Original Article

Dusk2Dawn: an EEGLAB plugin for automatic cleaning of whole-night sleep electroencephalogram using Artifact Subspace Reconstruction

Richard Somervail[1,2,*], Jacinthe Cataldi[3,4], Aurélie M. Stephan[3,4,5], Francesca Siclari[3,4,5,*], (iD) and Gian Domenico Iannetti[1,2,*], (iD)

[1]Neuroscience and Behaviour Laboratory, Italian Institute of Technology (IIT), Rome, Italy,
[2]Department of Neuroscience Physiology and Pharmacology, University College London (UCL), London, UK,
[3]Centre d'Investigation et de Recherche sur le Sommeil, Centre Hospitalier Universitaire Vaudois (CHUV), Lausanne, Switzerland,
[4]The Sense Innovation and Research Center, Lausanne and Sion, Switzerland,
[5]Netherlands Institute for Neuroscience, Amsterdam, The Netherlands
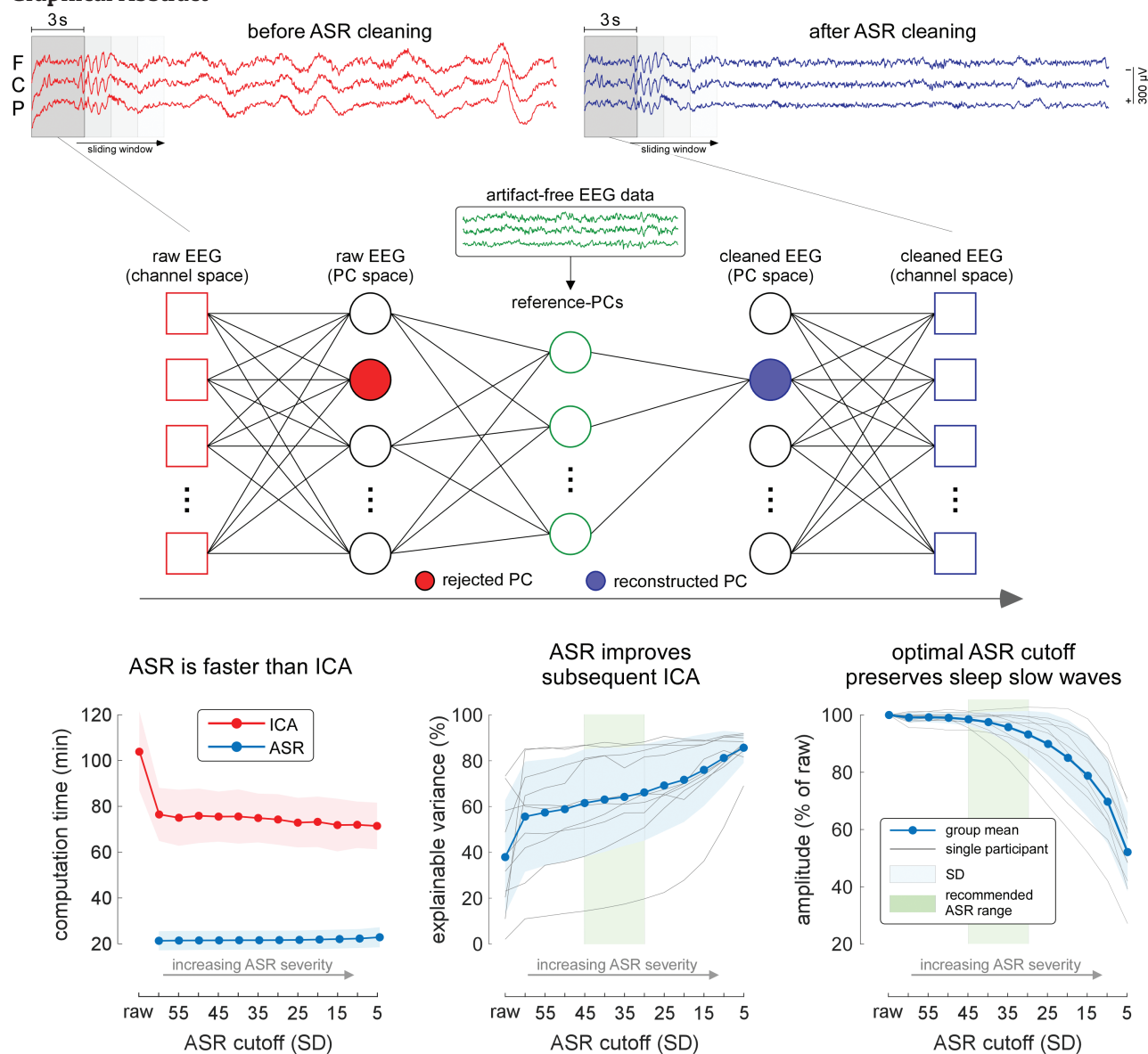
*Shared senior authorship.
*Corresponding author. Richard Somervail, Neuroscience and Behaviour Laboratory, Via Regina Elena 291, 00161 Rome, Italy. Email: r.somervail@gmail.com.

Abstract

Whole-night sleep electroencephalogram (EEG) is plagued by several types of large-amplitude artifacts. Common approaches to remove them are fraught with issues: channel interpolation, rejection of noisy intervals, and independent component analysis are time-consuming, rely on subjective user decisions, and result in signal loss. Artifact Subspace Reconstruction (ASR) is an increasingly popular approach to rapidly and automatically clean wake EEG data. Indeed, ASR adaptively removes large-amplitude artifacts regardless of their scalp topography or consistency throughout the recording. This makes ASR, at least in theory, a highly-promising tool to clean whole-night EEG. However, ASR crucially relies on calibration against a subset of relatively clean "baseline" data. This is problematic when the baseline changes substantially over time, as in whole-night EEG data. Here we tackled this issue and, for the first time, validated ASR for cleaning sleep EEG. We demonstrate that ASR applied out-of-the-box, with the parameters recommended for wake EEG, results in the dramatic removal of slow waves. We also provide an appropriate procedure to use ASR for automatic and rapid cleaning of whole-night sleep EEG data or any long EEG recording. Our procedure is freely available in *Dusk2Dawn,* an open-source plugin for EEGLAB.

Key words: EEG cleaning; EEG artifacts; EEG preprocessing; EEG analysis; event-related potentials; NREM-REM cycles; slow wave sleep; electrophysiology; neuroscience

Graphical Abstract

Statement of Significance

Whole-night sleep electroencephalogram (EEG) recordings are plagued by large-amplitude artifacts. Common approaches to remove these artifacts, such as rejection of noisy time intervals or channel interpolation, are time-consuming and result in signal loss. These problems become prohibitive with the recent trend towards large datasets comprising thousands of whole-night recordings. Artifact Subspace Reconstruction (ASR), widely popular in wake EEG, rapidly and automatically removes rare artifacts in long EEG recordings, and therefore seems particularly appropriate for whole-night EEG cleaning. However, its effectiveness for cleaning sleep EEG has never been systematically validated. Here we perform this validation and explain how to effectively clean whole-night EEG data with ASR. We also developed Dusk2Dawn—a freely-available open-source ASR plugin for the popular EEGLAB toolbox in MATLAB.

Introduction

Recording the electroencephalogram (EEG) has been central to understanding sleep, particularly its stages and disturbances [1–4]. Whole-night EEG data are however plagued by several types of large-amplitude artifacts, such as those caused by awakenings, whole-body movements, eye movements, sensor displacements and disconnections [5–7]. While simple approaches to automatically detect noisy data based on amplitude thresholds are popular in wake EEG, they are not suitable in sleep given that amplitude varies considerably across sleep stages [1–4, 8]. Consequently, methods that adapt the artifact detection threshold according to

sleep stage have been developed [9]. However, these artifacts are often removed manually via channel-wise interpolation or rejection of noisy segments of data, time-consuming processes resulting in lost brain signal [6, 7, 9–11]. This issue becomes even more pressing given the current trend towards exploring large datasets encompassing thousands of polysomnographies [12, 13]. For all these reasons there is an urgent need for an automatic, fast, and reliable approach to clean sleep EEG data.

The current gold standard for (semi)automatic cleaning of wake EEG data is independent component analysis (ICA). ICA is a spatial filter that removes artifacts of fixed scalp topography occurring regularly throughout the data, such as those due to eye movements and muscle contractions [14]. Conversely, ICA is much less appropriate to clean sleep EEG. The reason is that ICA is not effective at isolating rare, non-stationary artifacts or artifacts with variable scalp topography [14], such as those caused by unique, idiosyncratic movements occurring during sleep. As a result, aggressive manual removal of inherently ambiguous independent components (ICs) is often required to denoise the data—a process that, besides being time-consuming, results in unavoidable signal loss. Finally, traditional polysomnographic (PSG) sleep recordings only use 4–6 electrodes, which makes ICA cleaning unfeasible.

An increasingly used algorithm to clean EEG automatically is "Artifact Subspace Reconstruction" (ASR) [15–18]. ASR is becoming popular for its ability to effectively and automatically remove large-amplitude artifacts, regardless of their scalp topography or consistency throughout the data [17, 19–22]. ASR is widely used to clean wake EEG, complementing and often replacing standard ICA cleaning [23–30]. As the method can effectively clean even the noisiest EEG segments, ASR allows the user to retain and analyze continuous sleep EEG containing large artifacts associated with arousals and awakenings [31], rather than rejecting these segments. Although developed for high-density EEG (>64 electrodes [17]), recent work has demonstrated effective ASR cleaning with as few as 4–8 electrodes [32, 33] suggesting that the technique is suitable even for the low-density EEG montages used in clinical polysomnography.

Although these ASR features seem highly appropriate to deal with whole-night EEG data, the effectiveness of ASR in cleaning whole-night EEG has never been systematically validated. This neglect is important, given that ASR is starting to be used in sleep, but with the parameters recommended for wake EEG [31, 34–37]. This poses an urgent issue because this approach is likely to give incorrect results due to the important structural differences between wake and sleep EEG. For example, sleep graphoelements (e.g. slow waves [SWs], including delta-waves and K-Complexes [KCs] [38]) are far larger than those observed in wake. As a result, a typical ASR cutoff of 20–30 SD may result in the removal of SWs. Additionally, both the frequency of occurrence and the amplitude of these graphoelements vary substantially across sleep stages [38]. This makes standard ASR inappropriate because ASR requires a crucial calibration step in which relatively artifact-free data are used to define thresholds for the identification of artifactual EEG components to be removed [17]. As such, ASR calibrated with data coming from multiple sleep stages may lack the sensitivity to remove smaller-amplitude artifacts in REM sleep and, perhaps worse, may remove large-amplitude SWs characteristic of non-REM sleep.

To address these issues, we applied ASR to whole-night EEG recordings using a variety of parameters and approaches. We quantified the effects of a wide range of ASR cutoff values on SW amplitude and several other metrics. In Analysis 1, we tackled the problem of calibration by splitting the whole-night recording by sleep stage, before applying ASR independently to each stage. In Analysis 2, we implemented a more sophisticated ASR calibration procedure: segmenting the data into chunks of arbitrary length (from 2 min to the whole night) regardless of sleep stage and applying ASR to each chunk separately. Running the ASR on the whole-night recording likely results in inappropriate calibration across different sleep stages, while running the ASR on shorter chunks of data should minimize this problem and improve cleaning.

We used these results to define two cleaning pipelines appropriate for sleep, which we make freely available as a plugin for EEGLAB (*Dusk2Dawn*; available from github.com/rsomervail/dusk2dawn). The plugin includes an accessible graphical interface to allow users to apply these pipelines to their data and validate the results themselves.

Materials and Methods
Whole-night EEG data

We used approximately 70 h of high-density, whole-night EEG data from two previous studies [39, 40]. EEGs were recorded using a 256-channel EGI system sampled at 500 Hz from a whole night of natural sleep (with no interruptions or stimuli) in 10 healthy participants. Sleep stages were scored manually using consecutive 30 s epochs according to standard criteria [41] by author JC, and classified as belonging to one of the following five stages: wake, REM, N1, N2, or N3. Before ASR, EEG data were first re-referenced to the average of both mastoids. The data were then bandpass filtered between 0.5 and 40 Hz using a minimum-phase finite impulse response (FIR) filter, and the DC component was removed from each channel to ensure zero-mean signals (a prerequisite for both ASR and ICA decomposition).

The Artifact Subspace Reconstruction (ASR) algorithm

ASR relies on a preliminary calibration of segments of relatively clean baseline EEG data. These segments of baseline data can either be found automatically or chosen manually. Automatic identification is based on channel-wise signal variance [20]. Specifically, the algorithm computes the z-scored root-mean-square values (in 1 s windows) and checks that these z-scores fall within the –3.5 to 5.5 interval in more than 92.5% of channels [20].

Principal Component Analysis (PCA) is subsequently performed on these baseline data, and the means and standard deviation (SD) of the root-mean-square timecourses of the resulting "clean" PCs are used to define a threshold for each of these components [17, 20]. ASR then performs a series of PCA decompositions on small successive EEG segments of the data to be cleaned (these segments are typically 0.5 s, but their width must be larger than 1.5 × number of channels /sampling rate; see Table 1). In this second step, the PC thresholds defined in the initial baseline step are first projected onto the PC subspace of the segment to be cleaned. Then, the variance of the resulting PCs from the data to be cleaned is compared to the PC thresholds defined in the initial baseline step: if a PC exceeds those thresholds by a user-defined factor (e.g. 20 SDs) that PC is removed from the data. The non-artifactual underlying signal is finally reconstructed using the remaining components (i.e. the PCs not exceeding the thresholds identified in the initial baseline PCA performed on clean

Table 1. Summary of parameters used in each Analysis.

ASR parameters	Analysis 1	Analysis 2
Cutoff (SD)	{5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60}	{10, 20, 30, 40, 50, 60}
Max dimensions	2/3	2/3
Window length (s)	3	3
Step-size (samples)	768 (window length/2)	768 (window length/2)
Reference max bad channels	0.075	0.075
Reference tolerances (SD)	−3.5, 5.5	−3.5, 5.5
Reference window length (s)	2	2
Use Riemmanian modification	False	False
Use GPU	True	True
Max memory	5 GB	7 GB
Chunk length (min)	n/a	{2, 4, 8, 16, 32, 64, whole-night}
Chunk overlap (min)	n/a	1

data) and, crucially, the mixing matrix derived from the first PCA decomposition of the baseline data. Note that this two-layered structure of the algorithm, in which the data are reconstructed from both the remaining PCs and the mixing matrix from the initial PCA on the baseline data, means that ASR is not merely a sliding-window PC rejection algorithm. A more formal description of the function of the ASR algorithm is provided by Chang et al. (2020) [20].

ASR pipelines

We used custom MATLAB functions as well as the *clean_rawdata* EEGLAB plugin to apply two distinct ASR pipelines in Analyses 1 and 2, respectively. In each Analysis, we tested a range of ASR cutoffs (Analysis 1: from 5 to 60 SD, in increments of 5 SD; Analysis 2: from 10 to 60 SD, in increments of 10 SD; note that the lower the number the more severe the cutoff). In Analysis 1, we split the whole-night EEG data by sleep stage and performed ASR and validation separately on each of the three tested stages (N2, N3, and REM). Calibration data were chosen automatically by the ASR algorithm [20], and their average length across participants was 54 ± 9% (45–69%) [N2], 53 ± 9% (45–67%) [N3], and 53 ± 8% (45–66%) [REM]. In Analysis 2 we did not split the EEG data by sleep stage, but instead performed ASR on overlapping chunks ranging from 2 min to the whole night (six chunk lengths tested; 1 min overlap) without respecting sleep stage boundaries, and then averaged the resulting signals in the overlapping regions to reduce boundary discontinuities [42]. Note that these chunks do not refer to the sliding window of the ASR algorithm itself; rather the entire process of ASR calibration and cleaning was performed separately on each chunk. This approach has been used effectively to clean wake EEG with embedded episodes of microsleep [42–44] and has been shown to perform better than standard ASR with several wake EEG datasets [45]. The resulting cleaned data were finally split into sleep stages as in Analysis 1, and validation was performed on each stage separately. In both pipelines, we used an ASR window length (i.e. the window length of the ASR algorithm itself) of 3 s. This value is longer than what is typically used in wake (e.g. 0.5–1 s), and we chose it as a result of a preliminary analysis in which we varied its length systematically, under the reasoning that these typical window lengths were likely too short for sleep slow waves (Supplementary Figure S1). All other ASR parameters were kept at their default values (see Table 1).

Validation of ASR cleaning

In Analysis 1, we explored the effects of ASR on the quality of a subsequent ICA decomposition of the data. This was done to (1) quantify how effective ASR was at removing artifacts that are difficult to isolate with ICA, and (2) assess the ease of manual selection of artifactual ICs should the user choose to perform a subsequent ICA to remove remaining artifacts (e.g. muscle artifacts). This approach has been used previously to assess the performance of ASR cleaning in wake [20, 22]. We, therefore, computed ICs (using the *runica* algorithm implemented in EEGLAB) for each dataset before and after ASR. For each ICA output, we automatically labeled each of the resulting ICs using the *IClabel* plugin for EEGLAB [20, 46] as one of four categories (neural activity, physiological artifacts, recording artifacts, unknown). Finally, we calculated both the percentage of ICs belonging to, as well as the total variance explained by, each category.

In Analysis 1 we also explored the effect of ASR on brain oscillations and noise within different frequency bands. To this end, we computed the spectral decomposition using the Fast Fourier Transform, for each channel and sleep stage, before and after ASR. The resulting amplitudes were summarised as the mean of each of the following frequency bands: 1–4 Hz (delta), 4–8 Hz (theta), 8–12 Hz (alpha), 12–16 Hz (sigma), 18–30 Hz (beta), and 30–45 Hz (gamma). These amplitudes were finally averaged across channels and expressed as a percentage of the amplitude in the raw data. We defined the delta band as 1–4 Hz (rather than the 0.5–2 Hz range often used in clinical sleep research) as this range includes the frequency of K-complexes. Also, the 0.5–1 Hz frequency content is partly removed by the 0.5 Hz high-pass filter applied before ASR (due to the 0.25–0.75 Hz transition band).

In both Analysis 1 and 2, we explored the effects of ASR on the amplitude of large sleep graphoelements. To this end, we first extracted SWs from a version of each EEG dataset cleaned manually, by interpolating bad channels and performing ICA. SWs were identified using a validated algorithm as negative deflections between two consecutive zero-crossings separated by 0.25–1 s [31, 38, 47, 48]. We then used the timestamps of these validated SWs to isolate the same SW events in both raw and ASR-cleaned data. We quantified the SW amplitude by averaging all identified SWs and taking the mean of a 100-ms long time window centered on the negative peak. To quantify SW consistency (i.e. a proxy of

Figure 2. **Effects of ASR on subsequent ICA decomposition (Analysis 1).** Each IC was labeled automatically using the EEGLAB IClabel plug-in. The top plots of each panel show the total variance explained by identifiable ICs (left) and each IC category (right), for each ASR cutoff. The bottom plots of each panel show the percentage of identifiable ICs (left) and of ICs belonging to each category (right). Gray lines show individual participants and colored lines show the group means. Shaded areas represent SD across participants. Green boxes highlight

without also removing neural activity. Thus, ASR was not effective at totally removing artifacts caused by eye movements or single noisy electrodes.

Increasing the ASR severity also reduced the percentage of "unknown" ICs, although this effect was driven by increasing not only the percentage of "neural activity" ICs but also "recording arti-facts" ICs (espe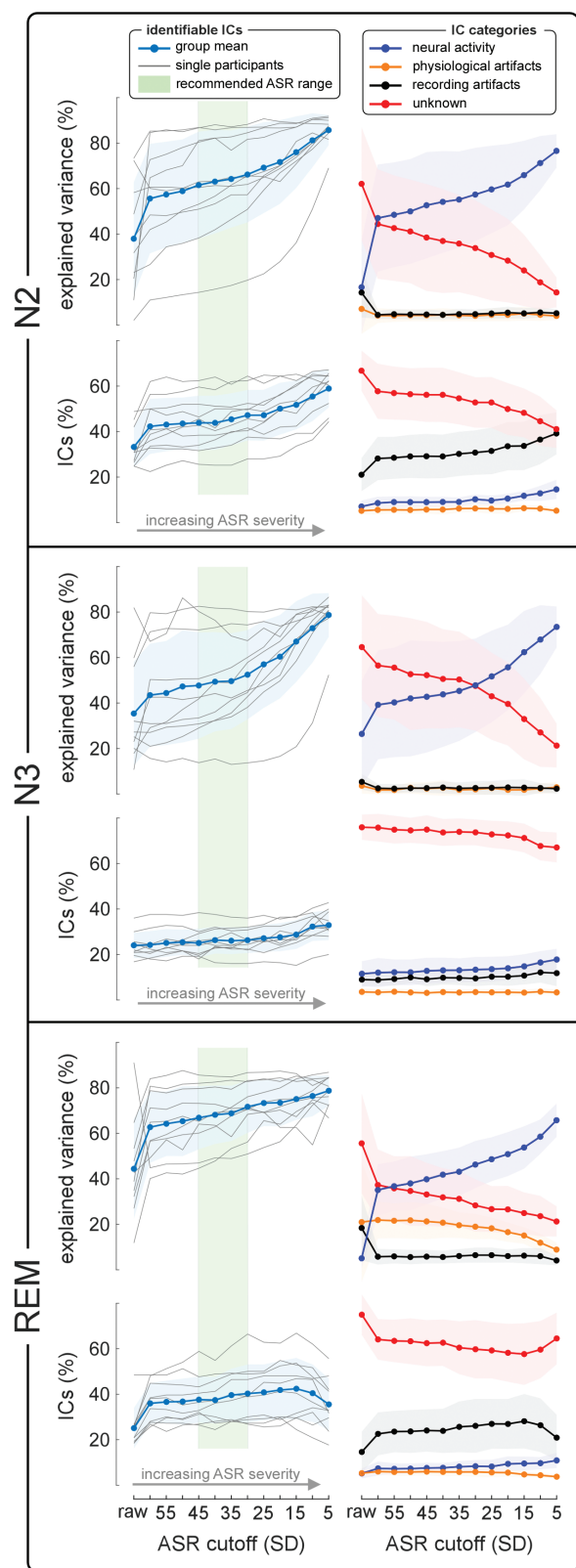cially in N2 sleep). This result indicates that, although the largest among these artifacts were removed by ASR, those that remained were more effectively isolated with ICA. This observation can also be explained by reduced ICA underfitting, due to artifacts consequent to noisy electrodes being more effectively isolated as separate ICs due to the lower dimensionality of the data.

ASR and ICA computation times

Figure 3 shows the time necessary to run ASR and ICA at each tested cutoff value. In all sleep stages and regardless of cutoff, ASR was faster than ICA by a factor of 2–5, taking only ~21 min for the whole N2 stage, and ~15 min each for N3 and REM. This shows that ASR does not add substantial time to the EEG pre-processing pipeline. This difference was more pronounced for the N2 stage, likely due to its longer length. Expectedly, when using the standard ASR window of 0.765 s (Supplementary Figure S2), ASR computation times were slightly faster than using a window of 3 s, albeit at the cost of worse SW reduction (Supplementary Figure S1). In general, cleaning with ASR also reduced the com-putation time of the subsequent ICA, with the only exception of an increase of computation time below ASR cutoffs of 30 SD in the REM stage of one participant. Note that the ICA computa-tion times shown in Figure 3 and Supplementary Figure S2 do not include the extensive additional time necessary to manually select the ICs to be removed from the data.

Effects of ASR on brain oscillations

Figure 4 shows, for each sleep stage and ASR cutoff, the change of power compared to raw data. ASR primarily removed power in the delta band (1–4 Hz) across sleep stages. In REM the removal of delta power was particularly dramatic (~50%) even at very mild levels of ASR cleaning (e.g. a cutoff of 60 SD; Figure 4, right plot), likely because REM delta power mostly reflects large artifacts. All other frequency bands were removed by 5–10% across sleep stages.

Effects of ASR on slow waves

To test the selectivity of ASR toward removing artifacts, we com-pared the amplitude and consistency of SWs before and after ASR. Figure 5 shows SW amplitude and consistency at each ASR cutoff value, separately for each sleep stage.

Applying ASR with the default cutoff of 5 SD resulted in a dra-matic reduction of SW amplitude in both N2 and N3 (Figure 5, top row). Even the fairly mild cutoff recommended for wake data

the recommended ASR cutoff range (see Discussion). Note how, in all sleep stages, increasing ASR severity drastically increased the variance explained by identifiable ICs. This effect was largely driven by the increased variance explained by "neural activity" ICs at the expense of "unknown" ICs. ASR also removed variance from the "physiological artifacts" and "recording artifacts" categories, especially in REM. These results show that ASR effectively cleans the data by removing large and unique artifacts with highly-variable scalp topographies and thereby reducing ICA underfitting. Consequently, ICA isolates better the remaining neural activity and artifacts.
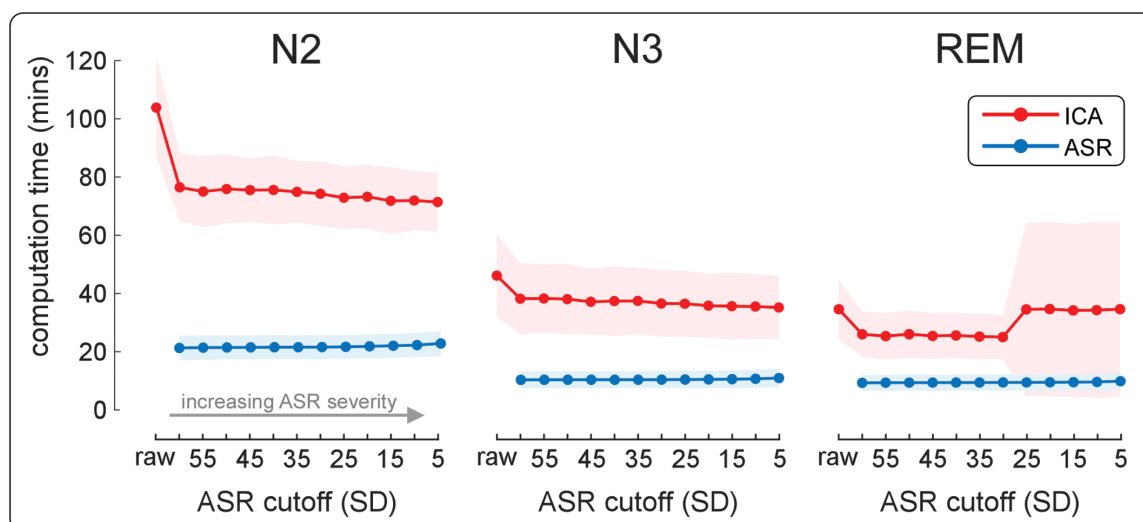
Figure 3. Computation times for ASR and ICA (Analysis 1). Average computation times for ICA (red) and ASR (blue) for each tested cutoff and sleep stage. Shaded areas represent the standard deviation across participants. ASR was 2–5 times faster than ICA at every cutoff and sleep stage. This difference was most pronounced in N2. Note that computation times shown here reflect 255-channel EEGs, and are therefore much longer than for most EEG recordings. Also, note that the increase of ICA computation time in REM at ASR cutoffs more severe than 30 SD was driven by a single participant.
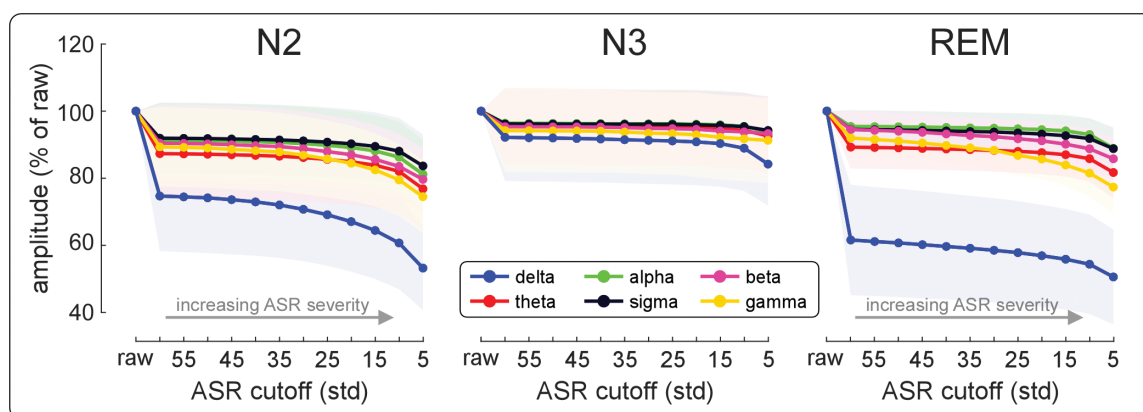

Figure 4. Effects of ASR on brain oscillations (Analysis 1). The amplitude of EEG frequency bands across ASR cutoffs is expressed as a percentage of raw data amplitude. In all sleep stages, ASR mostly reduced delta power (1–4 Hz), corresponding to large amplitude artifacts and slow waves. The removal of delta power in REM was especially dramatic (~50%) even at very mild levels of ASR cleaning (e.g. a cutoff of 60 SD), likely because in REM delta power mostly reflects large artifacts. All other frequency bands were removed by 5–10% across sleep stages.

(20–30 SD [20]) resulted in a substantial reduction of SW amplitude in N2. At more liberal cutoffs (>30 SD), however, SW amplitude was largely preserved, and the small decrease in amplitude at the highest cutoffs (50+ SD) was likely due to the removal of large-amplitude artifacts distorting the SW means.

Despite the reduction of SW amplitude with progressively more severe ASR cutoffs, their consistency increased (Figure 5, bottom row). Visual inspection suggests that this consistency increase was driven by the removal of positive amplitude artifacts coinciding with the negative SW peak (Figure 5, insets). In other words, the SW signal-to-noise ratio was increased after ASR despite the overall reduction of SW amplitude. However, at the most severe cutoffs (e.g. <30 SD) this SNR increase came at the cost of a more substantial SW removal.

Importantly, the amount of amplitude reduction varied across SWs: some individual SWs were entirely removed despite a fairly small reduction of the average SW amplitude. This finding is illustrated in Figure 6, which shows SWs from a single participant. Note

how the SW average is progressively reduced in amplitude with increasing ASR severity, while the exemplary individual SW abruptly halved its amplitude and then disappeared at ASR cutoffs of 5 and 25 SD, respectively. This indicates that even at mild cutoffs such as 45 SD, individual SWs can be dramatically reduced, while their across-trial average amplitude is only reduced by ~10–20%. This is problematic, given that individual SWs are far more informative than average SW amplitude, as they can reflect spatial and temporal dynamics throughout the night of sleep [38, 49–51].

Analysis 2—Calibration of ASR within a sliding window vs whole-night ASR

Length of ASR calibration data

In Analysis 2, we tackled the issue of calibration by performing ASR on overlapping chunks of data. We explored a wide range of chunk lengths (from 2 min to the whole night) and measured the resulting length of calibration data against the recommended
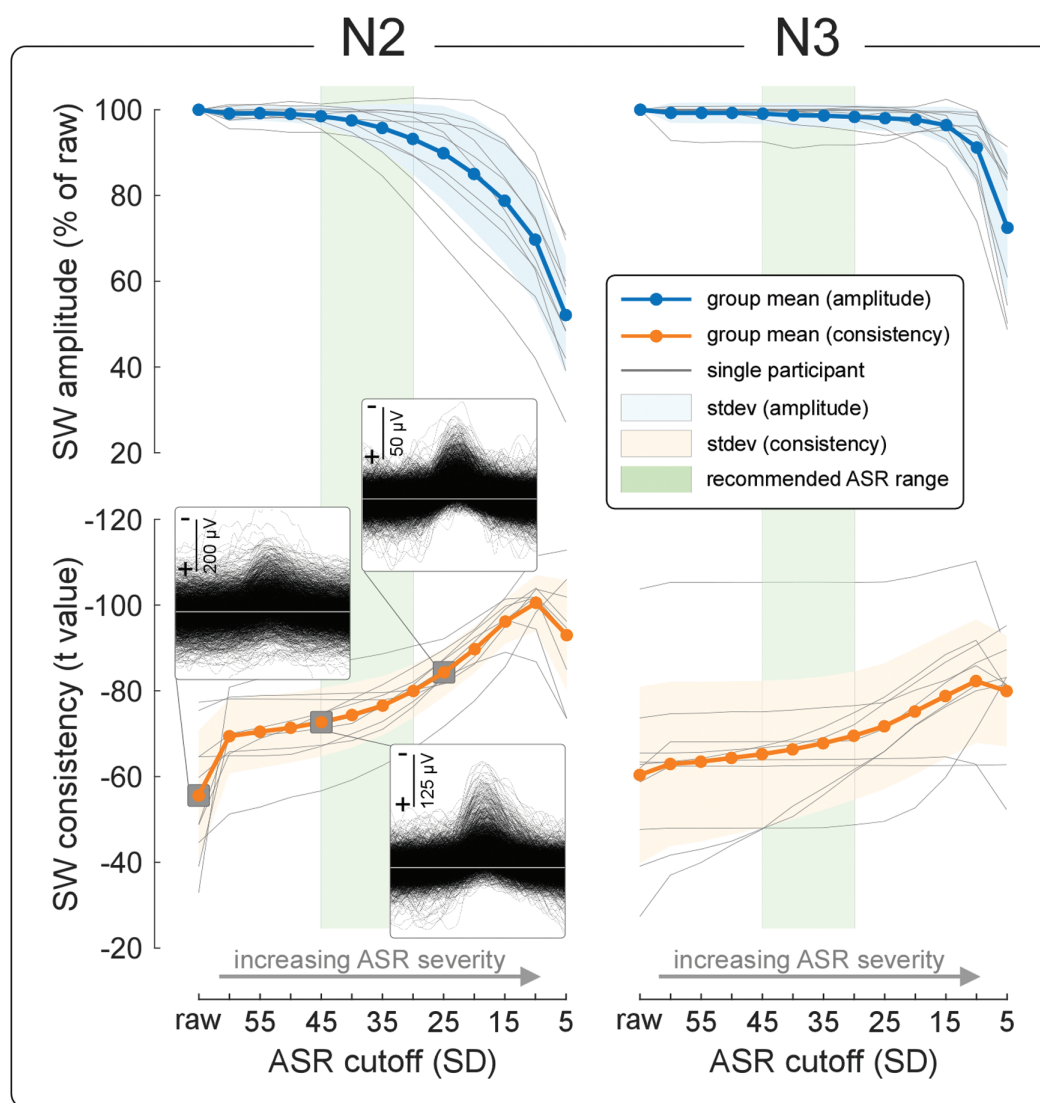
Figure 5. Effects of ASR on slow wave amplitude and consistency (Analysis 1). SW amplitude (top plots) and consistency (bottom plots) as a function of ASR severity. Colored lines show group means, and gray lines individual participants. Shaded areas show SD across participants. Green boxes indicate the recommended ASR range (see Discussion). Insets show representative single SWs from one participant. SW amplitude decreased non-linearly with increasing ASR severity. Very severe ASR resulted in substantial SW removal and loss of consistency in the N2 and N3 stages. However, at milder ASR severity (i.e. above 15 SD), SW consistency improved. Insets show that this was likely due to the removal of positive amplitude noise obscuring the negative SW peak. Thus, ASR improves the SW SNR, but at the cost of reducing the mean SW amplitude, especially at more severe ASR cutoffs.

duration of 1 min (see documentation of the *clean_rawdata* plugin). When using short chunks of 2–4 min, 14 ± 6% and 84 ± 9% of chunks (2- and 4-min long, respectively) did not provide >1 min of calibration data (Table 2). Thus, our results suggest that when calibrating using chunks of 2 or 4 min there is a danger of not obtaining sufficiently-long calibration data. This should be checked for each dataset.

Computation times of sliding-window ASR

In Analysis 2, the use of ASR on overlapping EEG chunks made computation times longer than in Analysis 1. Figure 7 shows the time necessary to run ASR for the whole night of sleep at each chunk length. Predictably, the smaller the chunk length, the longer the average computation time, from ~45 min to perform ASR on the whole night in a single chunk to ~130 min to perform ASR on all 2-min chunks with a 1-min overlap. Computation time

did not vary substantially between chunk lengths of 8 and 64 min. Despite the increase in computation time, sliding-window ASR was never longer to run than ICA run separately on N2, N3, and REM in Analysis 1 (~200 min; Figure 3).

Effects of sliding-window ASR on slow waves

In Analysis 2 we determined whether calibrating ASR using shorter sliding windows would improve its selectivity towards detecting artifacts, and thereby remove fewer SWs. We also explored how ASR on whole-night data compares to both the sliding window approach (Analysis 2) and the ASR performed separately on each sleep stage (Analysis 1).

Figure 8 shows the effects of sliding windows and whole-night ASR on SW amplitude and consistency. As in Analysis 1, mean SW amplitude decreased non-linearly with ASR severity, with smaller SW amplitudes for more severe cutoffs. In N2, there
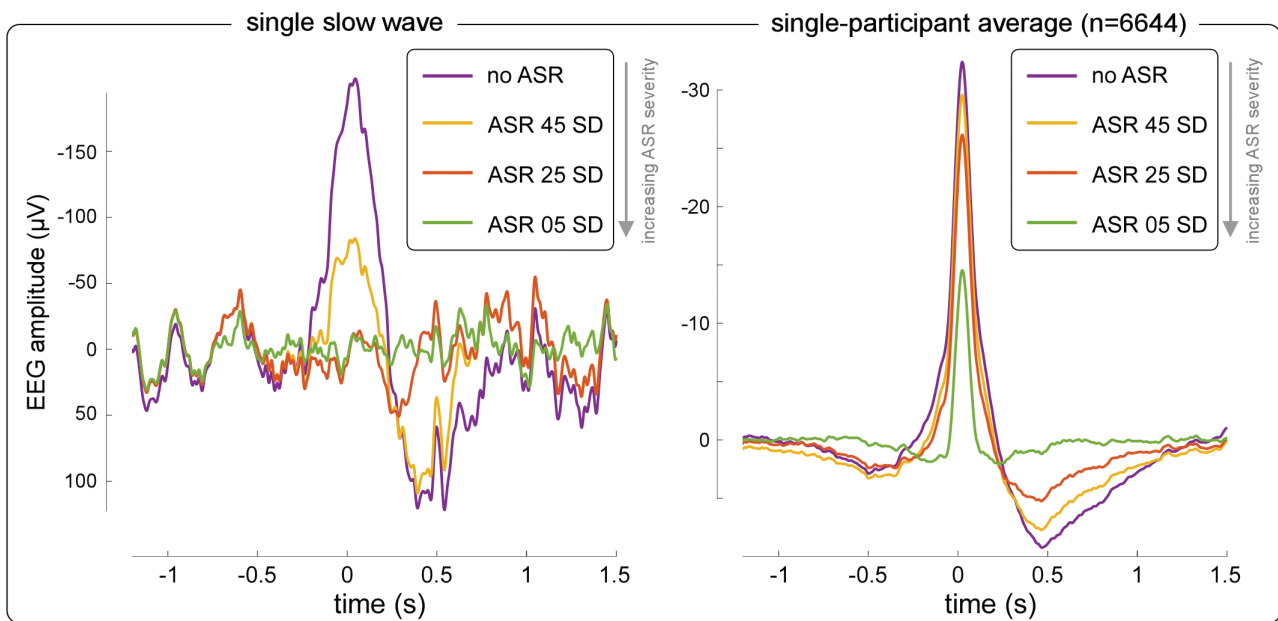
Figure 6. Exemplary single slow wave and average from the N2 stage of one participant (Analysis 1). Examples of individual (left) and average (right) SW waveforms from one participant, without ASR and at three levels of ASR severity (cutoffs: 5, 25, and 45 SD). Although the SW was still visible in the average even at the most severe ASR cutoff (right), an individual SW was already removed with less severe cutoffs (left). This sacrifice of individual SWs must be considered when aiming to perform single-trial SW analysis.

Table 2. Analysis 2—percentage of data chunks with an ideal minimum amount of clean data for ASR calibration (at least 1 min)

Subject	Chunk duration						
	Whole-night (%)	64-min (%)	32-min (%)	16-min (%)	8-min (%)	4-min (%)	2-min (%)
1	100	100	100	100	100	94	24
2	100	100	100	100	100	85	9
3	100	100	100	100	100	86	15
4	100	100	100	100	100	83	7
5	100	100	100	100	100	80	10
6	100	100	100	100	99	83	13
7	100	100	100	100	98	97	16
8	100	100	100	100	100	69	22
9	100	100	100	100	100	92	8
10	100	100	100	100	100	74	13
Mean	100	100	100	100	100	84	14
SD	0	0	0	0	1	9	6

was little variability in the relationship between SW amplitude and ASR cutoff across chunk lengths, probably because N2 comprised the majority of the data used for calibration, regardless of chunk length. In N3, however, the effect of chunk length was more drastic: when sliding-window ASR was applied to either the whole-night or longer chunks (i.e. 32–64 min) the SW amplitude reduction with increasingly-severe ASR cutoffs was more pronounced than for shorter chunks. This suggests that, in N3, running the sliding-window ASR with short chunks (2–16 min) resulted in better calibration than running the ASR for the whole night. Compared to Analysis 1, the ASR applied to short chunks resulted in a modest improvement of N2 and N3 SW amplitude preservation (compare Figure 5 with the 16- to 2-min chunk lengths columns of Figure 8).

Perhaps counterintuitively, in N3 ASR performed on either the whole-night or the longest chunk lengths resulted in enhanced SW consistency at middle cutoffs, compared to when ASR was performed separately by sleep stage (Analysis 1) or on shorter chunk lengths (Analysis 2). This increase in SW consistency mirrors the dissociation between SW amplitude and consistency observed in Analysis 1 (Figure 5). Specifically, in the whole-night and long-chunk ASR, the calibration was dominated by N2 sleep, which generally has lower-amplitude signals, and therefore ASR resulted in more severe thresholds for artifact removal at the same cutoff values. Consequently, a higher number of the larger SWs present in N3 were removed or reduced in amplitude, while artifacts of opposite polarity were also removed, leading to an overall SNR increase despite the
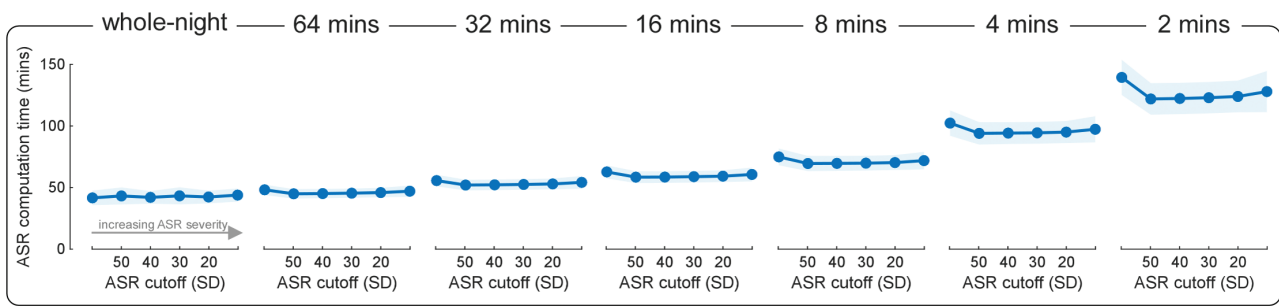
Figure 7. Computation times of sliding-window ASR (Analysis 2). Computation times for sliding-window ASR were performed with each chunk length. Shaded areas show standard deviation across participants. The smaller the chunk length, the longer the computation time, from ~45 min to perform ASR on a single whole-night chunk to ~130 min to perform ASR on 2-min chunks. At chunk lengths between 64 and 8 min, computation time did not vary substantially. Note that all computation times were measured for a 255-channel EEG dataset, while lower-density EEG data would entail much shorter computation times. Regardless of chunk duration, ASR was never longer to compute than ICA performed separately on N2, N3, and REM stages in Analysis 1 (~200 min; Figure 3).
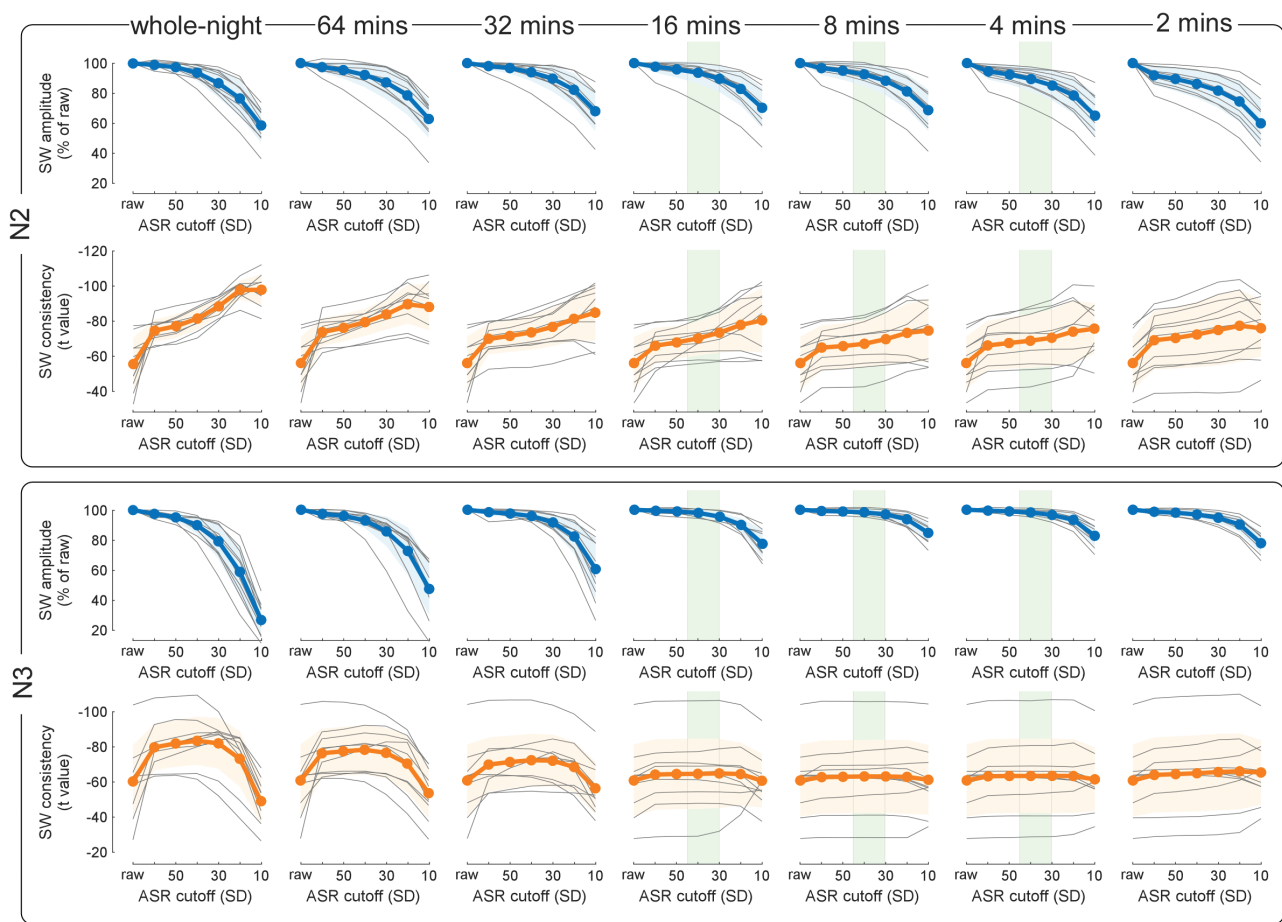

Figure 8. Effects of sliding-window ASR on slow wave amplitude and consistency (Analysis 2). Each panel shows SW amplitude (top rows) and consistency (bottom rows) after sliding-window ASR performed at different cutoffs (x-axis) and chunk lengths (columns). Colored lines show group means; gray lines show individual participants. Shaded areas show standard deviation (SD) across participants. Recommended ASR cutoff ranges are highlighted in green (see Discussion). As in Analysis 1, SW amplitude decreased non-linearly with increasing ASR severity. In N2, there was little variability in the relationship between SW amplitude and ASR cutoff across chunk lengths. In N3, the effect of chunk length was more drastic: when sliding-window ASR was applied to long chunks (e.g. 32–64 min) or the whole night, SW reduction with increasingly-severe ASR cutoffs was more pronounced compared to what was observed for shorter chunks. Thus, sliding-window ASR using shorter chunks yields a better calibration than running ASR on whole-night data. In N3, whole-night and long-chunk ASR enhanced SW consistency at middle cutoffs compared to shorter chunk lengths. This was caused by the calibration being dominated by the smaller voltage of N2, resulting in the removal of more noise but also of larger SWs.

reduction of SW mean amplitude (see insets of Figure 5). The results of Analysis 1 indicate that these smaller artifacts, whose amplitude is similar to that of the largest SWs, can be removed more selectively with ICA, and a more severe ASR cleaning consequent to using N2 as a source of baseline data is inappropriate to remove them.

Discussion

The study of electrical brain activity during sleep has produced a wealth of novel information about the function of sleep in health and disease. Many artifacts, however, contaminate whole-night sleep EEG. Their removal entails time-consuming user-dependent decisions and loss of potentially meaningful data. In this study, we explored the effects of an increasingly popular EEG cleaning method—ASR [17]—on whole-night sleep datasets. We describe an ASR implementation that allows a rapid, efficient, and unbiased cleaning of EEG artifacts, with minimal loss of sleep SWs. With the appropriate parameters, ASR adaptively removes artifacts that would otherwise demand the rejection of whole segments of data. These artifacts would be poorly isolated by ICA, the most frequently used approach for cleaning high-density EEG data.

Preliminary ASR drastically improves subsequent ICA cleaning

Traditional ICA-based cleaning of whole-night EEG is a laborious and time-consuming process, entailing a large number of subjective decisions [14]. This user dependence is a large source of variability in brain imaging data, and efforts to move towards more automatic preprocessing are increasing across fields and neural measuring techniques [9, 17, 35, 52, 53]. We demonstrate that when preceded by ASR, ICA decomposition yields ICs more easily classifiable as neural activity, physiological artifacts, and recording artifacts (Figure 2). This vast improvement in IC classifiability suggests that after preliminary ASR, ICA cleaning could be entirely automated. The two methods complement each other well, with the initial ASR removing the large and non-stationary artifacts, and the subsequent ICA removing smaller and more consistent artifacts caused by, for example, muscle activity, eye movements, and individual noisy electrodes. Given that after ASR the classification and removal of artifactual ICs can be easily automated, another advantage of ASR is the substantial time saved on difficult and inherently ambiguous user decisions. Ideally, automatically labeled ICs should be reviewed manually, although this procedure becomes far easier and more rapid when preceded by ASR cleaning. Unsurprisingly, ASR and ICA are starting to be combined in both wake [51, 52] and sleep [31, 32].

ASR must be performed before, not instead of or after ICA cleaning

We have demonstrated that ASR is a good complement to ICA. We note that ASR must not follow ICA, as the IC removal will reduce the data rank and interfere with the ASR algorithm. Furthermore, we do not recommend applying ASR alone, as many artifacts such as muscle activity, eye movements, or individual noisy channels, are not removed using the ASR cutoff we recommend (30–45 SD; see the following section), and even with more extreme cutoffs of 5–15 SD [the reader should remember that the smaller the cutoff, the more severe the ASR]. The application of ASR using severe cutoffs to avoid the subsequent ICA would not only fail to effectively remove small and consistent physiological and recording artifacts but would also result in a substantial signal loss (see following section). Although these artifacts only explain a small portion of the total variance (Figure 2), likely due to the high number of EEG electrodes and the choice of GFP as a variance calculation method, they can still obscure or distort many sleep-related signals of interest.

Preservation of sleep graphoelements relies on the correct choice of ASR cutoff and calibration data

Given that the range of amplitudes of physiologically-meaningful signals in sleep is far larger than in wake, we hypothesized that the ASR cutoffs used in wake may not be appropriate for preserving large sleep graphoelements such as SWs. When considering this issue it is important to remember that ASR cutoffs are expressed *relative to clean baseline data used for calibration*. Thus, both factors (the ASR cutoff and the amplitude of calibration data) determine the final absolute thresholds to identify artifacts. Since there is a vast amplitude variability both between and within sleep stages [1–3, 38], the choice of both factors must be carefully considered to prevent signal loss. We discuss the effects of these two factors in the following sections.

ASR cutoffs recommended for wake are unsuitable to process sleep EEG

Regardless of the choice of calibration data, the ASR cutoffs recommended for wake EEG (20–30 SD [20];) resulted in dramatic SW removal in both N2 and N3 sleep stages (Figures 5 and 8). This severe cleaning improved SNR (indexed by SW consistency; Figures 5 and 8) but at the cost of losing the larger SWs (see insets in the left panel of Figure 5). This SW sacrifice was, however, unnecessary since the small artifacts removed by ASR with the wake cutoffs could have been selectively removed by the subsequent ICA (see section *ASR must be performed before, not instead of or after ICA cleaning*). Thus, the ASR cutoffs of 20–30 SD recommended for wake should never be used when studying the large graphoelements of non-REM sleep. As detailed later, we recommend using milder cutoffs of 30–45 SD in non-REM sleep to optimally remove artifacts while preserving SW amplitudes. Conversely, ASR cutoffs closer to wake-like severity (e.g. 20–30 SDs) can be chosen to clean REM sleep.

More temporally local ASR calibration mitigates the removal of slow waves

As discussed above, the choice of calibration data also determines the thresholds to identify artifactual data. Consequently, applying ASR to a whole-night sleep dataset presents an issue: since there is substantial variability in the amplitudes of brain signals across different sleep stages, the thresholds determined from the calibration data may not be appropriate when applied to the whole night. To deal with this issue we explored two different strategies that varied how temporally-local the calibration data was concerning the data being cleaned. In Analysis 1, we applied ASR separately to each sleep stage. In Analysis 2, we applied ASR separately to overlapping chunks of data (duration ranging from 2 to 64 min), as well as to the whole night.

In N3, the SW loss was most pronounced when ASR was applied to the whole night or long chunks of 32–64 min (Figure 8, left columns), and it was substantially reduced when the calibration was performed either separately for that stage, or the whole night but using shorter chunks of 2–16 min. Conversely, the difference in SW loss between the one-calibration-for-each-stage approach of Analysis 1 and the shorter-chunks approach of Analysis 2 was small, with the short-chunk calibration resulting in only slightly better SW preservation (compare Figures 5 and 8). This improvement was due to the more temporally-local calibration (i.e. a calibration performed using baseline data closer in time to the time interval to clean)—which better accounts for the variability of EEG power within and between individual sleep

cycles. The impact of power variability *between* individual cycles of each given sleep stage can be ameliorated by simply running ASR on each individual cycle of the stage, while reducing the impact of power variability *within* each individual cycle of a stage is exclusively a benefit of the shorter-chunk method.

In N2, in contrast, the different options for choosing the calibration data had little effect on SW loss. This observation is explained by the fact that calibrating using either the whole night or long chunks implies that the calibration data are dominated by the N2 stage.

The SNR of SWs in stage N3 was improved using more severe cutoffs, as well as when ASR was calibrated using the whole night. However, this was again merely due to inappropriate calibration given that whole-night EEG is dominated by the N2 stage. Since N2 SWs have a smaller amplitude than N3 SWs, the resulting ASR rejection thresholds were lower, and more of the larger N3 SWs were removed.

In summary, the choice of calibration data should be made considering the large EEG variability throughout the night, to make ASR cleaning as adaptive and reliable as possible. This can be achieved either by running ASR on relatively short and overlapping chunks of data (4–16 min long) or separately for each sleep stage. Both approaches are preferable to running the ASR on the whole-night EEG.

Summary of recommended ASR parameters

Considering all the above, when using ASR to clean whole-night EEG we recommend using (1) a cutoff between 30 and 45 SD, i.e. much less severe than in wake, together with (2) a more temporally-local calibration, either applying ASR separately to each sleep stage or using a sliding window (4–16 min long). The temporally-local calibration yields more consistent results across sleep stages with the same cutoffs. When using the sliding-window method it is important to ensure sufficiently long calibration data (ideally >1 min). The recommended range of 30–45 SD strikes a good balance between effective artifact removal and sufficient SW preservation (Figures 2, 5, and 8).

Within this range, the deciding the best ASR cutoff depends on the physiological or clinical question. For example, should the assessment of SW amplitude be of particular importance, we suggest using milder cutoffs, which will still clean the largest EEG artifacts while leaving the SWs relatively unaffected (Figures 2 and 5). This reasoning is particularly relevant when specifically investigating very large amplitude SWs like KCs: compared to smaller types of SWs, KCs will be removed by milder ASR cutoffs. Furthermore, given the variability in SW amplitude across stages and participants, we emphasize the need to test the effects of a range of ASR parameters on an EEG sample containing known SW events (identified either manually or by using a well-validated algorithm). More generally, we recommend testing and visualizing the effects of ASR on a subset of data to ensure the balance of signal loss and noise removal is appropriate for each given research question.

We finally note two points. First, although all results presented here are based on data recorded from healthy adult participants, they also apply to the majority of sleep disorders. Indeed, only in individuals with drastic variations of EEG signal-to-noise ratio, such as children or patients with encephalopathies or focal brain lesions, the ASR performance will differ from our results. In these cases, we recommend tuning the cutoff to balance the preservation of large-amplitude brain signals and the removal of noise. Second, in addition to the parameters tested here (i.e. sliding window length, ASR cutoff, ASR window length) several other parameters (such as the tolerance thresholds for calibration data selection) could affect the ASR cleaning of sleep data, and specifically the removal of SWs. Although testing these other parameters is beyond the scope of the current work, users can exploit the flexibility of *Dusk2Dawn* to tweak them and thereby achieve optimal cleaning, as described in the following paragraph.

Dusk2Dawn—an open-source EEGLAB plugin to apply ASR to whole-night EEG data

To allow easy implementation of ASR with whole-night sleep data (orany long-duration EEG data, see below), we developed *Dusk2Dawn*, an open-source EEGLAB plugin with a simple and intuitive graphical interface (freely available from: github.com/rsomervail/dusk2dawn and the EEGLAB plugin downloader). The plugin contains a set of wrapper functions for the original ASR functions, allowing the user to run ASR using either of the two pipelines described above (1) splitting the data by sleep stage using imported stage scoring events (as in Analysis 1), or (2) blindly running the ASR in successive overlapping chunks (as in Analysis 2). A major advantage of this plugin is that it allows the user to test a range of parameters (with a maximum of three explored dimensions for any one run-through) and then validate the results of each run-through by testing the effects of ASR on (i) subsequent ICA decomposition, (ii) SW amplitude and consistency, and (iii) spectral power of cleaned data. Due to the flexibility of *Dusk2Dawn*, the user can explore a range of ASR parameters without requiring the coding expertise or time necessary to program custom pipelines manually. The plugin can also be extended to support custom validation functions, for instance allowing the user to explore the effects of ASR on sleep spindles. Importantly, the flexibility provided by this plugin makes it useful to clean *any* long-duration recordings comprising a variable range of brain states and signal-to-noise levels, such as whole-day EEG or electrospinogram (ESG) collected during naturalistic experiments entailing subject movement [54].

Conclusions

ASR can be a powerful tool for the automatic and rapid cleaning of whole-night sleep EEGs. However, given the large variability of amplitude throughout the night, ASR must not be used with the parameters recommended for wakefulness. We demonstrated a procedure that makes ASR suitable for sleep data, using milder cutoff values of 30–45 SD, and finding an appropriate baseline by separating the data into smaller chunks. These approaches are implemented into *Dusk2Dawn*, a user-friendly EEGLAB plugin with a graphical interface.

Supplementary material

Supplementary material is available at *SLEEP* online.

Data availability statement

The datasets for this manuscript will be made publicly available after the acceptance for publication of the main findings from the final dataset, and after the participants' privacy has been

protected in accord with applicable Swiss laws and regulations. Requests to access the datasets should be directed to Francesca Siclari (f.siclari@nin.knaw.nl).

Acknowledgments

Funding

Disclosure Statement

We declare no financial competing interests.

References

1. Sandro Lecci *et al*. Coordinated infraslow neural and cardiac oscillations mark fragility and offline periods in mammalian sleep.*Sci Adv*. 2017;**3**:e1602026. doi:10.1126/sciadv.1602026.

2. Borb AA, Achermann P. Sleep homeostasis and models of sleep regulation. *J Biol Rhythms*. 1999;**14**:559–570. doi:10.1177/074873099129000894.

3. Simor P, van der Wijk G, Nobili L, Peigneux P. The microstructure of REM sleep: why phasic and tonic? *Sleep Med Rev.* 2020;**52**:101305. doi:10.1016/j.smrv.2020.101305.

4. Sasidharan A, Kumar S, Nair AK, *et al*. Further evidences for sleep instability and impaired spindle-delta dynamics in schizophrenia: a whole-night polysomnography study with neuroloop-gain and sleep-cycle analysis. *Sleep Med*. 2017;**38**:1–13. doi:10.1016/j.sleep.2017.02.009.

5. Mumtaz W, Rasheed S, Irfan A. Review of challenges associated with the EEG artifact removal methods. *Biomed Signal Process Control*. 2021;**68**:102741. doi:10.1016/j.bspc.2021.102741.

6. Brunner DP, Vasko RC, Detka CS, Monahan JP, Reynolds CF, Kupfer DJ. Muscle artifacts in the sleep EEG: automated detection and effect on all-night EEC power spectra. *J Sleep Res*. 1996;**5**:155–164. doi:10.1046/j.1365-2869.1996.00009.x.

7. Motamedi-Fakhr S, Moshrefi-Torbati M, Hill M, Hill CM, White PR. Signal processing techniques applied to human sleep EEG signals—a review. *Biomed Signal Process Control*. 2014;**10**:21–33. doi: 10.1016/j.bspc.2013.12.003.

8. Boostani R, Karimzadeh F, Nami M. A comparative review on sleep stage classification methods in patients and healthy individuals. *Comput Methods Programs Biomed*. 2017;**140**:77–91. doi:10.1016/j.cmpb.2016.12.004.

9. Leach S, Sousouri G, Huber R. "High-Density-SleepCleaner": an open-source, semi-automatic artifact removal routine tailored to high-density sleep EEG. *J Neurosci Methods*. 2023;**391**:109849. doi:10.1016/j.jneumeth.2023.109849.

10. Saifutdinova E, Congedo M, Dudysova D, Lhotska L, Koprivova J, Gerla V. An unsupervised multichannel artifact detection method for sleep EEG based on riemannian geometry. *Sensors (Switzerland)*. 2019;**19**:602. doi:10.3390/s19030602.

11. Sprecher KE, Riedner BA, Smith RF, Tononi G, Davidson RJ, Benca RM. High resolution topography of age-related changes in non-rapid eye movement sleep electroencephalography. *PLoS One*. 2016;**11**:e0149770–e0149716. doi:10.1371/journal.pone.0149770.

12. Stephansen JB, Olesen AN, Olsen M, *et al*. Neural network analysis of sleep stages enables efficient diagnosis of narcolepsy. *Nat Commun*. 2018;**9**:1–15. doi: 10.1038/s41467-018-07229-3.

13. Luca G, Haba Rubio J, Andries D, *et al*. Age and gender variations of sleep in subjects without sleep disorders. *Ann Med*. 2015;**47**:482–491. doi:10.3109/07853890.2015.1074271.

14. Jung T-P, Makeig S, Humphries C, *et al*. Removing electroencephalographic artifacts by blind source separation. *Psychophysiology*. 2000;**37**:163–178. doi: 10.1111/1469-8986.3720163.

15. Kothe CA, Makeig S. BCILAB. A platform for brain–computer interface development. *J Neural Eng*. 2013;**10**:056014–. doi: 10.1088/1741-2560/10/5/056014.

16. Kothe CAE, Jung T-P. Artifact removal techniques with signal reconstruction (US PATENT: US20160113587A1), 2016.

17. Mullen TR, Kothe CAE, Chi YM, *et al*. Real-time neuroimaging and cognitive monitoring using wearable dry EEG. *IEEE Trans Biomed Eng*. 2015;**62**:2553–2567. doi:10.1109/TBME.2015.2481482.

18. Miyakoshi M, Jurgiel J, Dillon A, *et al*. Modulation of frontal oscillatory power during blink suppression in children: effects of premonitory urge and reward. *Cereb Cortex Commun*. 2020;**1**:1–11. doi: 10.1093/texcom/tgaa046.

19. Plechawska-Wojcik M, Kaczorowska M, Zapala D. *The Artifact Subspace Reconstruction (ASR) for EEG signal correction. A comparative study*. vol. **853**. Cham: Springer International Publishing; 2019. doi: 10.1007/978-3-319-99996-8_12.

20. Chang CY, Hsu SH, Pion-Tonachini L, Jung TP. Evaluation of artifact subspace reconstruction for automatic artifact components removal in multi-channel EEG recordings. *IEEE Trans Biomed Eng*. 2020;**67**:1114–1121. doi:10.1109/tbme.2019.2930186.

21. Robbins KA, Touryan J, Mullen T, Kothe C, Bigdely-Shamlo N. How sensitive are EEG results to preprocessing methods: a benchmarking study. *IEEE Trans Neural Syst Rehabil Eng*. 2020;**28**:1081–1090. doi:10.1109/TNSRE.2020.2980223.

22. Anders P, Müller H, Skjæret-Maroni N, Vereijken B, Baumeister J. The influence of motor tasks and cut-off parameter selection on artifact subspace reconstruction in EEG recordings. *Med Biol Eng Comput*. 2020;**58**:2673–2683. doi:10.1007/s11517-020-02252-3.

23. Tran XA, McDonald N, Dickinson A, *et al*. Functional connectivity during language processing in 3-month-old infants at familial risk for autism spectrum disorder. *Eur J Neurosci*. 2021;**53**:1621–1637. doi:10.1111/ejn.15005.

24. Pérez A, Dumas G, Karadag M, Duñabeitia JA. Differential brain-to-brain entrainment while speaking and listening in native and foreign languages. *Cortex*. 2019;**111**:303–315. doi:10.1016/j.cortex.2018.11.026.

25. Perera H, Shiratuddin MF, Wong KW, Fullarton K. EEG signal analysis of real-word reading and nonsense-word reading between adults with dyslexia and without dyslexia. *Proc IEEE Symp Comput Med Syst* 2017;**2017-June**:73–78. doi: 10.1109/CBMS.2017.108.

26. An J, Yoo D, Lee BC. Electrocortical activity changes in response to unpredictable trip perturbations induced by a split-belt treadmill. *Proc Annu Int Conf IEEE Eng Med Biol Soc EMBS* 2019;**2019**:110–113. doi:10.1109/EMBC.2019.8856762.

27. Setiawan H, Islamiyah WR, Wibawa AD, Purnomo MH. Identifying EEG parameters to monitor stroke rehabilitation using individual analysis. *Proc 2019 Int Semin Intell Technol Its Appl ISITIA* 2019;**2019**:337–342. doi: 10.1109/ISITIA.2019.8937238.

28. Bulea TC, Prasad S, Kilicarslan A, Contreras-Vidal JL. Sitting and standing intention can be decoded from scalp EEG recorded prior to movement execution. *Front Neurosci*. 2014;**8**:1–19. doi: 10.3389/fnins.2014.00376.

29. Suyasmad, Wibawa AD, Wulandari DP, Rahayu PS, Islamiyah WR. Statistical analysis of subject-specific EEG data during stroke rehabilitation monitoring. In: EECCIS 2020—2020 10th Electr Power, Electron Commun Control Informatics Semin 2020:168–72. doi: 10.1109/EECCIS49483.2020.9263462.

30. Jacobsen NSJ, Blum S, Witt K, Debener S. A walk in the park? Characterizing gait-related artifacts in mobile EEG recordings. *Eur J Neurosci.* 2021;**54**:8421–8440. doi:10.1111/ejn.14965.

31. Cataldi J, Stephan AM, Marchi NA, Haba-Rubio J, Siclari F. Abnormal timing of slow wave synchronization processes in non-rapid eye movement sleep parasomnias. *Sleep.* 2022;**45**:1–16. doi: 10.1093/sleep/zsac111.

32. Kumaravel VP, Kartsch V, Benatti S, Vallortigara G, Farella E, Buiatti M. Efficient artifact removal from low-density wearable eeg using artifacts subspace reconstruction. *Proc Annu Int Conf IEEE Eng Med Biol Soc EMBS* 2021;**2021**:333–336. doi:10.1109/EMBC46164.2021.9629771.

33. Cataldo A, Criscuolo S, De Benedetto E, *et al*. A method for optimizing the artifact subspace reconstruction performance in low-density EEG. *IEEE Sens J.* 2022;**22**:21257–21265. doi:10.1109/jsen.2022.3208768.

34. Saravanapandian V, Nadkarni D, Hsu SH, *et al*. Abnormal sleep physiology in children with 15q11.2-13.1 duplication (Dup15q) syndrome. *Mol Autism* 2021;**12**:1–14. doi: 10.1186/s13229-021-00460-8.

35. Miyakoshi M, Nariai H, Rajaraman RR, *et al*. Automated preprocessing and phase-amplitude coupling analysis of scalp EEG discriminates infantile spasms from controls during wakefulness. *Epilepsy Res.* 2021;**178**:106809. doi:10.1016/j.eplepsyres.2021.106809.

36. Donnelly NA, Bartsch U, Moulding HA, *et al*. Sleep EEG in young people with 22q11.2 deletion syndrome: a cross-sectional study of slow-waves, spindles, and correlations with memory and neurodevelopmental symptoms. *Elife.* 2022;**11**:1–32. doi: 10.7554/elife.75482.

37. Solano A, Riquelme LA, Perez-Chada D, Della-Maggiore V. Motor learning promotes the coupling between fast spindles and slow oscillations locally over the contralateral motor network. *Cereb Cortex.* 2022;**32**:2493–2507. doi:10.1093/cercor/bhab360.

38. Siclari F, Bernardi G, Riedner BA, LaRocque JJ, Benca RM, Tononi G. Two distinct synchronization processes in the transition to sleep: a high-density electroencephalographic study. *Sleep.* 2014;**37**:1621–1637. doi:10.5665/sleep.4070.

39. Stephan AM, Lecci S, Cataldi J, Siclari F. Conscious experiences and high-density EEG patterns predicting subjective sleep depth. *Curr Biol.* 2021;**31**:5487–5500.e3. doi:10.1016/j.cub.2021.10.012.

40. Lecci S, Cataldi J, Betta M, Bernardi G, Heinzer R, Siclari F. Electroencephalographic changes associated with subjective under- and overestimation of sleep duration. *Sleep.* 2020;**43**:1–12. doi: 10.1093/sleep/zsaa094.

41. Iber C, Ancoli-Israel S, Chesson A, Quan S. *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications.* 1st ed. Illinois: American Academy of Sleep Medicine; 2007.

42. Shoorangiz R, Weddell SJ, Jones RD. Prediction of microsleeps from EEG: preliminary results. *Proc Annu Int Conf IEEE Eng Med Biol Soc EMBS* 2016;**2016-October**:4650–4653. doi: 10.1109/EMBC.2016.7591764.

43. Shoorangiz R, Buriro AB, Weddell SJ, Jones RD. Detection and prediction of microsleeps from EEG using Spatio-Temporal Patterns. In: Proc Annu Int Conf IEEE Eng Med Biol Soc EMBS 2019:522–525. doi: 10.1109/EMBC.2019.8857962.

44. Shoorangiz R, Weddell SJ, Jones RD. Bayesian multi-subject factor analysis to predict microsleeps from EEG power spectral features. *Proc Annu Int Conf IEEE Eng Med Biol Soc EMBS* 2017;**2017**:4183–4186. doi:10.1109/EMBC.2017.8037778.

45. Tsai BY, Diddi SVS, Ko LW, Wang SJ, Chang CY, Jung TP. Development of an adaptive artifact subspace reconstruction based on hebbian/anti-hebbian learning networks for enhancing BCI performance. *IEEE Trans Neural Networks Learn Syst.* 2022;**1**:1–14. doi:10.1109/tnnls.2022.3174528.

46. Pion-Tonachini L, Kreutz-Delgado K, Makeig S. The ICLabel dataset of electroencephalographic (EEG) independent component (IC) features. *Data Br.* 2019;**25**:104101. doi:10.1016/j.dib.2019.104101.

47. Riedner BA, Vyazovskiy VV, Huber R, *et al*. Sleep homeostasis and cortical synchronization: III. A high-density EEG study of sleep slow waves in humans. *Sleep.* 2007;**30**:1643–1657. doi:10.1093/sleep/30.12.1643.

48. Siclari F, Bernardi G, Cataldi J, Tononi G. Dreaming in NREM sleep: a high-density EEG study of slow waves and spindles. *J Neurosci.* 2018;**38**:9175–9185. doi:10.1523/JNEUROSCI.0855-18.2018.

49. Massimini M, Huber R, Ferrarelli F, Hill S, Tononi G. The sleep slow oscillation as a traveling wave. *J Neurosci.* 2004;**24**:6862–6870. doi:10.1523/JNEUROSCI.1318-04.2004.

50. Bernardi G, Siclari F, Bellesi ME. Local aspects of sleep and wakefulness. *Front Neurosci.* 2020;**14**:1–3. doi: 10.3389/fnins.2020.00058.

51. Spiess M, Bernardi G, Kurth S, *et al*. How do children fall asleep? A high-density EEG study of slow waves in the transition from wake to sleep. *Neuroimage.* 2018;**178**:23–35. doi:10.1016/j.neuroimage.2018.05.024.

52. Hsu LM, Wang S, Ranadive P, *et al*. Automatic skull stripping of rat and mouse brain MRI data using U-Net. *Front Neurosci.* 2020;**14**:1–9. doi: 10.3389/fnins.2020.568614.

53. Salimi-Khorshidi G, Douaud G, Beckmann CF, Glasser MF, Griffanti L, Smith SM. Automatic denoising of functional MRI data: combining independent component analysis and hierarchical fusion of classifiers. *Neuroimage.* 2014;**90**:449–468. doi:10.1016/j.neuroimage.2013.11.046.

54. Gabrieli G, Lenoir C, Bufacchi RJ, Mouraux A, Iannetti GD; Federation of European Neuroscience Societies (FENS). 2022. doi:10.13140/RG.2.2.23649.99680