



Article

Unsupervised Multitaper Spectral Method for Identifying REM Sleep in Intracranial EEG Recordings Lacking EOG/EMG Data

Kyle Q. Lepage ^{1,*†}, Sparsh Jain ^{2,†}, Andrew Kvavilashvili ¹, Mark Witcher ³ and Sujith Vijayan ¹

¹ School of Neuroscience, Sandy Hall, Virginia Tech, 210 Drillfield Drive, Blacksburg, VA 24060, USA; akvavi4@vt.edu (A.K.); neuron99@vt.edu (S.V.)

² Department of Biomedical Engineering and Mechanics, Virginia Tech, 325 Stanger St., Blacksburg, VA 24061, USA; sparsh@vt.edu

³ Section of Neurosurgery, Carilion Clinic, Carilion Roanoke Memorial Hospital, 1906 Belleview Ave SE, Roanoke, VA 24014, USA; mrwitcher@carilionclinic.org

* Correspondence: kql@vt.edu

† These authors contributed equally to this work.

Abstract: A large number of human intracranial EEG (iEEG) recordings have been collected for clinical purposes, in institutions all over the world, but the vast majority of these are unaccompanied by EOG and EMG recordings which are required to separate Wake episodes from REM sleep using accepted methods. In order to make full use of this extremely valuable data, an accurate method of classifying sleep from iEEG recordings alone is required. Existing methods of sleep scoring using only iEEG recordings accurately classify all stages of sleep, with the exception that wake (W) and rapid-eye movement (REM) sleep are not well distinguished. A novel multitaper (Wake vs. REM) alpha-rhythm classifier is developed by generalizing K-means clustering for use with multitaper spectral eigencoefficients. The performance of this unsupervised method is assessed on eight subjects exhibiting normal sleep architecture in a hold-out analysis and is compared against a classical power detector. The proposed multitaper classifier correctly identifies 36 ± 6 min of REM in one night of recorded sleep, while incorrectly labeling less than 10% of all labeled 30 s epochs for all but one subject (human rater reliability is estimated to be near 80%), and outperforms the equivalent statistical-power classical test. Hold-out analysis indicates that when using one night's worth of data, an accurate generalization of the method on new data is likely. For the purpose of studying sleep, the introduced multitaper alpha-rhythm classifier further paves the way to making available a large quantity of otherwise unusable IEEG data.



Citation: Lepage, K.Q.; Jain, S.; Kvavilashvili, A.; Witcher, M.; Vijayan, S. Unsupervised Multitaper Spectral Method for Identifying REM Sleep in Intracranial EEG Recordings Lacking EOG/EMG Data. *Bioengineering* **2023**, *10*, 1009. <https://doi.org/10.3390/bioengineering10091009>

Academic Editor: Yunfeng Wu

Received: 28 June 2023

Revised: 10 August 2023

Accepted: 15 August 2023

Published: 25 August 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Intracranial electroencephalography (iEEG) recordings are routinely collected during the surgical treatment of epilepsy, while simultaneous electromyography (EMG) recordings and electrooculography (EOG) recordings are not [1,2]. Since the typical hospital stay of an epileptic patient undergoing invasive monitoring is 1 to 2 weeks, nearly every epilepsy unit in the US has iEEG sleep data that they have collected, and that they continue to collect, which is unaccompanied by electromyography (EMG) and electrooculography (EOG). These latter two recording modalities measure muscle tone and eye movements, respectively, and are required by expert human sleep scorers to separate wake from rapid eye movement (REM) sleep [3]. Reliable identification of REM in this data would rescue a large quantity of rare data collected during invasive surgery from living human brains. This data could be used to, for example, (i) assess the clinical use of REM dynamics to accurately identify the epileptic zone [4–6]; such identification is required to successfully treat focal epilepsy seizures surgically [7–9], and (ii) to further our understanding of brain dynamics during REM sleep (e.g., [10]).

Scientists have characterized sleep into four stages, namely: non-REM stage 1 (NREM1), non-REM stage 2 (NREM2), non-REM stage 3 (NREM3), and REM sleep. These stages are based, in part, upon electrophysiological features observed in recordings such as EEG and iEEG. Of all the sleep stages, REM sleep exhibits electrophysiological features that are most similar to those exhibited in recordings collected from awake subjects. As a consequence, the separation of REM from the awake state (hereafter labeled “Wake”), is the most difficult classification task when accurately labeling electrophysiological recordings (e.g., see [11,12]). This labeling process is sometimes referred to as “sleep scoring” by sleep scientists, and individuals sleep scoring data are colloquially known as “sleep scorers”. Since REM sleep is accompanied by low muscle tone and rapid eye movements, expert sleep scorers rely upon accompanying recordings of muscle activity (electromyography, EMG) and upon recordings that indicate eye movement (electrooculography, EOG) to separate Wake from REM sleep [13–17]. As previously stated, these recordings are absent from the majority of iEEG data, and a method capable of identifying REM based upon iEEG data alone will be of significant value.

Sleep scoring is a laborious and time-consuming process. As a consequence, many methods of automatically sleep scoring electrophysiological data exist; examples include refs. [18–21]. However, they have not been wholly successful, particularly when separating REM from Wake [11,12]. This limitation has not been alleviated with the advent of modern machine learning and artificial intelligence (AI) methods. In a review of 37 published deep learning sleep classifiers spanning the ten years from 2010 to 2020, it is recommended that EOG and EMG be analyzed in conjunction with EEG to achieve robust classification results [22]. More recently, classical machine learning methods have been compared against deep learning methods such as those reported in [23–25] for the purpose of scoring sleep and are found to perform similarly [26,27]. It is worth noting, however, that the performance reported in [24] on EEG is excellent, possibly due to the use of time-frequency distributions, possibly due to the use of a convolutional neural network, or possibly due to both.

Regardless, it is important to note that iEEG is not EEG, and it exhibits key differences. First, iEEG is significantly more invasive than EEG, with multi-contact electrodes penetrating through the cortex of the brain to clinically-indicated deep structures such as the hippocampus and the amygdala. With EEG, the electrodes lie upon the scalp with consistent, and more evenly spaced, coverage from individual to individual. Second, iEEG typically lacks occipital recordings. These two differences are substantial, and any sleep-scoring method that classifies sleep well using EEG recordings may not classify sleep well when applied to iEEG sleep data. For intracranial EEG, there is one published method of scoring sleep that does not use commensurate EMG and EOG recordings [28]. This method of sleep scoring also performs least well when identifying REM and N1 sleep stages (see Figure 3, [28]).

Our method uses alpha oscillations recorded intracranially and unaccompanied by EOG or EMG recordings to identify REM sleep episodes. Alpha (8–12 Hz) oscillations occur during eyes-closed wakefulness in humans but dissipate once the subject falls asleep [29–32]. This pattern of alpha activity has been used for the purpose of sleep scoring EEG data by human sleep scorers, especially to distinguish awake periods from periods of rapid eye movement (REM) sleep [4,33]. The alpha oscillations described above, as detected using standard EEG, are most prominent over the occipital cortex. However, intracranial patients rarely have electrodes in the occipital regions; instead, intracranial electrodes are commonly located in the temporal and frontal regions of the brain. Figure 1 displays an example of the difference in alpha oscillations between Wake and REM sleep exhibited in iEEG.

Unlike the other previously mentioned iEEG sleep scoring method (i.e., [28]), the proposed Wake vs. REM classifier is unsupervised (i.e., there is no training), its goal is to identify REM, it is restricted to a single 4 Hz frequency interval (i.e., alpha rhythm), and it uses K-means clustering modified to cluster multiple-channel, multitaper eigencoefficients.

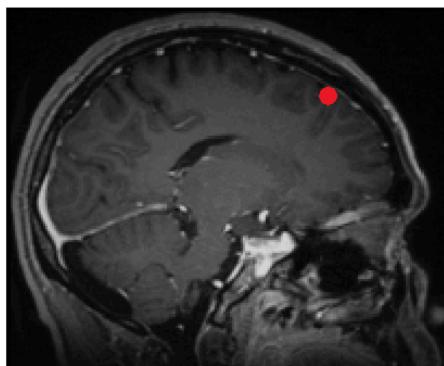
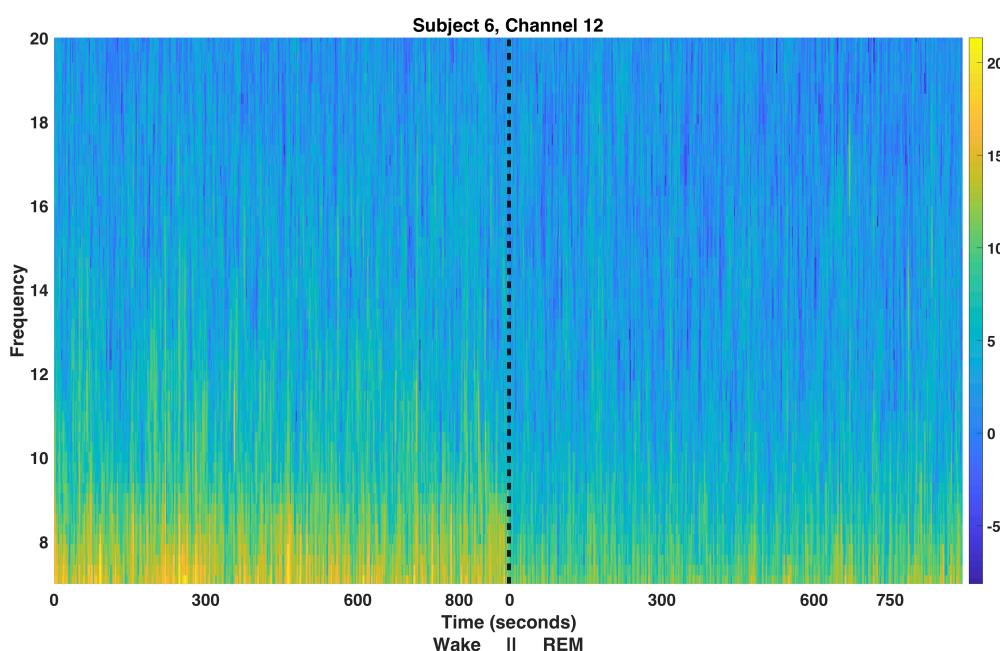
A.**B.**

Figure 1. Electrode localization and spectrogram. (A) A sagittal MRI image showing the location of the outermost cortical contact of a frontal electrode in the left hemisphere for subject 6; the contact is located in the left superior frontal gyrus. (B) Spectrogram showing the alpha activity pattern detected from this electrode contact during Wake and REM sleep. The black line indicates the beginning of the REM sleep episode. Note that not shown is a large lapse in time—the REM episode occurred much later in the night than the wake period. The frequency (vertical axis) is in units of Hz. The multitaper spectrogram is computed using the Chronux MATLAB software, version 2.10 [34] using a 3 second duration sliding window with 50% overlap, a time-bandwidth parameter of 3 and 5 DPSS tapers.

Similar to the sleep scoring method presented in ref. [28], the classifier proposed here requires only one night's worth of data. Applying the method introduced in ref. [28] to our data yields hypnograms that agree less with those identified by human sleep scorers than those depicted in Figure 4 in ref. [28]. This is the case despite trying all the software options provided in ref. [35] and selecting the analyzed channels using three different methods. Specifically, (i) selecting all available channels (this is the default “BEA” option in ref. [35]), (ii) selecting only the subset of iEEG channels used to compute the results in this paper, or (iii) selecting those channels that exhibit fewer than one ictal discharge per minute using the recommended ictal event detector [36]. These last two methods use the “EEA” option

specified in ref. [35]. Since we have not attained the performance reported in ref. [28], we leave further work directly comparing our method with those introduced in ref. [28] for a future study.

Numerous automatic EEG sleep scoring methods are reported in the literature [37,38]. A typical sleep scoring classifier involves feature selection, often through the use of clustering, followed by sleep stage classifier training. K-means clustering of EEG frequency domain features is reported in ref. [39–42]. In refs. [39,40], K-means clustering of reduced EEG spectral features is used to obtain feature weights to use with a K-nearest neighbor classifier and a decision tree classifier for the purpose of scoring sleep data. In ref. [42], K-means clustering is applied to the spectrogram of the Cz EEG channel for semi-automated sleep-scoring. In ref. [41], feature vectors containing EEG power spectra evaluated between the frequencies of 0.5 and 20 Hz are clustered to demonstrate subject-specific differences. Our method differs from these methods in that (i) its purpose is to identify REM from iEEG recordings, (ii) it is focused upon a single interval of frequencies, (iii) it is completely unsupervised and does not train a classifier, nor (iv) does it reduce the number of features in a feature-reduction step, (v) it uses aspects of the multitaper method of spectral time-series analysis to construct multi-channel alpha rhythm feature vectors, and (vi) the presented K-means clustering algorithm is updated from the original K-means algorithm to make use of the large-sample distribution of the resulting multitaper, multi-channel, alpha rhythm features.

2. Methods

2.1. Data Collection

Data were collected from nine patients (6 females and 3 males) between the age of 19 and 55 years (mean 33.11 ± 13.99 (SD) years) undergoing invasive monitoring using intracranial depth electrodes for intractable epilepsy (see Table 1). The study from which data was collected was approved by the Carilion Clinic IRB, IRB-20-1179, 8 February 2021, and was conducted in accordance with the Declaration of Helsinki. Patients signed informed consent prior to participating in the study. Two to sixteen depth electrodes, placed according to clinical criteria (see Table 2), each with 6 to 14 regularly-spaced macro contacts, were implanted in each subject, and the electric potential difference between each contact and the reference contact (a subdermal electrode contact placed near Fz), were sampled at 2 kHz. For each subject, recording started between 8 PM and midnight, depending on an estimated bedtime. EEG data from the raw Neuralynx files was converted to EDF files in blocks of 4 h periods. For each subject, between 6 and 12 h of sleep/wake data were recorded and considered for analysis. Electrophysiological readings from Ad-Tech electrodes are collected at 2 kHz with a 256-channel ATLAS digital acquisition system, (Neuralynx, Bozeman, MT), which was connected to a PC running Neuralynx, Pegasus Acquisition Software, version 2.1.1.

2.2. Electrode Localization

The locations of depth electrode contacts are determined from preoperative high-resolution magnetic resonance imaging (MRI) scans along with postoperative CT scans. For each subject, the Freesurfer program [43,44] is used to generate 3D RAS coordinates and Desikan-Killiany anatomical labels [45] for the electrode contacts, and then the post-operative CT scan is coregistered with these coordinates using the IELU pipeline and a MATLAB script [46,47] to obtain anatomical locations.

2.3. Manual Sleep Scoring

For the purposes of comparison, manual sleep scoring is conducted independently by two experienced sleep scorers using criteria specified by the American Association of Sleep Medicine for sleep scoring [4] using synchronized EEG (EEG is collected from a centrally located subdermal strip), iEEG, EMG, and EOG signals. All recordings are collected using the ATLAS recording system. Periods identified as REM sleep are further

verified by two additional sleep scorers. An epoch is labeled REM if low-amplitude mixed-frequency activity is observed along with rapid eye movements and a low chin EMG tone. Of the subjects examined, only subject 7 displays unusual electrophysiological phenomena. Though epochs for subject 7 labeled as NREM or REM exhibit many features of their respective sleep stages, all sleep stages also display unusual activity. For instance, during periods of low EMG tone accompanied by conjugate eye movements that would otherwise be labeled REM, high voltage, low frequency activity (but not low enough frequency to be N3) is often observed on some channels. This behavior is not typical for any of the other subjects. Subject 7 is included for completeness and also as an example of how the proposed REM/Wake classifier performs on unusual activity.

Table 1. Patient Information. The row for each patient provides demographic information, diagnostic information about their epilepsy, and brain imaging results.

Subject	Gender	Age	Handedness	Diagnosis	Imaging
1	F	24	R	Focal epilepsy, left and right temporal lobe	PET: Bilateral medial temporal hypometabolism. MRI: Normal
2	F	41	R	Focal epilepsy, left temporal lobe	PET: Borderline symmetric hypometabolism involving mesial temporal lobes. MRI: Normal
3	F	55	R	Medial refractory epilepsy with bilateral hippocampal foci	PET: Minimal hypometabolism of left medial and inferior temporal lobe. MRI: Normal
4	F	52	R	Intractable epilepsy, bilateral hippocampal	PET: Hypometabolism left para-hippocampal gyrus. MRI: Bilateral Hippocampal Cyst, 1-2 mm. Infarct, superior right cerebellum
5	M	29	R	Epilepsy of bilateral temporal origin.	PET: Hypometabolism bilaterally in the medial temporal lobes. MRI: Normal
6	M	19	L	Seizures of right temporal origin	PET: Hypometabolism in the right temporal lobe relative to left temporal lobe. MRI: Normal
7	F	20	R	Refractory epilepsy, left frontotemporal	PET: Small area of non-significant hypometabolism in left mesiotemporal region. MRI: Normal
8	M	38	R	Medically intractable epilepsy, left hemisphere	PET: Subtle hypometabolism predominantly in right parietal lobe. MRI: Normal
9	F	20	R	Medical refractory epilepsy with bilateral hippocampal foci	PET: Minimal hypometabolism of right lateral temporal lobe. MRI: Unremarkable (incidental finding of bilateral hippocampal cysts)

2.4. Data Selection

Analysis is restricted to at least one seizure-free night of sleep where the subject entered both NREM and REM sleep and to surface contacts on frontal and temporal depth electrodes. These contacts are more likely to be (i) seizure-free, (ii) collected by other iEEG data centers, (iii) may generalize to electrocorticography (ECoG) data, which records activity at the surface level of the brain, and (iv) are most likely to exhibit the neural REM/Wake alpha-rhythm dynamics observed in standard EEG recordings.

Electrode contacts are further removed from analysis if identified by epileptologists to lie in the epileptic focus or to exhibit interictal activity (see Tables 1 and 2).

Table 2. Number and placement of implanted electrodes. Each electrode had multiple macro contacts, which are the actual recording surfaces located along the electrode shaft. Therefore, some macro contacts were located closer to the surface of the brain. Note that the macro contacts (6–10 per electrode) were evenly spaced along the electrode shaft.

Subject Number	No. of Electrodes Implanted	No. of Electrodes Implanted, by Location			
		Left Hemisphere		Right Hemisphere	
		Frontal	Temporal	Frontal	Temporal
1	16	4	4	4	4
2	16	4	4	4	4
3	16	4	4	4	4
4	12	3	3	3	3
5	14	3	4	3	4
6	11	3	2	3	3
7	2	0	2	0	0
8	4	0	2	0	2
9	12	3	3	3	3

2.5. Filtering and Sectioning

To suppress power-line interference, selected data is filtered to remove 60 Hz and harmonic frequencies using a zero-phase FIR notch filter. Filtered data is partitioned into 30 s, non-overlapping sections. Each section is assigned a stage label by the manual sleep scoring procedure (see Section 2.3).

2.6. Multitaper Eigencoefficients

Each 30 s section of data is further subdivided into 5 non-overlapping sections, each of which is 6 s in duration. Let $c \in \mathbb{N}$ equal the electrode-contact index, $s \in \mathbb{N}$ indicate the subject, e indicate the 30 s episode (or section), $m \in \{1, 2, 3, 4, 5\}$ the subsection, and let t equal the subsection time-index. Further define, $d_{m,t}^{(c,s,e)}$ to be equal to the t^{th} centered sample of the m^{th} subsection of the e^{th} 30 s window for contact c and subject s . Note that the explicit dependence of c, e upon s is suppressed in the notation. Let $v_t^{(k)}$ be the t^{th} sample of the k^{th} -order discrete-prolate spheroidal sequence (DPSS); a DPSS is parameterized by order k , as well as by its length N and bandwidth W . The zero-th order DPSS possesses, among all length N sequences, the maximum signal energy within the frequency interval $(-W, W)$. The $k + 1^{\text{st}}$ order DPSS is, of all sequences orthogonal to $v^{(0)}$, maximally energy concentrated within $(-W, W)$ [48]. Here we focus on alpha-rhythm, which is classically considered bandlimited to the frequency interval (8 Hz, 12 Hz). For the remainder of this work f are equal to 10 Hz and W is equal to 2 Hz. For sample period Δ , the corresponding dimensionless DPSS time-bandwidth parameter is equal to ΔNW . It is equal to 12 for the chosen ΔN and W . Let F_N , equal to $1/2\Delta$, be the Nyquist frequency. The corresponding k^{th} order eigencoefficient, $Y_{e,m,k}$, evaluated at the frequency $f \in -f_N, f_N$, is equal to [49],

$$Y_{m,k}^{(c,s,e)}(f) = \sum_{t=0}^{N-1} v_t^{(k)} d_{m,t}^{(c,s,e)} e^{-i2\pi ft\Delta}. \quad (1)$$

Here, i^2 is equal to -1 and $Y_{m,k}^{(c,s,e)}(f) \in \mathbb{C}$ is a complex-valued number. A simple multitaper spectrum estimate, $\check{S}^{(c,s,e)}$, evaluated at f , is equal to,

$$\check{S}^{(c,s,e)}(f) = \sum_{m=1}^5 \sum_{k=1}^K |Y_{m,k}^{(c,s,e)}(f)|^2 / 5K. \quad (2)$$

The parameter K is chosen to be the number of DPSSs that possess signal-energy within the frequency interval $(-W, W)$ exceeding 0.98 ($K = 22$). The signal-energy of a sequence is equal to its squared ℓ_2 norm. The ℓ_2 norm of any DPSS is equal to 1. With this

specification, out-of-band bias due to spectral leakage is limited by the optimal in-band energy concentration properties of the DPSSs [34,49–51]. The resulting K is equal to 22 ($K = 2NW - 2$). Let N_c equal the number of electrode contacts for subject s . Form the matrix of eigencoefficients $\mathbf{Y}^{(s,e)} \in \mathbb{C}^{N_c \times 110}$ characterizing alpha rhythm for the e^{th} 30-s recording for subject s . The p^{th} row and q^{th} column of $\mathbf{Y}^{(s,e)}$ is equal to,

$$\left[\mathbf{Y}^{(s,e)} \right]_{p,q} = Y_{m_q, k_q}^{(p,s,e)}(f) \Big|_{f=10 \text{ Hz}}, \quad (3)$$

such that $q = 22(m_q - 1) + k_q$ and subject to the restriction $m_q \in 1, \dots, 5, k_q \in 1, \dots, 22$.

2.7. Multitaper Eigencoefficient Clustering

Cluster labels are assigned in a K-means clustering algorithm modified to cluster the multitaper eigencoefficients characterizing each 30 s recording episode. Specifically, the distance introduced in [52,53] is replaced with the cluster-conditioned probability density, $P(\mathbf{Y} | C = j)$, for cluster j . Furthermore, the probability density of observing the eigencoefficient feature matrix \mathbf{Y} , conditioned upon cluster j is,

$$\begin{aligned} \ln P(\mathbf{Y} | C = j) = & -N_c \ln (2\pi) + \\ & -\frac{N_c}{2} \left(\ln |\check{\mathbf{R}}_r^{(j)}| + \ln |\check{\mathbf{R}}_i^{(j)}| \right) - \\ & \frac{1}{2} \operatorname{tr} \left\{ \left(\mathbf{Y}_r - \check{\boldsymbol{\mu}}_r^{(j)} \mathbf{1}_{Nr}^T \right)^T \left[\check{\mathbf{R}}_r^{(j)} \right]^{-1} \left(\mathbf{Y}_r - \check{\boldsymbol{\mu}}_r^{(j)} \mathbf{1}_{Nr}^T \right) \right\} - \\ & \frac{1}{2} \operatorname{tr} \left\{ \left(\mathbf{Y}_i - \check{\boldsymbol{\mu}}_i^{(j)} \mathbf{1}_{Nr}^T \right)^T \left[\check{\mathbf{R}}_i^{(j)} \right]^{-1} \left(\mathbf{Y}_i - \check{\boldsymbol{\mu}}_i^{(j)} \mathbf{1}_{Nr}^T \right) \right\}. \end{aligned} \quad (4)$$

Here, N_r is equal to 110. It is equal to the number of subsections (i.e., 5) multiplied by the number of DPSS tapers, K (i.e., 22). The real-valued matrices \mathbf{Y}_r (\mathbf{Y}_i) are the real (imaginary) components of the eigencoefficient feature matrix \mathbf{Y} , $\check{\boldsymbol{\mu}}^{(j)} \in \mathbb{C}^{N_c}$ is the j^{th} cluster-conditional sample average, and $\check{\mathbf{R}}^{(j)} \in \mathbb{C}^{N_c \times N_c}$ is the j^{th} cluster-conditional sample covariance matrix. Let the j^{th} element of the set of labels $L_j \in \{1, \dots, N_{\text{cluster}}\}^{N_{\text{episode}}}$ indicate the cluster to which the eigencoefficient feature-vector $\mathbf{Y}^{(s,e)}$ is assigned. Here, N_{cluster} is the number of clusters and N_{episode} is the number of 30 s epochs. For subject s . Let Q_j be the number of episodes assigned to cluster j . Furthermore, for cluster j , the cluster-specific mean-vector estimate $\check{\boldsymbol{\mu}}^{(j)}$ is equal to,

$$\check{\boldsymbol{\mu}}^{(j)} = \sum_{e=1}^{Q_j} \sum_{j'=1}^{N_r} \left[\mathbf{Y}^{(s,e)} \right]_{.,j'} / (N_r Q_j). \quad (5)$$

Let the convenience variable $\tilde{\mathbf{Y}}^{(s,e,j)}$ equal

$$\tilde{\mathbf{Y}}^{(s,e,j)} = \mathbf{Y}^{(s,e)} - \check{\boldsymbol{\mu}}^{(j)} \mathbf{1}_{Nr}^T. \quad (6)$$

Here, $\mathbf{1}_{Nr}$ is a column-vector of N_r elements, each of which equals 1. The real part, $\check{\mathbf{R}}_r^{(j)}$, of the cluster specific covariance matrix estimate $\check{\mathbf{R}}^{(j)}$, is equal to,

$$\begin{aligned} \check{\mathbf{R}}_r^{(j)} = & \sum_{e=1}^{Q_j} \sum_{j'=1}^{N_r} \left[\tilde{\mathbf{Y}}_r^{(s,e,j)} \right]_{.,j'} \left[\tilde{\mathbf{Y}}_r^{(s,e,j)} \right]_{.,j'}^T \\ & / (N_r Q_j - 1). \end{aligned} \quad (7)$$

Similarly, the imaginary part, $\check{\mathbf{R}}_i^{(j)}$, is equal to,

$$\begin{aligned}\check{\mathbf{R}}_i^{(j)} = & \sum_{e=1}^{Q_j} \sum_{j'=1}^{N_r} \left[\tilde{\mathbf{Y}}_i^{(s,e,j)} \right]_{\cdot,j'} \left[\tilde{\mathbf{Y}}_i^{(s,e,j)} \right]_{\cdot,j'}^T \\ & / (N_r Q_j - 1).\end{aligned}\quad (8)$$

The cluster-conditional probability density function, Equation (4), is motivated by the asymptotic distribution of the tapered discrete-time Fourier transform of a multivariate, weak-sense stationary, random process (see, for example, Theorem 4.4.2, [54]), as well as the approximate independence of the multitaper eigencoefficients due to the orthogonality of the DPSSs [49]. The steps in the modified K-means algorithm are as follows:

1. Specify the number of clusters.
2. The initial observation labels are randomly guessed (as recommended in ref. [55]), and the cluster-specific parameters μ and \mathbf{R} are estimated for each cluster using Equations (5), (7) and (8), to obtain $\check{\mu}$, $\check{\mathbf{R}}$ for each cluster.
3. Given $\check{\mu}$, $\check{\mathbf{R}}$ for each cluster, the feature matrices, $\mathbf{Y}^{(s,e)}$, $e = 1, \dots, N_{episode}$ (i.e., one for every 30-s episode), are re-clustered by assigning them the label that maximizes Equation (4).
4. Using the new cluster assignments, $\check{\mu}$ and $\check{\mathbf{R}}$ are re-computed.
5. Steps 3 and 4 are repeated until an iteration occurs without a resulting change in the feature-matrix label. Once this occurs, subsequent iterations result in no change, and the algorithm has converged.

As previously stated, this iterative procedure is a variation on standard K-means clustering and is related to the K-means clustering algorithms discussed in [52,53,55]. For our data, by iteration 25, no changes to the eigencoefficient feature-matrix labels occur between one iteration and the next. The procedure consisting of steps 1 through to 5 above produces a single clustering. It remains to determine the number of clusters to use. Upon convergence, an approximate Akaike information criterion $A\tilde{C}$ is computed. It is equal to,

$$\begin{aligned}A\tilde{C}(N_{cluster}) = & N_{cluster} \left(N_c^2 + N_c \right) - \\ & 2 \sum_{j=1}^{N_{cluster}} \sum_{e=1}^{N_{episode}} \ln P(\mathbf{Y}^{(s,e)} | C = j).\end{aligned}\quad (9)$$

with, as before, N_c equal to the number of electrode contacts. As with the Akaike information criterion, for the purpose of model selection, models, or labelings, which minimize $A\tilde{C}$ promote models that result in probable labeling and penalize models with many clusters. Smaller values of AIC imply a lower Kullback-Liebler divergence between the unknown data-generating probability density function and the modeled generating probability density function [56,57]. In this work, steps 1 through 5 above are computed, starting with $N_{cluster}$ equal to 2 through 14. The chosen clustering, $\tilde{N}_{contact}$, minimizes the $A\tilde{C}$ specified in Equation (9). That is,

$$\tilde{N}_{contact} = \arg \min_{n' \in \{2, \dots, 14\}} A\tilde{C}(n'). \quad (10)$$

2.8. REM Cluster

The goal of this work is to reliably identify a subset of REM episodes. This is accomplished once cluster labels have been assigned to the episode feature matrices, $\mathbf{Y}^{(s,e)}$, $e = 1, \dots, N_{episode}$, by identifying, in a sense to be made precise below, the cluster exhibiting the lowest alpha spectral power. Let C_j be the set of episodes assigned to cluster j .

A cluster-conditional multitaper estimate of alpha-power for contact c and subject s is equal to,

$$\check{S}_j^{(c,s)} = \sum_{e \in C_j} \check{S}^{(c,s,e)}(f) \Big|_{f=10 \text{ Hz}} / |C_j|. \quad (11)$$

Next, for each contact, c , use $\check{S}_j^{(c,s)}$, $j = 1, \dots, N_{\text{cluster}}$, to rank the clusters from smallest $\check{S}_j^{(c,s)}$ to largest. Specifically, let the rank $r_{j,c}^{(s)}$ equal,

$$r_{j,c}^{(s)} = \underset{j' = \{1, \dots, N_{\text{cluster}}\}}{\text{rank}} \check{S}_{j'}^{(c,s)}. \quad (12)$$

The rank, $r_{j,c}^{(s)}$ is equal to 1 when $\check{S}_j^{(c,s)}$ is the minimum of the set $\{\check{S}_1^{(c,s)}, \dots, \check{S}_{N_{\text{cluster}}}^{(c,s)}\}$. Define the spectral rank, $\rho_j^{(s)}$ of cluster j , for subject s , to be the median of $r_{j,c}^{(s)}$ across electrode contacts:

$$\rho_j^{(s)} = \underset{c = \{1, \dots, N_c\}}{\text{median}} r_{j,c}^{(s)}. \quad (13)$$

Finally, let the REM cluster index $j_*^{(s)}$ for subject s equal,

$$j_*^{(s)} = \underset{j' = \{1, \dots, N_{\text{cluster}}\}}{\arg \min} \rho_{j'}^{(s)}. \quad (14)$$

The episodes belonging to C_{j_*} are identified as REM sleep. If $j_*^{(s)}$ is integer. In the situation where j_* is not integer-valued the analysis is deemed unreliable. For all data analyzed in this work, $j_*^{(s)}$ is integer.

2.9. Measures of Label Confidence

Two proxy measures of REM label identification specificity are computed. Let $j_{1,c}$ be equal to,

$$j_{1,c} = \underset{j \in \{1, \dots, N_{\text{cluster}}\}}{\arg \min} \check{S}_j^{(c,s)}, \quad (15)$$

and let $j_{2,c}$ equal,

$$j_{2,c} = \underset{\substack{j \in \{1, \dots, N_{\text{cluster}}\} \\ j \neq j_{1,c}}}{\arg \min} \check{S}_j^{(c,s)}. \quad (16)$$

Then, the minimum, per-contact, alpha-power difference, $d\alpha_c$ is equal to,

$$d\alpha_c = \check{S}_{j_{2,c}}^{(c,s)} - \check{S}_{j_{1,c}}^{(c,s)}. \quad (17)$$

The first figure of merit, FOM_1 , is equal to,

$$FOM_1 = \underset{c = 1, \dots, N_{\text{contact}}}{\arg \min} d\alpha_c. \quad (18)$$

Smaller FOM_1 indicates that, for at least one electrode contact, the difference in alpha power between the two clusters exhibiting lowest alpha power is small. It is listed in row 4 of Table 3 and is least for the two worst-performing subjects. A second figure of merit, FOM_2 , is equal to the fraction of the contacts, c , in cluster j_* (i.e., the REM cluster), for which $\check{S}_{j_*}^{(c,s)}$ is not minimum (see row 5, Table 3).

2.10. Hold-Out Analysis

The specificity of the proposed classifier depends on the sample size. Figure 2 shows the specificity of the proposed classifier (see Section 2.8) on random data splits. Data

is randomly divided into two equally sized sets of data (manually identified REM and Wake episodes are each randomly split apart from each other and then combined), and the classifier is applied to each set. This random splitting and classification is repeated 50 times. The resulting REM identification specificity is plotted in Figure 2.

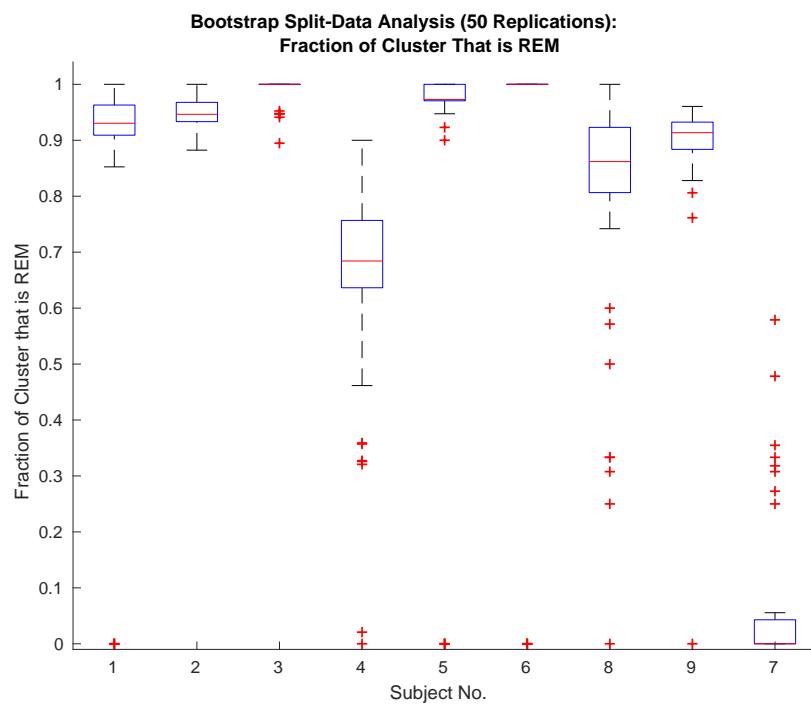


Figure 2. Performance of the proposed classifier is excellent with the exception of Subjects 4, 8 and 7. Subject 7 exhibits abnormal sleep and can be excluded along with Subject 4 based upon FOM_1 (see Table 3).

Dependence upon sample size is further quantified by repeating the above 50 random data splits, but in this case, changing the split to data divisions of varying size. The resulting boxplots of REM identification specificity are plotted in Figure 3.

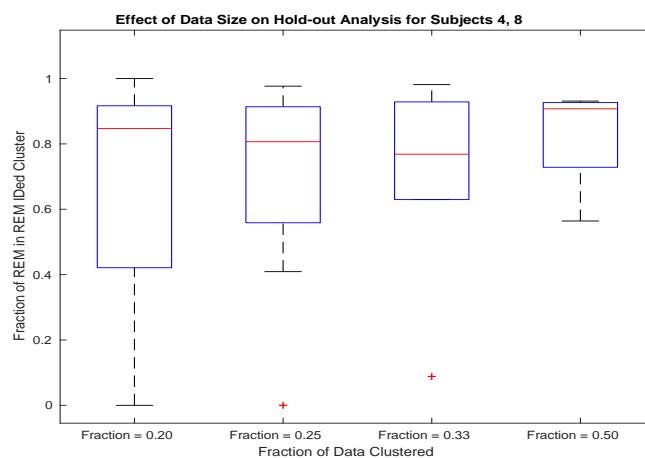


Figure 3. For subjects 4 and 8, to obtain specificity approaching full-data performance (see row 1 of Table 3), half of a night's worth of REM and Wake episodes are required.

2.11. Comparison with the Alpha Power Detector

The proposed unsupervised classifier introduced in Section 2.8 is an atypical classifier. It is insightful to compare this classifier with an alpha-power binary hypothesis test. To do this, the statistical power of the proposed classifier with the alpha-power test is matched by

setting the latter test's threshold such that the number of REM episodes identified by the two different methods is equal. The number of false positives resulting from classification using the two methods is compared in Figure 4. Of the subjects exhibiting normal sleep, the proposed classifier out-performs the equivalent statistical-powered hypothesis test based upon alpha power for five of eight subjects, and performs comparably on the remaining two subjects. For one subject, the specificity of the proposed classifier is 20% larger than the specificity demonstrated by the competing binary-hypothesis test.

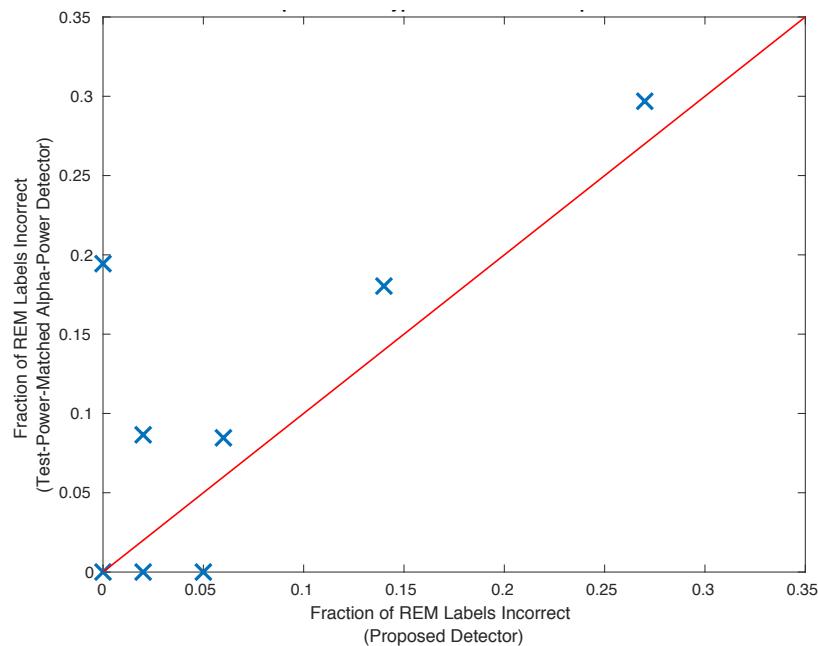


Figure 4. A matched-power comparison between the proposed classifier and a binary hypothesis test which compares alpha-power to reject a null hypothesis which is associated with REM sleep. The alpha-power threshold of the latter is set such that an equal number of REM episodes are identified by both classifiers. Of the subjects exhibiting normal sleep, the proposed classifier out-performs the equivalent statistical-powered hypothesis test based upon alpha power for six of eight subjects.

2.12. Clustering Pseudo-Code

The following pseudo-code illustrates the classification of the REM and Wake sleep stages. Given a night of 30 s, multi-channel, data section,

1. Divide the data into 30 s, non-overlapping sections of data.
2. Compute the $110N_c$ eigencoefficients characterizing each 30 s section of data.
3. Specify the number of clusters to be equal to 2.
4. Randomly assign a cluster label to each of the 30 s sections of data.
5. Compute each of the cluster statistics conditioned upon the cluster labels.
6. For each 30 s section of data, and for each cluster, compute the log-probability of the observed section data conditioned upon each of the cluster labels.
7. Assign to each 30 s section of data the cluster label that maximizes the log-probability of the observed 30 s section eigencoefficients.
8. Repeat steps 4 through 7 until the cluster labels do not change from one repetition to the next.
9. Compute the approximate AIC, $A\tilde{I}C$, for this number of clusters, using all of the log-probabilities.
10. Repeat steps 3 through 9 but increment the number of clusters on each repetition until 14 clusters are used, or until an empty cluster results.

Steps 1 through 10 above are used to estimate the cluster labels, which are taken to be the labels associated with the smallest $A\tilde{I}C$.

2.13. REM Identification Pseudo-Code

To identify the REM cluster, first perform the clustering estimation described in Section 2.12. Then:

1. Average the cluster-specific alpha spectral power across 30 s episodes.
2. For each electrode contact, order the average cluster alpha spectral power across clusters.
3. For each cluster, compute the median of the ranks across the contacts. This is the spectral rank.
4. The REM cluster is the cluster with the smallest spectral rank.

2.14. Software Overview

The Pegasus Acquisition Software, Neuralynx, version 2.1.1 is used for data acquisition. Except for the spectrogram depicted in Figure 1, custom MATLAB scripts are used for all spectral analysis, clustering, signal processing, and statistical analysis. These scripts are available upon request. Multitaper spectral analysis dates back to the seminal paper published in 1982 [49]. Software available to perform multitaper analysis is available in the following languages: Fortran, R, C, Python, Julia, and MATLAB [34,58–62].

3. Results

Table 3 shows the results of applying the proposed classifier introduced in Section 2.8 to recorded iEEG data. The unsupervised, multitaper, alpha rhythm classifier introduced in Section 2.8 outperforms an equivalently powered alpha-power hypothesis test (see Figure 4 in Section 2.11) and attains specificity comparable to that obtained by experienced human sleep-scorers. It requires one-half of a night’s worth of REM and Wake episodes of healthy sleep to obtain results comparable to those obtained with a full night of data (see Figure 3 in Section 2.10). Since the time at which these episodes are recorded spans an entire night, for practical purposes, an entire night’s worth of data is required. The provided figure of merit, FOM_1 , is lowest for the two worst-performing subjects (see row 3, Table 3), while FOM_2 does not correlate with performance (compare rows 2 and 4 of Table 3).

Table 3. The proposed method identifies REM in 6 out of 8 subjects exhibiting normal sleep with a specificity greater than or equal to 94 %. For the other two subjects, the specificity is 86% and 73%; respectively. It exhibits performance comparable to the inter-rater reliability exhibited by human sleep scorers [63].

Subject	1	2	3	4	5	6	8	9	7
Minutes of REM Correctly ID’ed	47.5	26.0	14.5	22.5	31.5	62.5	25.0	60.0	NA
Fraction of REM Labels Correct	0.98	0.95	1.00	0.73	0.98	1.00	0.86	0.94	0.00
Fraction of REM Correctly ID’ed	0.77	0.42	0.60	0.40	0.88	0.78	0.79	0.67	NA
Min. Alpha Power Diff. B/W Closest Competing Cluster and REM Cluster	0.79	0.07	0.09	2.00×10^{-3}	0.01	0.07	0.74	0.06	6.00×10^{-3}
Fraction of REM Cluster Electrode Alpha Power Not Minimum	0.00	0.10	0.20	0.08	0.09	0.00	0.00	0.18	0.00
Number of Clusters	14	12	10	14	8	6	14	14	14
Number of Electrode Contacts	12	9	10	12	11	12	3	11	2

4. Discussion

The proposed multitaper, multi-channel iEEG REM detector achieves a specificity greater than or equal to 94% on 6 out of 8 subjects exhibiting normal sleep. The figure-of-merit metric, FOM_1 , reported in Table 3, and labeled “Minimum alpha power difference between the closest competing cluster with the REM identified cluster” differs by over an order of magnitude between the typical subject and the two worst performing subjects.

This figure-of-merit responds to the separation in terms of alpha power between the closest cluster and the REM cluster itself. Subject selection based upon this criterion raises performance to a specificity of greater than or equal to 94% on 6 out of 7 subjects when identifying REM with the proposed method. Finally, it is worth noting that while the precision of the proposed detector is not high, it is sufficient to identify a median (minimum) minutes of REM equal to 28.75 (14.5) min per subject. This is identified from a single night of data when the typical hospital stay of a focal epilepsy patient is 1 to 2 weeks.

The proposed method is motivated by six main statistical signal processing considerations. The first (i) is to use sufficient information in the alpha frequency interval for the purpose of optimizing a binary detector (Wake vs. REM). Naively, this suggests the use of time-frequency distributions such as those employed in ref. [24] to classify sleep. It is important to note that there exists a bias/variance trade when temporally resolving frequency-domain structures. Resolution in time implies less temporal averaging. In many cases, this results in an increase in estimator variance, which in this case increases feature-domain variance, which can obscure clusters. That bias will move cluster centers; however, unless it moves them in an overlapping direction, this does not impact clustering performance for the purpose of detection. Finally, the potential information lost by not using a time-frequency distribution is the specific time within a 30 s episode at which a reproducible time-frequency event occurs. This information, if it exists, may further divide the REM cluster and necessitate an otherwise unnecessary cluster-merging step. Similarly, one might consider incorporating phase-phase and phase-amplitude spectral features (e.g., [51]). We seek to capture known average dynamics that occur in REM and, furthermore, the across-brain region timing of alpha oscillations. This suggests the use of multichannel multitaper spectral statistics [49,64], because (a) a multitaper spectrum estimator minimizes out-of-band bias while controlling the increase in variance associated with data tapering; (b) the multitaper eigencoefficients are concentrated within the frequency interval ($f - W, f + W$), are pairwise orthogonal on this interval, and form an approximate basis for this interval (and by symmetry for $(-f - W, -f + W)$, for real-valued time-series). The second statistical signal processing consideration (ii) is to use the likelihood-ratio. It is known by the Neyman-Pearson lemma to provide the optimal binary detector (see, for example, ref. [65]). The third, fourth, and fifth considerations (iii,iv,v) are to allow for subject specific idiosyncrasy (exhibited by ref. [41] and enhanced in iEEG due to variable epileptic foci), temporal changes of brain state, and corrupted data, such as that which results from recording during patient motion. These considerations suggest the use of unsupervised learning applied to the 30 s stretches of recording to be sleep-scored. Among the large number of unsupervised learning methods, K-means clustering is classical and is covered in books such as [66]. Finally (vi), there is a limited quantity of sleep-scored iEEG data from which to train a classifier. To generalize well, our detector must be simple and motivated by existing scientific and signal processing knowledge. These considerations lead us to select the minimum alpha power cluster as defined in Section 2.8 since Wake, motion artifacts, and other outliers all tend to increase the signal power observed in the alpha frequency interval.

The performance of the proposed method compared with that exhibited by a more simple alpha-power detector (see Figure 4, Section 2.11), suggests the use of the proposed detector. Specifically, the specificity is superior for 4 of 8 subjects, is the same for 1, slightly worse for 2, and exhibits a much greater specificity for one subject. Focusing on the impact of the performance improvement on this latter subject, should this performance improvement generalize, the proposed detector will avoid the expected 20% of incorrectly labeled REM episodes that would occur when using the simple alpha power detector on 1/8th of all iEEG data.

The bootstrap analysis (see Figure 2) indicates comparable detector performance across random data assignments for all but Subjects 4,8, and 7 (i.e., the worst-performing subjects). For subjects exhibiting normal sleep, the detector is applied in a hold-out analysis. For these worse-performing subjects, performance on half a night of data is comparable to

that attained from a night of data. See Figure 3 and Table 1. The bootstrap and hold-out analyses provide evidence that the specificity exhibited by the proposed multitaper detector will generalize.

This work can be extended by including the proposed REM detector as part of a more general sleep stage classifier. In this context, the sleep-stage transition probabilities may be used to potentially increase performance. The sample size of nine subjects is not large. It will be important to verify the performance of the proposed multitaper, multi-channel REM detector with larger, future iEEG data. Should the performance generalize, it provides a basis upon which REM can be identified in iEEG recordings unaccompanied by EOG and EMG. Identified REM activity can be correlated against, for example, the location of the epileptogenic zone for the purpose of improving the outcome of focal epilepsy surgery.

5. Conclusions

The proposed unsupervised multitaper classifier correctly identifies 36 ± 6 min of REM in one night of recorded sleep (row 1, Table 3) while incorrectly labeling less than 10% of all labeled REM episodes for all but one subject; human rater reliability is estimated at near 80% [63]. This classifier outperforms the equivalent statistical-power classical test. Hold-out analysis indicates that when using one night's worth of iEEG data, an accurate generalization of the method on new data is likely. The introduced multitaper multi-channel alpha-rhythm classifier further paves the way to making a large quantity of otherwise unusable iEEG data available for the purpose of studying sleep.

Author Contributions: Conceptualization, K.Q.L., S.J. and S.V.; Data curation, S.J., A.K., M.W. and S.V.; Formal analysis, K.Q.L.; Funding acquisition, S.V.; Investigation, K.Q.L. and S.V.; Methodology, K.Q.L.; Resources, M.W. and S.V.; Software, K.Q.L.; Supervision, S.V.; Validation, K.Q.L.; Writing—original draft, S.J.; Writing—review & editing, K.Q.L. and S.V. All authors have read and agreed to the published version of the manuscript.

Funding: SV was supported by funding from grants R21MH127567, BCS 2047529, and W911NF-23-1-0114 from the National Institutes of Health, the National Science Foundation, and the Army Research Office, respectively.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board of Carilion Clinic (20-1179, 8 February, 2021).

Informed Consent Statement: All subjects gave their informed consent for inclusion before they participated in the study. The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by Carilion Clinic IRB.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy concerns and regulations.

Acknowledgments: We thank Anna E Vijayan for carefully reading and editing the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

EEG	electroencephalography
iEEG	intracranial EEG
EOG	electrooculography
EMG	electromyography
REM	rapid eye movement
NREM	non rapid eye movement
AIC	Akaike information criterion

References

- Nobili, L.; Frauscher, B.; Eriksson, S.; Gibbs, S.A.; Halasz, P.; Lambert, I.; Manni, R.; Peter-Derex, L.; Proserpio, P.; Provini, F.; et al. Sleep and epilepsy: A snapshot of knowledge and future research lines. *J. Sleep Res.* **2022**, *31*, e13622. [[CrossRef](#)] [[PubMed](#)]
- Vakharia, V.N.; Duncan, J.S.; Witt, J.A.; Elger, C.E.; Staba, R.; Engel, J., Jr. Getting the best outcomes from epilepsy surgery. *Ann. Neurol.* **2018**, *83*, 676–690. [[CrossRef](#)] [[PubMed](#)]
- Demuru, M.; van Blooij, D.; Zweiphenning, W.; Hermes, D.; Leijten, F.; Zijlmans, M.; group, R. A practical workflow for organizing clinical intraoperative and long-term iEEG data in BIDS. *Neuroinformatics* **2022**, *20*, 727–736. [[CrossRef](#)] [[PubMed](#)]
- Iber, C.; Ancoli-Israel, S.; Chesson, A.L.; Quan, S.F. The new sleep scoring manual—the evidence behind the rules. *J. Clin. Sleep Med.* **2007**, *3*, 107. [[CrossRef](#)]
- McLeod, G.A.; Ghassemi, A.; Ng, M.C. Can REM sleep localize the epileptogenic zone? A systematic review and analysis. *Front. Neurol.* **2020**, *11*, 584. [[CrossRef](#)] [[PubMed](#)]
- Kang, X.; Boly, M.; Findlay, G.; Jones, B.; Gjini, K.; Maganti, R.; Struck, A.F. Quantitative spatio-temporal characterization of epileptic spikes using high density EEG: Differences between NREM sleep and REM sleep. *Sci. Rep.* **2020**, *10*, 1673. [[CrossRef](#)] [[PubMed](#)]
- Noachtar, S.; Borggraefe, I. Epilepsy surgery: A critical review. *Epilepsy Behav.* **2009**, *15*, 66–72. [[CrossRef](#)]
- Englot, D.J.; Han, S.J.; Rolston, J.D.; Ivan, M.E.; Kuperman, R.A.; Chang, E.F.; Gupta, N.; Sullivan, J.E.; Auguste, K.I. Epilepsy surgery failure in children: A quantitative and qualitative analysis. *J. Neurosurg. Pediatr.* **2014**, *14*, 386–395. [[CrossRef](#)]
- Englot, D.J.; Raygor, K.P.; Molinaro, A.M.; Garcia, P.A.; Knowlton, R.C.; Auguste, K.I.; Chang, E.F. Factors associated with failed focal neocortical epilepsy surgery. *Neurosurgery* **2014**, *75*, 648. [[CrossRef](#)]
- Vijayan, S.; Lepage, K.Q.; Kopell, N.J.; Cash, S.S. Frontal beta-theta network during REM sleep. *elife* **2017**, *6*, e18894. [[CrossRef](#)]
- Kubicki, S.; Höller, L.; Berg, I.; Pastelak-Price, C.; Dorow, R. Sleep EEG evaluation: A comparison of results obtained by visual scoring and automatic analysis with the Oxford sleep stager. *Sleep* **1989**, *12*, 140–149. [[CrossRef](#)]
- Flexerand, A.; Dorffner, G.; Sykacekand, P.; Rezek, I. An automatic, continuous and probabilistic sleep stager based on a hidden Markov model. *Appl. Artif. Intell.* **2002**, *16*, 199–207. [[CrossRef](#)]
- Björn, R.; Born, J. About sleep's role in memory. *Physiol. Rev.* **2013**, *93*, 681–766.
- Peever, J.; Fuller, P.M. The biology of REM sleep. *Curr. Biol.* **2017**, *27*, R1237–R1248. [[CrossRef](#)] [[PubMed](#)]
- Nir, Y.; Tononi, G. Dreaming and the brain: From phenomenology to neurophysiology. *Trends Cogn. Sci.* **2010**, *14*, 88–100. [[CrossRef](#)]
- Blumberg, M.S.; Lesku, J.A.; Libourel, P.A.; Schmidt, M.H.; Rattenborg, N.C. What is REM sleep? *Curr. Biol.* **2020**, *30*, R38–R49. [[CrossRef](#)]
- Siegel, J.M. REM sleep. *Princ. Pract. Sleep Med.* **2005**, *4*, 120–135.
- Gaillard, J. Principles of Automatic Analysis of sleep records with a hybrid system. *Comput. Biomed. Res.* **1973**, *6*, 1–13. [[CrossRef](#)]
- Lacroix, B.; Hanus, R. Online automatic sleep analysis. *JORBEL-Belg. J. Oper. Res. Stat. Comput. Sci.* **1984**, *24*, 39–47.
- Gath, I.; Bar-On, E. Computerized method for scoring of polygraphic sleep recordings. *Comput. Programs Biomed.* **1980**, *11*, 217–223. [[CrossRef](#)]
- Lim, A.J.; Winters, W.D. A practical method for automatic real-time EEG sleep state analysis. *IEEE Trans. Biomed. Eng.* **1980**, *BME-27*, 212–220. [[CrossRef](#)] [[PubMed](#)]
- Loh, H.W.; Ooi, C.P.; Vicnesh, J.; Oh, S.L.; Faust, O.; Gertych, A.; Acharya, U.R. Automated detection of sleep stages using deep learning techniques: A systematic review of the last decade (2010–2020). *Appl. Sci.* **2020**, *10*, 8963. [[CrossRef](#)]
- Michielli, N.; Acharya, U.R.; Molinari, F. Cascaded LSTM recurrent neural network for automated sleep stage classification using single-channel EEG signals. *Comput. Biol. Med.* **2019**, *106*, 71–81. [[CrossRef](#)] [[PubMed](#)]
- Khare, S.K.; Bajaj, V.; Taran, S.; Sinha, G. Multiclass sleep stage classification using artificial intelligence based time-frequency distribution and CNN. In *Artificial Intelligence-Based Brain-Computer Interface*; Elsevier: Amsterdam, The Netherlands, 2022; pp. 1–21.
- Eldele, E.; Chen, Z.; Liu, C.; Wu, M.; Kwoh, C.K.; Li, X.; Guan, C. An attention-based deep learning approach for sleep stage classification with single-channel EEG. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2021**, *29*, 809–818. [[CrossRef](#)] [[PubMed](#)]
- Sekkal, R.N.; Berekhi-Reguig, F.; Ruiz-Fernandez, D.; Dib, N.; Sekkal, S. Automatic sleep stage classification: From classical machine learning methods to deep learning. *Biomed. Signal Process. Control* **2022**, *77*, 103751. [[CrossRef](#)]
- Van Der Donckt, J.; Van Der Donckt, J.; Deprost, E.; Vandebussche, N.; Rademaker, M.; Vandewiele, G.; Van Hoecke, S. Do not sleep on traditional machine learning: Simple and interpretable techniques are competitive to deep learning for sleep scoring. *Biomed. Signal Process. Control* **2023**, *81*, 104429. [[CrossRef](#)]
- von Ellenrieder, N.; Peter-Derex, L.; Gotman, J.; Frauscher, B. SleepSEEG: Automatic sleep scoring using intracranial EEG recordings only. *J. Neural Eng.* **2022**, *19*, 026057. [[CrossRef](#)] [[PubMed](#)]
- Berger, H. On the electroencephalogram of man. *Electroencephalogr. Clin. Neurophysiol.* **1969**, (Suppl. 28), 37+.
- Wright, K.P., Jr.; Badia, P.; Wauquier, A. Topographical and temporal patterns of brain activity during the transition from wakefulness to sleep. *Sleep* **1995**, *18*, 880–889. [[CrossRef](#)]
- Putilov, A.A.; Donskaya, O.G. Alpha attenuation soon after closing the eyes as an objective indicator of sleepiness. *Clin. Exp. Pharmacol. Physiol.* **2014**, *41*, 956–964. [[CrossRef](#)]

32. De Gennaro, L.; Ferrara, M.; Bertini, M. The boundary between wakefulness and sleep: Quantitative electroencephalographic changes during the sleep onset period. *Neuroscience* **2001**, *107*, 1–11. [CrossRef]
33. Silber, M.H.; Ancoli-Israel, S.; Bonnet, M.H.; Chokroverty, S.; Grigg-Damberger, M.M.; Hirshkowitz, M.; Kapen, S.; Keenan, S.A.; Kryger, M.H.; Penzel, T.; et al. The visual scoring of sleep in adults. *J. Clin. Sleep Med.* **2007**, *3*, 121–131. [CrossRef] [PubMed]
34. Bokil, H.; Andrews, P.; Kulkarni, J.E.; Mehta, S.; Mitra, P.P. Chronux: A platform for analyzing neural signals. *J. Neurosci. Methods* **2010**, *192*, 146–151. [CrossRef]
35. von Ellenrieder, N. SleepSEEG: Matlab Code for Sleep Scoring Based on Intracranial EEG (v.2). 2022. Available online: <https://zenodo.org/record/7410501> (accessed on 5 June 2023).
36. Janca, R.; Jezdik, P.; Cmejla, R.; Tomasek, M.; Worrell, G.A.; Stead, M.; Wagenaar, J.; Jefferys, J.G.; Krsek, P.; Komarek, V.; et al. Detection of interictal epileptiform discharges using signal envelope distribution modelling: Application to epileptic and non-epileptic intracranial recordings. *Brain Topogr.* **2015**, *28*, 172–183. [CrossRef] [PubMed]
37. Gath, I.; Feuerstein, C.; Geva, A. Unsupervised classification and adaptive definition of sleep patterns. *Pattern Recognit. Lett.* **1994**, *15*, 977–984. [CrossRef]
38. Aboalayon, K.A.I.; Faezipour, M.; Almuhammadi, W.S.; Moslehpoor, S. Sleep Stage Classification Using EEG Signal Analysis: A Comprehensive Survey and New Investigation. *Entropy* **2016**, *18*, 272. [CrossRef]
39. Güneş, S.; Polat, K.; Yosunkaya, S.; Dursun, M. A novel data pre-processing method on automatic determining of sleep stages: K-means clustering based feature weighting. *Complex Syst. Appl. ICCSA* **2009**, *6*, 112–117.
40. Güneş, S.; Polat, K.; Yosunkaya, S. Efficient sleep stage recognition system based on EEG signal using k-means clustering based feature weighting. *Expert Syst. Appl.* **2010**, *37*, 7922–7928. [CrossRef]
41. Buckelmüller, J.; Landolt, H.P.; Stassen, H.; Achermann, P. Trait-like individual differences in the human sleep electroencephalogram. *Neuroscience* **2006**, *138*, 351–356. [CrossRef]
42. Koupparis, A.M.; Kokkinos, V.; Kostopoulos, G.K. Semi-automatic sleep EEG scoring based on the hypnospectrogram. *J. Neurosci. Methods* **2014**, *221*, 189–195. [CrossRef]
43. Dale, A.M.; Fischl, B.; Sereno, M.I. Cortical surface-based analysis: I. Segmentation and surface reconstruction. *Neuroimage* **1999**, *9*, 179–194. [CrossRef] [PubMed]
44. Fischl, B.; Salat, D.H.; Busa, E.; Albert, M.; Dieterich, M.; Haselgrove, C.; Van Der Kouwe, A.; Killiany, R.; Kennedy, D.; Klaveness, S.; et al. Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. *Neuron* **2002**, *33*, 341–355. [CrossRef] [PubMed]
45. Desikan, R.S.; Ségonne, F.; Fischl, B.; Quinn, B.T.; Dickerson, B.C.; Blacker, D.; Buckner, R.L.; Dale, A.M.; Maguire, R.P.; Hyman, B.T.; et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* **2006**, *31*, 968–980. [CrossRef]
46. LaPlante, R.A.; Tang, W.; Peled, N.; Vallejo, D.I.; Borzello, M.; Dougherty, D.D.; Eskandar, E.N.; Widge, A.S.; Cash, S.S.; Stufflebeam, S.M. The interactive electrode localization utility: Software for automatic sorting and labeling of intracranial subdural electrodes. *Int. J. Comput. Assist. Radiol. Surg.* **2017**, *12*, 1829–1837. [CrossRef] [PubMed]
47. Jain, S. Investigation of Sleep Neural Dynamics in Intracranial EEG Patients. Master’s Thesis, Virginia Tech, Blacksburg, VA, USA, 2021.
48. Slepian, D. Prolate spheroidal wave functions, Fourier analysis, and uncertainty V: The discrete case. *Bell Syst. Tech. J.* **1978**, *57*, 1371–1430. [CrossRef]
49. Thomson, D.J. Spectrum estimation and harmonic analysis. *Proc. IEEE* **1982**, *70*, 1055–1096. [CrossRef]
50. Lepage, K.Q.; Thomson, D.J. Reduced mean-square error quadratic inverse spectrum estimator. *IEEE Trans. Signal Process.* **2014**, *62*, 2958–2972.
51. Lepage, K.Q.; Fleming, C.N.; Witcher, M.; Vijayan, S. Multitaper estimates of phase-amplitude coupling. *J. Neural Eng.* **2021**, *18*, 056028. [CrossRef]
52. Cerioli, A. K-means cluster analysis and mahalanobis metrics: A problematic match or an overlooked opportunity. *Stat. Appl.* **2005**, *17*, 61–73.
53. Martino, A.; Ghiglietti, A.; Ieva, F.; Paganoni, A.M. A K-means procedure based on a Mahalanobis type distance for clustering multivariate functional data. *Stat. Methods Appl.* **2019**, *28*, 301–322. [CrossRef]
54. Brillinger, D.R. *Time Series: Data Analysis and Theory*; Society for Industrial and Applied Mathematics: Philadelphia, PA, USA, 2001.
55. Faber, V. Clustering and the continuous k-means algorithm. *Los Alamos Sci.* **1994**, *22*, 67.
56. Akaike, H. A new look at the statistical model identification. *IEEE Trans. Autom. Control* **1974**, *19*, 716–723. [CrossRef]
57. deLeeuw, J. Introduction to Akaike (1973) information theory and an extension of the maximum likelihood principle. In *Breakthroughs in Statistics*; Springer: New York, NY, USA, 1992; pp. 599–609.
58. Prieto, G.A.; Parker, R.; Vernon Iii, F. A Fortran 90 library for multitaper spectrum analysis. *Comput. Geosci.* **2009**, *35*, 1701–1710. [CrossRef]
59. Lees, J.M.; Park, J. Multiple-taper spectral analysis: A stand-alone C-subroutine. *Comput. Geosci.* **1995**, *21*, 199–236. [CrossRef]
60. Denovellis, E.L.; Myroshnychenko, M.; Sarmashghi, M.; Stephen, E.P. Spectral Connectivity: A python package for computing multitaper spectral estimates and frequency-domain brain connectivity measures on the CPU and GPU. *J. Open Source Softw.* **2022**, *7*, 4840. [CrossRef]

61. O'Brien, C.B.; Baghdoyan, H.A.; Lydic, R. Computer-based multitaper spectrogram program for electroencephalographic data. *JoVE J. Vis. Exp.* **2019**, *153*, e60333.
62. Rahim, K.; Burr, W. Multitaper: Spectral Analysis Tools Using the Multitaper Method. 2020. Available online: <https://cran.r-project.org/web/packages/multitaper/index.html> (accessed on 14 August 2023).
63. Danker-hopfe, H.; Anderer, P.; Zeitlhofer, J.; Boeck, M.; Dorn, H.; Gruber, G.; Heller, E.; Loretz, E.; Moser, D.; Parapatics, S.; et al. Interrater reliability for sleep scoring according to the Rechtschaffen & Kales and the new AASM standard. *J. Sleep Res.* **2009**, *18*, 74–84.
64. Percival, D.; Walden, A. *Spectral Analysis For Physical Applications*; Cambridge University Press: Cambridge, UK, 1993.
65. Scharf, L.L. *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*; Addison-Wesley Publishing Company, Inc.: Boston, MA, USA, 1991.
66. Hastie, T.; Tibshirani, R.; Friedman, J.H.; Friedman, J.H. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*; Springer: New York, NY, USA, 2009; Volume 2.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.