Sleep Medicine 140 (2026) 108733

Contents lists available at ScienceDirect

Sleep Medicine

journal homepage: www.elsevier.com/locate/sleep

Development of a probability model for high-resolution drowsiness detection using electroencephalogram

Ahnaf Rashik Hassan [a,b,1], Muammar Kabir [a,b,1], Shumit Saha [a,b,d], Behrang Keshavarz [b,c], Azadeh Yadollahi [a,b,*]

[a] Institute of Biomaterials and Biomedical Engineering, University of Toronto, Toronto, ON, Canada
[b] KITE, Toronto Rehabilitation Institute, University Health Network, Toronto, ON, Canada
[c] Department of Psychology, Toronto Metropolitan University, Toronto, ON, Canada
[d] Department of Information Technology, College of Computing and Software Engineering, Kennesaw State University, Georgia, USA

ARTICLE INFO

Keywords:
Classification algorithm
Clustering validation
Drowsiness
Electroencephalogram
Sleep
Wakefulness

ABSTRACT

Purpose: Capturing the dynamics of sleep onset process is fundamental to sleep medicine and circadian neurobiology. Even though wakefulness/sleep transition is a gradual and continuous process, it has been considered instantaneous and scored subjectively at low resolution. Therefore, a model to capture the dynamics of wakefulness to sleep transition is needed. The purpose of this study is to develop an efficient, high-resolution, and reliable model to quantitatively capture the dynamics of wakefulness/sleep transition using electroencephalogram (EEG).

Methods: We collected EEG signals from 53 subjects during an overnight sleep study. We extracted relative power features from EEG to develop a new model that yields the likelihood of wakefulness for each of the 3-s EEG segments. Furthermore, using the model, we identified three clusters, namely wakefulness, drowsiness and sleep, and employed statistical analyses, cluster quality evaluation, and graphical analysis for validation.

Results: The proposed method successfully separated three distinct cases of alertness. The mean silhouette value on the test data was 0.74 and the mean Davies-Bouldin index value was 0.43, which indicated that the three discovered clusters were compact. Based on the silhouette values, the detection accuracy was 93.21 %. One-way repeated measures analysis of variance results suggested that the feature values were significantly different ($p < .0001$) among the three detected clusters.

Conclusion: The proposed method was able to detect short episodes of wakefulness, drowsiness, and sleep with high accuracy in overnight polysomnography data. This proof-of-concept study suggests potential future applications in drowsiness detection, pending validation in relevant contexts such as driving simulators and workplace environments.

1. Introduction

Modeling the sleep onset process is essential to characterize and detect disorders affecting the transition from awake to sleep, such as narcolepsy, insomnia, and sleep deprivation [1,2]. Similarly, the model can be helpful to study the transition from wakefulness to sleep in a more granular way to assess drowsiness and daytime sleepiness, which may occur due to sleep apnea or medications [3]. A high-resolution and efficient awake/sleep transition model to detect and potentially prevent drowsiness will contribute to workplace and road safety. However, scoring wakefulness and different stages of sleep has been traditionally performed by subjective inspection of physiological signals in discrete epochs of 30s [4]. Even though previous research found that the 30s epoch-based sleep scoring is too coarse to track sleep onset [2,5,6], only a few studies attempted to capture the dynamics of sleep onset quantitatively at a higher resolution [5,7]. Taken together, a high-resolution, quantitative, and data-driven model to track the sleep onset process is missing.

* Corresponding author. University Health Network-Toronto Rehabilitation Institute, Institute of Biomaterials & Biomedical Engineering, University of Toronto, Room 12-106, 550 University Ave., Toronto, ON, M5G 2A2, Canada.
E-mail address: Azadeh.Yadollahi@uhn.ca (A. Yadollahi).
[1] These authors contributed equally.

https://doi.org/10.1016/j.sleep.2025.108733
Received 1 April 2025; Received in revised form 26 November 2025; Accepted 17 December 2025
Available online 12 January 2026
1389-9457/

Existing works in the literature have proposed metrics and algorithms to estimate sleep depth [7], to classify sleep stages [8–10], and detect drowsiness [11–24]. However, the majority of algorithms for sleep stage classification and sleep depth estimation is based on 30s epochs, which has no physiological basis [2,6]. Furthermore, researchers have developed drowsiness detection algorithms for monitoring applications [24,25]. One of the main applications of drowsiness detection is during driving. Systems based on the driver's behavioral patterns such as eye-blink/-closure or head-nodding, neurophysiological signal measurements, or vehicle-based performance have been developed [11,13,16,26–28]. However, these systems either have a low resolution of more than 30 s to detect drowsiness [12,29–31], have a low sensitivity of <75 % [32], or are intrusive [33], affected by external conditions such as road geometry [24], driving patterns [24,33], and lighting [15,24,34]. For example, drowsiness detection algorithms that involve analyzing facial recordings or eye tracking measures rely on lighting conditions [34]. Similarly, vehicle-based parameters such as lane deviation are affected by external factors such as weather, traffic, road markings, and lighting conditions [15]. Additionally, some of the car manufacturers have developed customized fatigue monitoring systems, which can only be integrated into their specific vehicles [15,24]. Therefore, a high-resolution and efficient drowsiness detection algorithm has not been established.

A high-resolution and efficient sleep onset process model can also be applied for sleep depth estimation and monitoring of the depth of anesthesia. The former application is essential for monitoring and diagnosing disorders of sleep onset such as insomnia, narcolepsy, hypersomnia, and sleep deprivation [1,5], whereas the latter is instrumental in administering optimal levels of anesthetic doses during surgery and unraveling how various anesthetic drugs induce and maintain general anesthesia [35]. A high-resolution model that quantitatively captures the likelihood of alertness can also be employed for computerized sleep scoring and sleep quality assessment and automated arousal detection which is currently performed manually, making the process time-consuming, burdensome, and error-prone due to fatigue [36,37]. Therefore, there is a dire need to develop an efficient and high-resolution drowsiness detection model that captures awake-to-sleep transition across individuals.

Electroencephalography (EEG) measures have been used as the cardinal physiological signals for detecting drowsiness [7,16,18,21,28,38]. EEG records the electrical activity of the brain and is often considered the “gold standard” for detecting wakefulness and different sleep stages in sleep studies [39,40]. Changes in wakefulness affect the alpha (8–13 Hz), delta (1–4 Hz), and beta (13–30 Hz) frequency bands of the EEG [1,2,5], and micro-sleep episodes that can be as short as 1s are often categorized as drowsiness [6]. Furthermore, most of the EEG-based studies [38,41] in the literature perform non-REM 1 detection, even though it is a sleep stage, as opposed to being a transitional state between wakefulness and sleep. However, most of the EEG-based algorithms that aim to measure drowsiness use longer (30s-15 min) signal episodes for drowsiness detection [28,30,42–44] and are therefore inappropriate to detect short episodes of drowsiness. Only a handful of existing studies [7,16,45,46] attempt to perform drowsiness detection at a higher resolution (<30s) Wei et al. [45] extracted EEG features in the frequency domain and performed detection at 6s resolution. However, the algorithm was validated against vehicle parameters which tend to show road and subject-related variability. Peiris et al. [26] and Nguyen et al. [16] used short 2s windows for EEG processing, but their results showed an accuracy of less than 80 % to detect drowsiness. Lees et al. (2023) analyzed 2-s EEG epochs in train and non-professional drivers during monotonous driving tasks and found that increased self-reported fatigue and sleepiness correlated with reductions in theta, alpha, and beta EEG activity. In addition, Younes et al. [7] developed a 3s-based sleep depth estimation metric (“odd ratio product”), but their study mainly focused on accurate sleep scoring and arousal detection rather than detecting drowsiness. Overall, there is still a need for a high-resolution and accurate model for drowsiness detection.

The primary aim of this study was to use EEG data collected from an overnight sleep study to develop a high-resolution and highly accurate model for detecting the transition from wakefulness to sleep. Compared to wakefulness, data collected during sleep are generally less noisy, with fewer artifacts such as eye movements. In addition, sleep data provide a large dataset for the initial development of drowsiness detection models which can later be trained and modified to be applied to other situations such as drowsy driving detection. Based on prior studies [2,5,47], we hypothesized that the changes in the alpha, beta, and delta frequency bands of the EEG can be used to develop a high-resolution and high-performance wake/sleep transition model.

2. Method

Data from 53 participants (26 females, Age: 49.58 ± 16.18), who were referred to the sleep laboratory of the Toronto Rehabilitation Institute-University Health Network for an overnight study to detect suspected sleep apnea were used. The study protocol was approved by the Institution's Research Ethics Board. Written informed consent was provided by all participants prior to participation in the study.

2.1. Sleep study protocol

The overnight study was conducted as part of clinical evaluation for suspected obstructive sleep apnea, with research data collected under informed consent. Participants attended the sleep laboratory at Toronto Rehabilitation Institute-University Health Network for a single overnight polysomnography study. Upon arrival (typically between 8:00–9:00 p.m.), participants completed intake questionnaires and the PSG electrodes and sensors were applied by trained sleep technicians. Participants were then allowed to prepare for sleep, with recordings continuing throughout the night until the participant's natural wake time in the morning (typically 6:00–7:00 a.m.). Continuous monitoring was maintained throughout the night by sleep technologists.

2.2. Data recording

Sleep was assessed with a full attended overnight polysomnograhpy (PSG) using Embla® N7000/S4500 (Natus Medical Incorporated). Standard surface electrodes were applied to record EEG, electrocardiogram (ECG), and electromyogram. EEG measures were recorded from six channels – two frontal (F3/F4), two central (C3/C4), and two occipital (O1/O2). The six channels were referenced against the mastoid electrodes (M1 and M2). Respiratory rate and volume were monitored using chest and abdominal respiratory inductance plethysmography bands, airflow was measured by nasal pressure cannula, and arterial oxyhemoglobin saturation (SaO2) was recorded using pulse oximetry. Sleep stages and arousals were scored by an experienced sleep technician in accordance with standard rules published by the American Academy of Sleep Medicine (AASM) [4]. Respiratory arousals occur after a respiratory event and are typically 3s long during which the participants are awake and alert [48].

2.3. Analysis of the EEG signal

2.3.1. Preprocessing

The EEG signal from all the channels were sampled at 128 Hz. First, the EEG data were bandpass filtered using a Butterworth filter with 0.5–30 Hz cut-off frequencies to eliminate low-frequency movement noises while retaining the essential EEG frequency bands [5]. Second, the data were segmented using 3s window with no overlap such that at least 2 cycles of the lowest frequency band (i.e., delta: 1–4Hz) were retained. According to the AASM guidelines, awake and sleep stages are scored on 30s basis, with more than 50 % of the epoch consisting of either wakefulness or one of the sleep stages, respectively [5]. Since we

were trying to detect the drowsiness and wakefulness to sleep transition, all episodes which were rated as wakefulness and followed by non-rapid eye movement (non-REM) stage 1 (N1) of sleep were considered for model development and validation. For model validation only, we also used arousal segments (transient periods of wakefulness that occur during sleep) and deep sleep (non-REM 2 and 3) segments. Arousal segments are extreme cases of alertness, and the deep sleep segments are extreme cases of non-alertness. Therefore, we hypothesize that the model will yield high probability of wakefulness for arousal segments and low probability of wakefulness for deep sleep segments.

2.3.2. Feature extraction from EEG frequency bands

The spectrogram in Fig. 1 shows an example of the changes in the alpha and delta band powers at the sleep onset for a subject. Accordingly, we defined the relative power of alpha, delta, and beta bands as:

$$\textit{Relative power of a band} = \frac{\text{Average power of the band}}{\text{Average power from } 0.5 - 30 \text{ Hz}} \quad (1)$$

2.3.3. Development of drowsiness detection model

We developed an awake probability model that used sigmoid functions to determine the probability of wakefulness ($Pr_F(W)$) based on the relative power values of the alpha, delta, and beta frequency bands extracted from the 3s signal segments. The sigmoid functions used in the model for each of the features are depicted in Fig. 2a. Since higher relative power of alpha and beta implies a higher $Pr_F(W)$, the red sigmoid function was used to define changes in alpha and beta. On the other hand, the black sigmoid function was used to define changes in delta. The parameters *a* and *b* of the sigmoid function were defined such that the feature values less than *a* and larger than *b* had very high probability of being awake or sleep. Therefore, to choose *a* and *b*, we selected EEG data during arousals and deep sleep stages (non-REM stages 2 and 3) and divided them into training and testing data sets. To estimate parameter *a*, the relative power values of the deep sleep segments were sorted and the maximum of the lower 80 % of data was set as *a*. Similarly, to estimate the parameter *b*, the relative power values of the arousal segments were sorted and the minimum of the higher 80 % of data was set as b.

After determining *a* and *b,* the relative power values of alpha, beta, and delta bands from the awake and N1 episodes in the training dataset were fed to the sigmoid model to estimate P_α, P_β, and P_δ, respectively. Subsequently, the probability of wakefulness, Pr(W), for each 3s of data was defined as:

$$\Pr(W) = w_1 \times P_\alpha + w_2 \times P_\beta + w_3 \times P_\delta \quad (2)$$

where weights w_1, w_2, and w_3 were computed using out-of-bag (OOB) permuted predictor delta error method [49] (detailed in supplementary material).

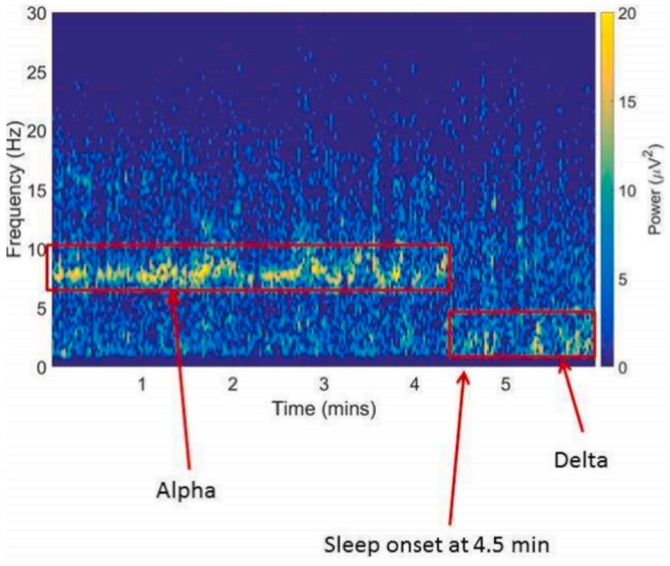

Fig. 1. Spectrogram (3s window, 50 % overlap) of the first few minutes of EEG recording at channel F4 of a single participant. The F4 electrode was referenced against the left mastoid (M1). Based on the scoring by sleep technicians, the transition from wakefulness to N1 occurs at 4.5 min.

2.3.4. Detection of drowsiness cluster

Since we employed overnight polysomnography data with expert sleep staging, our operational definition of drowsiness was based on EEG-derived sleep stages rather than independent behavioral measures. To differentiate wakefulness, drowsiness, and sleep episodes, the final step was to identify cut-off values of Pr(W) to determine the upper and lower bounds of the drowsiness cluster. Since we did not have independent behavioral gold standards (such as psychomotor vigilance test performance or subjective drowsiness ratings) to determine drowsy segments, we used cluster quality evaluation metrics, including Davies-Bouldin [50] and silhouette [51] indices to determine the upper and lower cut-offs of the drowsiness cluster. Davies-Bouldin index (DB) was defined as [50]:

$$DB = \frac{1}{n}\sum_{j=1}^{n}\sum_{i=1, i\neq j}^{n}\max\left(\frac{(\sigma_i + \sigma_j)}{d(c_i, c_j)}\right), \quad (3)$$

where n is the number of clusters, σ_i is the average distance of all points in cluster i to their cluster center c_i, σ_j is the average distance of all points in cluster j to their cluster center c_j, and $d(c_i, c_j)$ is the distance of cluster centers c_i and c_j. Small values of Davies-Bouldin index correspond to clusters that are compact, and whose centers are far away from each other.

In addition, for each data point k, we computed the silhouette value s (k) as [51]:

$$s(k) = \begin{cases} 1 - \dfrac{a(k)}{b(k)}, a(k) < b(k) \\ \dfrac{b(k)}{a(k)} - 1, a(k) \geq b(k) \end{cases}, \quad (4)$$

where a(k) is the average dissimilarity/distance of k to all other points within the same cluster, and b(k) is the lowest average dissimilarity of k to any other cluster, of which k is not a member. Silhouette value s(k) close to 1 suggests that the data point belongs to the proper cluster. On the other hand, silhouette value close to −1 suggests that the particular data point was assigned to the wrong cluster.

To determine the lower and upper bounds of the drowsiness cluster, the upper and lower cut-offs were changed from 1 % to 99 %, and the bounds that resulted in most different clusters based on Davies-Bouldin index and silhouette values were selected.

2.3.5. Validation and statistical analysis

For training and validation of the proposed method, we randomly selected half of the participants as training dataset (n = 26) and the other half as testing dataset (n = 27). Using episodes of the awake and N1 sleep stage from the testing dataset, we first visually inspected the relative power distributions of alpha, delta, and beta frequency bands for the identified awake, drowsiness, and sleep clusters. To quantitatively assess the clustering performance, we performed one-way repeated measures analysis of variance (ANOVA) to test if the three relative power feature values were significantly different among the three clusters. Significant main effects were followed up by post-hoc analysis (Tukey's multiple comparison tests). Furthermore, we computed the clustering quality evaluation metric values to validate the quality of the estimated clusters obtained from the awake and N1 sleep stage. Since there is no gold standard assessment of drowsiness for the

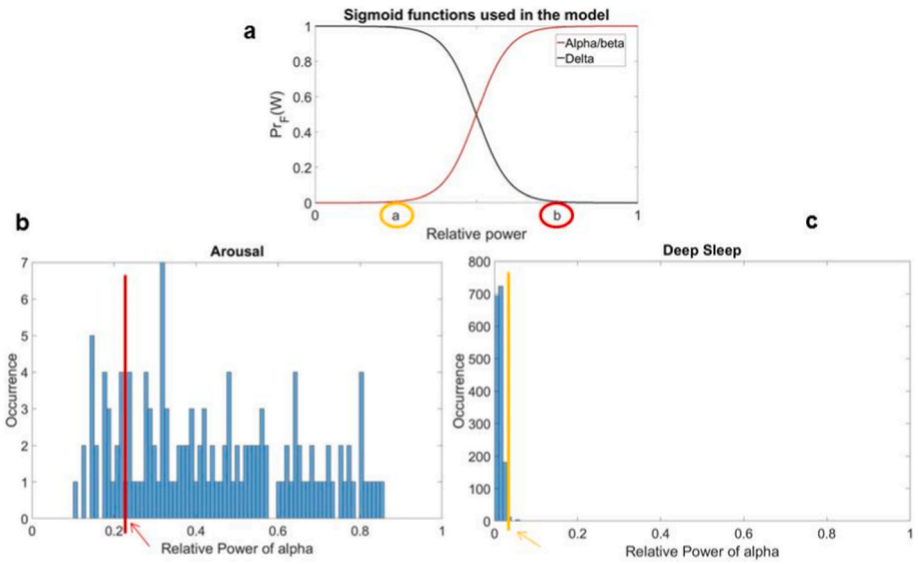

Fig. 2. (a) Sigmoid functions used in the proposed model. Probability of wakefulness for each feature ($Pr_F(W)$) should be high, if relative power values of alpha/beta are high. Therefore, the red curve is used to capture the changes in these two bands. The opposite scenarios are seen for delta band, which is why the black curve is used to compute $Pr_F(W)$ from delta band. Panel (b) Sigmoid parameter *b* is selected as the minimum feature value of the arousal distribution after removing outliers. Panel (c) Sigmoid parameter *a* is selected the maximum feature value of the deep sleep distribution after removing outliers. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

sleep data, we used the silhouette values to compute the average wakefulness and drowsiness detection accuracy. For every 3s segment, a positive silhouette value indicated the segment as being correctly classified while segments with negative silhouette values were considered as being misclassified. The data analyses were based on the EEG data obtained from the frontal F4-M1 electrodes, since wearable EEG headbands commonly use frontal EEG [52,53]. Results from other EEG electrodes were presented in supplementary S3.

3. Results

3.1. Demographic information and segment statistics

For this study, we evaluated the data of 53 patients (26 females, Age: 49.6 ± 16.2 years, BMI: 29.1 ± 6.2 kg/m^2). Table 1 summarizes the number of 3s segments that were used for model development and validation.

3.2. Parameter and weight selection

In order to select the sigmoid parameters *a* and *b*, and to validate the efficacy of the sigmoid awake probability model, 304 arousal segments and 1267 deep sleep segments were selected from all the participants.

Fig. 2a shows the distribution of the relative power of the alpha band for the arousal and deep sleep segments of the training data. Fig. 2b and c graphically illustrate the estimation of the sigmoid parameter *b* and *a* respectively ($b = 0.237$, Fig. 2b, marked by the red arrow; $a = 0.005$, Fig. 2c, marked by the yellow arrow). Table 2 shows the sigmoid parameter values obtained for all features.

Table 1
Average and standard deviation of the number of 3s segments from F4-M1 used in this study for model development and validation. Data are presented as mean ± standard deviation.

Stage	Number of Segments per Participant
Arousal	5 ± 1
Deep sleep	23 ± 14
N1	396 ± 112
Awake	1447 ± 337

Table 2
Sigmoid parameters computed from the training data for F4-M1.

Frequency band	*A*	*B*
Alpha	0.005	0.237
Beta	0.018	0.167
Delta	0.162	0.972

3.3. Cluster quality evaluation

For both training and test datasets, distribution of the estimated probabilities (Pr(W)) were significantly different for the arousal and deep sleep stages (Fig. 3a and b). In addition, the distributions could provide a rough estimate of the upper and lower bounds of the drowsiness cluster. For example, the segments with 28 %< Pr(W) < 65 % could belong to the drowsiness cluster. Therefore, the upper and lower boundaries of the clusters were validated by the Davies-Bouldin index (Eq (3)) and silhouette (Eq (4)) indices maps (Fig. 3c and d). Based on the training dataset, higher silhouette values were achieved if the lower cutoff of the drowsiness cluster was between 21 % and 30 % and the upper cutoff was between 54 % and 58 % (Fig. 3c). Similarly, in the test dataset, lower cutoff of 21 %–27 % and upper cutoff of 54 %–55 % resulted in the smaller Davies-Bouldin index (Fig. 3d). Therefore, the lower and upper cutoff of the drowsiness cluster were set to Pr(W) = 28 % and 55 %, respectively.

For sleep clusters, relative power values of alpha and beta bands were smaller (Fig. 4a and b), and for awake clusters, the relative power values of alpha and beta were higher. On the other hand, the relative delta power was higher during sleep than awake clusters (Fig. 4c). Furthermore, the mean silhouette value of the three clusters was 0.74 which is greater than the recommended value of 0.6 to consider the clusters separable [51]. Finally, the mean and maximum accuracies calculated using the silhouette values for detecting drowsiness were 93.21 % and 94.73 % respectively.

The relative power values of alpha, delta, and beta bands were significantly different among sleep, drowsiness, and awake clusters (Fig. 5). As expected, post hoc tests showed that the alpha and beta power values were significantly higher during wakefulness compared to drowsiness and sleep (Fig. 5a and b). Similarly, the delta power was

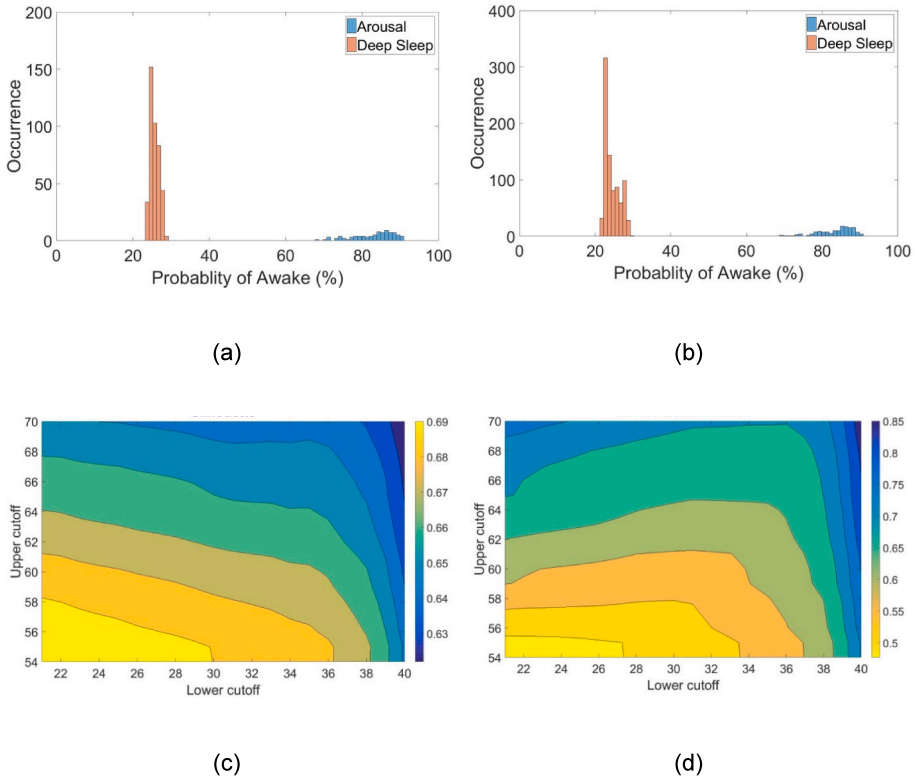

Fig. 3. Pr(W) distribution of arousal and deep sleep segments of (a) training data and (b) testing data obtained by weights and sigmoid parameters computed from the training data. The proposed model yields lower Pr(W) for deep sleep segments and higher Pr(W) for arousal segments in the testing data, which indicates its efficacy. (c) Silhouette and (d) Davies-Bouldin index values computed from the 3s segments of awake and N1 data obtained from the training participants. For both maps, the upper and lower bounds of the drowsy cluster have been varied and the two metrics were computed for the awake, sleep, and drowsy clusters.

significantly lower during wakefulness compared to drowsiness (0.15 ± 0.04 vs. 0.65 ± 0.02, $p < .0001$) and sleep (0.92 ± 0.03, $p < .0001$; Fig. 5c).

4. Discussions

The most important finding of our study was that, we achieved 94.7 % to detect drowsiness with a resolution of 3 s. Using features from three EEG frequency bands, the model calculated wakefulness likelihood (Pr (W)) for 3-s EEG segments. Results confirmed our hypothesis: arousal segments showed high Pr(W), while deep sleep segments showed low Pr (W), validating the power-based features. The thresholded wakefulness likelihood values formed three clusters—wakefulness, drowsiness, and sleep—with consistent, statistically significant power distributions (Figs. 3b, 4 and 5). Beyond drowsiness detection, the model has potential applications in sleep depth and anesthesia monitoring.

Our findings align with previous work demonstrating that EEG spectral features track behaviorally-relevant changes in alertness. Prerau et al.[47] showed that sleep onset involves parallel behavioral and physiological dynamics, with gradual changes in psychomotor vigilance task response times accompanying EEG spectral changes. Their work validates that EEG-based metrics can track behaviorally-measurable decrements in alertness, supporting our choice of spectral features (alpha, beta, and delta power). While our current study focuses on EEG-based classification validated against expert sleep staging, future work should incorporate behavioral performance measures (psychomotor vigilance tasks, subjective sleepiness scales, driving performance) to establish the relationship between our probability metric and functional impairment in alert-critical settings.

To the best of our knowledge, compared to the existing literature, our proposed algorithm had highest accuracy with a 3sec resolution to detect episodes of drowsiness only using EEG. For instance, previous studies using a combination of ocular features such as eye-closure and eye blink and EEG achieved very high accuracies of 92–99 %, but only with low resolutions of 1–5 min [18,42,54]. Similarly, combinations of EEG, EOG, and EMG signals showed high accuracies of 90–97 %, with 1–5 min resolutions [19,30,55,56]. ECG and questionnaire-based studies also reported high accuracies of 90–93 % [12,14,29,57], but these methods require at least 5 min of data to detect drowsiness and are not objective measures due to the nature of questionnaires. Furthermore, the studies that used sleep-EEG data to detect drowsiness had a resolution of 30s and accuracy values ranging from 83 to 97 % [10,28,58]. On the other hand, the few studies that have detected drowsiness with resolutions of less than 30s have only shown accuracy values ranging from 60 to 70 % [16,21,32].

Most of the available algorithms to detect drowsiness use multiple physiological signals [16,19,30] [15,28,32,59], which can be inconvenient for the user due to its intrusiveness. In this study, we have used a single frontal EEG electrode to detect drowsiness and have shown that similar results can be achieved for other EEG channels. Therefore, it may be concluded that the proposed method can be used to detect drowsiness using a single-channel EEG. Moreover, conventional clustering algorithms such as k-means or hierarchical clustering require hours of data to identify and separate groups or clusters of data [60]. In contrast, the sigmoid awake probability model, once trained, is capable of quantifying an arbitrary 3s episode as wakefulness, drowsiness or asleep – therefore making it a high-resolution algorithm. Furthermore, prior studies have shown that inter-rater disagreement for sleep scoring could be up to 20 % [2]. Also, a technician's scoring accuracy is subject to bias and error due to fatigue. Unlike the existing sleep study-based drowsiness detection algorithms, our proposed scheme, once trained, is independent of sleep technician's labels.

Despite the availability of various drowsiness detection measures, EEG has been the most widely used measure [56]. Previous studies

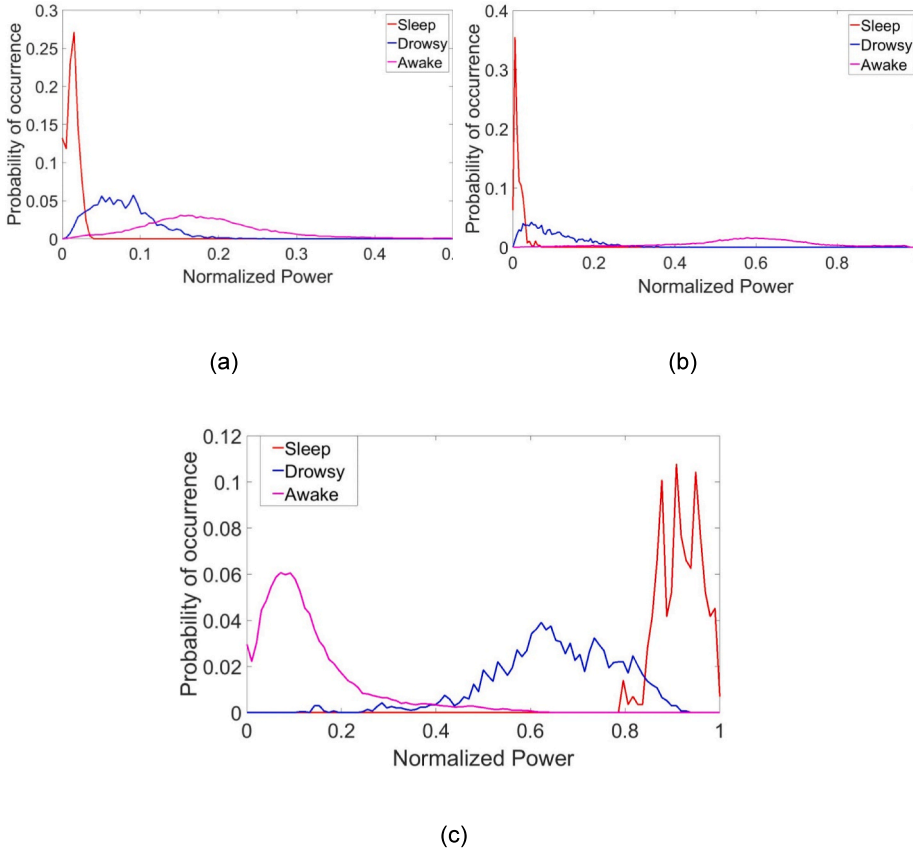

Fig. 4. Distributions of relative power of (a) alpha, (b) beta, and (c) delta for three clusters of awake, drowsy and sleep. Here, all episodes of N1 and wakefulness were considered. The sleep segments (Pr(W) < 28) have low alpha and beta power and high delta power, while awake segments (Pr(W) > 55) have high alpha and beta power and low delta power.

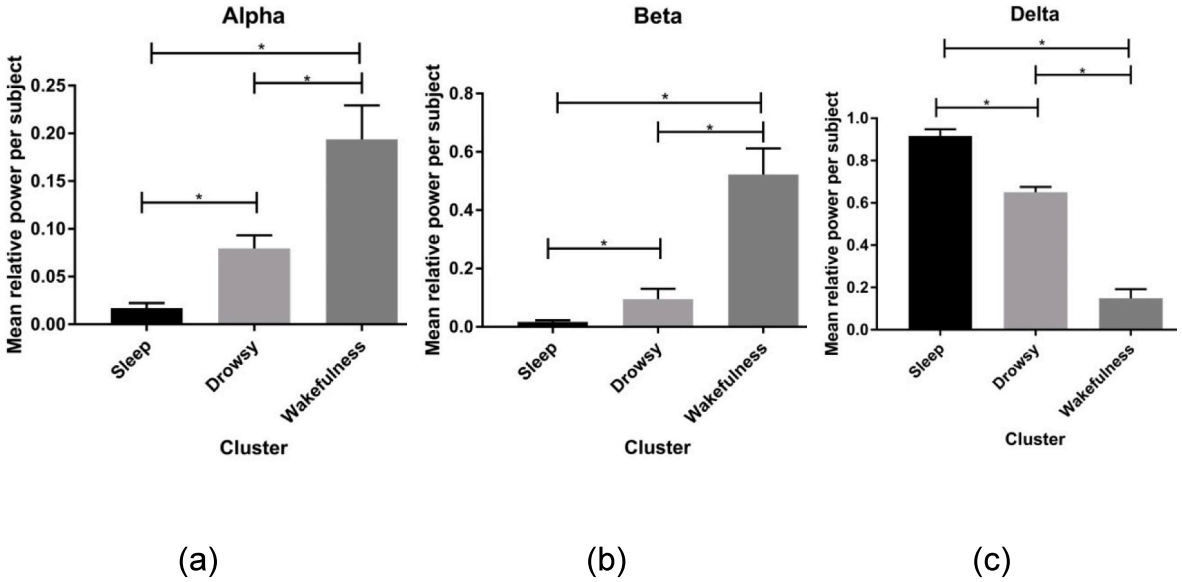

Fig. 5. Post hoc multiple comparison test suggests that (a) alpha, (b) beta, and (c) delta power features are significantly different between the clusters. Error bars indicate standard deviation. * indicates p < .0001.

found that the most important EEG features in detecting drowsiness are the powers of alpha and theta bands [28]. Furthermore, prior studies have shown that when an individual transitions from wakefulness to N1 sleep, the alpha and beta band powers decrease while the delta band power increases [1,5]. Accordingly, in this study, the relative powers of alpha, beta, and delta bands in the EEG were used to train our model in detecting drowsiness.

An important distinction between our study and recent work such as Lees et al. (2023) is that they validated their 2-s epoch EEG analysis against multiple independent behavioral measures during monotonous

driving tasks, finding negative correlations between EEG band activity and self-reported fatigue/sleepiness scores. This approach directly establishes the relationship between physiological markers and subjectively-experienced drowsiness states during active task performance. In contrast, the novelty of our proposed method is the use of the sigmoid function to model awake probability. The sigmoid function used in this study has been extensively used in neural networks and is often referred to as the special case of the logistic function [61]. It is non-linear in nature and gives an analog activation, unlike the step function. In between the sigmoid parameter values a and b, the output (prediction) changes significantly for a small change in input (feature values). Therefore, the selection of the sigmoid parameter values is very crucial. In this work, we utilized arousals as extreme cases of alertness and deep sleep segments as extreme cases of non-alertness to select the sigmoid parameters for our probability model. For a particular EEG frequency band, for example alpha, the parameters a (0.005) and b (0.237) were selected such that the probability of deep sleep and wakefulness below or above that value were the highest, respectively. The choices of the parameters were verified by the ability of the model to identify and separate the 3 clusters wakefulness, drowsiness, and sleep, as well as the statistically significant differences in the feature values among the clusters. Therefore, sigmoid function was able to capture the sleep onset dynamics empirically using the three relative power features computed from EEG.

Most of the existing works in the literature [16,30,32,42] do not randomize their training and testing data which might indicate that the reported performance metric values cannot be generalized. To overcome this limitation, in each iteration, we selected a percentage of the participants as training and the remainder as testing. This made sure that the results reported herein are robust, and the performance of the proposed model is generalizable. This data setting also has a benefit in terms of practical application. When implementing, the algorithm does not need to be trained but can be pre-trained on pre-existing data.

For model development, we also experimented with the theta (4–7 Hz) and gamma (30–100 Hz) oscillations, but their inclusion did not improve performance. We did not find consistent power changes in these bands at sleep onset. While some existing works [5,28] employ theta oscillations, our subjects were actively trying to fall asleep rather than stay alert, which may explain the lack of prominent theta band changes. Additionally, our 128 Hz sampling rate (64 Hz Nyquist frequency) did not capture the full gamma spectrum (30–100 Hz), limiting its utility as a feature.

This study has some limitations. Currently, there are no established guidelines on how to score drowsiness. Consequently, drowsy episodes were not scored by the technicians during the sleep study, but were instead defined as the period between wakefulness and N1 sleep. As a result, the exact start and end point of drowsy episodes remain unknown. In the future, a proposed model should be validated against video-based scales of drowsiness [62]. Furthermore, in active situations wherein the subject is trying to stay alert, the EEG data might be noisy due to movement and eye blink. Another limitation of the model is that it cannot detect drowsy episodes that are shorter than 3s. While the resolution is higher than most of the existing EEG-based works in the literature, lapses or microsleeps can be as short as 1s or even less [6]. The inclusion of features based on the delta band, which ranges from 1 to 4 Hz, necessitates the use of signal segments that are at least 2s or longer.

A significant limitation is that our drowsiness classification is defined and validated using EEG-based criteria (expert sleep staging) rather than independent behavioral or performance measures. The high accuracy we report (93.21 %) represents concordance with expert EEG-based sleep staging rather than prediction of independent behavioral outcomes. Future validation of drowsiness detection for safety-critical applications requires demonstration that our EEG-based probability metric predicts functionally-relevant outcomes such as impaired vigilance task performance, subjective sleepiness ratings, driving performance decrements, or real-world safety outcomes.

Our current study establishes proof-of-concept in a controlled clinical setting. However, validation in alert-critical environments is essential before clinical or commercial deployment. The overnight sleep study context provided advantages for initial model development, but wakefulness during daytime alert tasks differs from pre-sleep wakefulness. Future validation must occur in driving simulators, real-world scenarios, and populations with excessive daytime sleepiness, and the model may require recalibration for these contexts.

While we focused on frontal electrodes (F3-M2, F4-M1) to align with emerging wearable EEG devices, we did not systematically compare performance across different electrode montages. Additionally, approximately 10 % of the population exhibits low-amplitude or absent alpha rhythm, which may affect model performance and requires specific validation.

The practical implementation in real-world settings depends on wearable EEG technology advancement. Our frontal electrode focus was motivated by emerging wearable devices that can acquire frontal EEG in naturalistic settings. However, limitations must be addressed including signal quality in motion-heavy environments, user acceptance, power consumption, real-time processing capabilities, and regulatory approval. Successful translation will require co-development of algorithm and hardware solutions with validation in target environments.

5. Conclusion

In the present study, we have developed a single-channel EEG based high-resolution model to quantitatively capture the gradual transition from wakefulness to sleep using frontal EEG channels. The model demonstrated high accuracy in distinguishing wakefulness, drowsiness, and sleep states in 3-s epochs during overnight sleep studies. This approach offers a foundation for future drowsiness detection systems, though significant additional validation work is needed before deployment in applied settings.

Future research should focus on: (1) validating the model in alert-critical environments with concurrent behavioral measures, (2) testing generalizability across diverse populations including those with sleep disorders, (3) adapting the model for real-time wearable applications, and (4) investigating applications in clinical sleep medicine such as automated arousal detection and sleep depth monitoring. With appropriate validation, this high-resolution approach may eventually contribute to improved road safety, workplace accident prevention, and clinical sleep assessment.

CRediT authorship contribution statement

Ahnaf Rashik Hassan: Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis. **Muammar Kabir:** Writing – review & editing, Visualization, Validation, Methodology, Data curation. **Shumit Saha:** Writing – review & editing, Supervision. **Behrang Keshavarz:** Writing – review & editing, Writing – original draft, Supervision, Funding acquisition, Conceptualization. **Azadeh Yadollahi:** Writing – review & editing, Writing – original draft, Funding acquisition, Conceptualization.

Ethics approval

The study protocol was approved by the Institution's Research Ethics Board. Written informed consent was provided by all participants prior to participation in the study.

Funding

This work was supported by NSERC Discovery Grant RGPIN-2016-06549.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.sleep.2025.108733.

References

[1] Ogilvie RD. The process of falling asleep. Sleep Med Rev 2001;5(3):247–70. https://doi.org/10.1053/smrv.2001.0145.

[2] Prerau MJ, Brown RE, Bianchi MT, Ellenbogen JM, Purdon PL. Sleep neurophysiological dynamics through the lens of multitaper spectral analysis. Physiology 2017;32:60–92.

[3] Sforza E, Pichot V, Martin MS, Barthélémy JC, Roche F. Prevalence and determinants of subjective sleepiness in healthy elderly with unrecognized obstructive sleep apnea. Sleep Med 2015;16(8):981–6. https://doi.org/10.1016/j.sleep.2015.03.010.

[4] Silber MH, et al. The visual scoring of sleep in adults. J Clin Sleep Med 2007;3(2):121–31.

[5] Prerau MJ, et al. Tracking the sleep onset process: an empirical model of behavioral and physiological dynamics. PLoS Comput Biol Oct 2014;10(10):e1003866. https://doi.org/10.1371/journal.pcbi.1003866.

[6] Poudel GR, Innes CR, Bones PJ, Watts R, Jones RD. Losing the struggle to stay awake: divergent thalamic and cortical activity during microsleeps. Hum Brain Mapp 2014;35(1):257–69. https://doi.org/10.1002/hbm.22178.

[7] Younes M, et al. Odds ratio product of sleep EEG as a continuous measure of sleep state. Sleep 2015;38(4):641–54. https://doi.org/10.5665/sleep.4588.

[8] Dong H, Supratak A, Pan W, Wu C, Matthews PM, Guo Y. Mixed neural network approach for temporal sleep stage classification. IEEE Trans Neural Syst Rehabil Eng 2018;26(2):324–33. https://doi.org/10.1109/TNSRE.2017.2733220.

[9] Supratak A, Dong H, Wu C, Guo Y. DeepSleepNet: a model for automatic sleep stage scoring based on raw single-channel EEG. IEEE Trans Neural Syst Rehabil Eng 2017;25(11):1998–2008.

[10] Acharya UR, et al. Nonlinear dynamics measures for automated EEG-based sleep stage detection. Eur Neurol 2015;74(5–6):268–87.

[11] Wang Y-T, et al. Developing an EEG-based on-line closed-loop lapse detection and mitigation system. Front Neurosci 2014;8(321). https://doi.org/10.3389/fnins.2014.00321.

[12] Vicente J, Laguna P, Bartra A, Bailón R. Drowsiness detection using heart rate variability. Med Biol Eng Comput 2016;54(6):927–37. https://doi.org/10.1007/s11517-015-1448-7.

[13] Tripathy AK, Chinara S, Sarkar M. An application of wireless brain–computer interface for drowsiness detection. Biocybern Biomed Eng 2016;36(1):276–84. https://doi.org/10.1016/j.bbe.2015.08.001.

[14] Takahashi I, Takaishi T, Yokoyama K. Overcoming drowsiness by inducing cardiorespiratory phase synchronization. IEEE Trans Intell Transport Syst 2014;15(3):982–91. https://doi.org/10.1109/TITS.2013.2292115.

[15] Sahayadhas A, Sundaraj K, Murugappan M. Detecting driver drowsiness based on sensors: a review. Sensors (Basel) 2012;12(12):16937–53. https://doi.org/10.3390/s121216937.

[16] Nguyen T, Ahn S, Jang H, Jun SC, Kim JG. Utilization of a combined EEG/NIRS system to predict driver drowsiness. Sci Rep 2017;7:43933. https://doi.org/10.1038/srep43933.

[17] Liu CC, Hosking SG, Lenné MG. Predicting driver drowsiness using vehicle measures: recent insights and future challenges. J Saf Res 2009;40(4):239–45. https://doi.org/10.1016/j.jsr.2009.04.005.

[18] Lin FC, Ko LW, Chuang CH, Su TP, Lin CT. Generalized EEG-based drowsiness prediction system by using a self-organizing neural fuzzy system. IEEE Transactions on Circuits and Systems I: Regular Papers 2012;59(9):2044–55. https://doi.org/10.1109/TCSI.2012.2185290.

[19] Khushaba RN, Kodagoda S, Lal S, Dissanayake G. Driver drowsiness classification using fuzzy wavelet-packet-based feature-extraction algorithm. IEEE (Inst Electr Electron Eng) Trans Biomed Eng 2011;58(1):121–31. https://doi.org/10.1109/TBME.2010.2077291.

[20] Johnson RR, Popovic DP, Olmstead RE, Stikic M, Levendowski DJ, Berka C. Drowsiness/alertness algorithm development and validation using synchronized EEG and cognitive performance to individualize a generalized model. Biol Psychol 2011/05/01/2011;87(2):241–50. https://doi.org/10.1016/j.biopsycho.2011.03.003.

[21] Jagannathan SR, Ezquerro-Nassar A, Jachs B, Pustovaya OV, Bareham CA, Bekinschtein TA. Tracking wakefulness as it fades: Micro-measures of alertness. Neuroimage 2018;176:138–51. https://doi.org/10.1016/j.neuroimage.2018.04.046.

[22] Flores MJ, Armingol JM, de la Escalera A. Driver drowsiness detection system under infrared illumination for an intelligent vehicle. IET Intell Transp Syst 2011;5(4):241–51. https://doi.org/10.1049/iet-its.2009.0090.

[23] da Silveira TLT, Kozakevicius AJ, Rodrigues CR. Automated drowsiness detection through wavelet packet analysis of a single EEG channel. Expert Syst Appl 2016;55(Supplement C):559–65. https://doi.org/10.1016/j.eswa.2016.02.041.

[24] Chacon-Murguia MI, Prieto-Resendiz C. Detecting driver drowsiness: a survey of system designs and technology. IEEE Consumer Electronics Magazine 2015;4(4):107–19. https://doi.org/10.1109/MCE.2015.2463373.

[25] Aghaei AS, et al. Smart driver monitoring: when signal processing meets human factors: in the driver's seat. IEEE Signal Process Mag 2016;33(6):35–48. https://doi.org/10.1109/MSP.2016.2602379.

[26] Peiris MT, Davidson PR, Bones PJ, Jones RD. Detection of lapses in responsiveness from the EEG. J Neural Eng Feb 2011;8(1):016003. https://doi.org/10.1088/1741-2560/8/1/016003 (in eng).

[27] Lin CT, Chang CJ, Lin BS, Hung SH, Chao CF, Wang IJ. A real-time wireless brain-computer interface system for drowsiness detection. IEEE Transactions on Biomedical Circuits and Systems 2010;4(4):214–22. https://doi.org/10.1109/TBCAS.2010.2046415.

[28] Garces Correa A, Orosco L, Laciar E. Automatic detection of drowsiness in EEG records based on multimodal analysis. Med Eng Phys 2014;36(2):244–9. https://doi.org/10.1016/j.medengphy.2013.07.011.

[29] Vicente J, Laguna P, Bartra A, Bailón R. Detection of driver's drowsiness by means of HRV analysis. In: 2011 computing in cardiology; 2011. p. 89–92. Hangzhou.

[30] Akin M, Kurt MB, Sezgin N, Bayram M. Estimating vigilance level by using EEG and EMG signals. Neural Comput Appl 2008;17(3):227–36. https://doi.org/10.1007/s00521-007-0117-7.

[31] Subasi A. Automatic recognition of alertness level from EEG by using neural network and wavelet coefficients. Expert Syst Appl 2005/05/01/2005;28(4):701–11. https://doi.org/10.1016/j.eswa.2004.12.027.

[32] Malik TRP, Paul RD, Philip JB, Richard DJ. Detection of lapses in responsiveness from the EEG. J Neural Eng 2011;8(1):016003.

[33] Dong Y, Hu Z, Uchimura K, Murayama N. Driver inattention monitoring system for intelligent vehicles: a review. IEEE Trans Intell Transport Syst 2011;12(2):596–614. https://doi.org/10.1109/TITS.2010.2092770.

[34] Sommer D, Golz M. Evaluation of PERCLOS based current fatigue monitoring technologies. In: *2010 annual international conference of the IEEE engineering in medicine and biology*, Buenos Aires; 2010. p. 4456–9. https://doi.org/10.1109/IEMBS.2010.5625960 [Online]. Available: http://ieeexplore.ieee.org/document/5625960/.

[35] Brown EN, Lydic R, Schiff ND. General Anesthesia, sleep, and coma. N Engl J Med 2010/12/30 2010;363(27):2638–50. https://doi.org/10.1056/NEJMra0808281.

[36] Alvarez-Estevez D, Fernández-Varela I. Large-scale validation of an automatic EEG arousal detection algorithm using different heterogeneous databases. Sleep Med 2019/05/01/2019;57:6–14. https://doi.org/10.1016/j.sleep.2019.01.025.

[37] Li H, Guan Y. DeepSleep convolutional neural network allows accurate and fast detection of sleep arousal. Commun Biol 2021/01/04 2021;4(1):18. https://doi.org/10.1038/s42003-020-01542-8.

[38] Turkoglu M, Alcin OF, Aslan M, Al-Zebari A, Sengur A. Deep rhythm and long short term memory-based drowsiness detection. Biomed Signal Process Control 2021;65:102364. https://doi.org/10.1016/j.bspc.2020.102364. 2021/03/01.

[39] Vuckovic A, Radivojevic V, Chen ACN, Popovic D. Automatic recognition of alertness and drowsiness from EEG by an artificial neural network. Med Eng Phys 2002;24(5):349–60. https://doi.org/10.1016/S1350-4533(02)00030-9.

[40] Santamaria J, Chiappa KH. The EEG of drowsiness in normal adults 1987;4(4):327–82.

[41] Budak U, Bajaj V, Akbulut Y, Atila O, Sengur A. An effective hybrid model for EEG-based drowsiness detection. IEEE Sens J 2019;19(17):7624–31. https://doi.org/10.1109/JSEN.2019.2917850.

[42] Kurt MB, Sezgin N, Akin M, Kirbas G, Bayram M. The ANN-based computing of drowsy level. Expert Syst Appl 2009;36(2):2534–42. https://doi.org/10.1016/j.eswa.2008.01.085.

[43] Yeo MVM, Li X, Shen K, Wilder-Smith EPV. Can SVM be used for automatic EEG detection of drowsiness during car driving? Saf Sci 2009;47(1):115–24. https://doi.org/10.1016/j.ssci.2008.01.007.

[44] Shen M, Zou B, Li X, Zheng Y, Li L, Zhang L. Multi-source signal alignment and efficient multi-dimensional feature classification in the application of EEG-based subject-independent drowsiness detection. Biomed Signal Process Control 2021;70:103023. https://doi.org/10.1016/j.bspc.2021.103023. 2021/09/01.

[45] Wei C-S, Lin Y-P, Wang Y-T, Lin C-T, Jung T-P. A subject-transfer framework for obviating inter- and intra-subject variability in EEG-based drowsiness detection. Neuroimage 2018/07/01/2018;174:407–19. https://doi.org/10.1016/j.neuroimage.2018.03.032.

[46] Peiris MTR, Davidson PR, Bones PJ, Jones RD. Detection of lapses in responsiveness from the EEG. J Neural Eng 2011;8(1):016003. https://doi.org/10.1088/1741-2560/8/1/016003.

[47] Makeig S, Jung TP. Changes in alertness are a principal component of variance in the EEG spectrum. Neuroreport Dec 29 1995;7(1):213–6 (in eng).

[48] Azarbarzin A, Ostrowski M, Hanly P, Younes M. Relationship between arousal intensity and heart rate response to arousal. Sleep 2014;37(4):645–53. https://doi.org/10.5665/sleep.3560.

[49] Breiman L. Random forests. Mach Learn 2001;45:5–32.

[50] Davies DL, Bouldin DW. A cluster separation measure. IEEE Trans Pattern Anal Mach Intell 1979;1(2):224–7. https://doi.org/10.1109/TPAMI.1979.4766909.

[51] Rousseeuw PJ. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J Comput Appl Math 1987;20(Supplement C):53–65. https://doi.org/10.1016/0377-0427(87)90125-7.

[52] Lin C, et al. Wireless and wearable EEG system for evaluating driver vigilance. IEEE Transactions on Biomedical Circuits and Systems 2014;8(2):165–76. https://doi.org/10.1109/TBCAS.2014.2316224.
[53] Zhou DW, et al. Clustering analysis to identify distinct spectral components of encephalogram burst suppression in critically ill patients. In: 2015 37th annual international conference of the IEEE engineering in medicine and biology Society (EMBC); Aug. 2015 2015. p. 7258–61. https://doi.org/10.1109/EMBC.2015.7320067. Milan, 25-29, http://ieeexplore.ieee.org/document/7320067/?reload=true.
[54] Su H, Zheng G. A partial least squares regression-based fusion model for predicting the trend in drowsiness. IEEE Trans Syst Man Cybern Syst Hum 2008;38(5):1085–92. https://doi.org/10.1109/TSMCA.2008.2001067.
[55] Sommer D, Golz M, Trutschel U, Edwards D. Biosignal based discrimination between slight and strong driver hypovigilance by support-vector machines. In: Filipe J, Fred A, Sharp B, editors. Agents and Artificial Intelligence: international Conference, ICAART 2009, Porto, Portugal, January 19-21, 2009. Revised Selected Papers. Berlin, Heidelberg: Springer Berlin Heidelberg; 2010. p. 177–87.
[56] Papadelis C, et al. Monitoring sleepiness with on-board electrophysiological recordings for preventing sleep-deprived traffic accidents. Clin Neurophysiol 2007;118(9):1906–22. https://doi.org/10.1016/j.clinph.2007.04.031.
[57] Guaita M, et al. Regularity of cardiac rhythm as a marker of sleepiness in sleep disordered breathing. PLoS One 2015;10(4):e0122645. https://doi.org/10.1371/journal.pone.0122645.
[58] Asyali MH, Berry RB, Khoo MCK, Altinok A. Determining a continuous marker for sleep depth. Comput Biol Med 2007;37(11):1600–9. https://doi.org/10.1016/j.compbiomed.2007.03.001.
[59] Hu S, Zheng G. Driver drowsiness detection with eyelid related parameters by Support Vector Machine. Expert Syst Appl 2009;36(4):7651–8. https://doi.org/10.1016/j.eswa.2008.09.030.
[60] Bishop C. Pattern recognition and machine learning. New York: Springer-Verlag; 2006.
[61] Gibbs MN, Mackay DJC. Variational Gaussian process classifiers. IEEE Trans Neural Network 2000;11(6):1458–64. https://doi.org/10.1109/72.883477.
[62] Wierwille WW, Ellsworth LA. Evaluation of driver drowsiness by trained raters. Accid Anal Prev 1994;26(5):571–81. https://doi.org/10.1016/0001-4575(94)90019-1.