clean_artifacts

Clean Raw data version 2.10 is part of eeglab 2025

All-in-one function for artifact removal, including ASR.

[EEG,HP,BUR] = **clean_artifacts**(EEG, Options...)

This function removes flatline channels, low-frequency drifts, noisy channels, short-time bursts and incompletely repair segments from the data. Tip: Any of the core parameters can also be passed in as [] to use the respective default of the underlying functions, or as 'off' to disable it entirely.

Ideally, parameter tuning should only be necessary when using this function. However, there are three parameters that govern how aggressively bad channels, bursts and irrecoverable time windows are removed, as well as several detail parameters that only require tuning under special circumstances.

Notes:

  * This function uses the Signal Processing toolbox for pre- and post-processing of the data (removing drifts, channels and time windows); the core ASR method (clean_asr) does not require this toolbox but you will need high-pass filtered data if you use it directly.

  * By default this function will identify subsets of clean data from the given recording to enhance the robustness of the ASR calibration phase to strongly contaminated data; this uses the Statistics toolbox, but can be skipped/bypassed if needed (see documentation).

Input:

EEG : Raw continuous EEG recording to clean up (as EEGLAB dataset structure).

NOTE: The following parameters are the core parameters of the cleaning procedure; they should be passed in as Name-Value Pairs. If the method removes too many (or too few) channels, time windows, or general high-amplitude ("burst") artifacts, you will want to tune these values.
Hopefully you only need to do this in rare cases.

ChannelCriterion :

Minimum channel correlation. If a channel is correlated at less than this value to an estimate based on other channels, it is considered abnormal in the given time window. This method requires that channel locations are available and roughly correct; otherwise a fallback criterion will be used. (default: 0.85)

LineNoiseCriterion :

If a channel has more line noise relative to its signal than this value, in standard deviations based on the total channel population, it is considered abnormal. (default: 4)

BurstCriterion :

Standard deviation cutoff for removal of bursts (via ASR). Data portions whose variance is larger than this threshold relative to the calibration data are considered missing data and will be removed. The most aggressive value that can be used without losing much EEG is 3. For new users it is recommended to at first visually inspect the difference between the original and cleaned data to get a sense of the removed content at various levels. An agressive value is 5 and a quite conservative value is 20. Default: 5 (from the GUI, default is 20).

WindowCriterion :

Criterion for removing time windows that were not repaired completely. This may happen if the artifact in a window was composed of too many simultaneous uncorrelated sources (for example, extreme movements such as jumps). This is the maximum fraction of contaminated channels that are tolerated in the final output data for each considered window. Generally a lower value makes the criterion more aggressive. Default: 0.25. Reasonable range: 0.05 (very aggressive) to 0.3 (very lax).

Highpass :

Transition band for the initial high-pass filter in Hz. This is formatted as [transition-start, transition-end]. Default: [0.25 0.75].

ChannelCriterionMaxBadTime :

This is the maximum tolerated fraction of the recording duration during which a channel may be flagged as "bad" without being removed altogether. Generally a lower (shorter) value makes the criterion more aggresive. Reasonable range: 0.15 (very aggressive) to 0.6 (very lax). Default: 0.5.

BurstCriterionRefMaxBadChns:

If a number is passed in here, the ASR method will be calibrated based on sufficiently clean data that is extracted first from the recording that is then processed with ASR. This number is the maximum tolerated fraction of "bad" channels within a given time window of the recording that is considered acceptable for use as calibration data. Any data windows within the tolerance range are then used for calibrating the threshold statistics. Instead of a number

one may also directly pass in a data set that contains calibration data (for example a minute of resting EEG).

If this is set to 'off', all data is used for calibration. This will work as long as the fraction of contaminated data is lower than the the breakdown point of the robust statistics in the ASR calibration (50%, where 30% of clearly recognizable artifacts is a better estimate of the practical breakdown point).

A lower value makes this criterion more aggressive. Reasonable range: 0.05 (very aggressive) to 0.3 (quite lax). If you have lots of little glitches in a few channels that don't get entirely cleaned you might want to reduce this number so that they don't go into the calibration data. Default: 0.075.

BurstCriterionRefTolerances :
These are the power tolerances outside of which a channel in a given time window is considered "bad", in standard deviations relative to a robust EEG power distribution (lower and upper bound). Together with the previous parameter this determines how ASR calibration data is be extracted from a recording. Can also be specified as 'off' to achieve the same effect as in the previous parameter. Default: [-Inf 5.5].

BurstRejection :
'on' or 'off'. If 'on' reject portions of data containing burst instead of correcting them using ASR. Default is 'off'.

WindowCriterionTolerances :
These are the power tolerances outside of which a channel in the final output data is considered "bad", in standard deviations relative to a robust EEG power distribution (lower and upper bound). Any time window in the final (repaired) output which has more than the tolerated fraction (set by the WindowCriterion parameter) of channel with a power outside of this range will be considered incompletely repaired and will be removed from the output. This last stage can be skipped either by setting the WindowCriterion to 'off' or by taking the third output of this processing function (which does not include the last stage). Default: [-Inf 7].

FlatlineCriterion :
Maximum tolerated flatline duration. In seconds. If a channel has a longer flatline than this, it will be considered abnormal. Default: 5

NoLocsChannelCriterion :
Criterion for removing bad channels when no channel locations are present. This is a minimum correlation value that a given channel must have w.r.t. a fraction of other channels. A higher value makes the criterion more aggressive. Reasonable range: 0.4 (very lax) - 0.6 (quite aggressive). Default: 0.45.

NoLocsChannelCriterionExcluded :

The fraction of channels that must be sufficiently correlated with a given channel for it to be considered "good" in a given time window. Applies only to the NoLocsChannelCriterion. This adds robustness against pairs of channels that are shorted or other that are disconnected but record the same noise process.

Reasonable range: 0.1 (fairly lax) to 0.3 (very aggressive); note that increasing this value requires the ChannelCriterion to be relaxed in order to maintain the same overall amount of removed channels. Default: 0.1.

MaxMem :

The maximum amount of memory in MB used by the algorithm when processing. See function asr_process for more information. Default is 64.

Output:

EEG : Final cleaned EEG recording.

HP : Optionally just the high-pass filtered data.

BUR : Optionally the data without final removal of "irrecoverable" windows.

Examples:

% Load a recording, clean it, and visualize the difference (using the defaults)

raw = pop_loadset(...);

clean = **clean_artifacts**(raw);

vis_artifacts(clean,raw);

% Use a more aggressive threshold (passing the parameters in by position)

raw = pop_loadset(...);

clean = **clean_artifacts**(raw,[],2.5);

vis_artifacts(clean,raw);

% Passing some parameter by name (here making the WindowCriterion setting less picky)

raw = pop_loadset(...);

clean = **clean_artifacts**(raw,'WindowCriterion',0.25);

% Disabling the WindowCriterion and ChannelCriterion altogether

raw = pop_loadset(...);

clean = **clean_artifacts**(raw,'WindowCriterion','off','ChannelCriterion','off');

Christian Kothe, Swartz Center for Computational Neuroscience, UCSD

2012-09-04

NOTE

The following are detail parameters that may be tuned if one of the criteria does not seem to be doing the right thing. These basically amount to side assumptions about the data that usually do not change much across recordings, but sometimes do.