

Artifact Subspace Reconstruction (ASR) for cleaning continuous data. The primary challenge to analyze MoBI EEG data is to correct movement artifact. We developed a free open-source EEGLAB plugin, *clean_rawdata()*, a suit of EEG data preprocessing methods including Artifact Subspace Reconstruction (ASR) which plays a central role in correcting continuous data (Mullen *et al.*, 2015). The collection of the programs was originally developed for BCILAB by Christian Kothe.

First, *clean_rawdata()* performs bad channel rejection. The two definitions of bad channels are, 1) channels that have flat signal longer than 5 seconds, and 2) channels that are poorly correlated ($r < 0.85$) with adjacent channels.

Second, *clean_rawdata()* applies ASR, which is to remove short-time high-amplitude artifacts in the continuous data. The flow of the process is explained below. Let $X_{calibrate} \in \mathbb{R}^{Q \times M}$ be ‘the cleanest part of the channel data’ which ASR finds for calibration for Q-channel continuous EEG data. This is achieved by first applying IIR spectral weighting (8th order Yule-Walker) that is the inverse of an in-house developed heuristic model of EEG power spectral density that has 1/f curve with an attenuated lower-frequency end and a peak at around 8 Hz, then concatenating windows of time-series by z-scored RMS amplitude of the time series between -3.5 and 5.5 after. Let $\bar{M}$ the square root of the covariance matrix C of $X_{calibrate}$ such that $\bar{M}\bar{M}^T = Cov[X_{calibrate}]$ where $\bar{M}^T$ represents transpose of $\bar{M}$. Let $x_{(t)} \in \mathbb{R}^Q$ be a EEG sample at time point t , and let $X \in \mathbb{R}^{Q \times N}$ be the sliding window of data containing $x_{(t)}$. Principal component (PC) analysis (PCA) is applied to X to obtain PCs of the currently selected window $V = [v_1 \dots v_Q] \in \mathbb{R}^{Q \times Q}$. Those PCs V whose variance d_k exceeds an omnibus cutoff threshold $h(v_k)$ are determined as *artifact subspace*. The omnibus cutoff threshold h_k is defined as follows. Let $M = V^T \bar{M}$ be projection of $\bar{M}$ into PC space. Principal component analysis on $X_{calibrate}$ is computed to obtain $W = [w_1 \dots w_Q] \in \mathbb{R}^{Q \times Q}$ and activations $Y = W^T X_{calibrate}$. For

activation of each component Y_k , root-mean-square amplitude is calculated to obtain robust mean $\mu_{calibrate}$ and standard deviation $\sigma_{calibrate}$. The omnibus threshold is obtained as

$h_k = \mu_{calibrate_k} + \sigma_{calibrate_k} \times f_{user}$ where f_{user} represents user-specified factor for SD. For the current analysis, we chose $f_{user} = 20$ which is a very lax criterion, intending only to remove large outlier activations in the time-series data in order to let the subsequent ICA decompose artifacts that have physiological sources, such as eye blinks/movements, EMG, and ECG.

After artifact subspace is determined, the threshold operator $U \in \mathbb{R}^{Q \times Q}$ is built. For this, the omnibus threshold h_k in the channel space needs to be projected to the PC space $h(v_j)$ as

$$h(v_j) = \sum_k [(\mu_{calibrate_k} + \sigma_{calibrate_k} \times f_{user}) W_k^T V_j]^2 \text{ to be compared with } d_j. \text{ Now, } U_{jl} \text{ is built such}$$

that $U_{jl} = 0$ if $d_j > h(v_j)$ to remove them, otherwise $U_{jl} = 1$ to keep them. After removing

artifact subspace, it is interpolated, which completes the process *artifact subspace*

reconstruction. This process is represented by the linear operator $R = VM(M \circ U)^+ V^T$ such that

$\hat{x}_{(t)} = Rx_{(t)}$ produces corrected data $\hat{x}_{(t)}$ for input data $x_{(t)}$. $M \circ U$ represents element-wise

multiplication of M and U , and $(M \circ U)^+$ represents Moore-Penrose pseudoinverse of

$(M \circ U)$.

Finally, after applying ASR, 1-s sliding window rejection is applied, and those windows within which more than 25% of channels showed z-scored RMS amplitude outside the range of -3.5 and 7.0 after applying the aforementioned IIR filter are rejected for insufficient corrections.