ELSEVIER

Computers in Biology and Medicine 37 (2007) 1600–1609

Computers in Biology and Medicine

www.intl.elsevierhealth.com/journals/cobm

Determining a continuous marker for sleep depth

Musa H. Asyali[a,*], Richard B. Berry[b], Michael C.K. Khoo[c], Ayse Altinok[d]

[a]*Departmemt of Computer Engineering, Yasar University, Kazim Dirik Mah. 364 Sok. No. 5, Bornova 35500, Izmir, Turkey*
[b]*Department of Medicine, University of Florida, Gainesville, FL 32611, USA*
[c]*Department of Biomedical Engineering, University of Southern California, Los Angeles, CA 90089, USA*
[d]*Department of Ophthalmology, Ulucanlar Eye Education and Research Hospital, Ankara 06100, Turkey*

Received 11 April 2006; received in revised form 20 February 2007; accepted 5 March 2007

Abstract

Detection and quantification of sleep arousals is an important issue, as the frequent arousals are known to reduce the quality of sleep and cause daytime sleepiness. In typical sleep staging, electroencephalograph (EEG) is the core signal and based on the visual inspection of the frequency content of EEG, non-rapid eye movement sleep is staged into four somewhat rough categories. In this study, we aimed at developing a continuous marker based on a more rigorous spectral analysis of EEG to measure or quantify the depth of sleep. In order to develop such a marker, we obtained the time–frequency map of two EEG channels around sleep arousals and identified the frequency bands that show the most change during arousals. We then evaluated classification performance of the potential signals for representing the depth of sleep, using receiver operating characteristic analysis. Our comparisons based on the area under the curve values revealed that the sum of absolute powers in alpha and beta bands is a good continuous marker to represent the depth of sleep. Higher values of this marker indicate low-quality sleep and vice versa. We believe that use of this marker will lead to a better quantification of sleep quality.

Keywords: Cortical arousals; Sleep staging; Depth of sleep; Power spectral density analysis; Receiver operating characteristics

1. Introduction

Sleep fluctuates cyclically between two fundamentally distinct neurophysiological states known as Non-Rapid Eye Movement (NREM) and REM sleep. Identification of sleep stages or equivalently determining depth of sleep is important in the diagnosis of sleep disorders. Further, sleep studies revealed that sleep deprivation, whether or not it is caused by sleep disorders, results in daytime sleepiness which in turn increases risk of injury and reduce job performance. Therefore, assessment of sleep quality is a key issue for many purposes.

In order to assess sleep according to standard criteria [1], it is sufficient to record electroencephalograph (EEG), electromyograph (EMG), and electrooculogram (EOG) signals. In monitoring sleep, at least one set of central EEG leads, placed according to international 10–20 systems of electrode placement [2], is used. The use of an additional set of leads in the occipital area is recommended because alpha waves (EEG activity in 7.5–12.0 Hz range) are best recorded there. The EOG leads are positioned near the eyes to measure potential differences across each eyeball (positive anterior and negative interior). Eye movements result in potential difference fluctuations that are detected in the EOG leads. Surface EMG leads, placed usually in chin area, are used to assess muscle tone. The EMG signal is especially useful in detecting REM sleep, which is identified by low skeletal muscle tone, hence a low amplitude EMG signal.

In sleep staging, NREM sleep is further divided into four categories or stages in a subjective manner, on the basis of visual inspection of sleep records. Stages 1 and 2 correspond to light sleep whereas stages 3 and 4 correspond to deep or slow-wave sleep. A given night of sleep is divided into usually 30-s-long periods called *epochs*. The predominant stage in a given epoch names that epoch. This fragmented description of depth of sleep may be suitable for some purposes but is insufficient for many other purposes. In fact, division of depth

* Corresponding author. Tel.: +90 232 461 4111x302; fax: +90 232 461 4121.
E-mail address: musa.asyali@yasar.edu.tr (M.H. Asyali).

0010-4825/$ - see front matter © 2007 Published by Elsevier Ltd.
doi:10.1016/j.compbiomed.2007.03.001

Table 1
Description of frequency content of EEG signal and its relation to sleep analysis

Frequency band	Range (Hz)	Interpretation
δ (delta)	1.0–4.0	High power in δ band corresponds to deep sleep
θ (theta)	4.0–7.5	Falling asleep (transition from stage 1 or 2 sleep to stage 3 or 4 sleep) is accompanied by the occurrence of θ waves
α (alpha)	7.5–12.0	Generally high power in α band corresponds to (but not so much as much as β) alertness or awake state
Sleep spindles and K complexes (SS)	12.0–16.0	These signals occur mostly during stage 2 sleep and are generally considered as artifacts in sleep analysis
β (beta)	16.0–25.0	High power in β band corresponds to alertness or awake state

of sleep into four rough categories is a practical decision rather than a physiological one, as it may be difficult to define and identify too many sleep stages based on visual inspection. As sleep arousals represent an unambiguous change in sleep state, we hypothesized that through spectral analysis of EEG signals during sleep arousals, we can come up with a marker that will represent the depth of sleep accurately.

The organization of the paper is follows. In the remainder of the introduction, we will describe EEG signal and its use in sleep analysis in some detail. We will also discuss sleep arousals and how they are identified in this section. In the methods section, we will discuss our experiment design and data processing techniques, including the spectral analysis of EEG. Next, we will present the results of EEG spectral analysis around sleep arousals and show how we utilize this information in determining some candidate power signals to represent sleep state changes. We will then discuss how we further evaluate the classification performance of those signals by calculating their arousal detection capability on whole night's data by using the receiver operating characteristics (ROCs) curve analysis.

1.1. Sleep and the EEG signal

The cerebral cortex generates electric potentials that can be led off the skin covering the cranium and recorded as the EEG. The fluctuations of the EEG potentials normally depend upon the degree of alertness and vary in amplitude and in frequency. Other than its use in sleep-analysis, as for example in the assessment of the quality of sleep, in diagnosis, monitoring and managing sleep-related disorders, EEG is also an important clinical tool in general, as for example in epilepsy (localized or generalized *convulsion waves*), in judging the degree of maturity of the brain, in monitoring anesthesia, and in the diagnosis of brain death [3,4].

The EEG typically has amplitudes from 10 to 100 μV and a frequency content from 1 to 40 Hz. Signals of 10–30 μV are considered as low amplitude and potentials of 80–100 μV are considered as high amplitude. An alert adult displays a low-amplitude EEG of mixed frequencies in the 16.0–25.0 Hz range (beta), while a relaxed adult produces large amounts of sinusoidal waves, in the 7.5–12.0 Hz range (alpha), which is particularly prominent at the back of the head. In many disease states, EEG activity tends to be either in the 1.0–4.0 Hz range (delta) or the 4.0–7.5 Hz range (theta). Sleep is composed of a periodic sequence of states during which the organism displays physiological characteristics radically different from wakefulness. These include both transient and long-term changes in brain activity, body movement, cardiac function, and respiration [5]. In the context of sleep analysis, we assume that the relevant frequency content of the EEG signal is in the 1.0–25.0 Hz range. In Table 1, we show the scheme that we adhere to for the separation of this range into frequency bands and summarize the use of corresponding signals in sleep analysis. There is not a universally accepted standard for the definition of these frequency bands. However, various schemes that appear in the literature differ only slightly, hence we may assume that our analysis will not be significantly influenced by the particular choice of the frequency band definitions.

Sleep fluctuates cyclically between two fundamentally distinct neurophysiological states referred to as Non-Rapid Eye Movement (NREM) and REM, or active sleep. NREM sleep is further divided into four stages, namely stages I, II, III, and IV, which are distinguished from each other principally on the basis of EEG. Stages 1 and 2 correspond to light sleep whereas stages 3 and 4 correspond to deep or slow-wave sleep. A given night of sleep is divided into usually 30-s-long periods called epochs. The predominant stage in a given epoch names that epoch. Staging of sleep is relevant for the study of sleep because each stage has a characteristic impact on respiration. Furthermore, disease processes frequently alter not only the total sleep time but also the relative amount of time spent in different stages of the sleep.

Normal sleep progresses from light to deeper stages and returns to light stages typically in the following order: $1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 3 \rightarrow 2 \rightarrow$ REM. These cycles, lasting 90–120 min, are present in the first portion of the night and may repeat two to four times over the night. The remaining of the night is spent either in stage 2 or in REM sleep, with the REM episodes tending to increase in length. Periods of wakefulness during the night may also be observed.

Wakefulness and REM sleep correspond to cortical activation whereas NREM sleep corresponds to quiescence. Wakefulness is distinguished from sleep by awareness of the environment, capacity to develop meaningful responses to external stimuli, and ability to perform complex, coordinated sensorimotor tasks. In NREM sleep, mental activity is minimal, if any, and sensorimotor responses to external stimuli are generally limited to simple reflexes, such as withdrawal from pain. In addition,

the threshold for eliciting response is increased compared to its wakefulness level. However, sleep differs from other states of diminished consciousness, such as anesthesia and coma, by the ability to awake quickly in response to sufficient stimulation.

1.2. Sleep arousals

Transient arousals during sleep can lead to disruption of the normal sleep and may cause excessive daytime sleepiness [6–8]. During NREM sleep, these brief arousals can be identified on the standard sleep recordings (polysomnogram) and are characterized by an abrupt increase in EEG frequency (suggestive of an awake state). Arousals can be scored from either the central or occipital derivation EEG [9].

According to Atlas Task Force of the American Sleep Disorders Association (ASDA) [9] an EEG arousal is an abrupt shift in EEG frequency, which may include theta, alpha and/or frequencies greater than 16.0 Hz but not spindles, subject to the following rules and conditions:

- Subjects must be asleep, defined as 10 continuous seconds or more of the indications of any stage of sleep, before an EEG arousal can be scored.
- A minimum of 10 continuous seconds intervening sleep is necessary to score a second arousal.
- The EEG frequency shift must be 3 s or greater in duration to be scored as an arousal.
- Arousals in NREM sleep may occur without concurrent increases in sub-mental electromyographic (EMG) amplitude.
- Arousals are scored in REM sleep only when accompanied by concurrent increases in submittal EMG amplitude.
- Artifacts, K complexes or delta waves are not scored as arousals unless accompanied by an EEG frequency shift (as previously defined) in at least one derivation. If such activity precedes an EEG frequency shift, it is not included in reaching 3-s duration criteria. When occurring within the EEG frequency shift, artifacts or delta wave activity are included in meeting the duration criteria.
- Non-concurrent, but contiguous, EEG and EMG changes, which were individually less than 3 s in duration, are not scored as arousals.
- Intrusion of alpha activity of less than 3-s duration into NREM sleep at a rate greater than one burst per 10 s is not scored as an EEG arousal. Three seconds of alpha sleep is not scored as an arousal unless a 10-s episode of alpha-free sleep precedes.
- Transitions from one stage to another are not sufficient of themselves to be scored as EEG arousals unless they meet the criteria indicated above.

From the above criteria, it is clear that the arousal scoring criteria are based on the EEG signal alone with only one exception: scoring of arousal during REM sleep requires the presence of a simultaneous increase in the EMG amplitude. The presence of bursts of alpha or theta activity in REM sleep EEG is a common phenomenon; however, not all of these events reflect physiological arousal from REM sleep. Therefore, to reliably score arousal from REM sleep, we need the additional requirement of EMG amplitude increase. The principle of defining an arousal as being 3 s or greater in duration is a methodological decision rather than a physiological one, as the identification of events of shorter duration may be difficult to achieve practically.

2. Methods

2.1. Experimental procedures

We recruited three normal/healthy and two obstructive sleep apnea (OSA) subjects and carried out sleep studies to determine a continuous marker that would accurately represent the depth of sleep. We conducted the experiments at the Sleep Laboratory of the Long Beach Veterans Administration Hospital, in Long Beach, California, USA. Written informed consent was obtained from all subjects prior to their participation in the study. Table 2 summarizes the profiles of our subjects. The last column in the table reports the number of arousal segments we analyzed from each subject. As sleep is fragmented by frequent arousals following airway obstructions in OSA subjects, we thought that the nature (i.e., changes in the EEG spectral content) of the sleep arousals could be different in these subjects compared to normal subjects. This is why we included the two OSA subjects in the study.

Along with the spontaneous sleep arousals, we wanted to have and analyze arousals of known time and cause. To this end, we used a brief acoustic signal to cause transient arousals in our sleep subjects. When the subject was in a stable sleep stage, we remotely turned on speakers placed in the subject's room driven by a signal generator and exposed the subject to a sound stimulus (beep signal) of 1000 Hz for 5 s. We electronically marked starting time of the acoustic stimulus and also stored this time marker information in our sleep data file. Throughout the night, we repeated this procedure 10–15 times depending on the sleeping condition of the subject. Early in the night, we experimented with the level (amplitude) of the stimulus and set it to a level which was high enough to briefly awake (or arouse) the subject. This level varied between 60 and 90 dB, depending on the subject.

During the experiments, we recorded the following physiological variables or signals: partial pressure of alveolar (end tidal) CO_2 (P_{CO_2}), arterial O_2 saturation (Sa_{O_2}), mouth pressure (P_{mask}), respiratory airflow, two EEG derivations, one central (C3-A2) and one occipital (O1-A2), chin EMG, EOG, ECG

Table 2
Subject profiles

Subject	Condition	Age	Weight (lbs.)	Number of arousal segments analyzed
1	Normal	35	185	55
2	Normal	27	180	132
3	Normal	41	118	98
4	OSA	52	190	70
5	OSA	45	152	110

(electrocardiograph), and respitraces (for measuring rib-cage and abdominal movements which indicate respiratory effort). To measure airflow and mask pressure, we had our subject wear a facemask, held in place by head straps. A pneumotacograph was connected to the facemask for measuring airflow. P_{CO_2} was sampled near the expiratory port of the mask by using a CO_2 analyzer that detected the CO_2 content in the exhaled gas. Sa_{O_2} level was measured by pulse oximetry using a finger probe.

We sampled and converted these continuous signals into discrete time by using a plug-in analog to digital converter board (Dataq Instruments, Akron, OH) and its associated data acquisition software (Windaq Software, Dataq Instruments). The sampling rate and conversion resolution were 100 samples per second per channel and 12 bit per sample, respectively. Two EEG, the EOG, and the EMG signals constitute the standard set-up [1,2] for monitoring and staging sleep.

In order to convert the measured signals into physical units, we made test or calibration recordings before each sleep study, by supplying two-level calibration signals of known physical units for each channel. We have calibrated the signals by simply multiplying them with their respective calibration coefficients, which are calculated as:

$$\text{Calibration coefficient} = \frac{\text{High physical value} - \text{low physical value}}{\text{High calibration reading} - \text{low calibration reading}}.$$

2.2. Spectral analysis of EEG and data processing

The analysis of EEG signal has been the subject of extensive research [10,11], as EEG is an important clinical tool in general, aside from its use in sleep studies. The EEG signal, like most biomedical signals, is highly non-stationary, i.e., its statistical characteristics change with time. In practice, the stationarity can artificially be obtained on a non-stationary signal by dividing the signal into short consecutive segments. These segments can assumed to be stationary and the well-known spectrum analysis techniques like the Fourier transform can be applied. This method, called short-time Fourier transform (STFT), was first proposed by Gabor [12]. The basic idea of STFT is that if we want to know what frequencies exist at a particular time, we take a small/short segment of the signal around that time and Fourier analyze it, neglecting the rest of the signal. The length of the time interval and the way we weigh the signal segment (for example, we may want to emphasize the central region) are decided by the choice of the window function. STFT (also known as spectrogram) is the prototype of all other time–frequency distributions and has been an extremely powerful tool in many areas. The main advantage of the STFT is that it is easy to obtain and interpret. The disadvantage of STFT is that when we want to get better resolution in time, we have to choose a shorter window and this causes poor frequency resolution. On the other hand, using a long analysis window will improve the frequency resolution, however, at the expense of reducing/loosing time resolution and compromising the assumption of stationarity of the signal within the window.

We have utilized autoregressive (AR) power spectral density (PSD) estimation to obtain time–frequency map of the EEG signal. In this technique, as in the case of STFT, we take small segments of the signal around the time of interest and assume stationarity over those segments. Instead of FFT, we use AR model fitting approach to estimate the spectrum of the signal segments. With this technique, we utilize the simple yet powerful idea behind the STFT and at the same time we overcome frequency and time resolution tradeoff associated with the STFT. With AR-PSD estimation, we eliminate the frequency resolution problem, because after fitting an AR model to signal segments we can practically compute the spectrum at any point/frequency we like. For a detailed discussion of AR model fitting-based PSD estimation, we refer the reader to the excellent references in the literature, such as [13–15].

There are various approaches for estimating the parameters of the AR model. In this study, we used *Yule–Walker Algorithm* which is intuitively simple and capable of producing good estimates when a reasonable number of samples is available. There are two variants of the Yule–Walker approach to estimate the AR model parameters, namely *autocorrelation* and *covariance* methods [16,17]. While estimating the autocorrelation of the underlying signal for different lags, the autocorrelation method assumes that data are zero padded from both ends whereas the covariance method assumes no zero padding and uses only available portion of the data. We have generated some random data (white Gaussian noise) and inputted these data to an AR system of known parameters and computed the response (output) of the system. We then estimated the AR model parameters from the output data using the two techniques and compared their performance by using "mean square error" between the correct and estimated AR model parameters. After extensive simulations of this sort, we concluded that the autocorrelation technique produces slightly better estimates of the model parameters compared to the covariance method. Furthermore, the autocorrelation method has some technical (computational) advantages over the covariance method. In autocorrelation method, the "autocorrelation matrix" that needs to be inverted to solve for the model parameters is Toeplitz and symmetric, and therefore positive semi-definite. This property of the autocorrelation matrix leads to more stable solutions, i.e., estimation error variances for the model parameters are smaller. Due to these considerations, we chose to implement the *autocorrelation technique* to estimate the power spectrum of EEG.

We followed the procedure outlined below to obtain the time–frequency map of EEG data segments around arousals.

1. Get a chunk or segment of EEG data of length 100 points (i.e., 1-s long at 100 samples per second sampling rate) around the time point of interest and detrend (i.e., remove linear trends from) this data segment by subtracting the best-fit line from it. The rationale behind detrending the data before spectral analysis is that DC and linear shifts in data are mostly due to artifacts such as movement rather than real physiological events.
2. Find an AR model for the 1-s long EEG data segment: start with a first-order AR model, estimate AR model parameters

and compute the residual error sequence. If residuals are *not white*, i.e., if there is a high correlation between the residuals and the data, increase AR model order and repeat this step until residuals are *white* or model order exceeds 15. We imposed such a limit on the model order, because if we keep increasing it without a bound, after a certain point we will be attempting to model even the inherent noise in the data, which is not desirable. (Besides, for a data of length 100 points, estimation of more than 15 model parameters will be numerically troublesome.) We tested the residual error sequence for *whiteness* by computing its autocorrelation and finding out the percentage of the autocorrelation lags which are outside the 95% confidence limits: $\pm 1.96\hat{\sigma}/\sqrt{N}$ [18]. Here, N is the length and $\hat{\sigma}$ is the standard deviation of the residual error sequence. If the percentage of the lags outside the bounds is $< 5\%$, we consider the sequence to be white.

3. Compute the PSD estimate for the 1-s long EEG data segment from its AR model. The PSD estimate is calculated at 250 frequency points from 0 to 25 Hz, that is, we set the frequency resolution as 0.1 Hz. By summing the values of the PSD estimate at corresponding frequencies, we compute the power in different frequency bands. For instance, to compute the delta power, we sum the values of the PSD estimate corresponding to the frequencies in 1.0–4.0 Hz region; for the total power, the summation goes from 1.0 to 25.0 Hz.
4. Go to the next analysis point in time, which is 100 points (1 s) away. This way, i.e., by taking 1-s long non-overlapping segments, we set our time resolution as 1 s.

We have carried out the EEG spectral analysis and other data processing operations using in house code that we developed under Matlab^{TM} (Mathworks Inc., Natick, Massachusetts).

3. Results

3.1. Changes in the EEG spectral content during arousals

We have looked at the changes in EEG frequency content for many spontaneous and acoustically induced sleep arousals from the same subject and observed that their characteristics are quite similar. We therefore did not differentiate the type of arousal in the analysis and processed spontaneous and acoustically induced sleep arousals all together. (The number of arousals that we analyzed from each subject is reported in Table 2.)

An expert sleep physician went through the whole night's sleep data and marked spontaneous arousals for three normal and two OSA subjects. Although the time of the acoustically induced arousals was already known, the physician also checked/confirmed whether there was really an arousal in agreement with the ASDA definition of sleep arousal [9], since sometimes, depending on the depth of sleep, the acoustic stimulus may not be strong enough to induce an arousal. We then focused on the shifts in the power levels of different EEG frequency bands around these arousal events in order to understand the nature of these abrupt sleep state changes. For instance, Fig. 1 shows a sample arousal event for Subject 2 on the two EEG channels. When we compute the powers in different frequency bands, we obtain the graphs given in Fig. 2. The ASDA criteria [9] suggest that the information in the SS band should be excluded in the arousal analysis, we therefore studied the changes in delta, theta, alpha, and beta bands only.

Although the sample spectral analysis that we presented in Fig. 2 was carried on 30-s long EEG segments, we used 13-s long EEG signal segments corresponding to arousal events in the rest of the analysis to assess the changes in the EEG spectral content during arousals. These arousal segments are formed by going backward 10 s and forward 3 s, in reference to the starting time of the arousals. The selection of pre- and post-arousal event time intervals as 10 and 3 s, respectively, is again based on the ASDA criteria [9].

In order to compare pre- and post-arousal power levels in different EEG frequency bands, we computed the average changes listed in Tables 3–6. The convention we used to define the *change in power levels* is post- minus pre-arousal power level; therefore, positive sign indicates an increase in power and vice versa. The values reported in the tables are in mean ± standard error of the mean (SEM) format. Tables 3 (EEG1) and 4 (EEG2) and Tables 5 (EEG1) and 6 (EEG2) report the changes in relative power (expressed as a percentage of total power) and absolute power levels, respectively. The numbers in parenthesis are the p-values obtained using Student's two tailed t-test indicating how significant the changes in the mean power levels are. We observe that most of the changes are significant ($p < 0.05$, a few exceptions are indicated by the * marks in the tables); however, we should look for consistency across different subjects and EEG channels in order to identify the best candidate signals for further analysis.

3.2. Obtaining ROC curves for different criteria

When we analyze the information in Tables 3–6, we note that there is a consistent significant decrease in relative theta and increases in relative beta, and absolute alpha and beta powers. By consistent we mean that the observations are valid across all subjects in both EEG channels. We therefore consider the corresponding four power signals (relative theta and beta, and absolute alpha and beta) for further analysis. In order to assess the discrimination or classification capability of these different candidate signals statistically, we have established their ROC curves.

The ROC curve is a graphical representation of the trade off between the false negative and false positive rates for every possible cutoff or threshold. By tradition, the plot shows the false positive rate (1−specificity) on the x-axis and the true positive rate (sensitivity or 1−the false negative rate) on the y-axis. In this context, *sensitivity* is the proportion of true positives (i.e., arousals as determined manually by the physician) that are correctly detected by the criterion and *specificity* is the proportion of true negatives (i.e., events that are correctly identified by the criterion as non-arousals). Therefore, in this analysis, the information contained in the potential criterion signals over the whole night (not only around arousals) is tested. The steps of

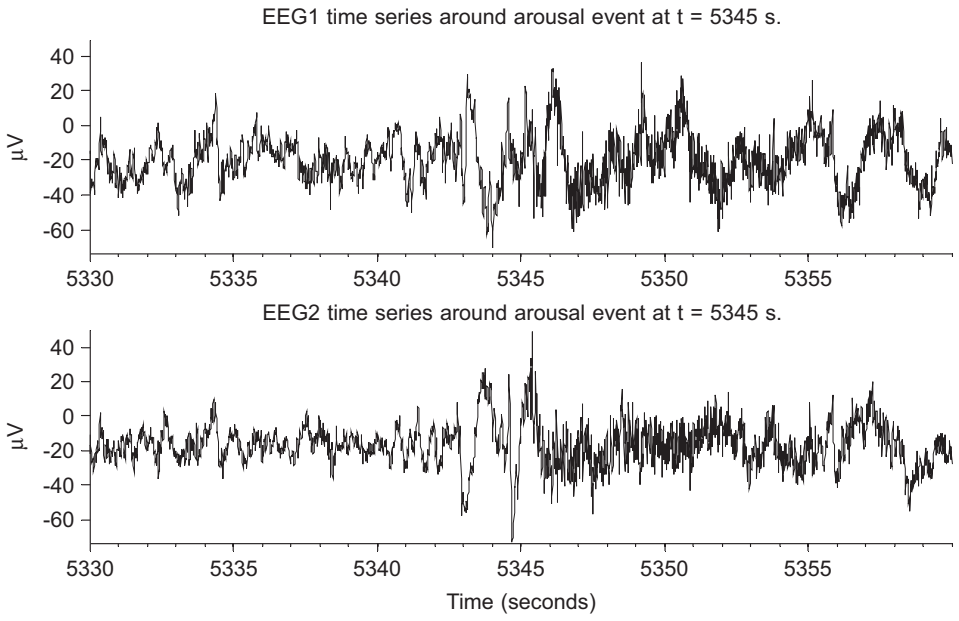

Fig. 1. EEG signals around a sample arousal event from Subject 2. Upper panel: EEG1 (central derivation). Lower panel: EEG2 (occipital derivation).

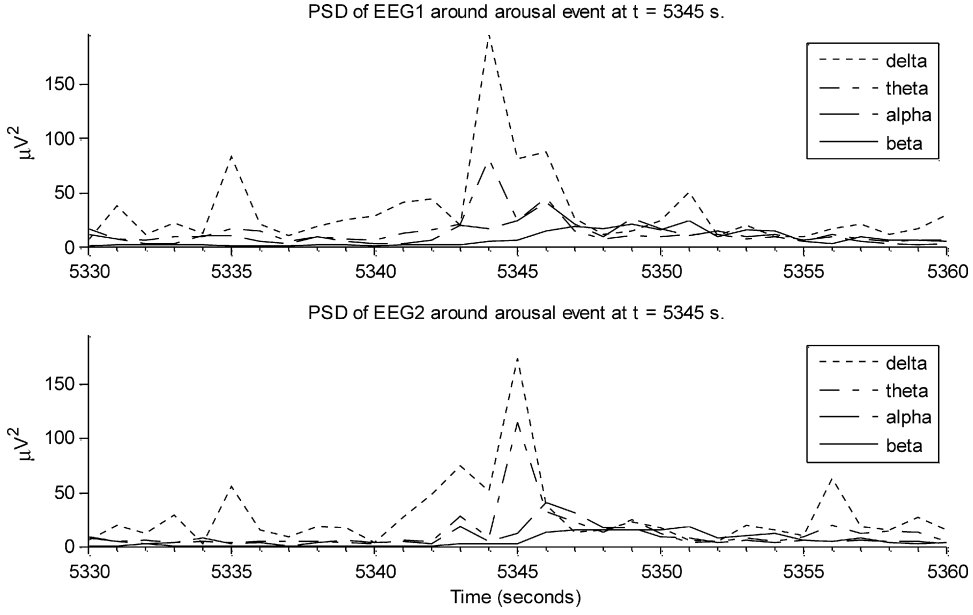

Fig. 2. Time variation of power in different frequency bands during the arousal event of Fig. 1. Upper panel: EEG1 (central derivation). Lower panel: EEG2 (occipital derivation). (PSD: power spectral density.)

the ROC analysis that we have repeated for each different criterion can be outlined as follows:

1. Select suitable thresholds (based on our spectral analysis results) and calculate the sensitivity and specificity for each data set.
2. Change the threshold by a certain amount, and repeat the calculation in step (1) to get new values of sensitivity and specificity.

We have used a total of 500 threshold values/stops and plot the ROC curve using the corresponding 500 combinations of sensitivity versus specificity. Fig. 3 shows two sample ROC curves (one for each EEG channel) obtained from Subject 2. Here, the criterion function is the absolute power in the beta band.

Instead of presenting many ROC curves obtained for different criterions, EEG channels, and data sets, we choose to report the *area under the ROC curve* value, which is simply known as

Table 3
Change (in percent) in relative power levels in different frequency bands from pre- to post-arousal averages for EEG1 (central derivation)

Subject	Frequency band			
	Delta	Theta	Alpha	Beta
1	4.123 ± 2.230 (=0.07)*	-7.599 ± 0.714 (<0.001)	-5.213 ± 1.163 (<0.001)	8.169 ± 1.039 (<0.001)
2	-6.923 ± 1.045 (<0.001)	-5.252 ± 0.613 (<0.001)	5.034 ± 0.715 (<0.001)	5.697 ± 0.495 (<0.001)
3	-11.038 ± 1.251 (<0.001)	-9.203 ± 0.610 (<0.001)	3.679 ± 0.654 (<0.001)	10.360 ± 0.947 (<0.001)
4	-14.508 ± 1.894 (<0.001)	-4.843 ± 0.754 (<0.001)	12.311 ± 1.561 (<0.001)	4.475 ± 0.844 (<0.001)
5	-4.175 ± 1.380 (<0.05)	-7.056 ± 0.545 (<0.001)	2.047 ± 0.679 (<0.05)	4.925 ± 0.858 (<0.001)

Table 4
Change (in percent) in relative power levels in different frequency bands from pre- to post-arousal averages for EEG2 (occipital derivation)

Subject	Frequency band			
	Delta	Theta	Alpha	Beta
1	-1.326 ± 1.998 (=0.51)*	-7.508 ± 0.639 (<0.001)	-4.336 ± 0.724 (<0.001)	11.961 ± 1.104 (<0.001)
2	-4.306 ± 1.260 (<0.001)	-2.285 ± 0.553 (<0.001)	1.177 ± 0.547 (<0.05)	4.489 ± 0.600 (<0.001)
3	-6.625 ± 1.400 (<0.001)	-6.339 ± 0.659 (<0.001)	4.600 ± 0.996 (<0.001)	3.495 ± 0.963 (<0.001)
4	-4.179 ± 1.296 (<0.05)	-10.067 ± 1.094 (<0.001)	9.565 ± 1.257 (<0.001)	3.419 ± 0.841 (<0.001)
5	-2.806 ± 1.366 (<0.05)	-10.946 ± 0.741 (<0.001)	2.929 ± 0.760 (<0.001)	5.195 ± 0.730 (<0.001)

Table 5
Change (in μV^2) in absolute power levels in different frequency bands from pre- to post-arousal averages for EEG1 (central derivation)

Subject	Frequency band			
	Delta	Theta	Alpha	Beta
1	81.137 ± 21.925 (<0.001)	16.669 ± 4.248 (<0.001)	10.649 ± 2.140 (<0.001)	17.559 ± 2.215 (<0.001)
2	79.718 ± 25.791 (<0.05)	21.821 ± 5.959 (<0.001)	30.963 ± 6.954 (<0.001)	37.047 ± 9.182 (<0.001)
3	-10.247 ± 11.653 (=0.38)*	-8.381 ± 2.480 (<0.05)	6.396 ± 1.301 (<0.001)	20.464 ± 5.266 (<0.001)
4	4.061 ± 3.661 (=0.27)*	6.715 ± 2.000 (<0.05)	15.431 ± 2.054 (<0.001)	14.117 ± 4.169 (<0.05)
5	68.047 ± 19.254 (<0.001)	13.401 ± 2.770 (<0.001)	19.738 ± 4.489 (<0.001)	27.746 ± 5.780 (<0.001)

Table 6
Change (in μV^2) in absolute power levels in different frequency bands from pre- to post-arousal averages for EEG2 (occipital derivation)

Subject	Frequency band			
	Delta	Theta	Alpha	Beta
1	115.316 ± 38.524 (<0.05)	30.121 ± 5.254 (<0.001)	18.807 ± 2.460 (<0.001)	46.449 ± 3.640 (<0.001)
2	53.846 ± 9.224 (<0.001)	22.496 ± 3.181 (<0.001)	30.100 ± 7.640 (<0.001)	36.430 ± 8.930 (<0.001)
3	19.167 ± 5.569 (<0.001)	4.817 ± 1.178 (<0.001)	8.078 ± 0.923 (<0.001)	10.331 ± 2.504 (<0.001)
4	13.017 ± 10.773 (=0.23)*	-3.513 ± 2.034 (=0.09)*	8.747 ± 1.320 (<0.001)	6.586 ± 1.482 (<0.001)
5	131.192 ± 40.233 (<0.05)	7.490 ± 3.321 (<0.05)	14.284 ± 1.630 (<0.001)	18.136 ± 2.812 (<0.001)

the AUC, for each case. The AUC practically summarizes the information in the ROC curve, i.e., the discriminative quality of the criterion, in one number. Since there is a decrease in relative theta power level during arousals (see Tables 3 and 4), we inverted this signal before we use it as the classification criterion.

In Tables 7 and 8, we present the AUC results obtained using the four different criterions on EEG1 and EEG2. These results indicate that the absolute alpha and beta power signals are the best performing criterions on average. Therefore, as the next logical step we consider a combination of these signals, i.e., the sum of absolute power in alpha and beta bands as the criterion. The AUC results obtained for this new criterion on both EEG channels are given in Table 9.

4. Discussion and conclusion

The currently established method of scoring sleep (i.e., determining sleep stages), is based merely on visual inspection of the frequency content of the EEG signal. Other than being very subject and scorer dependent, such a sleep state assessment procedure can only produce four discrete sleep states, stages

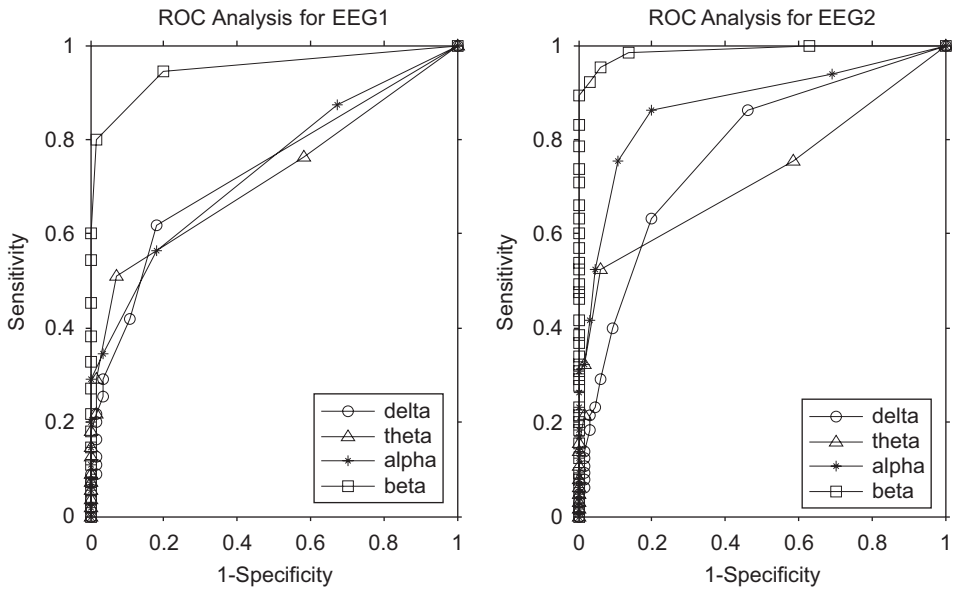

Fig. 3. Sample ROC curves for Subject 1 using absolute power in beta band obtained from EEG1 (left panel) and EEG2 (right panel). The AUC values are 0.9496 and 0.9891 for EEG1 and EEG2, respectively.

1–4. Such a categorization of sleep depth may be practical for physicians but it is not suitable for adequate description of the underlying physiological phenomenon. Further, for ease of processing, sleep data for the whole night's sleep are divided into 30-s long periods called epochs and the predominant stage in a given epoch names that epoch. For instance, if the physician or sleep technician identifies 20 s of Stage 3 sleep in an epoch, the epoch is marked as Stage 3 completely. If the numbers of different stage epochs are approximately equal on average, the error introduced by this rough quantification may cancel out. However, this assumption is not necessarily true. Additionally, in modeling studies where the influence of sleep depth on the respiratory control is investigated, a continuous signal representing level brain alertness is needed [19–21].

Therefore, we carefully looked at the frequency content of EEG to produce a marker that will quantify the sleep state or depth in a continuous and methodologically uniform manner. We concentrated on the changes of EEG spectrum that occur during sleep arousals, as they correspond to unambiguous changes in sleep depth. We developed a systematic approach based on ASDA definition of sleep arousal [9] and AR model fitting-based PSD estimation on short (1-s long) EEG segments. For each frequency band, we established pre- and post-arousal power levels based on 10-s pre-event and 3-s post-event time averages, respectively. We performed the analysis on central and occipital EEG derivations.

The information presented in Tables 3–6 helped us determine which power signals could be better representative of changes in sleep depth. Out of this analysis we have identified four candidate signals or criterions, namely relative theta and beta, and absolute alpha and beta power signals and used these as criterions in the ROC analysis. This compressive analysis (see Tables 7 and 8) revealed that the absolute powers in beta (top performer) and alpha (second performer) bands are the most effective criterions in terms of identifying sleep arousals. When we combined these two functions, we obtained the results shown in Table 9. The AUC results for the combined criterions (i.e., the sum of absolute powers in alpha and beta bands) indicate that there is a performance increase, with respect to the use of absolute alpha or beta alone, for both EEG channels.

These findings led us to propose the sum of absolute powers in alpha and beta bands as the marker for depth of sleep. Lower values of this marker correspond to deeper sleep and conversely higher values indicate lighter sleep or brain activation. Our proposed marker is very practical as it only requires one EEG channel to compute and is totally non-invasive. As we have used information obtained from both normal and OSA subjects while developing our marker, we can stipulate that it can be used on any subject to accurately describe the sleep or alertness status. We also observe that either of the two EEG channels we used in this study is suitable and enough to obtain our proposed marker. However, in order to address the issue of standardization and identify the best candidate channel for continuous quantification of sleep depth, further studies involving more than two EEG channels are needed.

Using our proposed sleep depth marker, carrying out an analysis to determine the whole night's sleep quality would be very practical and simple compared to tedious manual/visual analysis. All we need to do is to process one EEG channel to obtain its second-by-second spectral content and compute our proposed marker as the sum of absolute powers in alpha and beta bands. Then, many interesting variables can be produced from this continuous marker of sleep depth. For instance, the average of this marker over the whole night or certain periods could be used to assess the quality of sleep. The only critical point about our proposed marker is that the use of absolute powers requires firm calibration of EEG recordings, for accurate comparison across different subjects or recording platforms.

Table 7
AUC values obtained using the four different criterion functions on EEG1 (central derivation)

Subject	Classification criterion			
	Relative theta	Relative beta	Absolute alpha	Absolute beta
1	0.7048	0.7555	0.7367	0.9496
2	0.5948	0.6962	0.7825	0.8696
3	0.7402	0.7553	0.5805	0.8300
4	0.5582	0.6563	0.8252	0.7856
5	0.7135	0.5790	0.7520	0.8001
Mean ± SEM	0.6623 ± 0.0360	0.6885 ± 0.0332	0.7354 ± 0.0416	0.8470 ± 0.0294

Table 8
AUC values obtained using the four different criterion functions on EEG2 (central derivation)

Subject	Classification criterion			
	Relative theta	Relative beta	Absolute alpha	Absolute beta
1	0.7737	0.8208	0.8729	0.9891
2	0.4601	0.5723	0.8376	0.8527
3	0.5941	0.5408	0.5608	0.7073
4	0.7177	0.6284	0.6721	0.6024
5	0.7690	0.5954	0.7146	0.7782
Mean ± SEM	0.6629 ± 0.0602	0.6315 ± 0.0494	0.7316 ± 0.0567	0.7859 ± 0.0654

Table 9
AUC results when the classification criterion is the sum of absolute powers in alpha and beta bands for EEG1 and EEG2

Subject	Classification criterion: absolute alpha + beta	
	EEG1	EEG2
1	0.9250	0.9618
2	0.8749	0.9001
3	0.8301	0.7035
4	0.8459	0.6975
5	0.8321	0.8120
Mean ± SEM	0.8616 ± 0.0178	0.8150 ± 0.0525

References

[1] A. Rechtschaffen, A. Kales (Eds.), A Manual of Standardized Terminology, Techniques, and Scoring System for Sleep Stages of Human Sleep, Brain Information Service/Brain Research Institute, UCLA, Los Angeles, 1968.

[2] H.H. Jasper, The ten twenty electrode system of the International Federation, Electroencephalogr. Clin. Neurophysiol. 10 (1958) 371–375.

[3] A. Despopoulos, S. Silbernagl, Color Atlas of Physiology Thieme, Stuttgart, 1991, pp. 292–293.

[4] J.R. William, I. Karacan, C.J. Hursch, Electroencephalography (EEG) of Human Sleep: Clinical Applications, Wiley, New York, 1974.

[5] R.M. Harper, R.J. Sclabassi, T. Estrin, Time series analysis and sleep research, IEEE Trans. Autom. Control AC-19 (6) (1974) 932–943.

[6] P.C. Deegan, W.T. McNicholas, Pathophysiology of obstructive sleep apnea, Eur. Respir. J. 8 (1995) 1161–1178.

[7] M.H. Bonnet, Performance during sleep, in: W.B. Webb (Ed.), Biological Rhythms, Sleep and Performance, Wiley, New York, 1982, pp. 205–237.

[8] P.A. Cistulli, C.E. Sullivan, Pathophysiology of sleep apnea, in: N.A. Sunders, C.E. Sullivan (Eds.), Sleep and Breathing, second ed., Marcel Dekker, New York, 1994, pp. 405–448.

[9] Atlas Task Force of the American Sleep Disorders Association, EEG arousals: scoring rules and examples, Sleep 15(2) (1992) 174–184.

[10] J.S. Barlow, Methods of analysis of nonstationary EEGs, with emphasis on segmentation techniques: a comparative review, J. Clin. Neurophysiol. 2 (3) (1985) 267–304.

[11] A. Isaksson, A. Wennberg, L.H. Zetterberg, Computer analysis of EEG signals with parametric models, Proc. IEEE 69 (4) (1981) 451–461.

[12] L. Cohen, Introduction: a primer on time–frequency analysis, in: B. Boashash (Ed.), Time–Frequency Signal Analysis, Longman Cheshire, Melbourne, 1992, pp. 3–42.

[13] S.M. Kay, S.T. Marple, Spectrum estimation—a modern perspective, Proc. IEEE 69 (11) (1981) 1380–1419.

[14] H. Akaike, Power spectrum estimation through autoregressive model fitting, Ann. Inst. Stat. Math. 21 (1969) 407–419.

[15] J.G. Proakis, D.G. Manolakis, Digital Signal Processing: Principles, Algorithms and Applications, Prentice-Hall, Englewood Cliffs, NJ, 1996.

[16] R. Shiavi, An Introduction to Applied Statistical Signal Analysis, Aksen Associates, Homewood, 1991, pp. 341–386.

[17] P.M. Clarkson, Optimal and Adaptive Signal Processing, CRC Press, Boca Raton, 1993.

[18] G. Jenkins, D. Watts, Spectral Analysis and its Applications, Holden-Day, San Francisco, 1968.
[19] M.H. Asyali, M.C.K. Khoo, R.B. Berry, Assessment of closed-loop ventilatory stability in obstructive sleep apnea, IEEE Trans. Biomed. Eng. 49 (3) (2002) 1434–1441.
[20] M.C.K. Khoo, J.J. Shin, M.H. Asyali, T.S. Kim, R.B. Berry, Ventilatory dynamics of transient arousal in patients with obstructive sleep apnea, Respir. Physiol. 85 (4) (1998) 1434–1441.
[21] M.H. Asyali, Sleep and respiratory variability in obstructive sleep apnea: a modeling perspective, Ph.D. Dissertation, University of Southern California, Los Angeles, CA, 1998.

Musa Hakan Asyali was born in 1969, in Ankara, Turkey. He received his B.S. degree in Electrical & Electronics Engineering in 1990 from Bilkent University in Ankara, Turkey. Between 1990 and 1992 he was a research assistant at the same department. He received his Ph.D. degree in biomedical engineering from the University of Southern California in Los Angeles, USA, and joined the Electrical & Electronics Engineering Department at Ege University, Izmir, Turkey, as an assistant professor, in 1998. He became an associate professor at the same department in 2000. He assumed a research scientist position at the Biostatistics and Scientific Computing Department of King Faisal Specialist Hospital & Research Center in Riyadh, Saudi Arabia, in 2001. He became a full professor at the Computer Engineering Department at Yasar University, Izmir, Turkey, in 2005.
Prof. Dr. Asyali's research interests include mathematical modeling of physiological systems, biomedical signal processing, and bioinformatics.

Dr. Richard B. Berry, M.D., is a Professor of Medicine and a member of the Pulmonary, Critical Care, and Sleep Medicine Division at the University of Florida, Gainesville, Florida. Dr. Berry is the medical director of the Shands at UF Sleep Disorders Center and chief of the pulmonary section at the Malcom Randall VA Medical Center. He is a member of the Sleep Medicine Board Writing Committee of the American Board of Internal Medicine and President of the American Sleep Medicine Foundation. Dr. Berry is an Associate Editor of the Journal of Clinical Sleep Medicine and a member of the Board of Directors of the American Academy of Sleep Medicine. His research interests include sleep apnea, mechanisms of arousal from sleep, and upper airway muscle physiology.

Michael C.K. Khoo received the B.Sc.(Eng.) degree in Mechanical Engineering from Imperial College of Science and Technology, University of London, in 1976, and the M.S. and Ph.D. degrees in bioengineering from Harvard University, Cambridge, MA, in 1977 and 1981, respectively. From 1981 to 1983, he was a Research Associate at the V.A. Hospital, West Roxbury, MA, and the Brigham and Women's Hospital, Boston, MA. Since September 1983, he has been on the faculty of the University of Southern California, Los Angeles, where he is currently Professor and Chairman of Biomedical Engineering. He is also Co-Director of Education and Outreach in the Biomimetic Microelectronic Systems Engineering Research Center at USC. His current research interests include cardiorespiratory regulation and variability in sleep apnea, physiological modeling, and biomedical signal processing. Dr. Khoo is a Fellow of the Biomedical Engineering Society and the American Institute of Medical and Biological Engineering. He is a member of the American Physiological Society, IEEE, Sleep Research Society and the American Heart Association. He is also on the editorial board of the IEEE Press Book Series on Biomedical Engineering. He is the author of the biomedical engineering textbook: Physiological Control Systems: Analysis, Simulation and Estimation (Piscataway, NJ: Wiley–IEEE Press, 2000).

Ayse Altinok was born in Ankara, Turkey, in 1967. She studied medicine at Ankara University, School of Medicine, between 1984 and 1990. She completed her residency in ophthalmology at SSK Diskapi Hospital, in Ankara, in 1995. After one year of mandatory service at a state hospital in Bozhuyuk, Bilecik, she joined SSK Ulucanlar Eye Research and Education Hospital as an ophthalmology specialist, in 1996. She has been leading the electrophysiology and corneal surgery units at this hospital since 1996 and 1997 respectively. In 2005, she assumed the deputy chief of medical staff position at the same hospital. Dr. Altinok's research interests include corneal surgery and electrophysiology (EEG, EOG and ERG) especially visual evoked potentials.