ISRUC-Sleep: A comprehensive public dataset for sleep researchers

Article *in* Computer Methods and Programs in Biomedicine · November 2015

DOI: 10.1016/j.cmpb.2015.10.013

4 authors:

Sirvan Khalighi
Emory University
26 PUBLICATIONS **485** CITATIONS
SEE PROFILE

Teresa Sousa
University of Coimbra
33 PUBLICATIONS **480** CITATIONS
SEE PROFILE

José Moutinho Dos Santos
Centro Hospitalar de Coimbra
63 PUBLICATIONS **1,244** CITATIONS
SEE PROFILE

Urbano J. Nunes
University of Coimbra
297 PUBLICATIONS **7,713** CITATIONS
SEE PROFILE

COMPUTER METHODS AND PROGRAMS IN BIOMEDICINE 124 (2016) 180–192

journal homepage: www.intl.elsevierhealth.com/journals/cmpb

ISRUC-Sleep: A comprehensive public dataset for sleep researchers

Sirvan Khalighi[a,*], Teresa Sousa[a], José Moutinho Santos[b], Urbano Nunes[a]

[a] Institute of Systems and Robotics (ISR-UC), Department of Electrical and Computer Engineering, University of Coimbra, Portugal

[b] Sleep Medicine Centre, The Central Hospital of University of Coimbra (CHUC), Portugal

ARTICLE INFO

Article history:
Received 6 May 2015
Received in revised form
6 September 2015
Accepted 5 October 2015

Keywords:
Sleep dataset
Automatic sleep stage classification
Polysomnographic signals
Effects of sleep disorder
Medication effects
Feature selection

ABSTRACT

To facilitate the performance comparison of new methods for sleep patterns analysis, datasets with quality content, publicly-available, are very important and useful.

We introduce an open-access comprehensive sleep dataset, called ISRUC-Sleep. The data were obtained from human adults, including healthy subjects, subjects with sleep disorders, and subjects under the effect of sleep medication. Each recording was randomly selected between PSG recordings that were acquired by the Sleep Medicine Centre of the Hospital of Coimbra University (CHUC). The dataset comprises three groups of data: (1) data concerning 100 subjects, with one recording session per subject; (2) data gathered from 8 subjects; two recording sessions were performed per subject, and (3) data collected from one recording session related to 10 healthy subjects. The polysomnography (PSG) recordings, associated with each subject, were visually scored by two human experts.

Comparing the existing sleep-related public datasets, ISRUC-Sleep provides data of a reasonable number of subjects with different characteristics such as: data useful for studies involving changes in the PSG signals over time; and data of healthy subjects useful for studies involving comparison of healthy subjects with the patients, suffering from sleep disorders.

This dataset was created aiming to complement existing datasets by providing easy-to-apply data collection with some characteristics not covered yet. ISRUC-Sleep can be useful for analysis of new contributions: (i) in biomedical signal processing; (ii) in development of ASSC methods; and (iii) on sleep physiology studies. To evaluate and compare new contributions, which use this dataset as a benchmark, results of applying a subject-independent automatic sleep stage classification (ASSC) method on ISRUC-Sleep dataset are presented.

1. Introduction

Sleep is an active and regulated process with an essential restorative function for physical and mental health [1]. Quality of sleep and sleep disorders have an important effect on the health and quality of life. The study of individual behaviors during sleep through all-night PSG recordings has consistently been an important research topic.

* *Corresponding author.* Tel.: +351 239796201.
E-mail address: skhalighi@isr.uc.pt (S. Khalighi).

http://dx.doi.org/10.1016/j.cmpb.2015.10.013
0169-2607/

Numerous methods have been developed for automatic detection of arousals, apnea, and sleep stages [2–5]. These methods often use PSG recordings, including electrophysiological signals (electrocardiographic activity, brain-wave patterns, eye movements, and activation signal of muscles), pneumological signals (airflow, blood oxygen level, and movement of respiratory muscles), and other contextual information (body position, lights, snore recording, etc.) [6]. These signals have been collected from human individuals using noninvasive surface electrodes. To evaluate the efficiency of automatic sleep pattern analysis methods, non-public and few existing public datasets have been used. Rigorous comparisons between the developed methods cannot be done since the used datasets differ in recording conditions, physiological conditions of subjects and number of assessed subjects. Datasets with quality content, publicly available, are an important vehicle for accelerating research, since they facilitate the performance comparison of new approaches and methods. This paper presents three main contributions:

- We introduce a publicly-available comprehensive sleep dataset, called ISRUC-Sleep, which comprises three subgroups as illustrated in Fig. 1. The subgroups of the dataset contain PSG signals of different adult individuals, including healthy subjects, subjects with sleep disorders, and subjects under the effect of sleep medication. Sleep stages were labeled by two sleep experts. Furthermore, for 8 subjects (subgroup-II), two sets of PSG data, which have been recorded at different time dates, are provided.
- Aiming to help sleep researchers in their analysis and inferences using this dataset, for each subject, useful and complementary information related to sleep disorders, used medications, and their effect on sleep patterns, are presented.
- Aiming to evaluate and compare of new contributions, which will use this dataset as a benchmark, results of applying a subject-independent ASSC method on ISRUC-Sleep dataset are presented. This supervised-learning based method, detailed in Khalighi et al. [7], is henceforth named SSM4S.[1]

2. Terminology and definitions

Background material, terms definition (Table 1), and effects of sleep disorders and medications on sleep patterns are summarized in the next subsections.

2.1. Background

- The Rechtschaffen and Kales standard (R&K) rules are the basis of a consensus scoring procedure for adults [8]. The American academy of sleep medicine (AASM) defined new criteria for sleep scoring based on the R&K rules. In adults, sleep-wake cycle is categorized in awake, non-rapid eye movement (NREM) and rapid eye movement (REM) sleep stages. NREM sleep is further divided into three stages: N1 (drowsiness/transitional sleep), N2 (light sleep) and N3 (deep sleep) [9], the last of which is also called delta sleep or slow wave sleep (SWS). The 2007 AASM visual scoring rules recommend a frontal electrode for best detecting K-complexes, a central electrode for spindles, and an occipital electrode for alpha waves [10]. Based on both scoring rulesets (R&K and AASM), epochs of 30 s (more rarely 20 s) are defined for the PSG signals scoring [11]. Difficulties in sleep scoring arise when sleep does not behave in accordance with the *normal*/expected way as a consequence of sleep disorders, medication or in face of individual specific characteristics of sleep Electroencephalogram (EEG).
- Since the collected PSG signals are characterized by low signal-to-noise ratio (SNR), a preprocessing stage is applied to improve the quality of the signals; i.e. some channels of the recorded signals are filtered to eliminate noise and undesired background EEG, aiming to enhance the PSG signal quality and increase the SNR. The filtering stage comprises: (1) a notch filter to eliminate the 50 Hz electrical noise; (2) a bandpass Butterworth filter with a lower cutoff of frequency 0.3 Hz and higher cutoff of frequency 35 Hz for EEG and EOG channels, and a lower cutoff of frequency 10 Hz and higher cutoff of frequency 70 Hz for EMG channels. More details are presented in Table 3.
- The common EEG frequency bands are: low delta 0.3–1 Hz, delta 1–4 Hz, theta 4–8 Hz, alpha 8–12 Hz, sigma 12–15 Hz, and beta 15–30 Hz. Different EEG waves (alpha, beta, sigma, delta, and theta) characterize different sleep stages. Low amplitude, mixed EEG frequency, saw-tooth pattern, low amplitude Electromyogram (EMG) and high level Electrooculogram (EOG) signals from both eyes, are apparent during the REM stage. In stage N1, waves with high amplitude and frequency range of 2–7 Hz together with the existence of alpha waves are found in EEG signal. Still regarding N1, EMG level is lower when compared to the awake stage. Sleep spindles (12–14 Hz) and K-complexes are observed during N2. N3 (deep sleep) consists of low-frequency high-amplitude waves with frequencies of 2–4 Hz.

2.2. Effect of sleep related disorders and sleep pathology on sleep stage patterns

Sleep apnea is the most frequent sleep disorder seen in sleep medicine centers. The syndrome is characterized by repetitive episodes of upper airway obstruction that occur during sleep and are usually associated with a reduction in blood oxygen saturation. These nocturnal respiratory disturbances result in brief arousals in sleep, which promotes sleep fragmentation that typically disturbs sleep architecture with reduction or even complete deprivation of REM sleep and N3 sleep. An increase of arousals of different length together with an increase in sleep stage changes is a feature of the syndrome. This fragmentation of sleep, inhibiting cortical synchronization, would be responsible for the lower amount of slow wave sequences of the deep sleep [12]. On the other hand, transient experimental hypoxia induced abnormal posterior resting state delta and alpha rhythms in healthy volunteers, and EEG slowing during awake with an increase in relative theta and delta power in occipital, temporal and parietal areas

[1] Sirvan Supervised Method for Sleep Staging.

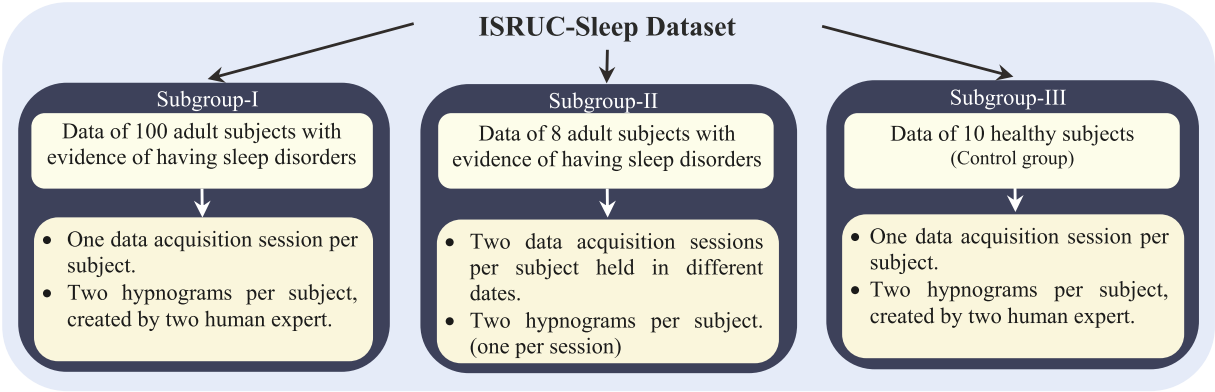

Fig. 1 – Details of ISRUC-Sleep dataset.

was observed in sleep apnea subjects [13], which can be correlated with sleepiness in these patients.

Also, moderate obstructive sleep apnea (OSA) patients have a lower percentage of slow spindles with deceleration compared to mild OSA or normal groups in frontal and parietal regions, which may represent a disruption of thalamo-cortical loops generating spindle oscillations [14]. All together, these findings can contribute to significant differences in agreement among multiple raters of sleep studies in sleep apnea patients [15].

Affective disorders (depression/anxiety disorder) can induce sleep-EEG changes. In depression, an increase in sleep latency and an increase in sleep fragmentation by arousals or intermittent awakenings, and early-morning awakening can be seen. Frequently a shortened REM latency including sleep onset REM periods (REM latency <20 min), prolonged first REM periods with an increase of REM density is present in all age groups [16]; NREM-sleep changes include a decreased SWS, EEG delta power throughout the night and increase of stage N2 sleep together with a shift of EEG-delta power from the first to the second sleep cycle in younger patients [17]. In generalized anxiety disorder (GAD) patients no clear differences with control subjects in SWS and REM sleep are seen, being the differences confined to insomnia-like symptoms [18]. Nevertheless, some studies have shown increased REM latency [19], or increased mean REM latency over consecutive nights [20], in GAD patients compared to depressed patients, which could be useful to distinguish both disorders.

In primary insomnia (psychophysiological insomnia) the hyperarousal model suggests that a deficit of reduce arousal during sleep may be responsible for non-restorative sleep. It has been shown that in these patients there is elevated spectral power values in the EEG beta (cortical arousal) and sigma (spindle) frequency band during N2 sleep stage with no differences in other frequency bands. This increase in cortical arousal and in an index of sleep protective mechanisms (spindles) may provide further evidence for the concept that a simultaneous activation of wake-promoting and sleep-protecting neural activity patterns contributes to the experience of non-restorative sleep in primary insomnia [21].

In dementia, sleep is characteristically grossly disrupted with lower sleep efficiency, higher percentage of N1 and increase of arousals and awakenings. A decrease of SWS can be expected but the hallmark of the sleep EEG in these patients is a slowing EEG activity with spindles reduction. Therefore, the scoring of sleep stages may be challenging [18].

2.3. *Effect of medications on sleep stage patterns*

The medication can also affect the EEG sleep patterns. Chronic use of benzodiazepines has shown to induce an increase of

Table 1 – List of acronyms.

Acronym	Description	Acronym	Description
AASM	The American Academy of Sleep Medicine	LOOCV	Leave-one subject-out cross-validation
ACC	Accuracy	MODWT	Maximal overlap discrete wavelet transform
AR	Auto regressive	NRI	Noradrenaline reuptake inhibitor
ASSC	Automatic sleep stage classification	NREM	Non-rapid eye movement
BCR	Balanced classification rate	PSG	Polysomnography
CAP	The cyclic alternating pattern	REM	Rapid eye movement
CV	Cross validation	R&K	Rechtschaffen and Kales standard
ECG	Electrocardiogram	SAS	Sleep apnea syndrome
EEG	Electroencephalogram	SENS	Sensitivity
EMG	Electromyogram	SNR	Signal-to-noise ratio
EOG	Electrooculogram	SNRI	Selective noradrenaline reuptake inhibitors
GAD	Generalized anxiety disorders	SPEC	Specificity
ISRUC	Institute of Systems and Robotics, University of Coimbra	SSRI	Selective serotonin reuptake inhibitors
		SWS	Slow wave sleep

sleep stage N2, decrease of N3 (lower delta activity and theta activity), and an increase of arousals [22]; increase of spindle activity that can intrude in REM sleep is also described [23].

REM sleep reduction or even complete suppression was reported in humans after an administration of tricyclic or tetracyclic antidepressants [24], monoaminoxidase inhibitors [25,26], SSRIs [27], selective noradrenaline reuptake inhibitors (NRI) [28], SSRIs and selective noradrenaline reuptake inhibitors (SNRI) [29]. A few exception should be the noradrenergic and specific serotoninergic antidepressant mirtazapine [30] which also increase total sleep time and sleep efficiency after four weeks of administration.

Most tricyclic antidepressants increase SWS [24], but there is evidence that selective serotonin reuptake inhibitors (SSRI) impairs sleep continuity by increasing of intermittent wakefulness [18]. On the other side an increase of sleep stage N3 was found during treatment of depressed patients with the SNRI duloxetine [29] and a decrease in REM sleep and increase in REM sleep latency were observed with the SNRI venlafaxine [31]. An increase in the total sleep time and sleep efficiency, and a decrease in the time spent awake, were verified in patients with depression, under mirtazapine medication. These changes persisted after four weeks [30].

Trazodone, a triazolopyridine antidepressant weak, is a specific inhibitor of serotonin (5-HT) reuptake. The use of this medication showed increases in sleep efficiency, total sleep time, total sleep period, N3 and REM duration, as well as decreases in wakefulness during the total sleep period, early morning awakening, and N2 [32]. Besides alterations in sleep EEG induced by diseases or medications in healthy people, there is specific individual sleep patterns that can be seen as a fingerprint that allows a correct discrimination between individuals with a probability of 92% [33]. This EEG fingerprint is genetically determined as shown by studies on monozygotic and dizygotic twins, particularly in the range of alpha and sigma bands [33].

The above observations can explain why human-experts analysis of sleep EEG can be, so far, superior to computer analysis by a better adaptation to the individual characteristics of the electrophysiologic signals, recognizing specific sleep-related characteristics constructing their own, subjective and patient specific PSG pattern, which allows to decide what the most likely sleep stage is [23]. On the other hand, this also explains why as much as 25% of overall disagreement can be seen between two human-experts sleep scorings of the same recording [34]. The disagreement is particularly seen in N1 where the AASM definition include attenuation or slowing of the alpha rhythm and the presence of slow eye movements, 4–7 Hz EEG and vertex sharp waves. Many subjects do not generate some or even any of these waves as is stated in the manual [23].

3. Sleep datasets

To assess the efficiency of sleep pattern analysis methods, each research team collects their own test data with expenditure of time and/or financial resources [47,48]. These datasets, mainly used in the context of their own research, often lack several relevant information details regarding acquisition and subject pathological conditions (neural, cardiorespiratory, medication effects). Some of these datasets [49,50] also lack statistical significance and just recorded some of the PSG channels. Therefore, an accurate and comparative evaluation of the performances of these methods with new methods cannot be done effectively.

Recognizing the need and usefulness of publicly available sleep datasets, which can be used as a common reference for researchers, some sleep-related datasets were developed by sleep research groups. As shown in Table 2, these datasets contain multiple signals from some healthy and patient subjects. The sleep datasets of PhysioBank [35] have been used in a few works (see Table 2). Even though *MIT-BIH*, *Sleep-EDF* and *Extended Sleep-EDF* are general purpose datasets, these do not have enough subjects for generalization purposes. *CAP-Sleep* dataset is an exception in PhysioBank repository, containing 108 recordings, however it consists of the specific data useful for studies related to CAP.

The *sleep heart health study (SHHS)* dataset [43], which has a convenient number of recordings, is not a completely public dataset. It is available only upon special request and approval. On the other hand, due to providing just the signals of two EEG (C3-A2 and C4-A1) channels, *SHHS* has limitations for general-purpose sleep research. In fact, it is a specific purpose dataset useful in research studies involving relationships between sleep-disordered breathing and heart diseases.

Recently, *Montreal archive of sleep study (MASS)*, which is an open-access sleep dataset collected from healthy subjects, was proposed by O'Reilly et al. [45]. Although it is reported that the dataset contains data of 200 participants, it is a collection of five different subgroup of data. These subgroups were pooled from 8 different research protocols performed in 3 different hospital-based sleep laboratories. Furthermore, there exist some access restrictions regarding different kinds of information of the dataset. As detailed in Table 2, the subgroups of this dataset have significant differences in terms of number of channels, filtering methods applied to the signals, acquisition software, annotations, scoring criteria and epoch size.

In summary, all the dataset detailed in Table 2 have limitations in some aspects and as far as we know, except Sleep-EDF dataset (expanded), which were recorded in two subsequent day-night at the subject's home, in the others only one acquisition session (one recording) per subject is available.

4. ISRUC-Sleep dataset

ISRUC-Sleep dataset contains data collected from all-night PSG recordings with duration around eight hours. Each recording was randomly selected between PSG recordings that were acquired by the Sleep Medicine Centre of the Hospital of Coimbra University (CHUC), in the period 2009–2013. Overall standard setup setting for data acquisition, comprised a biosignal acquisition equipment (a SomnoStar Pro sleep system which is a multi-channel ambulatory recording device), and a set of sensors collecting data in a non-invasive way, according to the international 10–20 standard electrode placement (Fig. 2) [52]. With regard to the arrangements, the subject

Table 2 – Some of the public sleep datasets.

Dataset	Subjects/sampling rate	Recorded channels	Recording duration	Purpose of creating the dataset	Subjects age	Literatures cited dataset
MIT-BIH [35]	18 recording of 16 subjects with or without sleep apnea syn-drome(SAS)/250 Hz	Four-, six-, and seven-channel recordings of ECG signal, an invasive blood pressure signal, an EEG signal, a respiration signal and a text. Some records contain other signals such as, respiratory effort signal., an EOG signal, an EMG (from the chin) signal, a stroke volume signal	8–10 h	General purpose	32–56, Avg. = 43	Adnane et al. [36], Nicolaou and Georgiou [37], Fraiwan et al. [38]
Sleep-EDF [35]	8 subjects without any sleep-related medication, scored based on R&K/100 Hz	EEG (Fpz-Cz and Pz-Oz), Horizontal EOG, submental-EMG envelope, oro-nasal airflow, rectal body temperature and an event marker	1.25–6.5 h	General purpose	21–35	Bajaj and Pachori [4], Ronzhina et al. [39]
Expanded Sleep-EDF [35]	61 recordings from healthy subjects, without any sleep-related medication/100 Hz	EEG(Fpz-Cz and Pz-Oz), Horizontal EOG, submental chin EMG, and an event marker	Around 9 h	Study of age effects on sleep; study of temazepam effect on sleep	25–101	Kemp et al. [40], Yaghouby et al. [41]
CAP-Sleep [35]	108 recordings scored based on R&K/512 Hz	3 EEG channels (F3 or F4, C3 or C4 and O1 or O2, referred to A1 or A2), 2 EOG channels, EMG of the submental is muscle, bilateral anterior tibial EMG, respiration signals (airflow, abdominal and thoracic effort and SaO2) and ECG	8–10 h	Study of the cyclic alternating pattern (CAP)	30–75	Terzano et al. [42]
SHHS-1, -2 [43]	9736 recordings scored based on R&K/125 Hz	2 EEG channels (C3 or C4, referred to A1 or A2), 2 EOG channels, EMG of the submental, bilateral anterior tibial EMG, respiration signals (airflow, abdominal and thoracic effort) and ECG	Overnight	Study of OSA, sleep-disordered breathing, and heart diseases	40 and older	Ebrahimi et al. [44]
MASS [45]	Collection of 200 recordings (5 different subsets) scored based on R&K or AASM/256 Hz	4–20 EEG channels, 2–4 EOG channels, 1–5 EMG of the submental, bilateral anterior tibial EMG, sometimes with respiration signals (airflow, abdominal and thoracic effort and SaO2) and ECG	Overnight	Study of the sleep spindles and general purpose	18–76	Tsanas et al. [46]

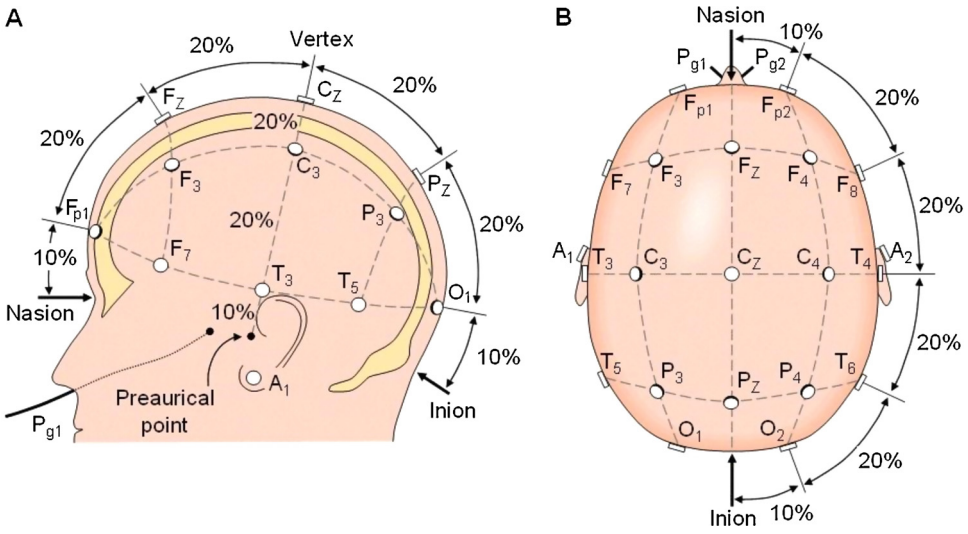

Fig. 2 – The international 10–20 system seen from (A) left and (B) above the head. A = ear lobe, C = central, Pg = nasopharyngeal, P = parietal, F = frontal, Fp = frontal polar, O = occipital [51].

sleeps in a bed in a patient's room, and the experts and technicians stay in a separate room. All patients referred were submitted to an initial briefing with the support of an informed consent document. The ethics committee of CHUC approved the use of the data of the referred patients as anonymous for the research purposes.

The PSG signals were recorded according to the recommendations of the AASM manual. As described in Table 3, each recording consists of signals from 19 channels. All EEG, EOG, and chin EMG signals were sampled at 200 Hz and stored using the standard EDF+ data formats with .REC extension [53]. All recordings of the dataset were segmented into epochs of 30 s and visually scored by two different sleep experts in CHUC according to the guidelines of AASM [9], with the stages: awake, NREM (N1, N2, and N3) and REM sleep. Calculating Cohen's kappa index between two experts over the subjects of the subgroups yields to the following Kappa indexes: overall kappa index of 0.87 ± 0.09 for subgroup-I, 0.82 ± 0.15 for subgroup-II, and 0.9 ± 0.06 for subgroup-III.[2] The labels are stored in standard text file format, where each line corresponds to one epoch; Moreover, the gender, height, weight, age and date of recording, of individuals tested are recorded in the header of each text.

Further analysis such as sleep events, sleep related disorders, other diseases, sleep pathology, used medications, EEG pattern alterations, and percentage of each sleep stage for each subject are presented.

ISRUC-Sleep dataset comprises three subgroups of data[3] as described in Table 4.

[2] Distribution of the individual kappa indexes over the recordings are presented in the result files, which are available via http://sleeptight.isr.uc.pt/ISRUC_Sleep/.

[3] Recordings, summary of the characteristics, and clinical information for each subject of the dataset are available via http://sleeptight.isr.uc.pt/ISRUC_Sleep/.

5. Application of ISRUC-Sleep in ASSC

Some data of Subgroup-I was already used in a few works [7,54–57]. Aiming to improve the applicability of automatic sleep staging, we proposed the *SSM4S* classification method with main blocks depicted in Fig. 3. In this ASSC method, after applying common preprocessing, and segmentation of the signals in 30 s epochs, some features are extracted using several methods in the temporal, frequency and time-frequency domains. PSG signals are traditionally analyzed in the frequency domain, since each sleep stage is characterized by a specific pattern of frequency contents. Moreover, PSG signals are non-stationary; therefore time-frequency transformations like wavelets are very useful. Due to superiority of the maximal overlap discrete wavelet transform (MODWT) [7,58] versus discrete wavelet transform, a MODWT of depth 6 with Daubechies order four (db4) is applied to every 30 s epochs with a sampling rate of 200 Hz. The frequency ranges are broken down into δ range (<4 Hz), θ range (4–8 Hz), α range (8–13 Hz) and β range (13–30 Hz). To represent the time-frequency distribution of the EEG, EOG and EMG signals, features such as energy, percent of energy [55], mean and standard deviation are extracted from each sub-band.

Furthermore, due to the importance of spectral and temporal analysis, features such as relative spectral power, peak to peak amplitude of two EOGs, Tsallis ($q = 2$), Renyi ($\alpha = 2$), Shannon entropy, Hjorth parameters, harmonic parameters, percentile 25, 50, 75, autoregressive coefficients (order 3), slow wave index (SWI), kurtosis and skewness [1] are extracted from EEG and EOG channels.

To reduce the influence of extreme values, the matrix of features $\mathbf{Y}$ is transformed as follows:

$$\mathbf{X} = \arcsin(\sqrt{\mathbf{Y}}) \quad (1)$$

Table 3 – Details of recorded signals of ISRUC-Sleep dataset.

Channel number	Type of the signal	Label	Frequency rate/Hz	Butterworth	Notch filter	Description
1	EOG	LOC-A2	200	0.3 Hz–35 Hz	50 Hz	Left eyes movements
2		ROC-A1				Right eyes movements
3	EEG	F3-A2	200	0.3 Hz–35 Hz	50 Hz	Brain channels with the references A1 and A2, which placed in the left and right ear-lobes
4		C3-A2				
5		O1-A2				
6		F4-A1				
7		C4-A1				
8		O2-A1				
9	Chin EMG	X1	200	10 Hz–70 Hz	50 Hz	Chin EMG, placed between the chin and the lower lip
10	ECG (EKG)	X2	200		50 Hz	Electrocardiographic
11	Leg-1 EMG	X3	200	10 Hz–70 Hz	50 Hz	Left leg movement
12	Leg-2 EMG	X4				Right leg movement
13	Snore	X5	200	10 Hz–70 Hz	50 Hz	Snore (derived)
14	Flow-1	X6	12.5			Airflow (pressure based)
15	Flow-2	DC3	25			
16	Abdominal	X7	25			Abdominal efforts
17		X8				
18	Pulse oximetry	SaO2	12.5			Pulse oximetry (SaO2)
19	Body position	DC8	25			Body position (BPOS)

where $\mathbf{Y}$ denotes the feature matrix, and

$$\mathbf{X} = \{x_{ij};\quad i = 1, 2, \ldots, N \quad and \quad j = 1, 2, \ldots, M\} \tag{2}$$

is the transformed feature matrix, where N and M denote the number of subjects and the number of features, respectively, and then to avoid features in greater numeric ranges dominating those in smaller numeric ranges, as well as numerical difficulties during classification; each feature of the transformed matrix $\mathbf{X}$ is independently normalized to the [0,1] range by applying

$$\bar{x}_{ij} = x_{ij}/(\max(\mathbf{x}_j) - \min(\mathbf{x}_j)) \tag{3}$$

where $\mathbf{x}_j$ is a vector of each independent feature. Next, a two-step feature selection process that consists on a filtering and a wrapper phases is performed: firstly, as detailed in [7], the less discriminative feature-types are removed and features such as relative spectral power, harmonic parameters, percentile 75, autoregressive coefficients (order 3), kurtosis and skewness [1] are selected. Then, in the second step, to select the best elements of each feature-type, resulted feature vector is fed into a mRMR feature selector. Finally the features are classified based on classical supervised learning by a SVM classifier.

6. Performance evaluation

We briefly present the average performances of *SSM4S* method, using different subgroups of the ISRUC-Sleep dataset. Two types of experiments have been carried out: sleep-wake detection and multiclass sleep staging based on AASM. The purpose of these experiments is to provide the reader with sufficient evidence to consider ISRUC-Sleep as a working

Table 4 – Characteristics of ISRUC-Sleep dataset.

Dataset	Subjects	Number of recording per subject	Subject characteristics	Subjects age
Subgroup-I	100 subjects (55 male, 45 female) with evidence of having sleep disorders	One data acquisition session per subject	Most of the subjects have detected sleep apnea events; the subjects could be under medication, but all were in position to breathe without the help of machine	20–85, Avg. = 51, std. = 16 years
Subgroup-II	8 subjects (6 male, 2 female) with evidence of having sleep disorders	Two data acquisition sessions were performed in two different dates	Detected sleep apnea events; the subjects could be under medication, but all were in position to breathe without the help of machine	26–79, Avg. = 46.87, std. = 18.7 years
Subgroup-III	10 subjects (9 male, 1 female)	One data acquisition session per subject	Healthy subjects (control group)	30–58, Avg. = 40, std. = 10 years

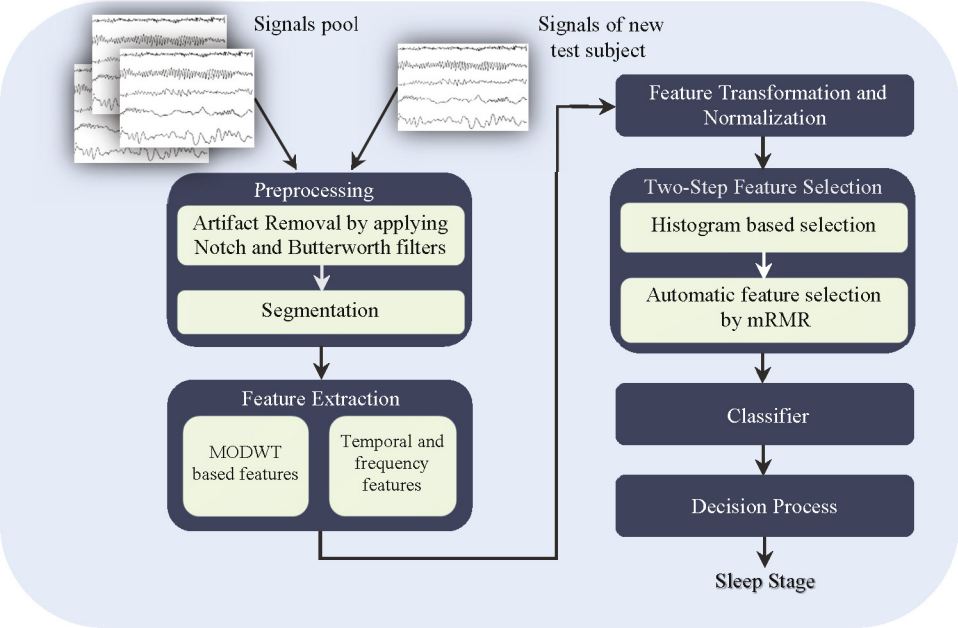

Fig. 3 – Structure of the automatic sleep stage classification method SSM4S.

dataset for the development of sleep monitoring and analysis systems. To evaluate the system the following performance measures were used: accuracy (ACC), specificity (SPEC), sensitivity (SENS). In addition to these common performance measures, the balanced correction rate (BCR), which is the average of the proportion of correct classifications in each class, was also calculated.

The ASSC method was applied on data of subgroup-I, -II and -III. Due to the high number of subjects, all the assessments with the data of subgroup-I, were determined by five- and ten-fold cross validation. However, since there are eight and ten subjects in subgroup-II and -III, respectively, to verify reliability of the results, all the experiments with these two subgroup were done using leave-one subject-out cross-validation (LOOCV) strategy. The experimental results for sleep-wake detection are summarized in Table 5. From the analysis of detailed results,[4] it was verified that, the higher-performance values were attained with the subjects with longer periods of awake stage during the all-night recording (approximately 8 h of data collection). As expected, the ASSC method of SSM4S, achieved worse average and standard deviation, with data of the subjects with suspected sleep disorders (subgroup-I). The ambiguous patterns on PSG recordings, mainly due to sleep disorders and artifacts, can affect the performance of the method.

Moreover, the recordings of subgroup-I, -II and -III were used to evaluate SSM4S, in multiclass sleep staging. Based on average/std., shown in Table 6, the best discrimination were achieved for awake, N3, REM, and N2 stages, respectively. The lowest average performance resides in the classification of stage N1.

7. Analysis of ISRUC-Sleep dataset for ASSC

This section summarizes the main conclusions derived from applying SSM4S over ISRUC-Sleep dataset. To analyze the relation of experts agreement and classification performance two measures were calculated. Auto-regressive coefficients, which is a representation of a time series such that it specifies that output variable depends linearly on its own previous values, and balanced correlation rate (BCR) were evaluated.

7.1. *For sleep-wake detection*

- There is a remarkable correlation between the agreement levels of two experts and the classification performance (Fig. 4). Despite of this inference, a few exception (recordings related to subjects 12 and 40 of 100 subjects of subgroup-I) were found with high agreement level of experts and

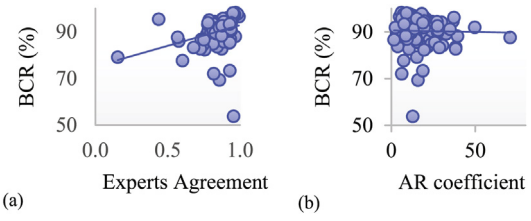

Fig. 4 – (a) Distribution of balanced classification rate (BCR) values concerning two experts agreement in sleep-wake visual scoring; (b) distribution of BCR values corresponding to auto regressive (AR)-coefficient, in automatic sleep-wake detection.

[4] The detailed results of the overall performance of SSM4S, associated with the dataset are available via http://sleeptight.isr.uc.pt/ISRUC_Sleep/Results.

Table 5 – Average results of the ASSC method *SSM4S* for sleep-wake detection. Balanced classification rate (BCR), specificity (SPEC), sensitivity (SENS), accuracy (ACC) are calculated for sleep and awake stages.

Method	Applying method on subgroup-I		Applying method on subgroup-II	Applying method on subgroup-III
Cross validation	Five-fold CV	Ten-fold CV	LOOCV	LOOCV
Average BCR	90.16 ± 06.88	90.30 ± 06.29	82.84 ± 10.26	91.19 ± 03.15
Average SENS	83.98 ± 14.99	84.13 ± 14.67	69.11 ± 23.07	85.03 ± 07.05
Average SPEC	96.34 ± 06.88	96.47 ± 03.50	96.57 ± 05.12	97.35 ± 01.37
Average ACC	93.97 ± 06.32	94.10 ± 06.19	92.40 ± 05.15	95.39 ± 01.10

very low classification performance. Since alpha activity is one of the relevant patterns in awake stage, this factor can affect the performance of awake detection. The observed low amplitude of the alpha activity in the EEG signals of subject 12 of subgroup-I can be the main reason of performance degradation. For subject 40 of subgroup-I, the artifacts in EEG signals, which resulted from the low quality of data acquisition, affected the classification performance.

- There is also a correlation between degradation of the classification performance, and the increase of the number of arousals and awakens.
- There is none significant relation between characteristics such as age, gender, diagnosis and medication, and the classification performance.
- Unusual patterns of alpha activity and rapid activity affected the classification performance.

7.2. *For multiclass sleep staging*

- Similar to sleep-wake detection, a direct relation between experts agreement, AR coefficients and the classification performance, was detected (Fig. 5).
- Regarding misclassification of stage N2, most of the time, this stage was misclassified as stage N1. Moreover, in most of the epochs with high AR coefficients and low level of experts agreement, stage N2 was misclassified (Fig. 6a and b).
- In subjects with higher number of transitions between sleep stages it was verified a trend to lower performances. Since the possibility of stage transition from N2 to the other stages (awake, N1, N3 or REM) is high, once stage N2 is more prevalent in a PSG signal, the classification performance of this stage is affected by the number of stage transitions (Fig. 6c).
- As mentioned in state of the art, the lowest classification performance was related to N1. Since stage N1 makes a link between the wakefulness and the sleep, it has common transition characteristics. Moreover, in subjects with high misclassification of stage N1, it was verified that, the EEG alterations such as artifacts, paroxystic activity and rapid activity affected the automatic classification.
- For some subjects, the standard deviation from average recognition rate of N3, is too high. For example, in subject 10 of subgroup-I, cardiac and sweat artifacts affected the

Table 6 – Average results of the ASSC method *SSM4S* for multiclass sleep staging. Balanced classification rate (BCR), specificity (SPEC), sensitivity (SENS), accuracy (ACC) are calculated for each stage.

Method	Applying method on subgroup-I		Applying method on subgroup-II	Applying method on subgroup-III
Cross validation	Five-fold CV	Ten-fold CV	LOOCV	LOOCV
Average BCR-Awake	91.63 ± 5.57	91.75 ± 5.29	85.82 ± 07.41	92.77 ± 03.09
Average BCR-N1	67.34 ± 8.08	67.12 ± 7.94	63.32 ± 10.35	72.35 ± 23.57
Average BCR-N2	83.93 ± 6.64	83.86 ± 6.91	78.06 ± 08.32	79.58 ± 15.26
Average BCR-N3	89.83 ± 8.08	90.11 ± 7.77	87.26 ± 23.56	91.41 ± 05.61
Average BCR-REM	91.63 ± 5.57	92.28 ± 3.55	85.00 ± 23.02	86.50 ± 06.29
Average SPEC-Awake	95.53 ± 4.26	95.23 ± 4.58	95.34 ± 06.26	93.03 ± 03.78
Average SPEC-N1	95.07 ± 3.90	94.92 ± 4.37	91.24 ± 06.47	79.71 ± 13.56
Average SPEC-N2	86.80 ± 7.14	87.50 ± 6.91	81.98 ± 15.98	82.61 ± 04.56
Average SPEC-N3	96.49 ± 4.43	96.70 ± 4.11	87.26 ± 23.56	83.37 ± 17.95
Average SPEC-REM	97.36 ± 2.53	97.04 ± 4.26	95.61 ± 05.16	87.86 ± 10.77
Average SENS-Awake	87.68 ± 12.06	88.28 ± 11.72	76.30 ± 17.58	93.36 ± 04.25
Average SENS-N1	39.61 ± 17.59	39.32 ± 17.51	35.40 ± 24.22	73.41 ± 20.27
Average SENS-N2	81.06 ± 13.14	80.22 ± 14.54	74.14 ± 10.34	86.86 ± 03.78
Average SENS-N3	83.18 ± 17.33	83.52 ± 16.66	78.41 ± 28.35	91.93 ± 06.44
Average SENS-REM	81.10 ± 21.73	81.76 ± 20.33	73.39 ± 27.00	89.19 ± 07.64
Average ACC-Awake	94.15 ± 04.97	94.07 ± 05.05	92.61 ± 04.71	93.43 ± 06.85
Average ACC-N1	88.26 ± 04.50	88.10 ± 04.69	83.18 ± 06.33	75.91 ± 22.10
Average ACC-N2	85.28 ± 05.29	85.49 ± 05.20	77.00 ± 14.03	87.15 ± 06.14
Average ACC-N3	94.00 ± 03.82	94.18 ± 07.77	91.37 ± 05.45	94.38 ± 03.14
Average ACC-REM	89.22 ± 15.43	89.37 ± 15.19	92.64 ± 03.79	91.46 ± 05.44

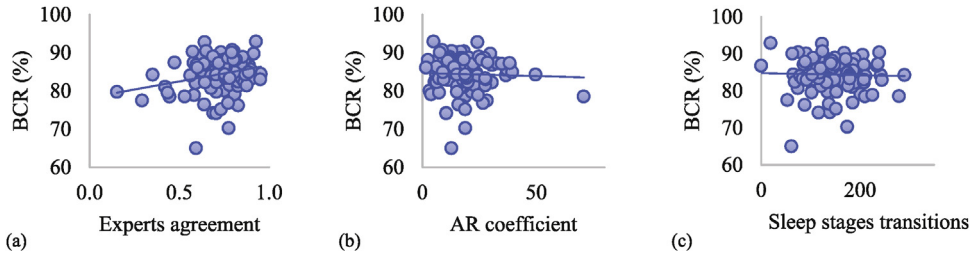

Fig. 5 – (a) Distribution of balanced classification rate (BCR) values concerning two experts agreement in multistage visual scoring; (b) distribution of BCR values corresponding to auto regressive (AR)-coefficient, in automatic multiclass sleep staging; (c) distribution of BCR values corresponding to sleep stage transitions.

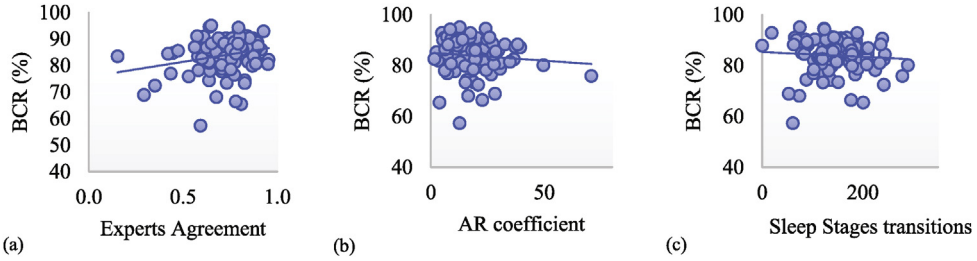

Fig. 6 – (a) Distribution of balanced classification rate (BCR) values concerning two experts agreement in detection N2; (b) distribution of BCR values corresponding to auto regressive (AR)-coefficient, in detection N2; (c) distribution of BCR values corresponding to the number of sleep stage transitions.

results. Furthermore, muscle activities, were extensively observed in subjects with higher misclassified N3.

- For REM sleep stage, the ambiguous EEG patterns are the most influencing reason of misclassification.

8. Conclusion

The ISRUC-Sleep dataset, which contains PSG recordings of different subjects, is introduced. This dataset was created aiming to complement existing datasets by providing easy-to-apply data collection with some characteristics not covered yet. In addition, a set of scripts was developed and turn publicly available allowing to test new algorithms and experiments replication. ISRUC-Sleep dataset is useful for research: (i) in biomedical signal processing; (ii) in development of new ASSC methods; and (iii) on sleep physiology studies.

Even though other publicly available datasets exist, to the best of our knowledge, except Sleep-EDF dataset (expanded), which was recorded during years 1987–1991 in two subsequent day-night periods at the subjects' homes, there is no traceable public dataset providing two recordings for the same subject at different time dates. On the other hand, comparing to the other datasets such as *MASS*, which mostly contains data of healthy subjects, the ISRUC-Sleep dataset includes data of healthy subjects, subjects with sleep disorders, and subjects under the effect of sleep medication. This variety of data can be useful for generalization purposes. In ISRUC-Sleep dataset, for each subject two hypnograms, created independently by two human-experts, are provided.

The details, benefits and characteristics of different subgroups of the dataset were illustrated by analyzing the performance of the *SSM4S* method. According to the results, there is a direct relation between the disagreement level of two experts and degradation of the classification performance. As expected, PSG recordings affected by artifacts and sleep related disorders, contain the most challenging patterns for sleep PSG analysis. Furthermore, some specific events and variations in usual neurophysiological patterns, such as variations in the alpha brain activities, are the main causes of performance drop in ASSC.

Acknowledgments

This work has been supported by the Portuguese Foundation for Science and Technology (FCT) under Ph.D. Grants SFRH/BD/81828/2011 and SFRH/BD/80735/2011 and by QREN funded project SLEEPTIGHT, with FEDER reference CENTRO-01-0202-FEDER-011530. The authors would also like to acknowledge sleep experts from the Sleep Medicine Centre of the Hospital of Coimbra University (CHUC), for their invaluable support in data acquisition and visually scoring the PSG recordings of ISRUC-Sleep dataset.

REFERENCES

[1] L. Zoubek, S. Charbonnier, S. Lesecq, A. Buguet, F. Chapotot, Feature selection for sleep/wake stages classification using data driven methods, Biomed. Signal Process. Control 2 (3) (2007) 171–179.

[2] T. Lajnef, S. Chaibi, P. Ruby, P.-E. Aguera, J.-B. Eichenlaub, M. Samet, A. Kachouri, K. Jerbi, Learning machines and sleeping brains: automatic sleep stage classification using

decision-tree multi-class support vector machines, J. Neurosci. Methods (2015).
[3] C. O'Reilly, T. Nielsen, Assessing EEG sleep spindle propagation. Part 1: Theory and proposed methodology, J. Neurosci. Methods 221 (2014) 202–214.
[4] V. Bajaj, R.B. Pachori, Automatic classification of sleep stages based on the time-frequency image of EEG signals, Comput. Methods Progr. Biomed. 112 (3) (2013) 320–328.
[5] L. Fraiwan, K. Lweesy, N. Khasawneh, H. Wenz, H. Dickhaus, Automated sleep stage identification system based on time-frequency analysis of a single EEG channel and random forest classifier, Comput. Methods Progr. Biomed., 2011.
[6] D. Alvarez-Estevez, V. Moret-Bonillo, Identification of electroencephalographic arousals in multichannel sleep recordings, IEEE Trans. Biomed. Eng. 58 (1) (2011) 54–63.
[7] S. Khalighi, T. Sousa, G. Pires, U. Nunes, Automatic sleep staging: a computer assisted approach for optimal combination of features and polysomnographic channels, Expert Syst. Appl. 40 (17) (2013) 7046–7059.
[8] T. Hori, Y. Sugita, E. Koga, S. Shirakawa, K. Inoue, S. Uchida, H. Kuwahara, M. Kousaka, T. Kobayashi, Y. Tsuji, M. Terashima, K. Fukuda, N. Fukuda, Sleep Computing Committee of the Japanese Society of Sleep Research, Proposed supplements and amendments to 'a manual of standardized terminology, techniques and scoring system for sleep stages of human subjects', the Rechtschaffen & Kales (1968) standard, Psychiatry Clin. Neurosci. 55 (3) (2001) 305–310.
[9] C. Iber, S. Ancoli-Israel, A. Chesson, S. Quan, The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications, 2007.
[10] M.H. Silber, S. Ancoli-Israel, M.H. Bonnet, S. Chokroverty, M.M. Grigg-Damberger, M. Hirshkowitz, S. Kapen, S.A. Keenan, M.H. Kryger, T. Penzel, M.R. Pressman, C. Iber, The visual scoring of sleep in adults, J. Clin. Sleep Med. 3 (2) (2007) 121–131.
[11] I.M. Colrain, Sleep and the brain, Neuropsychol. Rev. 21 (1) (2011) 1–4.
[12] S.L. Himanen, A. Joutsen, J. Virkkala, Visual assessment of selected high amplitude frontopolar slow waves of sleep: differences between healthy subjects and apnea patients, Clin. EEG Neurosci. 35 (3) (2004) 125–131.
[13] A.G. Xiromeritis, A.A. Hatziefthimiou, G.M. Hadjigeorgiou, K.I. Gourgoulianis, D.N. Anagnostopoulou, N.V. Angelopoulos, Quantitative spectral analysis of vigilance EEG in patients with obstructive sleep apnoea syndrome, Sleep Breath. 15 (1) (2011) 121–128.
[14] D.Z. Carvalho, G.J.L. Gerhardt, G. Dellagustin, E.L. de Santa-Helena, N. Lemke, A.Z. Segal, S.V. Schonwald, Loss of sleep spindle frequency deceleration in obstructive sleep apnea, Clin. Neurophysiol. 125 (2) (2014) 306–312.
[15] C. Stepnowsky, D. Levendowski, D. Popovic, I. Ayappa, D.M. Rapoport, Scoring accuracy of automated sleep staging from a bipolar electrooocular recording compared to manual scoring by multiple raters, Sleep Med. 14 (11) (2013) 1199–1207.
[16] C.J. Lauer, D. Riemann, M. Wiegand, M. Berger, From early to late adulthood – changes in EEG sleep of depressed-patients and healthy-volunteers, Biol. Psychiatry 29 (10) (1991) 979–993.
[17] R. Armitage, Sleep and circadian rhythms in mood disorders, Acta Psychiatr. Scand. 115 (2007) 104–115.
[18] M. Dresler, V.I. Spoormaker, P. Beitinger, M. Czisch, M. Kimura, A. Steiger, F. Holsboer, Neuroscience-driven discovery and development of sleep therapeutics, Pharmacol. Ther. 141 (3) (2014) 300–334.
[19] C.F. Reynolds, D.H. Shaw, T.F. Newton, P.A. Coble, D.J. Kupfer, EEG sleep in outpatients with generalized anxiety – a preliminary comparison with depressed outpatients, Psychiatry Res. 8 (2) (1983) 81–89.
[20] G.N. Papadimitriou, M. Kerkhofs, C. Kempenaers, J. Mendlewicz, EEG sleep studies in patients with generalized anxiety disorder, Psychiatry Res. 26 (2) (1988) 183–190.
[21] K. Spiegelhalder, W. Regen, B. Feige, J. Holz, H. Piosczyk, C. Baglioni, D. Riemann, C. Nissen, Increased EEG sigma and beta power during NREM sleep in primary insomnia, Biol. Psychol. 91 (3) (2012) 329–333.
[22] C.H. Bastien, M. LeBlanc, J. Carrier, C.M. Morin, Sleep EEG power spectra, insomnia, and chronic use of benzodiazepines, Sleep 26 (3) (2003) 313–317.
[23] G.A. Kerkhof, H.P.A. Van Dongen, Measurement of Sleep. Human Sleep and Cognition Part I: Basic Research, vol. 185, 2010, pp. 21–35.
[24] M.E. Thase, Depression, sleep, and antidepressants, J. Clin. Psychiatry 59 (Suppl. 4) (1998) 55–65.
[25] H.P. Landolt, E.B. Raimo, B.J. Schnierow, J.R. Kelsoe, M.H. Rapaport, J.C. Gillin, Sleep and sleep electroencephalogram in depressed patients treated with phenelzine, Arch. Gen. Psychiatry 58 (3) (2001) 268–276.
[26] A. Steiger, O. Benkert, F. Holsboer, Effects of long-term treatment with the MAO-A inhibitor moclobemide on sleep EEG and nocturnal hormonal secretion in normal men, Neuropsychobiology 30 (2-3) (1994) 101–105.
[27] U. Vonbardeleben, A. Steiger, A. Gerken, F. Holsboer, Effects of fluoxetine upon pharmacoendocrine and sleep-EEG parameters in normal controls, Int. Clin. Psychopharmacol. 4 (1989) 1–5.
[28] H.E. Kuenzel, H. Murck, K. Held, M. Ziegenbein, A. Steiger, Reboxetine induces similar sleep-EEG changes like SSRI's in patients with depression, Pharmacopsychiatry 37 (5) (2004) 193–195.
[29] M. Kluge, P. Schussler, A. Steiger, Duloxetine increases stage 3 sleep and suppresses rapid eye movement (REM) sleep in patients with major depression, Eur. Neuropsychopharmacol. 17 (8) (2007) 527–531.
[30] D.A. Schmid, A. Wichniak, M. Uhr, M. Ising, H. Brunner, K. Held, J.C. Weikel, A. Sonntag, A. Steiger, Changes of sleep architecture, spectral composition of sleep EEG, the nocturnal secretion of cortisol, acth, gh, prolactin, melatonin, ghrelin, and leptin, and the DEX-CRH test in depressed patients during treatment with mirtazapine, Neuropsychopharmacology 31 (4) (2006) 832–844.
[31] R. Luthringer, M. Toussaint, N. Schaltenbrand, P. Bailey, P.G. Danjou, D. Hackett, J.Y. Guichoux, J.P. Macher, A double-blind, placebo-controlled evaluation of the effects of orally administered venlafaxine on sleep in inpatients with major depression, Psychopharmacol. Bull. 32 (4) (1996) 637–646.
[32] G.M. Saletu-Zyhlarz, M.H. Abu-Bakr, P. Anderer, G. Gruber, M. Mandl, R. Strobl, D. Gollner, W. Prause, B. Saletu, Insomnia in depression: differences in objective and subjective sleep and awakening quality to normal controls and acute effects of trazodone, Prog. Neuropsychopharmacol. Biol. Psychiatry 26 (2) (2002) 249–260.
[33] M. Ferrara, L. De Gennaro, Going local: insights from EEG and stereo-EEG studies of the human sleep-wake cycle, Curr. Top. Med. Chem. 11 (19) (2011) 2423–2437.
[34] P. Anderer, G. Gruber, S. Parapatics, M. Woertz, T. Miazhynskaia, G. Klosch, B. Saletu, J. Zeitlhofer, M.J. Barbanoj, H. Danker-Hopfe, S.L. Himanen, B. Kemp, T. Penzel, M. Grozinger, D. Kunz, P. Rappelsberger, A. Schlogl, G. Dorffner, 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) (2005) 115–133.

[35] A.L. Goldberger, L.A.N. Amaral, L. Glass, J.M. Hausdorff, P.C. Ivanov, R.G. Mark, J.E. Mietus, G.B. Moody, C.K. Peng, H.E. Stanley, Physiobank, physiotoolkit, and physionet – components of a new research resource for complex physiologic signals, Circulation 101 (23) (2000) E215–E220.
[36] M. Adnane, Z.W. Jiang, Z.H. Yan, Sleep-wake stages classification and sleep efficiency estimation using single-lead electrocardiogram, Expert Syst. Appl. 39 (1) (2012) 1401–1413.
[37] N. Nicolaou, J. Georgiou, Towards automatic sleep staging via cross-recurrence rate of EEG and ECG activity, in: 2013 IEEE Biomedical Circuits and Systems Conference (Biocas), 2013, pp. 198–201.
[38] L. Fraiwan, K. Lweesy, N. Khasawneh, M. Fraiwan, H. Wenz, H. Dickhaus, Classification of sleep stages using multi-wavelet time frequency entropy and LDA, Methods Inf. Med. 49 (3) (2010) 230–237.
[39] M. Ronzhina, O. Janousek, J. Kolarova, M. Novakova, P. Honzik, I. Provaznik, Sleep scoring using artificial neural networks, Sleep Med. Rev. 16 (3) (2012) 251–263.
[40] B. Kemp, A.H. Zwinderman, B. Tuk, H.A.C. Kamphuisen, J.J.L. Oberye, Analysis of a sleep-dependent neuronal feedback loop: the slow-wave microcontinuity of the EEG, IEEE Trans. Biomed. Eng. 47 (9) (2000) 1185–1194.
[41] F. Yaghouby, S. Sunderam, Quasi-supervised scoring of human sleep in polysomnograms using augmented input variables, Comput. Biol. Med. (2015).
[42] M.G. Terzano, L. Parrino, A. Smerieri, R. Chervin, S. Chokroverty, C. Guilleminault, M. Hirschkowitz, M. Mahowald, H. Moldofsky, A. Rosa, R. Thomas, A. Walters, Atlas, rules and recording techniques for the scoring of cyclic alternating pattern (CAP) in human sleep (vol. 2, pg. 537, 2001), Sleep Med. 3 (2) (2002) 185.
[43] S.F. Quan, B.V. Howard, C. Iber, J.P. Kiley, F.J. Nieto, G.T. O'Connor, D.M. Rapoport, S. Redline, J. Robbins, J. Samet, et al., The sleep heart health study: design, rationale, and methods, Sleep 20 (1998) 1077–1085.
[44] F. Ebrahimi, S.-K. Setarehdan, J. Ayala-Moyeda, H. Nazeran, Automatic sleep staging using empirical mode decomposition, discrete wavelet transform, time-domain, and nonlinear dynamics features of heart rate variability signals, Comput. Methods Progr. Biomed. 112 (1) (2013) 47–57.
[45] C. O'Reilly, N. Gosselin, J. Carrier, T. Nielsen, Montreal archive of sleep studies: an open-access resource for instrument benchmarking and exploratory research, J. Sleep Res. 23 (6) (2014) 628–635.
[46] A. Tsanas, G. Clifford, Stage-independent, single lead EEG sleep spindle detection using the continuous wavelet transform and local weighted smoothing, Front. Hum. Neurosci. 9 (181) (2015).
[47] A. Krakovska, K. Mezeiova, Automatic sleep scoring: a search for an optimal combination of measures, Artif. Intell. Med. 53 (1) (2011) 25–33.
[48] H. Koch, J.A. Christensen, R. Frandsen, M. Zoetmulder, L. Arvastson, S.R. Christensen, P. Jennum, H.B. Sorensen, Automatic sleep classification using a data-driven topic model reveals latent sleep states, J. Neurosci. Methods 235 (2014) 130–137.
[49] A.M. Koupparis, V. Kokkinos, G.K. Kostopoulos, Semi-automatic sleep EEG scoring based on the hypnospectrogram, J. Neurosci. Methods 221 (2014) 189–195.
[50] A. Brignol, T. Al-Ani, X. Drouot, Phase space and power spectral approaches for EEG-based automatic sleep-wake classification in humans: a comparative study using short and standard epoch lengths, Comput. Methods Progr. Biomed. 109 (3) (2013) 227–238.
[51] J.M. Plonsey, Robert, Bioelectromagnetism – Principles and Applications of Bioelectric and Biomagnetic Fields, Oxford University Press, New York, 1995.
[52] H. Jasper, Appendix to report to committee on clinical examination in EEG: the tentwenty electrode system of the international federation, Electroencephalogr. Clin. Neurophysiol. (1958) 371–375.
[53] B. Kemp, J. Olivan, European data format 'plus' (EDF+), an EDF alike standard format for the exchange of physiological data, Clin. Neurophysiol. 114 (9) (2003) 1755–1761.
[54] S. Khalighi, T. Sousa, U. Nunes, Adaptive automatic sleep stage classification under covariate shift, in: 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) 2012, 2012, pp. 2259–2262.
[55] S. Khalighi, T. Sousa, D. Oliveira, G. Pires, U. Nunes, Efficient feature selection for sleep staging based on maximal overlap discrete wavelet transform and SVM, in: 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2011, pp. 3306–3309.
[56] T. Sousa, D. Oliveria, S. Khalighi, G. Pires, U. Nunes, Neurophysiologic and statistical analysis of failures in automatic sleep stage classification, in: International Conference on bio-inspired systems and signal processing (Biosignals 2012) (Vilamoura, Portugal), 2012.
[57] T. Sousa, A. Cruz, S. Khalighi, G. Pires, U. Nunes, A two-step automatic sleep stage classification method with dubious range detection, Comput. Biol. Med. (2015).
[58] B.D. Percival, T.A. Walden, Wavelet Methods for Time Series Analysis, Cambridge University Press, 2000.

Sirvan Khalighi received M.Sc. degree in artificial intelligence and robotics from Sharif University of Technology, Iran, in 2005. He was a lecturer and faculty member of electrical and computer engineering department in AZAD University of Qazvin, Iran during the years 2005–2010. He is currently a Ph.D. student in electrical and computer engineering department of university of Coimbra, Portugal. He is also a researcher in institute of systems and robotic (ISR). He has been involved with projects in the area of automatic sleep staging and brain computer interfaces. His research interests include biomedical signal processing, statistical pattern recognition, transfer learning, and domain adaptation.

Teresa Sousa received the M.Sc. degree on Biomedical Engineering from School of Engineering of the Minho University, Portugal, in 2010. She has research works in automatic detection of electroencephalographic seizures and in automatic sleep scoring. Her current research interests are related with Medical Electronics and Neuroengineering. She is currently a Ph.D. student in Coimbra University, with a project of automatic pattern recognition algorithms applied to brain signals for BCI and Neurofeedback.

José Moutinho dos Santos received the Medicine degree from the University of Coimbra, Portugal, in 1975, and is pneumology specialist from Centro Hospitalar de Coimbra, Portugal since 1985. He has been working in sleep field for the last 24 years and was Director of Centro de Medicina do Sono (Sleep Medicine Center) in the Centro Hospitalar Universitrio de Coimbra until June 2014. He has been supervisor or co-supervisor in academic dissertations for university or master degrees in sleep themes. He was President of the Portuguese Sleep Association, 2005–2007 and was appointed as member of the Committee created by the Portuguese Medical Association to define consensus criteria for admission in Sleep Medicine Competence. His research interests include clinical research in sleep apnea, sleepiness technologic investigation associated to sleep.

Urbano Nunes received the Ph.D. degree in electrical engineering from the University of Coimbra, Coimbra, Portugal, in 1995. Prof. Nunes is a Full Professor with the Electrical and Computer Engineering Department of Coimbra University. He is the coordinator of the Automation and Robotics for Human Life (AR4LIFE) Group of the Institute for Systems and Robotics (ISR-UC). He has been involved with/responsible for several funded projects at both national and international levels in the areas of human-centered mobile robotics, biomedical signal processing, brain computer interfaces and intelligent vehicles. Prof. Nunes is an Associate Editor of IEEE Transactions on Intelligent Transportation Systems. He was a General Chair of the 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems.