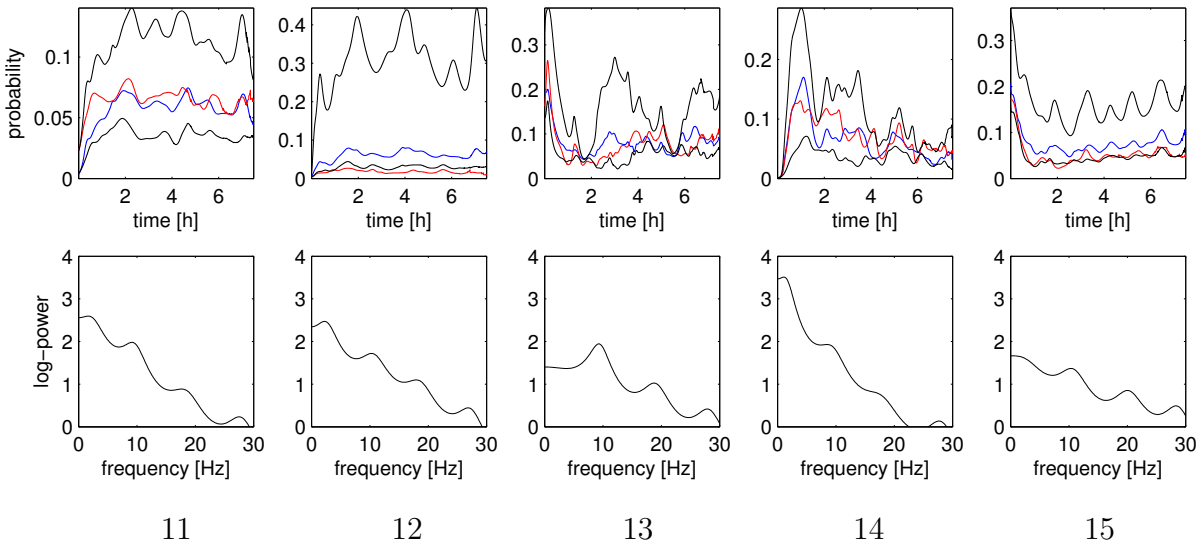

Figure 3: Microstates 11 to 15. From top to bottom: Two-way table of counts, with rows corresponding to clinical classification and colums representing clusters. Mean posterior curves with blue, red and black representing healthy subjects, stroke subjects and the two clusters respectively. Microstate frequency spectrum.

Conclusion

In this thesis, we analysed and compared sleep patterns of healthy sleepers and subjects with stroke. We used the probabilistic sleep model to obtain a tractable but exhaustive representation of sleep and proposed two ways of proceeding with the analysis. While the first approach involved repeated model fitting and feature extraction from the resultant RK-like posterior curves, the other approach combined functional principal component analysis with 2-means clustering and the pure microstate view. The two concepts complement each other. The former allowed us to see the salient differences in aggregate measures of sleep, whereas the latter provided a more fine-grained picture.

Our search did lead to interesting findings. We detected notable differences in sleep microstructure of the studied populations and these differences were also reflected in macro-level sleep descriptors. We recognised considerable within-group variation of the sleep process, especially in the case of healthy subjects, and achieved preliminary results pointing to a non-trivial relationship between the type of stroke and its effect on sleep architecture. No obvious pathologies in the sleep of stroke patients were detected.

In addition to the obtained results, we see the contribution of the thesis in the following. First, we demonstrated the descriptive power of the PSM, either as a continuous alternative to the RK hypnogram, or as a standalone detailed transcription of sleep. Second, we developed the subset-conditional bootstrap, a modification of the bootstrap for schemes involving repeated model fitting and testing. We illustrated its use in a two-sample problem and saw that the absence of similar techniques for more-sample problems led to problematic evaluation. Last but not least, we hinted at the potential of further research in this area of neuroscience, and at the multitude of ways to tackle problems involving neurophysiological signals.

Still, our study leaves open a number of questions. There is still much to be learnt from our dataset and we will follow the presented work with a more in-depth investigation. A project is under way to ascertain whether the differences in sleep architecture affect stroke patients' day-time neurocognitive performance. Also, it would be interesting to study a large enough stroke sample that would permit a classification

based on the type, position and severity of stroke, and a whole new realm of possibilities would open up by incorporating other polysomnographic channels, most notably the electromyogram and electrooculogram.

References

[1] P. Anderer, G. Gruber, S. Parapatics, M. Woertz, T. Miazhynskaia, G. Klösch, B. Saletu, J. Zeitlhofer, M. J. Barbanoj, H. Danker-Hopfe, et al. An E-health solution for automatic sleep classification according to Rechtschaffen and Kales: validation study of the Somnolyzer 24×7 utilizing the SIESTA database. *Neuropsychobiology*, 51(3):115–133, 2005.

[2] D. J. Buysse, C. F. Reynolds III, T. H. Monk, S. R. Berman, and D. J. Kupfer. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry research*, 28(2):193–213, 1989.

[3] A. C. Davison and D. V. Hinkley. *Bootstrap methods and their application*, volume 1. Cambridge University Press, 1997.

[4] C. Iber, S. Ancoli-Israel, A. Chesson, and S. Quan. *The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications*. American Academy of Sleep Medicine, 2007.

[5] B. Jiang, C. Ding, G. Yao, C. Yao, Y. Zhang, J. Ge, E. Qiu, M. Elia, and R. Ferri. Polysomnographic abnormalities in patients with vascular cognitive impairment - no dementia. *Sleep medicine*, 14(11):1071–1075, 2013.

[6] S. Kapen, A. Park, J. Goldberg, and J. Whynter. The incidence and severity of obstructive sleep apnea in ischemic cerebrovascular disease. *Neurology*, 41(Suppl 1):125, 1991.

[7] G. Klösch, B. Kemp, T. Penzel, A. Schlögl, P. Rappelsberger, E. Trenker, G. Gruber, J. Zeithofer, B. Saletu, W. Herrmann, et al. The SIESTA project polygraphic and clinical database. *Engineering in medicine and biology magazine*, 20(3):51–57, 2001.

[8] A. Lewandowski, R. Rosipal, and G. Dorffner. Extracting more information from EEG recordings for a better description of sleep. *Computer methods and programs in biomedicine*, 108(3):961–972, 2012.

[9] D. J. Miller and H. S. Uyar. A mixture of experts classifier with learning based on both labelled and unlabelled data. In *Advances in neural information processing systems*, pages 571–577, 1996.

[10] J. O. Ramsay. *Functional data analysis*. Wiley Online Library, 2006.

[11] A. Rechtschaffen and A. Kales. A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects. 1968.

[12] R. Rosipal, A. Lewandowski, and G. Dorffner. In search of objective components for sleep quality indexing in normal sleep. *Biological psychology*, 94(1):210–220, 2013.

[13] M. Seeger et al. Learning with labeled and unlabeled data. Technical report, University of Edinburgh, 2001.

[14] F. Yao, H. G. Müller, A. J. Clifford, S. R. Dueker, J. Follett, Y. Lin, B. A. Buchholz, and J. S. Vogel. Shrinkage estimation for functional principal component scores with application to the population kinetics of plasma folate. *Biometrics*, 59(3):676–685, 2003.