ICACLEANEEG VERSION 1.3

MATLAB-Based EEG Preprocessing Using Independent Component Analysis

Developed by Seung Suk Kang, Ph.D.
kangseung@umkc.edu

Cognition and Brain Laboratory
Minneapolis VAHCS | University of Minnesota
http://cablab.umn.edu/

Document feedback and typos: **marqu310@umn.edu**
Version – February 2020 (draft)

Acknowledgements

ICAcleanEEG has been programmed and supported by Seung Suk Kang who is currently assistant professor at the University of Missouri-Kansas City School of Medicine. ICAcleanEEG is for research and offers no guarantee of suitability for any particular purpose. Under no circumstances should ICAcleanEEG be applied for clinical use.

This manual is the product of contributions from many people affiliated with the Cognition and Brain (CAB) Laboratory operated through the Minneapolis VAHCS and the University of Minnesota Department of Psychiatry and Behavioral Sciences. These include, but are not limited to Christopher Hollowell, Julia Longenecker, Peter Lynn, Craig Marquardt, Victor Pokorny, Scott Sponheim, Elle Stahura, Joshua Stim, Collin Teich, and Abraham Van Voorhis. Opinions expressed in this manual are those of the authors, and do not necessarily reflect the views of the United States Department of Defense, Department of Veterans Affairs, or the Department of Health and Human Services.

Contents

1. Introduction to EEG Preprocessing

Applying the ICA technique using ICAcleanEEG

After electroencephalography (EEG) recordings have been collected, there are multiple techniques available to prepare the data for analysis. One sophisticated and increasingly popular technique is to apply independent component analysis (ICA) for the purpose of separating signal from noise. EEG signal recorded at the scalp is not a pure measure of neurophysiological activity from the brain and is contaminated with noise from a variety of sources. It is common to observe muscle contractions, eye blinks, heartbeats, sweat-related signal alterations, environmental noise, and equipment malfunctions in EEG recordings. These noise sources can make it difficult to measure subtle changes in brain activity related to experimental task manipulations or individual differences. ICA can be applied to identify these noise sources in the data. The end goal is to separate out and keep brain-related "components" while discarding as much of everything else as practically possible. (For more background information, see: Hyvarinen & Oja 2000; Neural Networks).

Multiple EEG software packages and procedures with ICA functionality are currently available for researchers. ICAcleanEEG offers several advantages over other ICA-based methods for preprocessing. Some software packages are entirely automated and lack tools for visualizing the impact of the ICA procedure on your data. Other methods require substantial experience in computer programing or a significant background in EEG research methods. The semi-automated procedures of ICAcleanEEG address these limitations by leveraging sophisticated algorithms for automatic component classification while also giving users multiple ways to visualize the preprocessing at every stage. We believe this reduces inter-rater inconsistencies in identifying signal and noise, and supports flexibility when encountering idiosyncratic artifacts in EEG recordings that require special treatment.

At present, there is no established consensus on how to best apply ICA for preprocessing EEG recordings. EEG data are often quite complex and ICA is not a perfect solution to the challenges involved with separating neural signal from electroencephalographic noise. ICAcleanEEG is no exception to this. This manual was designed to provide potential users of ICAcleanEEG with a step-by-step guide of how to apply the technique in the MATLAB environment. Throughout, we have also supplemented this with information about best practices in ICA classification gained through countless hours of preprocessing EEG recordings by multiple people. Understanding and adopting these principles is important for maintaining methodological consistency across projects and maximizing the potential quality of your post-processed recordings. We hope these materials will be useful as you begin your research with EEG.

Collecting and preprocessing EEG recordings

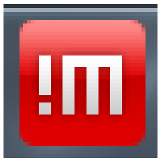
ICAcleanEEG was originally developed to preprocess EEG recordings collected for research projects at the Minneapolis Veterans Affairs Health Care Systems (VAHCS) by the Cognition and Brain (CAB) Laboratory. These “CAB Lab” studies typically involve participants arriving for 2-3 different appointments, one of which is used for administering various EEG tasks. EEG appointments often last between 2-4 hours depending on the study.

Most CAB Lab research recordings are conducted using “high density” EEG systems with 64 or 128 electrodes spread across the scalp. Each recording is processed differently depending on the number of electrodes used, the type of tasks administered, and the analyses you plan to conduct on the post-processed data. Although a larger study protocol may involve many EEG tasks, it is typical to only analyze one task at a time. As such, preprocessing usually involves repeatedly applying ICAcleanEEG to one particular type of task recording collected across many participants. It is important to apply ICA consistently, which will be discussed in greater detail below. Preprocessing some participants differently than others could introduce systematic differences into the data.

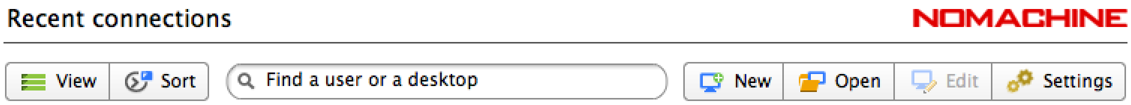
NoMachine and MATLAB basics

Remote access to the servers

In the CAB Lab, ICAcleanEEG is accessed using MATLAB via a remote connection to University of Minnesota computing infrastructure. This can be achieved using NoMachine software (https://www.nomachine.com/download; http://latisresearch.umn.edu/connecting-nomachine-ltsclaumnedu). NoMachine should already be downloaded and accessible on all CAB Lab computers. The icon looks like this:



Log in using your U of MN credentials (most likely your x500 username and password). You may need to request a Linux account, which can be initiated by CAB Lab staff members. Create a new connection the first time you log in by clicking the '**New**' button to the right of the '**Find a user or a desktop**' field.



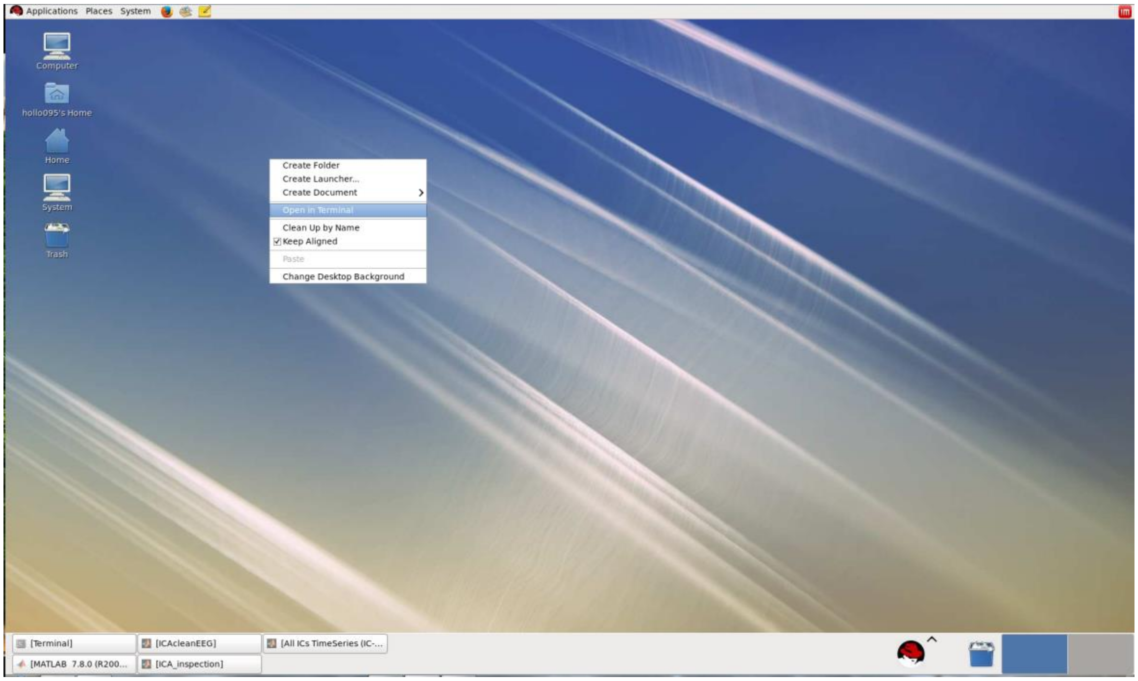
Select SSH from the 'Protocol' drop down menu. The host is **'lts.cla.umn.edu'** and the port is **'22'**. Select the **'use the system login'** option to authenticate the host. For the authentication method, select the **'password'** option. Select **'Don't use a proxy**' on the Proxy page. Finally, save the connection as **'UMN'** and click the option to save a shortcut on the desktop.

You will then click the shortcut you created to connect to NoMachine. Enter your x500 and password. Click on the '**New Desktop**' button. Then, click **'Create a new virtual desktop'** to complete the configuration of your NoMachine desktop environment.

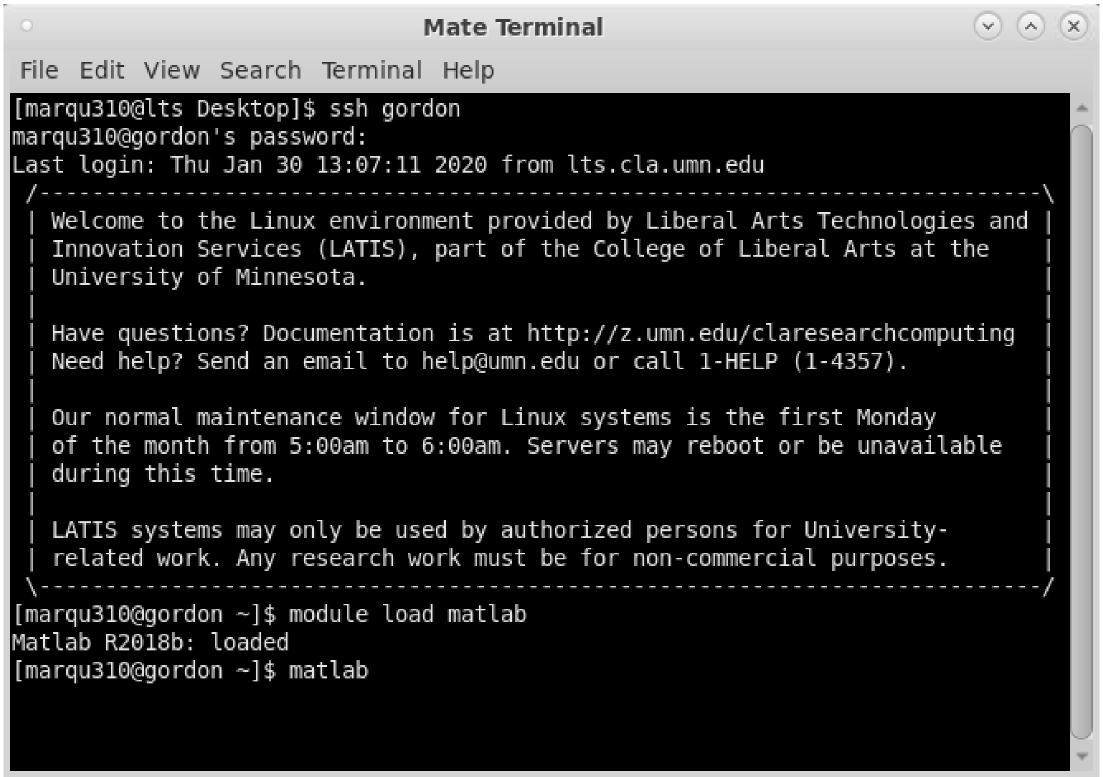
Starting ICAcleanEEG

To operate ICAcleanEEG effectively, it is recommended that you run your remote desktop using a larger computer monitor (i.e., 22+ inches in diameter). Many of the descriptive plots contain dense depictions of various indices, which can be hard to distinguish using a small display.

When you first open NoMachine, you will see a blank home screen. Right click and select **'Open in Terminal'**. It is through this terminal that you can access the software and files you need to conduct your analyses.



A command window will appear. To access MATLAB, you will need to log-in to specific servers. Type **'ssh gordon'** and the command window will prompt you for your X500 password. Enter the same log-in password that you used to open NoMachine. Once you are logged in, type **'module load matlab'.** Then, type **'matlab'**.



```
[marqu310@lts Desktop]$ ssh gordon
marqu310@gordon's password:
Last login: Thu Jan 30 13:07:11 2020 from lts.cla.umn.edu
 /--------------------------------------------------------------------------\
 | Welcome to the Linux environment provided by Liberal Arts Technologies and |
 | Innovation Services (LATIS), part of the College of Liberal Arts at the    |
 | University of Minnesota.                                                   |
 |                                                                            |
 | Have questions? Documentation is at http://z.umn.edu/claresearchcomputing  |
 | Need help? Send an email to help@umn.edu or call 1-HELP (1-4357).          |
 |                                                                            |
 | Our normal maintenance window for Linux systems is the first Monday        |
 | of the month from 5:00am to 6:00am. Servers may reboot or be unavailable   |
 | during this time.                                                          |
 |                                                                            |
 | LATIS systems may only be used by authorized persons for University-       |
 | related work. Any research work must be for non-commercial purposes.       |
 \--------------------------------------------------------------------------/
[marqu310@gordon ~]$ module load matlab
Matlab R2018b: loaded
[marqu310@gordon ~]$ matlab
```

Brief tour of the ‘srslab’ folder layout

Nearly all EEG recording storage, preprocessing, and analyses occur on laboratory data servers. This keeps the data secure and properly archived while also making it easier for multiple people to collaborate on EEG analysis projects. Raw data, analysis files, and custom processing scripts from multiple studies are stored in this location. *Please exercise caution.* As a rule, make sure everything is properly backed up before making changes.

/data_staging/ – This is where *raw* data is stored. We always store a copy of the original, unprocessed data as backup. Do not conduct analyses in these folders. The first step of any EEG project should be to copy the desired files to other locations on the server for analysis.

/data_main/ – This is where *processed* data is stored and where analyses are conducted. You can find ongoing preprocessing projects here. You will be working primarily within folders located in **/data_main/**. Each project has their own primary investigator who has often invested a great deal of time and effort into the analyses. Check with staff before investigating or altering files within project folders.

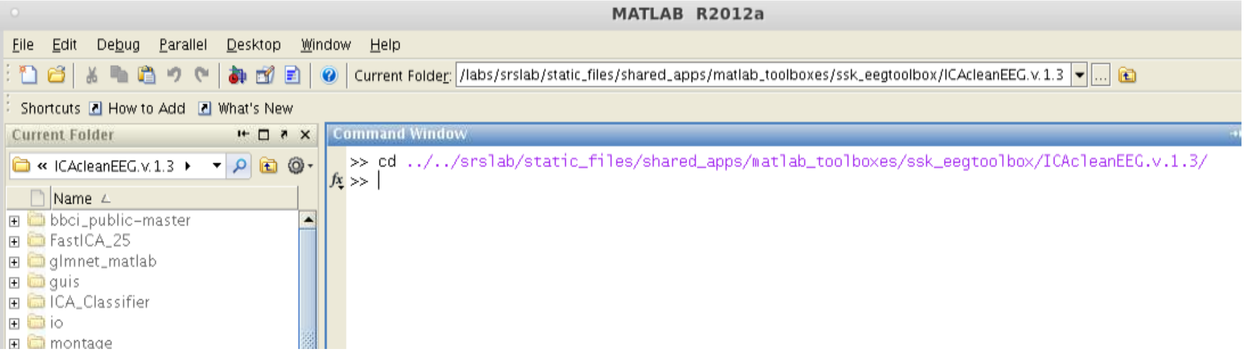
/data_public/ – Some staff members have personal folders for saving assorted project data on the server rather than their own personal computer. Please do not add your own folder without permission.

/static_files/ – Deep within this giant folder are the MATLAB-based preprocessing scripts that power ICAcleanEEG. These include all code for the graphical user interfaces (GUIs) used to conduct the preprocessing. **Do not alter these files.** Seung Suk Kang should be the only person to alter these scripts because they are critical for nearly all EEG projects ongoing within the CAB Lab.

To load the ICAcleanEEG GUI, navigate to the following folder: **'/labs/srslab/static_files/shared_apps/matlab_toolboxes/ssk_eegtoolbox/ICAcleanEEG.v.1.3/'**. This can be done by clicking through the folders in the directory on the left side of the MATLAB interface. Alternatively, you can navigate to this location using the command line. Some basic MATLAB commands are below, which are similar to the command line codes accepted in UNIX/LINUX environments.

pwd	Display the present working directory
ls	List the folder contents
cd	Change directory
edit	Create new MATLAB script
mkdir	Create a new folder
clear	Clear your current variables
.	Current directory
..	Directory up one in your directory tree
[TAB]	Auto-complete your command

For example, you might type the following to get to the ICAcleanEEG software folder:

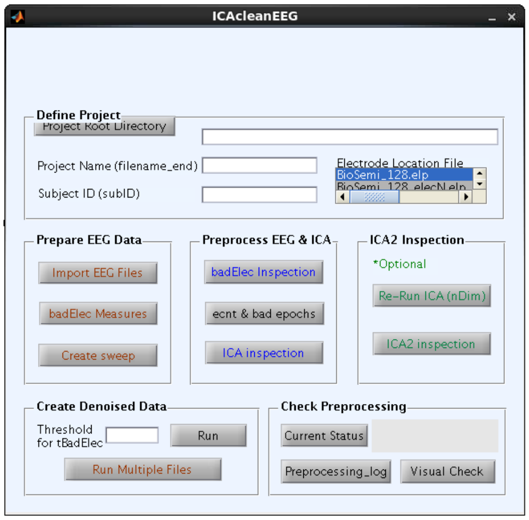


Current Folder: /labs/srslab/static_files/shared_apps/matlab_toolboxes/ssk_eegtoolbox/ICAcleanEEG.v.1.3

« ICAcleanEEG.v.1.3 ▸

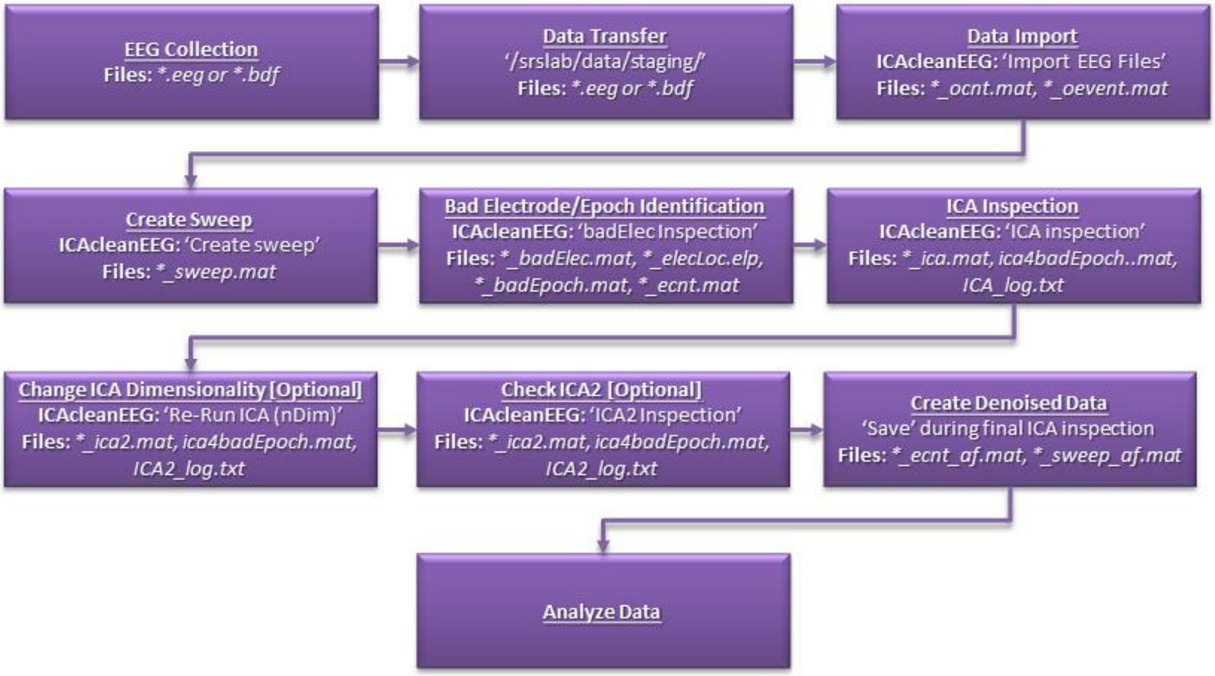
```
>> cd ../../srslab/static_files/shared_apps/matlab_toolboxes/ssk_eegtoolbox/ICAcleanEEG.v.1.3/
>> |
```

Now, type **ICAcleanEEG** to run the MATLAB script that loads the preprocessing GUI. You should now be viewing the main menu of the ICAcleanEEG interface.


ICAcleanEEG
Define Project
Project Root Directory
Project Name (filename_end)
Subject ID (subID)
Electrode Location File
BioSemi_128.elp
BioSemi_128_elecN.elp
Prepare EEG Data
Import EEG Files
badElec Measures
Create sweep
Preprocess EEG & ICA
badElec Inspection
ecnt & bad epochs
ICA inspection
ICA2 Inspection
*Optional
Re-Run ICA (nDim)
ICA2 inspection
Create Denoised Data
Threshold for tBadElec
Run
Run Multiple Files
Check Preprocessing
Current Status
Preprocessing_log
Visual Check

2. Overview of Preprocessing with ICAcleanEEG

There are multiple steps involved in taking electrical activity recorded at the scalp and converting it into data that clearly depict meaningful brain signal. Below is a flow chart of the preprocessing "pipeline" steps supported by ICAcleanEEG Version 1.3.


EEG Collection
Files: *.eeg or *.bdf
Data Transfer
'/srslab/data/staging/'
Files: *.eeg or *.bdf
Data Import
ICAcleanEEG: 'Import EEG Files'
Files: *_ocnt.mat, *_oevent.mat
Create Sweep
ICAcleanEEG: 'Create sweep'
Files: *_sweep.mat
Bad Electrode/Epoch Identification
ICAcleanEEG: 'badElec Inspection'
Files: *_badElec.mat, *_elecLoc.elp, *_badEpoch.mat, *_ecnt.mat
ICA Inspection
ICAcleanEEG: 'ICA inspection'
Files: *_ica.mat, ica4badEpoch..mat, ICA_log.txt
Change ICA Dimensionality [Optional]
ICAcleanEEG: 'Re-Run ICA (nDim)'
Files: *_ica2.mat, ica4badEpoch.mat, ICA2_log.txt
Check ICA2 [Optional]
ICAcleanEEG: 'ICA2 Inspection'
Files: *_ica2.mat, ica4badEpoch.mat, ICA2_log.txt
Create Denoised Data
'Save' during final ICA inspection
Files: *_ecnt_af.mat, *_sweep_af.mat
Analyze Data

EEG collection

There are a variety of best practices for EEG collection, which are beyond the scope of this tutorial. Interested readers are strongly encouraged to review "An Introduction to the Event-Related Potential Technique" by Steven Luck (2014; MIT Press).

Data transfer

After each EEG recording session is completed, task files are transferred for storage in CAB Lab server space (i.e., **/data_staging/**). Files for recordings need to be duplicated/copied to **/data_main/** before preprocessing can begin. Seek guidance and approval before beginning a data transfer.

Data import

Raw EEG recordings are commonly stored in either ***.bdf** or ***.eeg** format depending on the collection system. Neither format is directly compatible with MATLAB. The import step in ICAcleanEEG converts raw recording files to ***.mat** format.

Next, recordings are down-sampled to decrease file size. EEG signal is often measured 1024 times per second (“online” sampling rate). This temporal resolution is unnecessary for most conventional EEG studies. Resampling to 256 measurements per second (“offline” sampling rate) still ensures that brain signal as fast as 128 Hz (i.e., 128 cycles per second) can be quantified. In addition to resampling, high- and low-pass filters are applied. For high-pass filters, signal at frequencies *above* your designated cut-off remain unchanged (i.e., “passing over” the higher frequency signal) and signal *below* your cut-off frequency will be greatly attenuated (vice-versa for low-pass filters). It is typical to apply a 0.5 Hz high-pass filter to reduce the impact of low frequency noise and a 110 Hz low-pass filter to exclude non-neural high frequency signal while preserving brain signal in the higher frequency ranges (avoids violating the Nyquist sampling criterion, removes harmonic oscillations of the 60 Hz powerline at 120 Hz).

For visualization and analysis with ICAcleanEEG, electrode signals across the scalp are referenced to the “linked earlobe” signal. An average virtual channel is calculated from the two electrodes placed on participants’ earlobes. Correcting for this earlobe signal in each of the scalp electrodes provides a quick initial fix for simple sources of noise that may have been present at the time of the recording.

Several MATLAB files are produced at this step including an “original continuous” version of the recording (***_ocnt.mat)** and an "original event" file containing task event information (***_oevent.mat)**.

badElec Measures

badElec Measures is the next step after importing raw EEG data. This script will run some preliminary algorithms and create a file called ***_badElec.mat**, which will enable the user to progress to **badElec Inspection.**

Create sweep

Before examining data using ICAcleanEEG visualization tools, it is important to check the task event codes. Event codes are numbers inserted into the recordings to indicate the onset of visual or auditory stimuli, eye-tracking events, and button presses. These provide valuable information about the timing, duration, and structure of the task. If the event codes are correct for a particular file, no modifications will be required to the ***_oevent.mat** file created during import.

Event code errors are common, unfortunately. Check to see that the expected codes are visible, no codes are missing, and erroneous codes are not included in your recordings. In the majority of cases, small event code problems can be fixed by modifying the ***_oevent.mat** file. However, larger and more pervasive problems may indicate that a file has been corrupted or incorrectly recorded. In these cases, participant files may need to be excluded from further processing. Seek guidance before excluding a participant for reasons related to event codes because multiple solutions are available.

ICAcleanEEG relies on “sweep” variables (***_sweep.mat**) for each task to interpret the event codes. Building this sweep file is an important initial step for preprocessing, and requires some prior knowledge about the structure and programming of your specific EEG task. Accurate sweep files will be critical for later preprocessing such as when the continuous EEG file is restructured into discrete time segments of interest. The sweep variables should be consistent from participant to participant in your study. Be sure to fix event code errors prior to applying the sweep variable during preprocessing.

Bad electrode/epoch identification

Clicking **‘badElec Inspection’** in the main ICAcleanEEG GUI begins the process of interacting with various visualizations of the electrical activity recorded at the scalp. Users are tasked with reviewing each EEG recording for prominent artifacts that may interfere with application of ICA during later preprocessing steps. In particular, one should be vigilant for corrupted individual electrode channels and brief periods of time (i.e., “epochs”) dominated by noise artifacts. Several interactive figures are displayed to depict the degree to which each electrode is correlated with its neighboring electrodes on the scalp, the amplitude strength of each electrode, and the frequency spectrum present within each electrode channel. Additionally, the time series for individual electrodes (referenced to the linked ears signal) can be depicted to better appreciate the nature of any potential artifacts.

Decision-making regarding artifacts observed using badElec has changed over time within the CAB Lab. Individual researchers were initially encouraged to be aggressive in their exclusion of potential artifacts. However, this approach resulted in negligible increases to the quality of the post-processed data. ICA is capable of detecting many of the same sources of noise that were previously excluded manually during the badElec preprocessing step. Instead, users should focus on using badElec to exclude obvious phenomena in the recording that could interfere with the effectiveness of ICA. Of particular note are slow wave (< 2 Hz) artifacts originating from non-brain sources and epochs with broadband frequency noise with indistinguishable amounts of brain activity.

Once inspection is complete, badElec produces a summary of the electrodes and epochs deleted from the recording. ***_BadEpoch.mat** is also updated to save the results. Next, this information is combined with the sweep file information to segment the continuous EEG recording around each valid event code trigger. Event codes located within epochs that were excluded during the badElec step are also excluded when segmenting the recording. Additionally, excluded electrodes are interpolated using spherical spline interpolation. The end result is an “epoched” version of the EEG recording, which is saved as a ***_ecnt.mat** file.

ICA inspection

At this stage, each participant’s epoched recording file still represents a mixture of brain signal and noise. After inspecting the data with badElec, it is possible to apply ICA to each individual

***_ecnt.mat** file. ICA is a dimensionality reduction technique that can decompose multivariate datasets into independent source signals. The challenge of understanding EEG recordings is similar to the difficulty associated with understanding all the noise coming from a crowded room with people talking (i.e., “cocktail party” analogy). If you were to set up a microphone outside the room, each individual voice would be indistinguishable. However, it would be much easier to identify the voices of individual people if multiple microphones were placed around the room. Although each microphone would record slightly different sounds, the patterns observed across time and across all of the microphones together could provide insight into which individual people might be contributing to the noise and what might be the content of their conversations. EEG recordings are analogous, but the microphones are substituted for electrodes and the noisy room is substituted for electrical activity at the scalp. ICA is applied to identify the potential contributors of activity in the EEG recordings, which could be from various brain regions “talking” to each other or from other extra-brain sources.

ICA uses multiple assumptions to identify constituent signals in recordings. One such assumption is statistical “independence” such that the activity estimated for one component does not affect the probability of activity in the other components. Furthermore, each resulting component is additive. It is possible to sort through the components produced after running ICA and only retain the ones you want. Subsets of components can be added together to produce a new dataset emphasizing particular features in the original recording. Many researchers are enthusiastic about ICA because of the potential to identify and exclude sources of noise in EEG signals, and reconstitute the remaining components into a recording with only the signals we believe are neural in nature.

This manual includes multiple examples of both brain and noise components produced using ICAcleanEEG. The tutorial below will provide you with guidance about how to reliably distinguish between typical noise sources such as muscle contractions, eye movements, electrocardiograph-related activity, electrical noise, slow waves, and bad electrodes. After each user inspects and checks each component classification, ***_BadEpoch.mat** is updated and a ***_ICA_log.txt** file is created. **ICA_log.txt** is a text file that contains a record of your final classifications for each component. ICAcleanEEG chooses many of the ICA parameters for you automatically based on observed features of the data, but this does not always produce optimal characterizations of the EEG signals. As needed, users are able to delete additional electrodes or epochs, and manually rerun ICA using the optional “ICA2” functionality of ICAcleanEEG.

Create denoised data

Based on classifications from the previous ICA inspection, brain components are reconstructed while sources of artifact are excluded. Additionally, any additional filtering of the signal you may have requested is applied at this step. What remains should be signal suitable for event-related potential (ERP) or time-frequency (TF) analysis. Consequentially, ***_sweep_af.mat** and

***_ecnt_af.mat** files are created, and can be used for generating dependent variables that quantify the neural phenomena of interest to you.

3. Preparation for Preprocessing

Things to keep in mind:

1) Rule of thumb—*don't make the data worse*

It is important to take appropriate steps to minimize noise when collecting EEG recordings. Although ICA is a powerful technique, it is not able to identify neural signals that were never successfully measured in the original recordings. Preprocessing is done to highlight and clean the useful data that already exists in recordings. Users are encouraged to utilize the badElec tool to delete electrodes and epochs that are problematic. However, it is generally recommended to not remove more than **30% of the time series**. Additionally, it is not recommended to delete more than **10% of the electrodes** in a given file. If this amount of removal is necessary, it may be indicative of larger problems in the recording that are not fixable using ICA.

2) Rule of forefinger—*ask questions*

All CAB Lab staff members working with EEG have gone through this same training. While this document outlines the main steps of the processing "pipeline," a didactic overview can feel very different from actually working with individual files. Much of the real learning is experiential. For that reason, please take your time and do not hesitate to ask questions. Preprocessing becomes easier and clearer over time, but there will always be odd files. All lab personnel working with EEG recordings have encountered data that do not make sense. In those cases, it is best to consult with others rather than make guesses. **Consensus is strongly encouraged.** In the CAB Lab, we consider consensus to be best practice when making difficult decisions about challenging recordings.

3) Avoid rater drift

ICAcleanEEG is semi-automated, which has both advantages and disadvantages. On the one hand, this provides users with the flexibility to apply expert judgment. On the other hand, this can lead to biased and unreliable decision-making. People generally believe they are much more reliable than they actually are (Meehl, 1954; Grove et al., 2000; Dawes, Faust, & Meehl, 1989). Users of ICAcleanEEG should be vigilant to avoid complacency in this regard. Once one gets familiar with the software, it is easy to forget the preprocessing guidelines. This leads to systematic differences over time within a person or between people. **It is more important to be reliable than to be perfect.** When reliability decreases, it lessens our ability to associate the waveforms with other measures (i.e., reliability is necessary for validity). The most effective users of ICAcleanEEG are people who work hard to be consistent from recording to recording. Users are encouraged to keep this manual open during preprocessing. Refer to the example ICs and reread the guidelines simultaneously while making decisions.

4. Beginning the Preprocessing: Importing EEG files

Defining your project

Within the main ICAcleanEEG GUI, begin by typing in identifying information for your project files ("Define Project" box).

Project root directory: Click the button and browse through the folders to the location where preprocessed data is stored and analyzed for your specific project. All data for preprocessing should be within /**data_main/**.

Project name: One case sensitive word or abbreviation is chosen for each task during data collection, which is used to name the task recordings (e.g. SRRT, Dichotic, VerbMem). Use this same file name to reference the recordings during preprocessing.

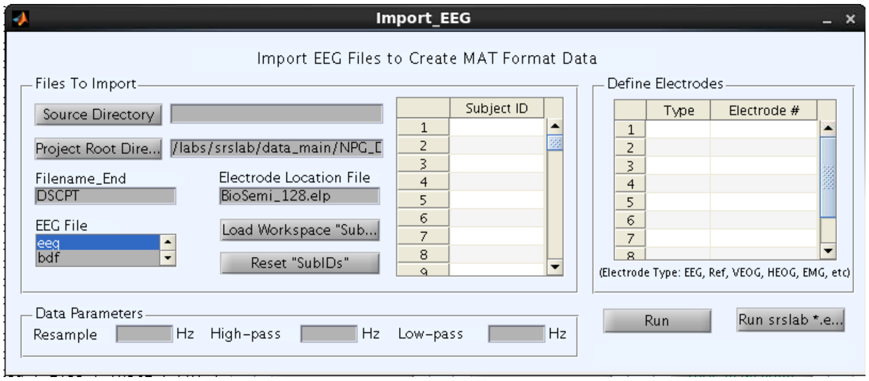
Subject ID: Enter this information to preprocess a single subject from start to finish (i.e., all ICAcleanEEG pipeline steps). If you plan on completing one preprocessing step for multiple subjects at a time, leave this field blank.

Electrode Location File: This file allows MATLAB to recognize the electrode channels (via BESA software plugins) and label them according to each manufacturer's predefined montage layout scheme. Common montage files are listed below, but other files may be used for select projects in the CAB Lab. You may need assistance to find the correct file.

BioSemi_128_elecN.elp	BioSemi 128-channel system
BioSemi_64_elecN.elp	BioSemi 64-channel system
BrainVision_125_elecN.elp	Brain Products 128-channel system

Data Importing

After clicking 'Import EEG Files', the following GUI will open.



Import_EEG

Import EEG Files to Create MAT Format Data

Files To Import

Project Root Dire... /labs/srslab/data_main/NPG_[

Filename_End: DSCPT

Electrode Location File: BioSemi_128.elp

EEG File: eeg, bdf

(Electrode Type: EEG, Ref, VEOG, HEOG, EMG, etc)

Data Parameters

Resample Hz High-pass Hz Low-pass Hz

Importing EEG files typically takes a long time. For this reason, it is recommended to import all files slated for preprocessing together in batch. Depending on the number of files, this step may take several hours.

Source Directory: Select where your unprocessed recordings are stored. This will likely be the folder specific to your project on /**data_staging/**.

Project Root Directory: Automatically populated from the main ICAcleanEEG GUI.

Filename_End: Type the same name used within the main ICAcleanEEG GUI.

EEG File Format: For BioSemi systems, **'bdf'**. For Brain Products systems, **'eeg'**.

Electrode Location File: Automatically populated from the main ICAcleanEEG GUI.

Data Parameters: Recommended parameters are listed below.

Resample Rate = 256
High-pass Filter = .5
Low-pass Filter = 110

Subject ID: Either manually type subject IDs into the cells or click **'Load Workspace "SubIDs"'**. This latter option works if a MATLAB array has already been defined with your subject list. To create this array labeled **'SubIDs'**, you can use the code below and fill in your own subject numbers. If you need to clear the list of subject IDs, click **'Reset "SubIDs"'**.

```
subIDs = {'10012100';
          '51312100';
          …};
```

Define Electrodes: Use these fields to specify the nature of each of your electrodes (e.g., scalp EEG electrodes, reference, vertical electrooculography, etc.). If you're working with data recorded using BrainVision hardware, nothing needs to be entered; simply press the "run srslab *.eeg" button. For a 128-channel BioSemi system, input the numbers below. Then, click **'Run'**.

	Type	Electrode #
1	EEG	1:128
2	Ref	129:130
3	VEOG	131:132
4	HEOG	133:134
5	rEKG	135:136
6	lEKG	137:138

badElec Measures

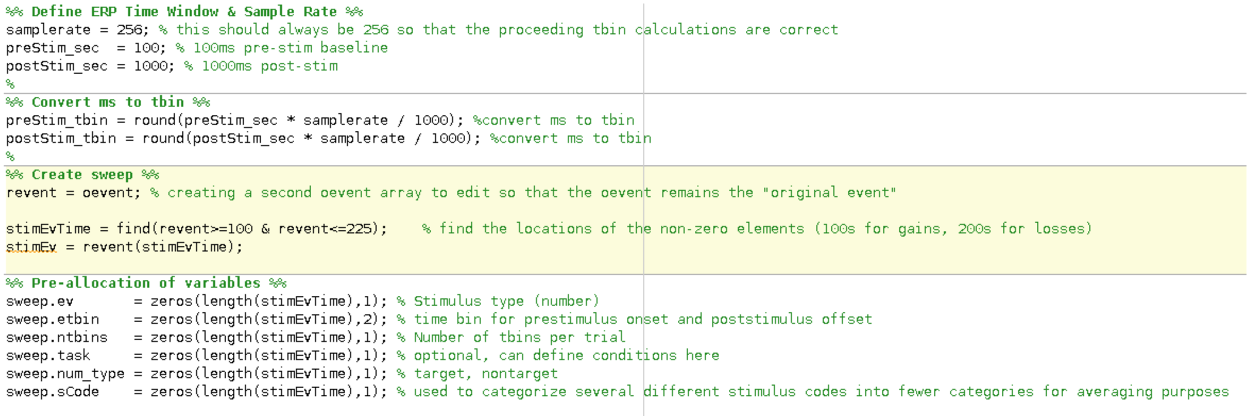
After importing, click on the **'badElec Measures'** button in the GUI. This step creates the ***_badElec.mat** file and therefore needs to be completed before loading data in **'badElec Inspection'**.

Create sweep

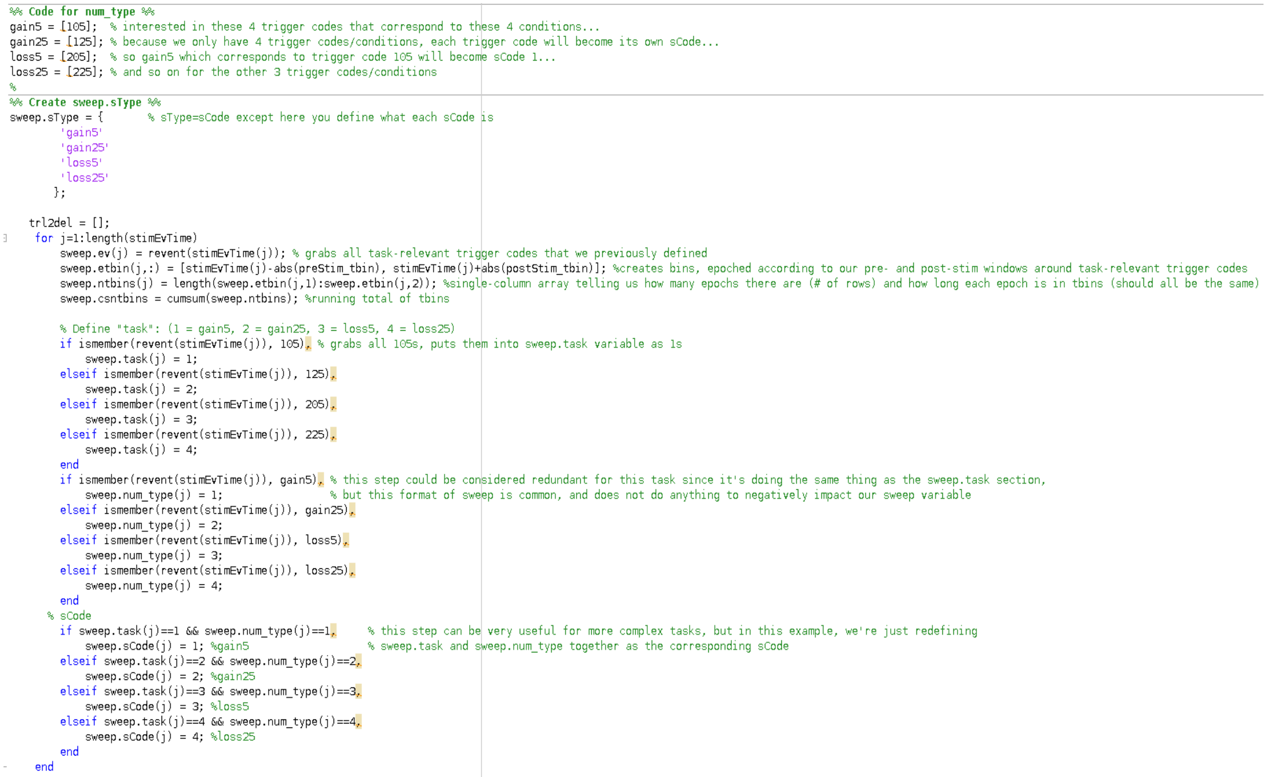
At this stage, the data is still in continuous format, and ICAcleanEEG functions do not recognize what events are happening when, and which event codes are important. To manually inspect the event codes, enter the commands below. This will display all of the relevant trigger codes in the oevent matrix.

plot(oevent(oevent~=-1))	Plot all oevent codes without the -1 placeholder values
unique(oevent)	List all oevent codes without duplicates

After confirming that the ***_oevent.mat** file is correct, we can run a "sweep" on the file. This will epoch the data around the trigger codes identified as being task-relevant using whatever pre- and post-stimulus window is chosen. Note that your epoch windows need to be **at least** 1 second long (e.g. -100ms pre-stim, 900ms post-stim) to be compatible in ICAcleanEEG. Below are screenshots from an example sweep script with comments.



```
%% Define ERP Time Window & Sample Rate %%
samplerate = 256; % this should always be 256 so that the proceeding tbin calculations are correct
preStim_sec  = 100; % 100ms pre-stim baseline
postStim_sec = 1000; % 1000ms post-stim
%
%% Convert ms to tbin %%
preStim_tbin = round(preStim_sec * samplerate / 1000); %convert ms to tbin
postStim_tbin = round(postStim_sec * samplerate / 1000); %convert ms to tbin
%
%% Create sweep %%
revent = oevent; % creating a second oevent array to edit so that the oevent remains the "original event"

stimEvTime = find(revent>=100 & revent<=225);    % find the locations of the non-zero elements (100s for gains, 200s for losses)
stimEv = revent(stimEvTime);

%% Pre-allocation of variables %%
sweep.ev       = zeros(length(stimEvTime),1); % Stimulus type (number)
sweep.etbin    = zeros(length(stimEvTime),2); % time bin for prestimulus onset and poststimulus offset
sweep.ntbins   = zeros(length(stimEvTime),1); % Number of tbins per trial
sweep.task     = zeros(length(stimEvTime),1); % optional, can define conditions here
sweep.num_type = zeros(length(stimEvTime),1); % target, nontarget
sweep.sCode    = zeros(length(stimEvTime),1); % used to categorize several different stimulus codes into fewer categories for averaging purposes
```

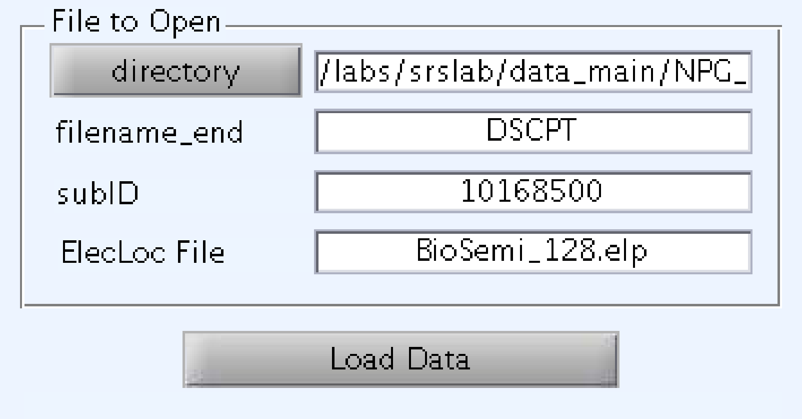


```
%% Code for num_type %%
gain5 = [105];  % interested in these 4 trigger codes that correspond to these 4 conditions...
gain25 = [125]; % because we only have 4 trigger codes/conditions, each trigger code will become its own sCode...
loss5 = [205];  % so gain5 which corresponds to trigger code 105 will become sCode 1...
loss25 = [225]; % and so on for the other 3 trigger codes/conditions
%
%% Create sweep.sType %%
sweep.sType = {       % sType=sCode except here you define what each sCode is
        'gain5'
        'gain25'
        'loss5'
        'loss25'
       };

   trl2del = [];
    for j=1:length(stimEvTime)
        sweep.ev(j) = revent(stimEvTime(j)); % grabs all task-relevant trigger codes that we previously defined
        sweep.etbin(j,:) = [stimEvTime(j)-abs(preStim_tbin), stimEvTime(j)+abs(postStim_tbin)]; %creates bins, epoched according to our pre- and post-stim windows around task-relevant trigger codes
        sweep.ntbins(j) = length(sweep.etbin(j,1):sweep.etbin(j,2)); %single-column array telling us how many epochs there are (# of rows) and how long each epoch is in tbins (should all be the same)
        sweep.csntbins = cumsum(sweep.ntbins); %running total of tbins

        % Define "task": (1 = gain5, 2 = gain25, 3 = loss5, 4 = loss25)
        if ismember(revent(stimEvTime(j)), 105), % grabs all 105s, puts them into sweep.task variable as 1s
            sweep.task(j) = 1;
        elseif ismember(revent(stimEvTime(j)), 125),
            sweep.task(j) = 2;
        elseif ismember(revent(stimEvTime(j)), 205),
            sweep.task(j) = 3;
        elseif ismember(revent(stimEvTime(j)), 225),
            sweep.task(j) = 4;
        end
        if ismember(revent(stimEvTime(j)), gain5), % this step could be considered redundant for this task since it's doing the same thing as the sweep.task section,
            sweep.num_type(j) = 1;                 % but this format of sweep is common, and does not do anything to negatively impact our sweep variable
        elseif ismember(revent(stimEvTime(j)), gain25),
            sweep.num_type(j) = 2;
        elseif ismember(revent(stimEvTime(j)), loss5),
            sweep.num_type(j) = 3;
        elseif ismember(revent(stimEvTime(j)), loss25),
            sweep.num_type(j) = 4;
        end
      % sCode
        if sweep.task(j)==1 && sweep.num_type(j)==1,     % this step can be very useful for more complex tasks, but in this example, we're just redefining
            sweep.sCode(j) = 1; %gain5                   % sweep.task and sweep.num_type together as the corresponding sCode
        elseif sweep.task(j)==2 && sweep.num_type(j)==2,
            sweep.sCode(j) = 2; %gain25
        elseif sweep.task(j)==3 && sweep.num_type(j)==3,
            sweep.sCode(j) = 3; %loss5
        elseif sweep.task(j)==4 && sweep.num_type(j)==4,
            sweep.sCode(j) = 4; %loss25
        end
    end
```

5. BadElec Inspection

Opening BadElec

Click on **'badElec Inspection'** and the following BadElec GUI will open.



directory: /labs/srslab/data_main/NPG_

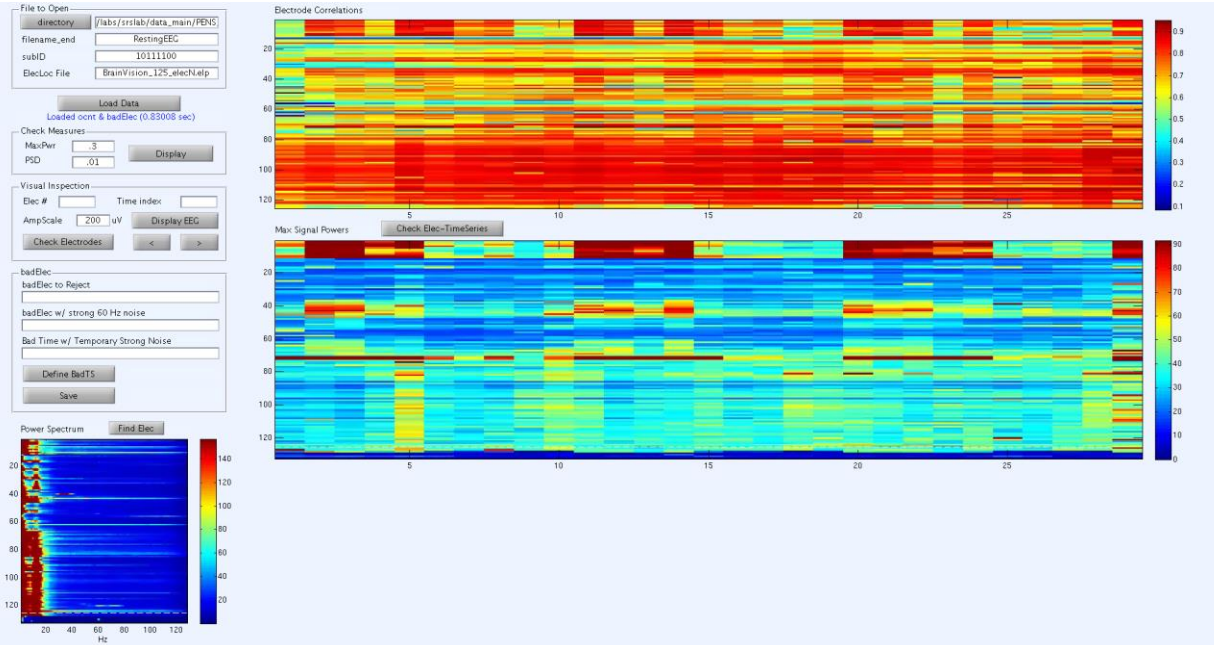
filename_end: DSCPT

subID: 10168500

ElecLoc File: BioSemi_128.elp

Ensure all fields in the **'File to Open'** section contain the correct information. The **'directory'** field should match the **'Project Root Directory'** from the main ICAcleanEEG GUI. Similarly, **'filename_end'** should match **'Project Name'**, and **'ElecLoc File'** should match **'Electrode Location File'**. Type in the desired subject ID number in **'subID'** and then click **'Load Data'**. Depending on the file size, it may take 2-20 seconds to load. If it does not load properly, an error will be displayed in the MATLAB command window.

Once loaded, click **'Display'** to view several different plots.



directory: /labs/srslab/data_main/PENS

filename_end: RestingEEG

subID: 10111100

ElecLoc File: BrainVision_125_elecN.elp

Loaded ocnt & badElec (0.83008 sec)

MaxPwr: .3

PSD: .01

Elec #:

Time index:

AmpScale: 200 uV

badElec to Reject:

badElec w/ strong 60 Hz noise:

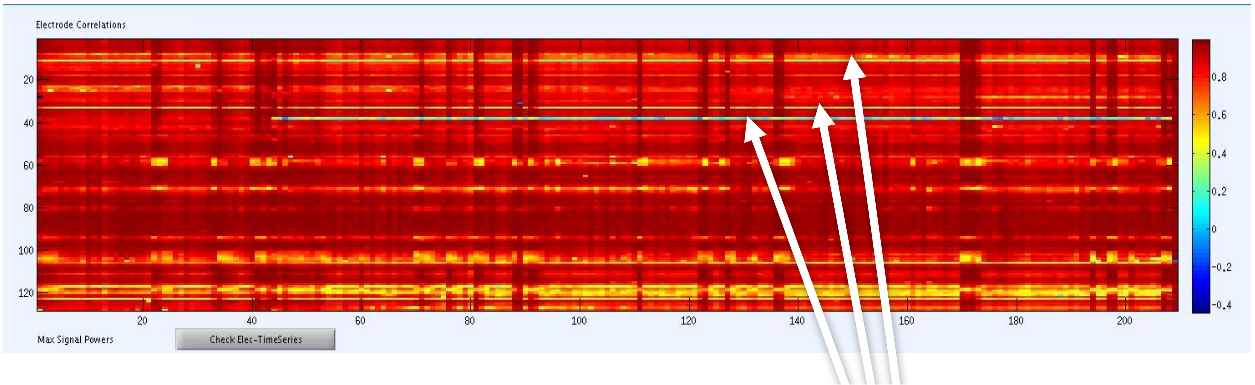
Bad Time w/ Temporary Strong Noise:

BadElec features

Several visualization tools are available to identify isolated bad electrodes, bad epochs, or 60 Hz power line noise contamination. These include displays of electrode correlations, max signal power, power spectrum densities, a “Check Electrodes” tool, and a time series plot. Each visualization tool is explained in more detail below.

Electrode correlation plot

The degree to which signal in an electrode is correlated with its nearest neighboring electrodes on the scalp is plotted below. Strong positive correlations are in red, strong negative correlations are in dark blue, and zero correlation is in light blue/turquoise. Due to signal conduction across the skull, electrophysiological signals from the brain are blurred in space across many electrodes. Strong positive correlations (i.e., red color) are expected for all electrodes. Deviation from that pattern could be an indication of isolated noisy electrodes due to non-physiological reasons.

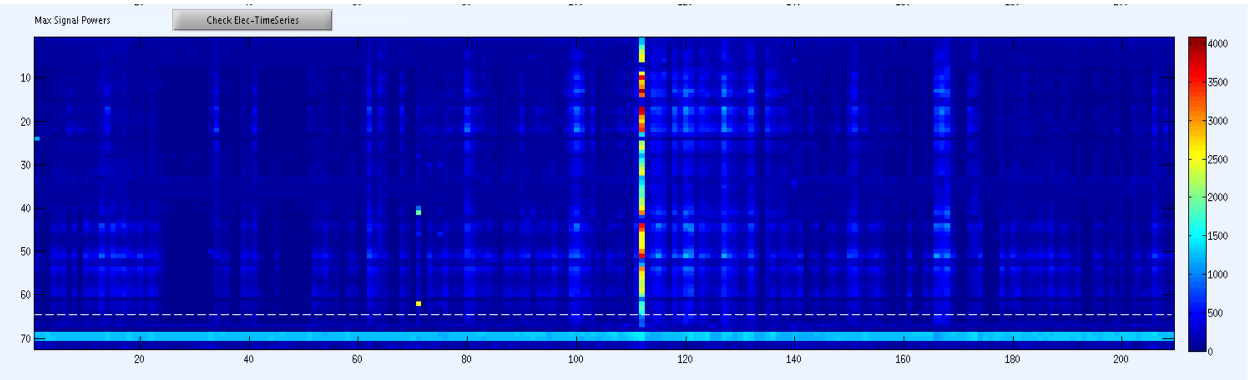


(Electrodes that need to be more closely inspected.)

Large visible bands of low correlation values in either the vertical or horizontal directions could indicate problematic signal that needs to be deleted or filtered. On this plot, the horizontal axis shows time in seconds and the vertical axis shows electrodes (electrode #1 at the top, electrode #128 at the bottom). Any plot areas that are not red should be inspected further to see if there are bad epochs or bad electrodes present. In the example above, there are at least 3 electrodes that need to be inspected for deletion/filtering.

Maximum signal power plot

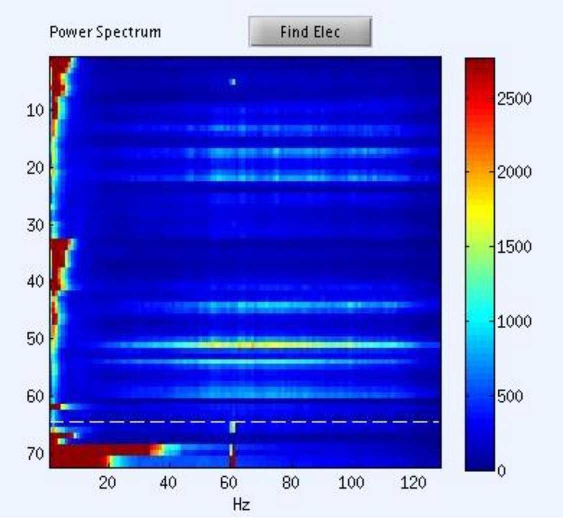
Another useful figure for visualizing your data is the display of maximum signal power. Like the electrode correlations plot, the horizontal axis depicts time in seconds while the vertical axis corresponds to different electrodes. Power is mathematically defined as the square of the signal amplitude (i.e., microvolts^2). Large positive or large negative deflections in signal amplitude will be equally represented by the power index. Maximum power of each electrode within ten second recording epochs is depicted with a color scale. Physiological signal is unlikely to be reflected as large maximum amplitude deflections and should appear blue. Larger amplitude deflections that are less plausibly the result of physiological sources will be displayed as red. Thus, visual bands of warmer color in either the horizontal or vertical directions should be inspected further. In this example recording (below), there is one time segment that needs to be examined.



At the bottom of this plot, there is a dotted line. This section of the plot contains information from the 10 peripheral (i.e., external, "EX") electrodes measuring signal from the face and body. These 10 EX channels are not included in the electrode correlations plot because they are not situated near one another on the body. We do not expect EX channels be strongly correlated.

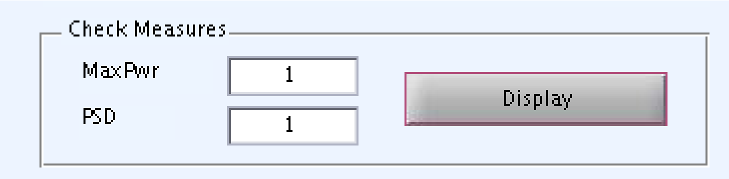
Power spectrum plot

Examination of the power spectrum across a range of signal frequencies can reveal similar information as the Max Signal Power plot. An important difference is that the Power Spectrum figure displays frequency (in Hertz) along the horizontal axis. Examination of the frequencies represented in each electrode can be useful for detecting noise isolated to particular frequency ranges. For clean recordings of neurophysiological activity, power should be strongest below approximately 15 Hz. It is common to observe 60 Hz noise in some or all of the electrodes due to properties of electrical current in the United States. 10 EX electrodes are again included in the bottom portion of the plot.



Adjusting the scaling of individual plots

On the left side of the screen within the **'Check Measures'** box, there are two fields: **'MaxPwr'** and **'PSD'**. These parameters allow for adjustment of the displayed range in the Max Signal Powers and Power Spectrum plots, respectively. Default values of '1' will appear upon loading data into badElec. Users may change the fields to whatever values work best for visualizing differences across time and across various electrodes.

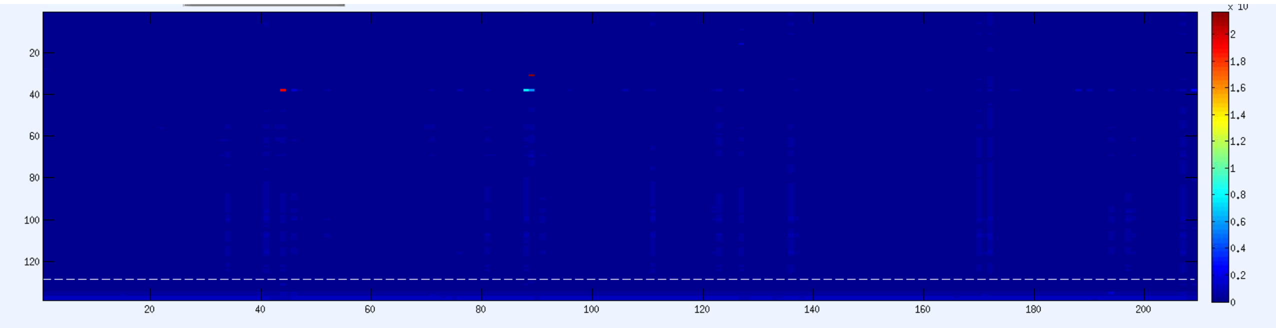


Check Measures

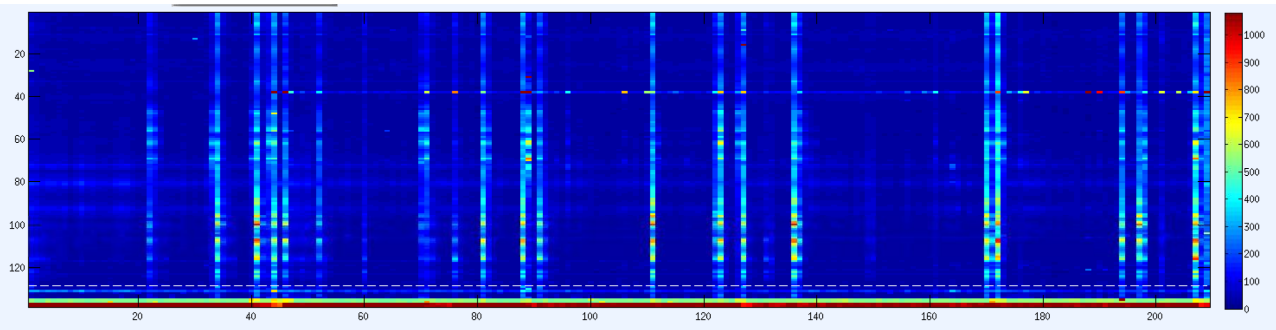
MaxPwr 1

PSD 1

For example, here is the Max Signal Powers plot when the **'MaxPwr'** is set to 1:



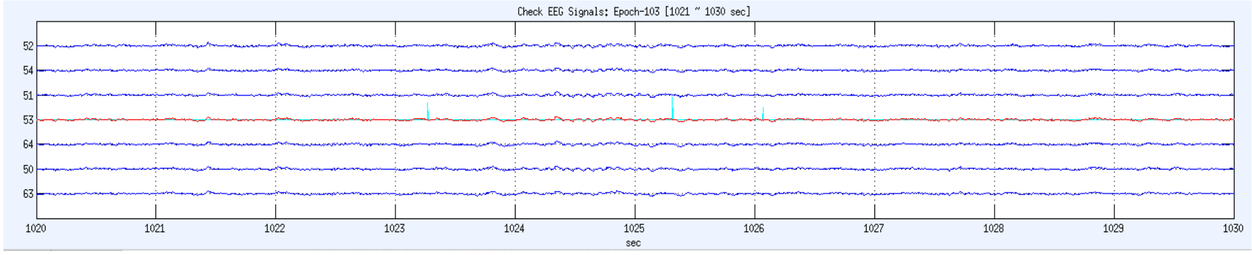
And here is the same data with a **'MaxPwr'** value of .05:



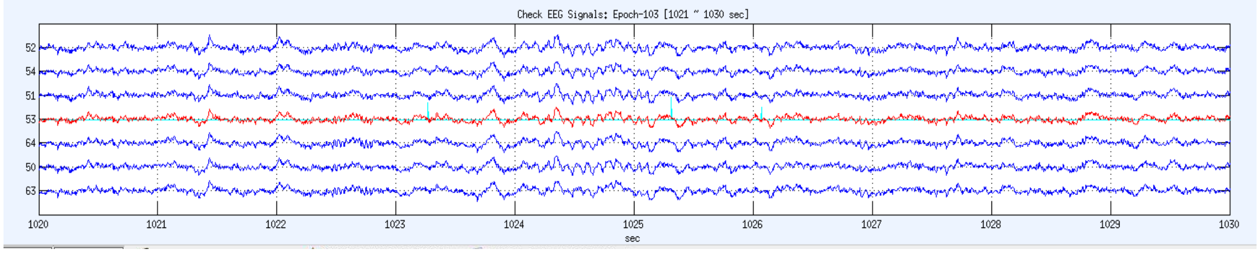
By changing the scaling, it became easier to detect time points or electrodes that may require special attention. Generally, it is helpful to decrease these values sequentially by an order of 10 (e.g., 1, .1, .01, .001, etc.). Values between .1 - .00001 are commonly applied for typical EEG recordings. Each file is different. Users may have to try different values to see what works best. For most files, **.1 for the max signal powers** and **.001 for the power spectrum density** work well. Note that bottom EX channels (especially ECG electrodes) depict very prominent maximum power values in red regardless of the scalp electrode scaling. This is to be expected.

Time series display

It is good practice to use multiple badElec descriptive plots together for exploring possible artifacts in the recordings. However, it is important to confirm your determinations through actual examination of the recorded signal. By clicking **'Check Elec-Time Series'**, the mouse icon will change into two intersecting lines resembling crosshairs. Users can now point and click onto any area within the Electrode Correlations or Max Signal Powers plots to view the actual channel electrical signal. Each time series window contains ten second segments of the recording with time labeled on the x-axis and electrode numbers displayed on the y-axis.



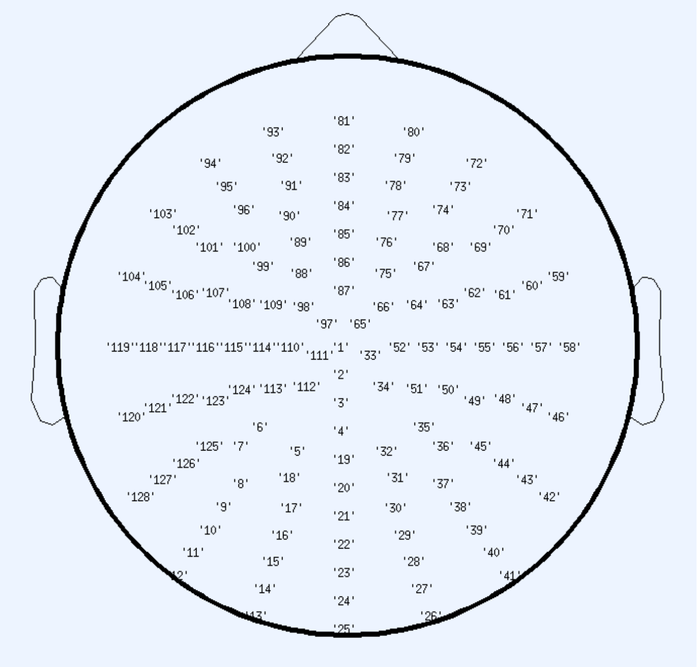
It is often helpful to change the scaling of the waveforms until discernible neurophysiological signal can be observed (i.e., not horizontal lines). If the channels look too smooth, try adjusting the value of **'AmpScale'** in the 'Visual Inspection' box on the left of the screen. Smaller values result in more accentuated depictions of the peaks while larger values will diminish the appearance of the peaks. Once the scaling is adjusted, neural signal may become easier to distinguish. For the majority of cases, setting the scale to **100uV** will suffice.



Each time series plot shows signal from one primary electrode channel (red color) as well as signal from the nearest neighboring electrodes on the scalp (blue color). Users can scroll backwards and forwards through the entire time series using the right and left arrow buttons in the 'Visual Inspection' box. The primary electrode displayed can be changed by entering a new value into the **'Elec #'** field. The **'Time index'** field can also be changed to jump to a new section in the recording. Press the **'Display EEG'** button to activate these changes in the display.

To see where a particular electrode is located on the scalp, click **'Check Electrodes'**. This will bring up a map of the electrode locations. Appreciating the topographical locations of the recording electrodes may help clarify why an electrode is consistently noisy or the reason behind consistent muscle artifact in the signal (e.g., clenching jaw muscles distorting left/right temporal

EEG signal). Here is an example electrode location topographical map for the 128-channel Brain Products system:



Check electrodes tool

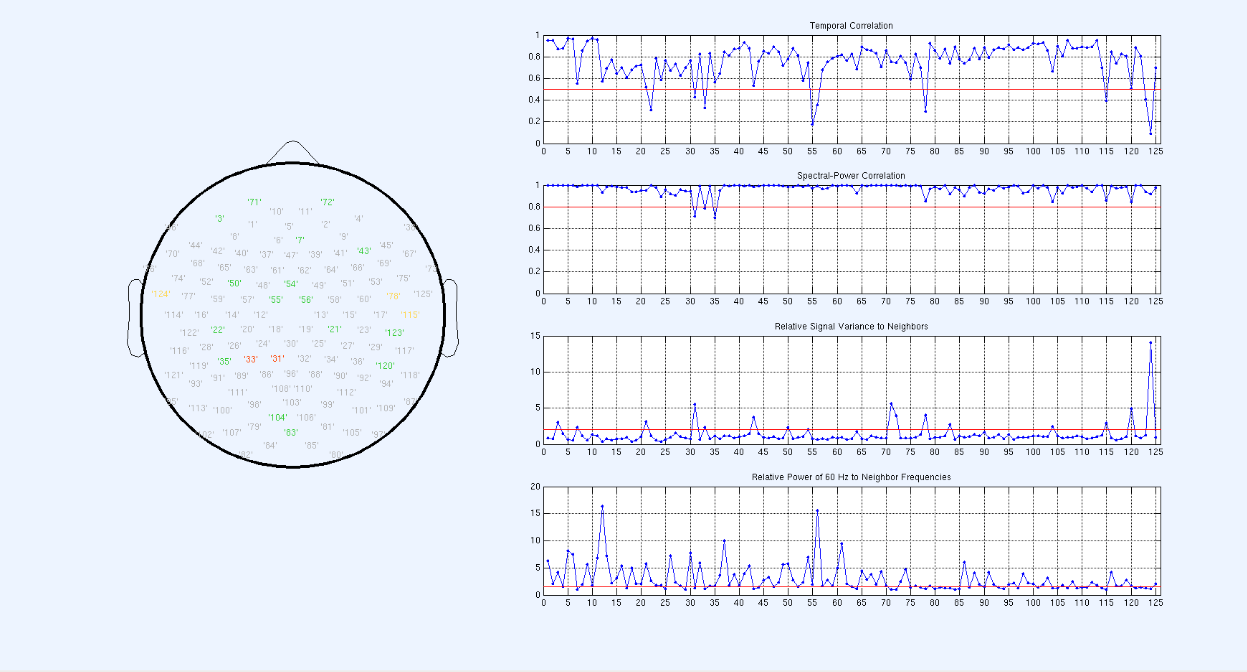
Another method for identifying bad electrodes is the 'Check Electrodes' tool. Clusters of neighboring electrodes on the scalp should exhibit similar patterns in electrical activity. This tool displays four different calculations of similarity. Each has a statistically meaningful threshold, which may (or may not) be helpful for determining when an electrode is problematic. No single metric is necessarily indicative of electrode noise. As with the main 'badElec' plots, multiple Check Electrodes indices should be used together to make determinations. When in doubt, always refer back to the time series display of the original channels.

'Temporal Correlation': The degree to which amplitude changes over time in one electrode are similar to neighboring electrodes within the recording.

'Spectral-Power': Similar to the "Temporal Correlation" index, but instead assesses similarity in power across the frequency spectrum observable in the recording relative to electrode neighbors.

'Relative Signal Variance to Neighbors': An isolated bad electrode may produce signal that fluctuates widely relative to the average variance in signal of neighboring electrodes. Isolated elevations in this metric may be justification for electrode exclusions.

'Relative Power of 60 Hz to Neighbor Frequencies': 60 Hz noise produces an isolated increase in power not observable in other portions of the frequency spectrum within the same electrode channel. An increase in 60 Hz noise relative to frequencies just above and below 60 Hz will be indicative of this type of noise and will be reflected as an increase in this index.

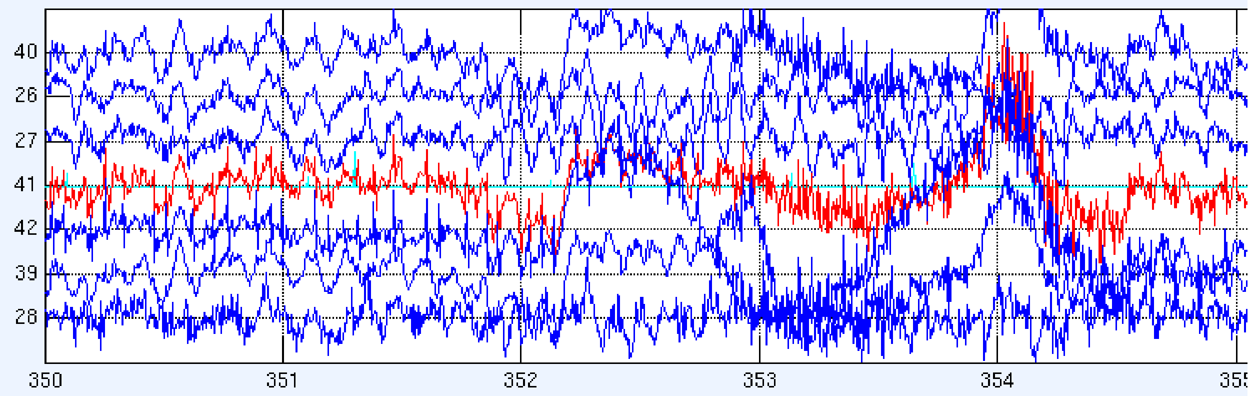


On the right side of the ‘Check Electrode’ display are the four indices. On each plot, the x-axis lists each electrode and the y-axis shows the range for each parameter. A red line indicates a statistically meaningful threshold. The electrode topographical map is shown on the left. Numbers on the map are displayed in color when one or more of the statistical thresholds have been exceeded. The color progresses towards red based on the number of indices that exceed the thresholds: 1 = lime green, 2 = gold, 3 = orange-red, 4 = dark-red (less common, not pictured above). Thresholds should be considered crude starting points for interpreting these indices, and might be justifiably ignored in some cases. Further research is needed about these chosen thresholds.

Marking bad electrodes/time segments for deletion or filtering

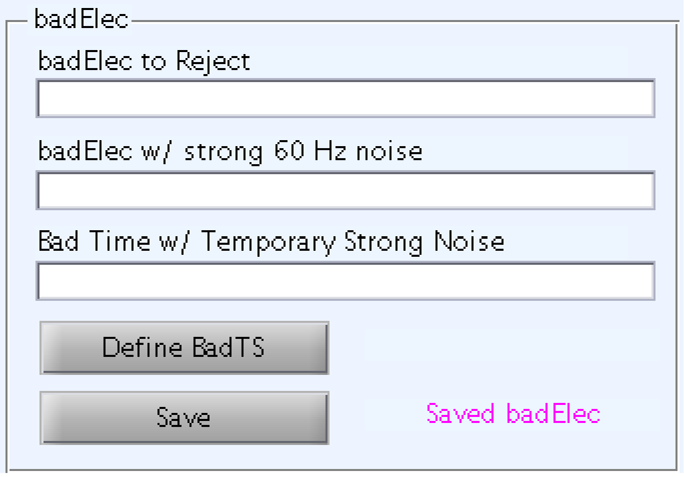
ICA is a powerful tool for isolating subtle sources of noise across time and space in recordings. However, ICA can have difficulty detecting these subtleties when strong sources of noise are not removed first. Deleting or filtering electrodes and time segments with obvious sources of noise is important. This will increase the amount of signal and decrease the noise submitted to ICA, which enhances your probability of getting a useful decomposition of the data.

What makes an electrode or time segment “bad?” The following example time segment shows several electrodes that differs from their surrounding electrodes:

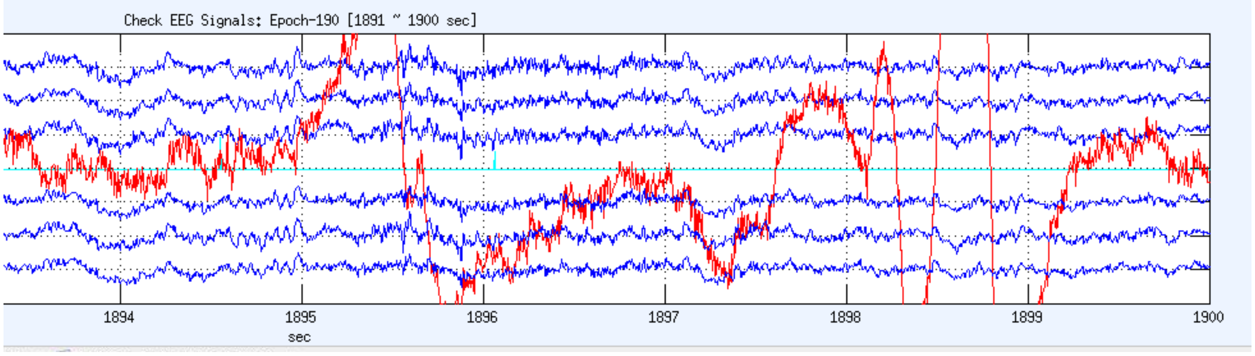


ICAcleanEEG allows users to delete and interpolate electrodes from the entire recording, but this is disadvantageous when the noise is isolated to a brief period of time. Electrodes can temporarily 'act up' for a variety of reasons. Electrode 39, 41, and 42 above are certainly noisy, but it may be excessive to exclude them outright. In cases like this when the noise is isolated to specific points in time, it makes more sense to simply remove the time segment rather than sacrificing the entire channels.

ICAcleanEEG denotes epochs based on the number at the end of each one second time segment. Epoch 353 contains the waveforms to the left of the line marked 353 (i.e., between “352” and “353” on the x-axis). For this current example, time segments 353-355 would be deleted by typing ‘**353:355**’ into the “Bad Time w/ Temporary Strong Noise” field. Additional time segments could be entered into this field using the space bar (e.g., ‘**353:355 467 654**’).

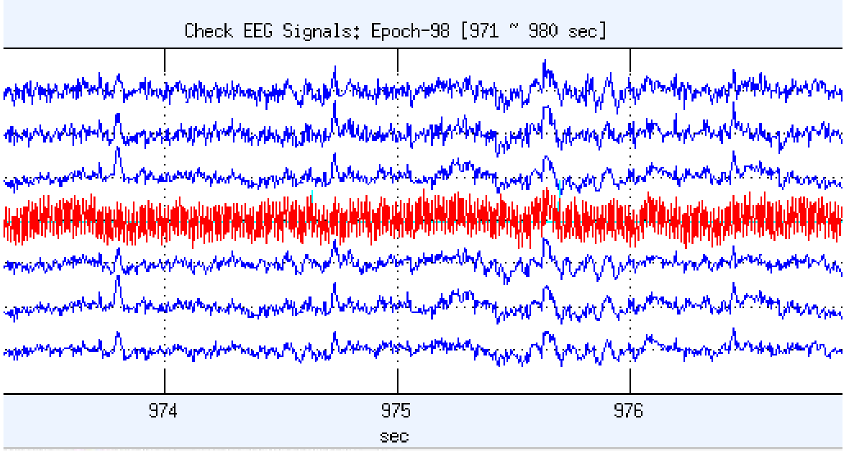

badElec
badElec to Reject
badElec w/ strong 60 Hz noise
Bad Time w/ Temporary Strong Noise
Define BadTS
Save
Saved badElec

If an electrode consistently displays wild fluctuations in its signal across the recording, it is likely a bad electrode. Below is an example of what the time series of a bad electrode would look like:

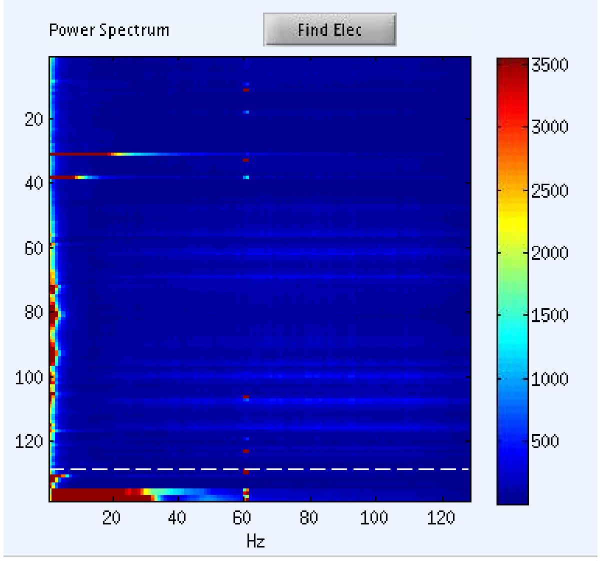


This electrode should be added to the **'BadElec to Reject'** field. The CAB Lab typically excludes individual electrodes when noise is noticeable **for at least half of a recording**. Bad electrodes with shorter segments of noise are sometimes deleted as well. For the example above, this significant noise would likely be isolated to its own component (even if the duration was only several seconds). Users should note electrodes like these and be mindful of their effects on ICA performance. Users may later decide to delete such electrodes and perform ICA2.

60 Hz noise from the surrounding environment is sometimes observed in CAB Lab recordings. Electrodes exhibiting this noise are sometimes easy to spot when the noise is strong. As seen below, 60 Hz noise contamination makes the signal appear denser or more saturated. This contamination would make neural signal difficult or impossible to detect. To fix this, set aside this electrode for "notch" filtering by listing it within the **'BadElec w/ strong 60 Hz noise'** field.



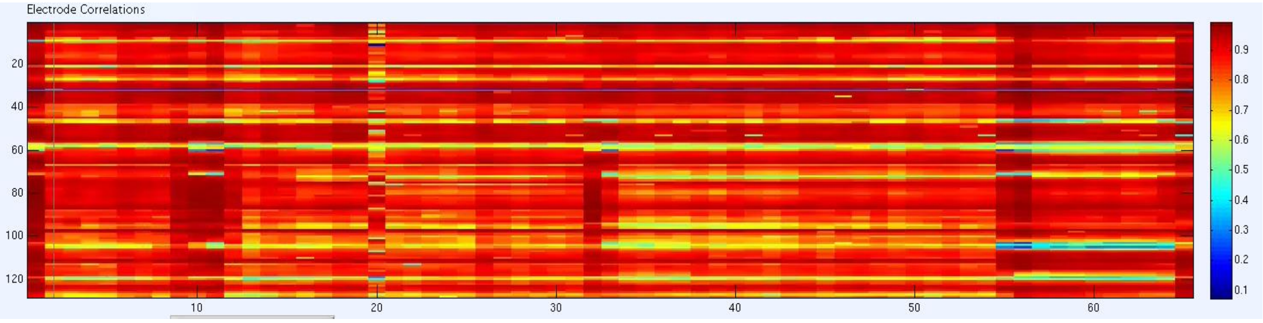
Other instances of 60 Hz noise are subtler. The Power Spectrum figure makes it easier to find affected electrodes. For example, locate 60 Hz on the horizontal axis within the screenshot below. As you look vertically around 60 Hz, you can see several electrodes with non-blue color.



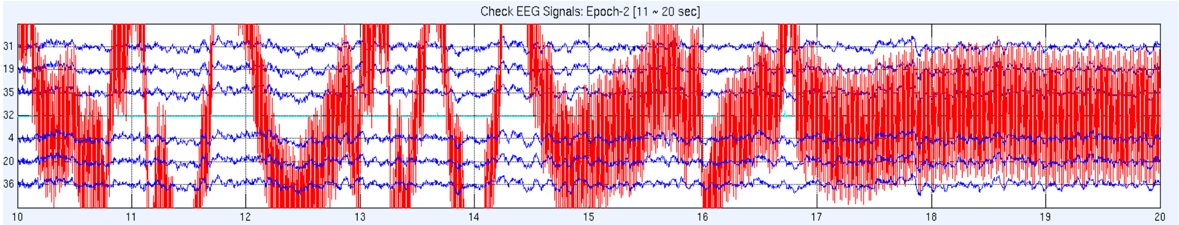
This could mean that several electrodes have been contaminated by 60 Hz noise, which needs to be filtered out. To pinpoint electrodes precisely, click ‘Find Elec’ and select the red color square of interest. The electrode number selected will be displayed next to the ‘Find Elec’ button in blue. Type that number into the ‘Elec #’ field in the ‘Visual Inspection’ box to check the electrode time series. Only then can you make a confident decision about the use of a 60 Hz notch filter.

Notice underneath the dashed lines that ‘Ref1’ and ‘Ref2’ EX electrodes (two reference electrodes attached to the participant’s earlobes) also contain a fair amount of 60 Hz noise. This is *good*. We want these electrodes to pick up environmental noise so that we can filtered it out from the EEG electrodes using either linked ear or average head referencing. Any additional 60 Hz noise isolated to specific electrodes needs to be filtered out because the ‘Ref1’ and ‘Ref2’ electrodes may not be adequately capturing that noise.

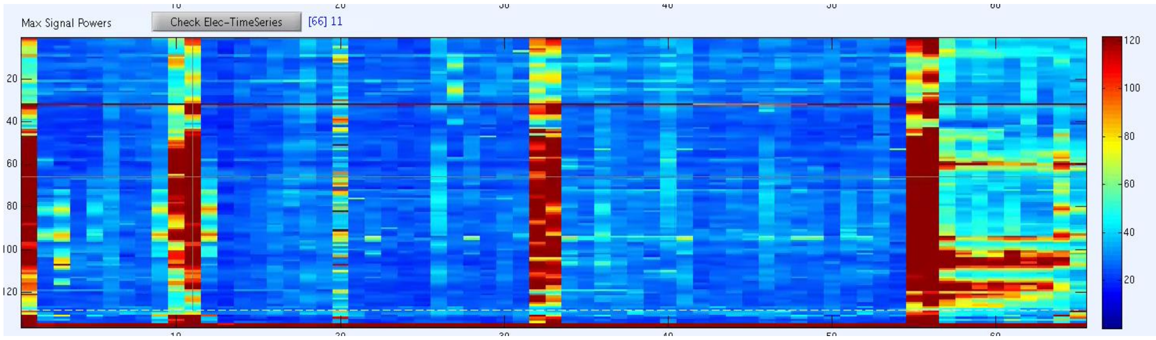
The ‘Electrode Correlations’ and ‘Max Signal Powers’ plots greatly enhance searches for problematic electrodes and time segments. Within the following ‘Electrode Correlations’ example, you can see a dark blue horizontal band at electrode 32. This electrode is only weakly correlated with nearby electrodes for the entirety of the recording.



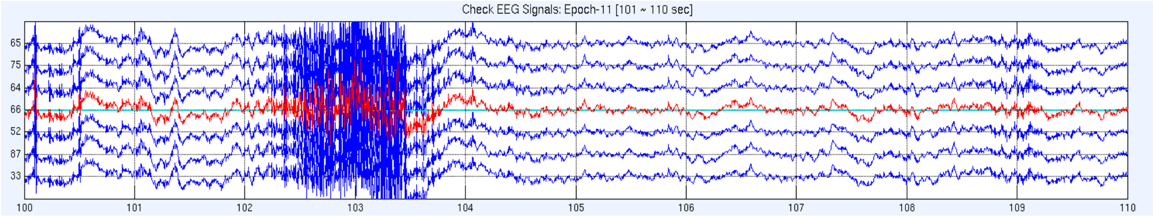
The ‘Check Elec-Time series’ cursor can be used to select the electrode in the ‘Electrode Correlations’ plot. The time series below shows the extent of the noise isolated to electrode 32.



The ‘Max Signal Powers’ plot also confirms the presence of this noise at electrode 32. Furthermore, there appear to be multiple problematic time segments as evidenced by the vertical stripes of red color in the plot. These time segments should be investigated further.

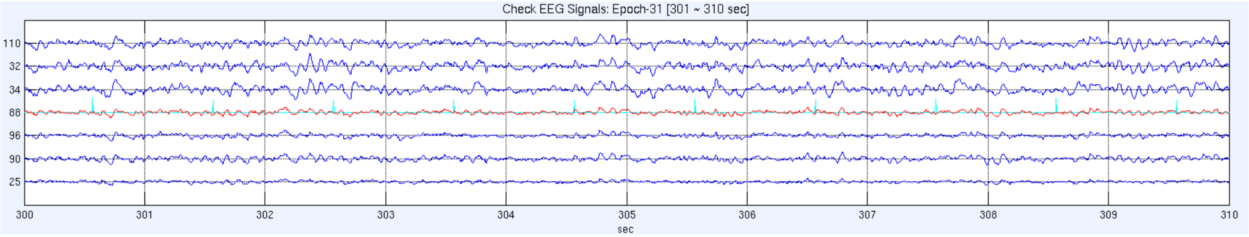


The ‘Check Elec-Time series’ function can also be used to click on these red vertical stripes in the ‘Max Signal Powers’ plot. For this example, the time series between 100-110 seconds was selected.



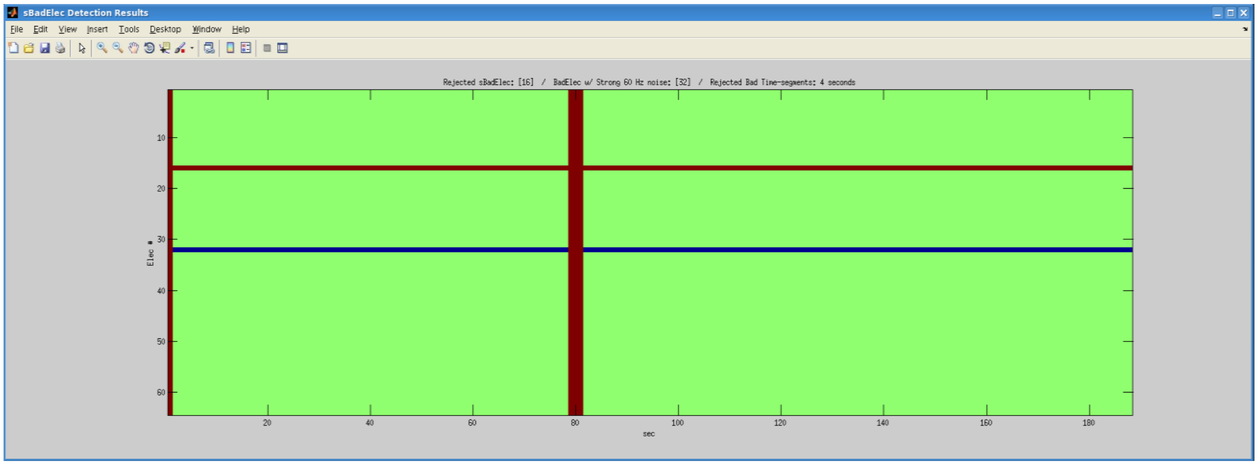
Time segments 103 and 104 contain a high energy burst. Users are faced with a difficult decision regarding this time segment. Although this segment is indeed noisy, there is a good probability that the ICA decomposition will be able to isolate this muscle noise from the underlying neural signal. CAB Lab personnel have developed a set of guidelines to assist with these choices and enhance the performance of ICA (see **Appendix A**). In this particular case, the noise is not strong enough to make the signal across the electrodes indistinguishable from one another with the Amp Scale being set to 100uV. If the signal visually appeared like one solid mass of blue color and was too difficult to distinguish between the electrodes, it would be appropriate to delete this isolated time segment.

It is common to see very high frequency noise during the first few epochs of files. These time segments should not necessarily be deleted. These artifacts occur when recordings are started or paused to give instructions to participants. In the screenshot below, small cyan-colored vertical tick-marks are visible in the middle of each epoch. These tick-marks denote where events/trigger codes occur in the file. If there are noisy time segments located far away from tick-marks (e.g., ±1 second), they do not need to be deleted. Those time segments will not be epoched when the sweep variable is applied. Therefore, the time segments are irrelevant and will not be included during any ICA.



Summary screen

Once you have entered all the necessary information for the bad electrodes, bad time segments, and electrodes needing 60 Hz filter, click **‘Save’** in the ‘badElec’ box. A summary screen will appear and show what parts of the recording have been excluded.


sBadElec Detection Results
File Edit View Insert Tools Desktop Window Help
Rejected sBadElec: [16] / BadElec w/ Strong 60 Hz noise: [32] / Rejected Bad Time-segments: 4 seconds
10
20
30
40
50
60
Elec #
20
40
60
80
100
120
140
160
180
sec

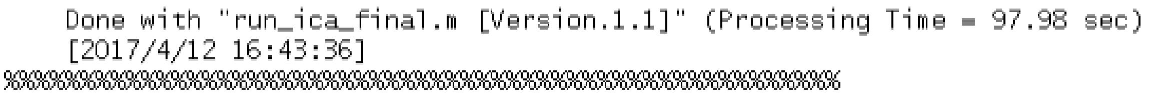
Vertical columns in crimson are time segments that have been deleted from the file. Horizontal lines in maroon are electrodes that have been deleted. Horizontal blue lines are electrodes that have been filtered for 60 Hz noise. If this is accurate, close both this window and the ‘BadElec_inspection’ GUI to continue with the next steps in preprocessing.

Important note

This previous section went into detail on using the ‘BadElec’ GUI to mark bad electrodes and time segments. The reality of preprocessing is more complicated. There are many ambiguous cases requiring more thorough deliberation. Potential users are strongly encouraged to become familiar with the examples and guidelines in **Appendix A**. In the CAB Lab, thoroughly understanding these examples is required to maintain consistency across users. There may be strange files with unusual issues that are not addressed in this manual. Ask questions to people more experienced with ICAcleanEEG rather than guessing. This will save time and help you learn more quickly.

6. Ecnt and bad epochs

Following badElec inspection, ICAcleanEEG will create the ***_ecnt.mat** file by separating the recording into time-locked epochs positioned on the onset of task events identified by the sweep variable. Additionally, it will calculate and automatically detect bad epochs to be deleted for ICA. After clicking **'ecnt & bad epochs'**, this process will run automatically. It usually takes a few minutes per subject. You will know when the process is completed when a message similar to the one below is displayed.



```
    Done with "run_ica_final.m [Version.1.1]" (Processing Time = 97.98 sec)
    [2017/4/12 16:43:36]
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

If you receive error messages at this step, you are most likely using a version of MATLAB that is too new. ICAcleanEEG was initially developed in MATLAB version 2012a. We are in the processing of updating the code for MATLAB version 2018. Many ICAcleanEEG functionalities will work in newer versions of MATLAB, but the code is less reliable. There are other things that could be contributing to error messages at this stage. If you have access, visit the 'Encountered MATLAB Errors and Possible Fixes' Google Doc saved within 'Sponheim Lab Documents' for more information.

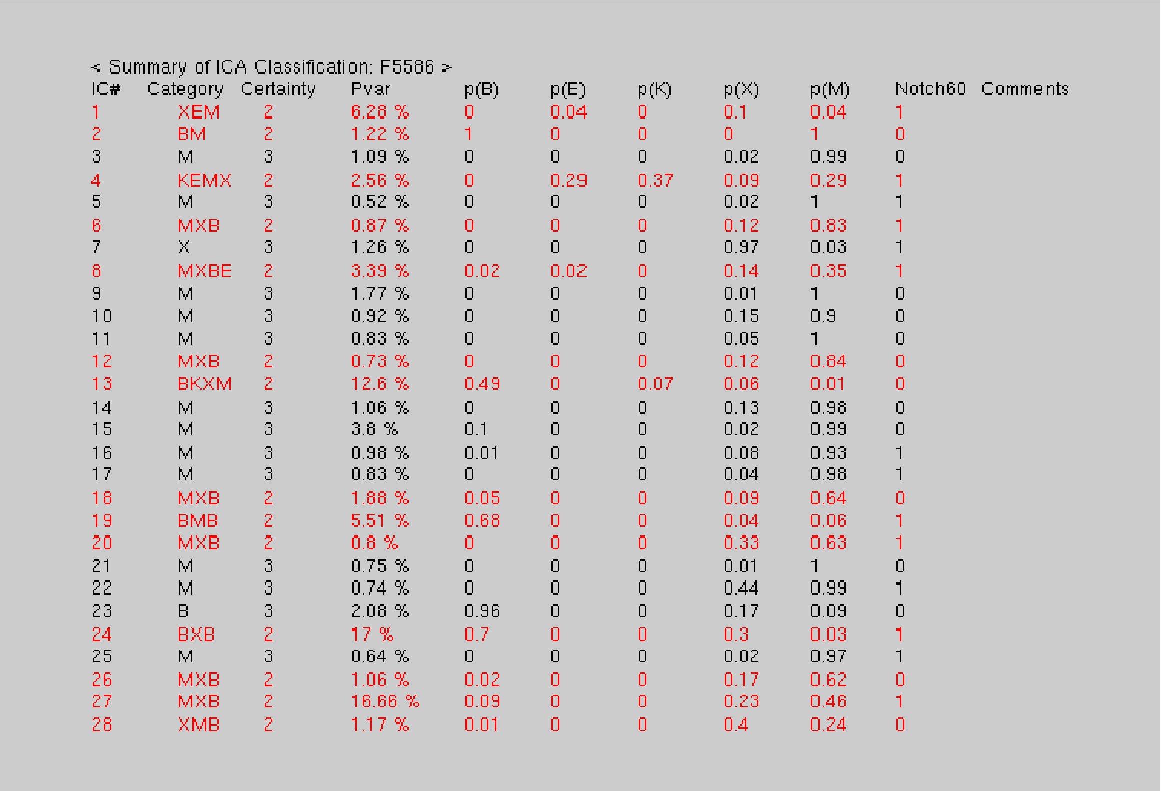
7. ICA Inspection

The last step of preprocessing involves classifying the components produced by ICA performed upon the epoched individual subject recordings. ICAcleanEEG can perform this decomposition with little input from users, but it is necessary to confirm or adjust the automatic classifications generated from the included classification algorithms. This is the step that requires the most expertise from users.

The ICA Inspection GUI will be explained first followed by a brief tutorial on visual identification of common components. Example ICA decompositions will be depicted to demonstrate the procedures. It is common for CAB Lab users to reference this section while making judgments about components. You are encouraged to keep this manual open on a separate computer monitor while you make decisions about ICs.

ICA Inspection Features

Files are opened in the ICA Inspection GUI by clicking the **'Load Data'** button. A summary window appears with a complete listing of the automatic component classifications generated by ICAcleanEEG.



< Summary of ICA Classification: F5586 >

IC#	Category	Certainty	Pvar	p(B)	p(E)	p(K)	p(X)	p(M)	Notch60	Comments
1	XEM	2	6.28 %	0	0.04	0	0.1	0.04	1	
2	BM	2	1.22 %	1	0	0	0	1	0	
3	M	3	1.09 %	0	0	0	0.02	0.99	0	
4	KEMX	2	2.56 %	0	0.29	0.37	0.09	0.29	1	
5	M	3	0.52 %	0	0	0	0.02	1	1	
6	MXB	2	0.87 %	0	0	0	0.12	0.83	1	
7	X	3	1.26 %	0	0	0	0.97	0.03	1	
8	MXBE	2	3.39 %	0.02	0.02	0	0.14	0.35	1	
9	M	3	1.77 %	0	0	0	0.01	1	0	
10	M	3	0.92 %	0	0	0	0.15	0.9	0	
11	M	3	0.83 %	0	0	0	0.05	1	0	
12	MXB	2	0.73 %	0	0	0	0.12	0.84	0	
13	BKXM	2	12.6 %	0.49	0	0.07	0.06	0.01	0	
14	M	3	1.06 %	0	0	0	0.13	0.98	0	
15	M	3	3.8 %	0.1	0	0	0.02	0.99	0	
16	M	3	0.98 %	0.01	0	0	0.08	0.93	1	
17	M	3	0.83 %	0	0	0	0.04	0.98	1	
18	MXB	2	1.88 %	0.05	0	0	0.09	0.64	0	
19	BMB	2	5.51 %	0.68	0	0	0.04	0.06	1	
20	MXB	2	0.8 %	0	0	0	0.33	0.63	1	
21	M	3	0.75 %	0	0	0	0.01	1	0	
22	M	3	0.74 %	0	0	0	0.44	0.99	1	
23	B	3	2.08 %	0.96	0	0	0.17	0.09	0	
24	BXB	2	17 %	0.7	0	0	0.3	0.03	1	
25	M	3	0.64 %	0	0	0	0.02	0.97	1	
26	MXB	2	1.06 %	0.02	0	0	0.17	0.62	0	
27	MXB	2	16.66 %	0.09	0	0	0.23	0.46	1	
28	XMB	2	1.17 %	0.01	0	0	0.4	0.24	0	

ICAcleanEEG employs algorithms that were trained using an “in-house” library of ICA component classifications produced by multiple research assistants. Components from your particular subject are given a probability score reflecting the degree of similarity to these exemplar components from past studies. The automatic classification can serve as a starting point, but users are required to visually inspect each component themselves. The classification designation may change depending on what you observe within the ICA output plots. Below is a description of each summary table column.

'IC#': "Independent component" number. The ordering of each IC can be considered arbitrary for the purposes of EEG preprocessing. Users can generally expect between 9 and 40 ICs depending on the length of the recording, the number of electrodes, the complexity of the signal or noise sources, and idiosyncratic factors related to the automated decomposition of your particular data file. Assigning numbers to ICs makes it easier to communicate with colleagues about ambiguous ICs. Numbers of ambiguous ICs requiring a consensus decision are often copied onto preprocessing tracker spreadsheets and shared by individuals involved with the project.

'Category': A letter labeling system is employed to classify ICs using ICAcleanEEG. Nearly all ICs can be identified using a single letter (e.g., IC is clearly only muscle-related = 'M') or two letters (e.g., IC is primarily brain, but there is some slight muscle signal = 'BM'). If the IC is categorized with two letters, always place the dominant signal source first followed by the secondary signal source. For more complex ICs, user may use three letters to designate the relative prominence of the observed signal sources (e.g., IC mixed with brain, muscle, and bad electrode(s) = 'BMX'). Ideally, this latter case should be less common; most ICs should be suitably classified with two or less letters. The nuance between 'B', 'B_', and 'BB_' is important. Specific techniques such as time-frequency analysis sometimes require a higher threshold of IC "cleanliness" in terms of signal-to-noise. These various brain IC designations are meant to communicate the relative quantity of recorded signal for investigators performing later analyses on post-processed data.

'Certainty': Users should include a confidence rating for each IC: '3' signifies strong confidence in the classification, '2' signifies moderate confidence, and '1' signifies uncertainty and the need for a consensus decision. Data cannot move forward through the ICAcleanEEG preprocessing pipeline when certainty levels of '1' are used for any ICs. In other words, '1' functions as a "red flag" and not a permanent value. For mixed ICs (i.e., two or more component labels), we do not give a certainty level greater than 2. Certainty of '3' is saved for single source IC classifications (e.g., 'B').

'Pvar': Percent of total variance in the epoched recording file. This value signifies the relative importance of each IC in accounting for signal fluctuations in the original data. User decisions for ICs become more consequential as percent of total variance increases. If, in the screenshot above, IC 19 were mislabeled 'M' instead of 'B', then we would ultimately be removing a large portion (i.e., 5.51%) of the variance in the file from our analyses. Careful attention to IC classification ensures users are able to maximize the amount of brain signal extracted from the original recording.

'p(B)/p(E)/p(K)/p(X)/p(M)': These columns provide information that tells you how the pipeline algorithm generated the default classifications. The automatic detection program classifies ICs based on how likely they are to meet certain characteristics of a component. These columns tell you, in percent, an estimate of how likely the IC may have that component type within its signal. If the percentage is greater than a particular threshold, the IC is classified as a one-component IC. For example, look at IC 1 in the picture above. It is 93% likely to contain muscle signal and the IC was classified as 'M' with a certainty of 3.

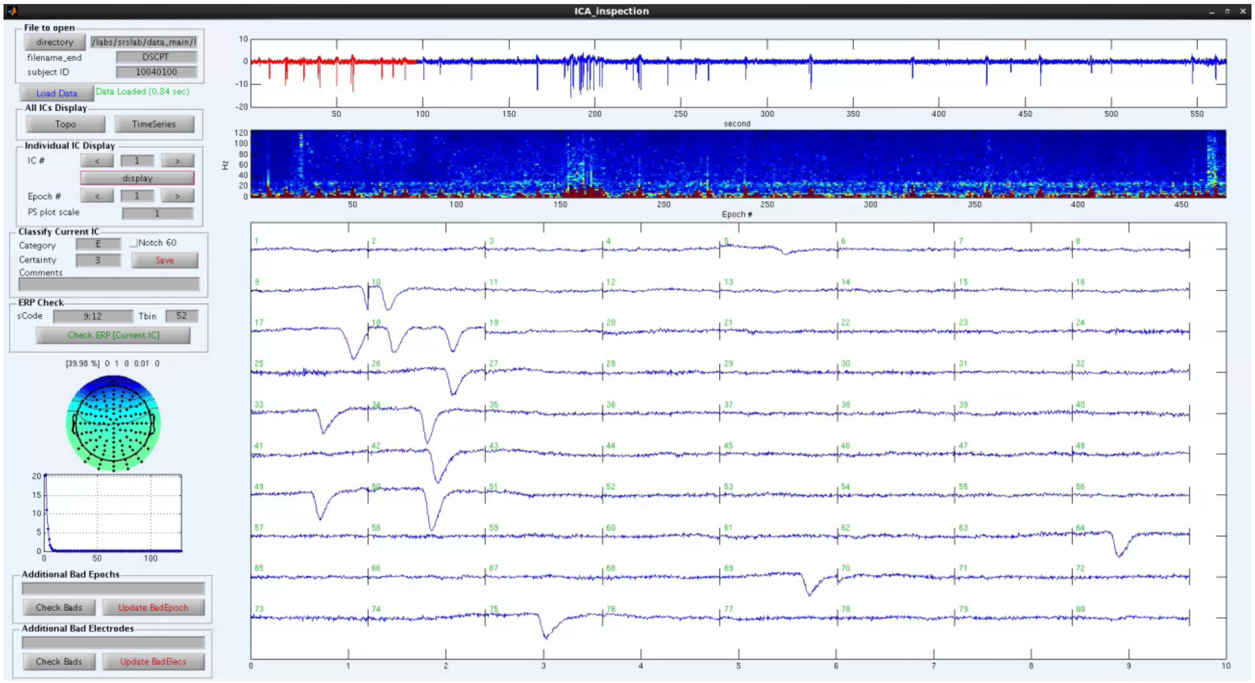
'Notch60': Sometimes 60 Hz or other forms of environmental noise are still noticeable at the time of IC classification. If you observe 60 Hz noise in the ICA inspection using power spectrum density and time series plots, it is important to notch filter out this noise using the 'Notch60' function. Only apply 'Notch60' to ICs that are either solely or primarily brain (e.g., 'B', 'BM'). We are ultimately analyzing neural signal, so it is unnecessary to correct for environmental noise in other ICs (e.g., 'M'). Noise ICs are discarded, anyway.

'Comments': The automated classification functions performed by ICAcleanEEG will always leave the comments section blank. Whenever you deviate from default results, make a note explaining why you changed the classification, certainty, or 'Notch60' filter.

Note that some of the ICs appear highlighted in red color. These ICs were determined to be ambiguous by the automatic classification algorithms. Users should review all ICs, but it is important to pay particular attention to red ICs.

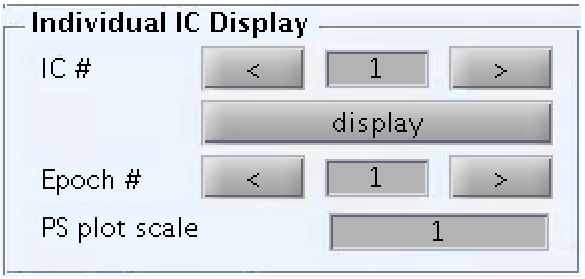
Individual IC displays

Using the ICA_inspection GUI, click **'Load Data'** and then **'Display'** to take advantage of additional descriptive plots for each IC. Below is an example eye blink IC.



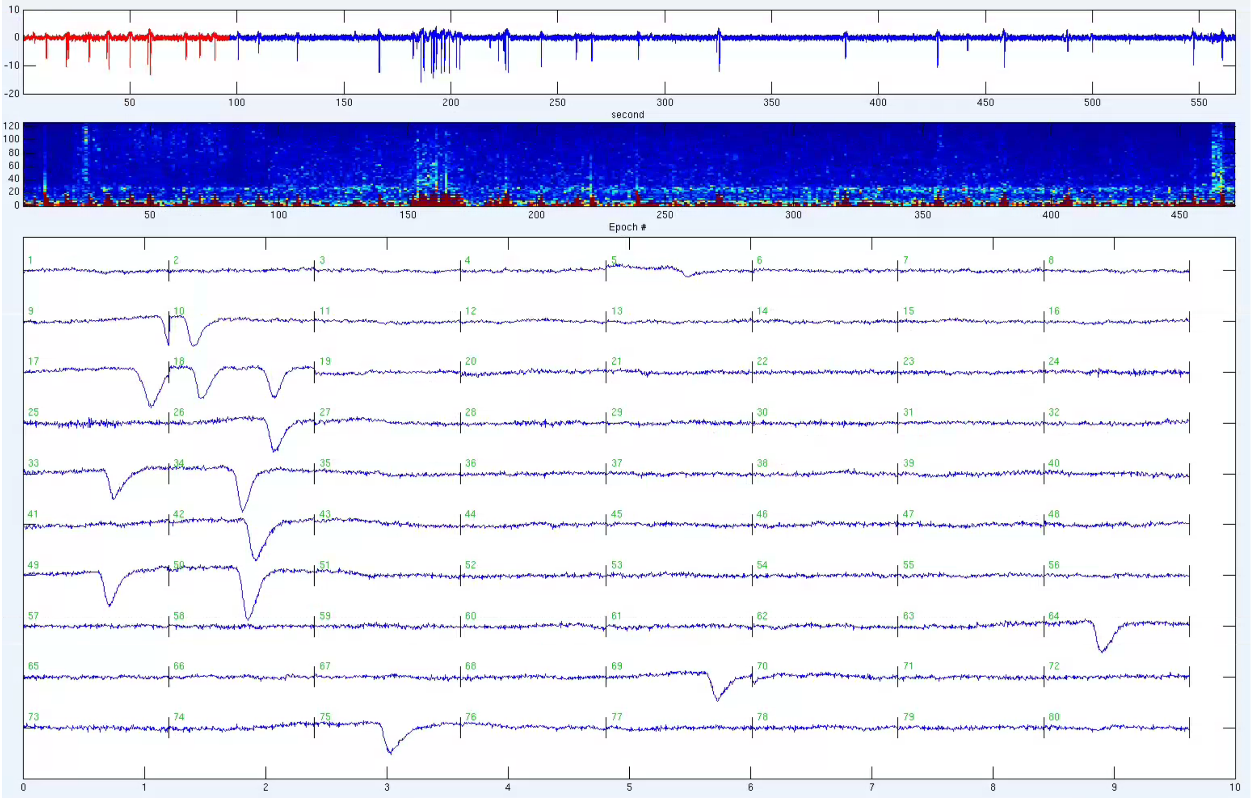
directory /labs/srslab/data_main/l
filename_end DSCPT
subject ID 10040100
Data Loaded (0.84 sec)
IC # 1
Epoch # 1
PS plot scale 1
Category E
Certainty 3
sCode 9:12
Tbin 52
[39.98 %] 0 1 0 0.01 0

The ‘Individual IC Display’ box can be used to change the IC displayed and epoch page depicted in the plots. If needed, users can also adjust the scale of the power spectrum plot middle colored plot to the right of the ‘Individual ICs Display’ box) using the ‘PS plot scale’ field. In most cases, the default value of 1 is sufficient.



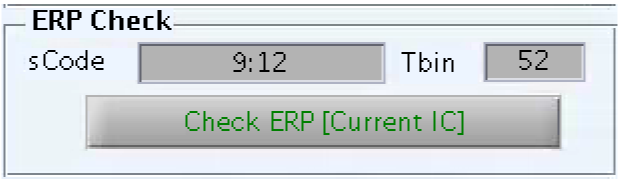
IC # 1
Epoch # 1
PS plot scale 1

The top plot displayed in the ‘ICA_inspection’ window is the time series of the IC with time on the x-axis and amplitude on the y-axis. All valid epochs across the entire recording are depicted together and displayed in sequence. Below the time series, the middle plot shows the power spectrum of this individual IC with time on the x-axis and frequency on the y-axis. The bottom plot shows the time series of the current IC with each epoch numbered successively and demarcated with vertical lines. If there are many epochs, it may be necessary to scroll with the **‘Epoch #’** function.



ERP check for individual ICs

The ERP technique is a simple, yet powerful tool for characterizing signal changes that are *time-locked* to the onset of task events. By examining many recording epochs derived from similar events, it is possible to leverage the power of averaging to "cancel out" the non-time-locked noise and leave behind responses synchronized to the onset of task events. Use of the ERP technique in tandem with the other 'ICA_inspection' plots can often be critical for identifying the subtle influence of brain signal within ambiguous ICs. "An Introduction to the Event-Related Potential Technique" (Luck, 2014; MIT Press) provides the authoritative beginner's guide to ERP methodology. ICAcleanEEG includes an 'ERP Check' function to assist with IC classification.


ERP Check

sCode 9:12 Tbin 52

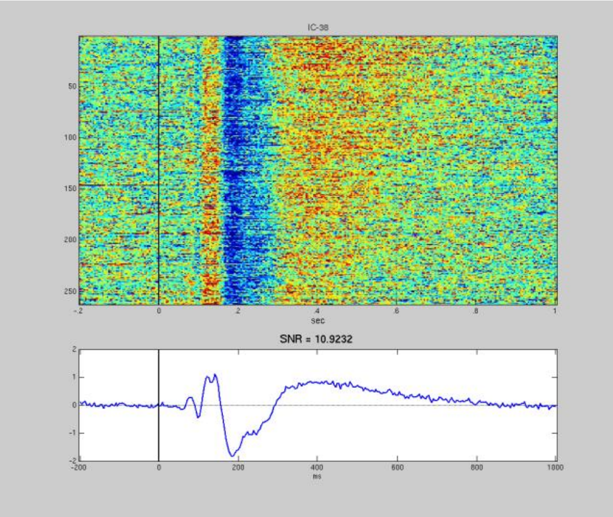
'tbin' ("time bin") designates the onset of the event(s) of interest in your recording. Users can find the appropriate tbin value within the 'sweep.tbin' variable of 'sweep.mat', or by calculating it using the following formula:

$$prestim_{ms} x \frac{sample\ rate}{1000} = tbin$$

The ‘prestim’ time duration (“prestimulus time;” in milliseconds) can be found at the beginning of the sweep file for each task. The ‘sample rate’ for epoched files is generally 256 Hz. Round the value you get from the above equation.

‘sCode’ (“stimulus code”) values are unique to each trial type and indicate which category the stimulus belongs to (e.g., target or non-target, button press, etc.). The ‘sCode’ values you select will compile these trials into an ERP waveform. Users may only want to examine particular trial types. For example, the DS-CPT task used within the CAB Lab has 12 possible ‘sCode’ values, but the only ‘sCode’ values relevant for the “degraded stimulus” condition are 9:12. Determine which ‘sCode’ values are relevant for your analyses. The sweep script is a good place to determine the relevant codes.

Clicking **‘Check ERP [Current IC]’** generates ERP waveforms for the current IC selected within the ‘ICA_inspection’ window.

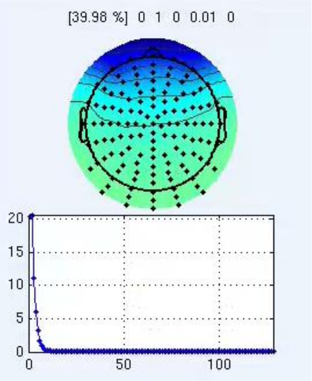


Once the ‘ERP Check’ function is initiated, two plots depicting ERP amplitude become visible. A blue line in the lower plot represents average ERP signal across epochs. A vertical black line at 0 ms represents the stimulus onset for the ‘sCode’ values selected. ‘SNR’ is the signal to noise ratio metric with a higher ratio indicating more signal (post-stimulus amplitudes) relative to noise (pre-stimulus amplitudes). This particular ERP was generated from an IC consisting of neural signal originating from the primary visual cortex to the onset of visual stimulation. The top plot also displays amplitude by way of a color scale. Red color would indicate a positive amplitude deflection and blue color would indicate a negative deflection of the signal. Rather than an average across all epoched trials, the top plot depicts each epoch individually as its own row on the y-axis. Consistency across trials (i.e., vertical colored stripes) is arguably the most important indication of time-locked brain activity.

Brain activity in task recording epochs is not always synchronized to the onset of task events (i.e., non-time-locked brain activity). Although the ERP Check is frequently very useful, ICA can also be sensitive to brain activity that is not cleanly evoked by the stimulus onset. For example, induced alpha activity (approx. 8 – 15 Hz) may be evident in the time series plots, but may not be appreciable in the ERP depictions. Research assistants in the CAB Lab have begun to position both the 'ICA_inspection' and 'ERP check' GUI windows side-by-side on their computer monitors while preprocessing so all of the available information across plots can be considered when identifying brain ICs. The ability to make these comparisons is another advantage of conducting preprocessing using larger computer monitors rather than laptop computer screens.

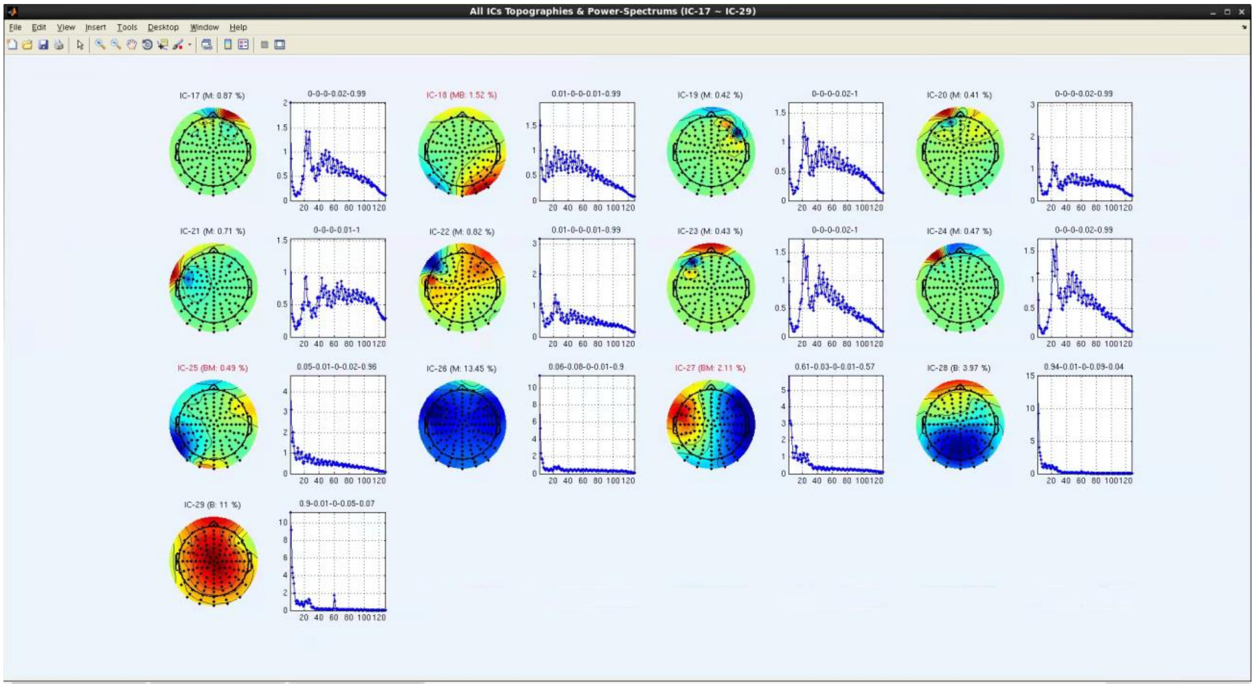
Power Density and Topography

Positioned below the 'ERP Check' box are a topographical plot and a power density spectrum for the current IC. Above these plots, percent variance and IC classification probabilities are listed. From left to right, you can see the following metrics: percent variance, probability of brain, probability of ocular, probability of ECG/EKG, probability bad electrode, and probability muscle.



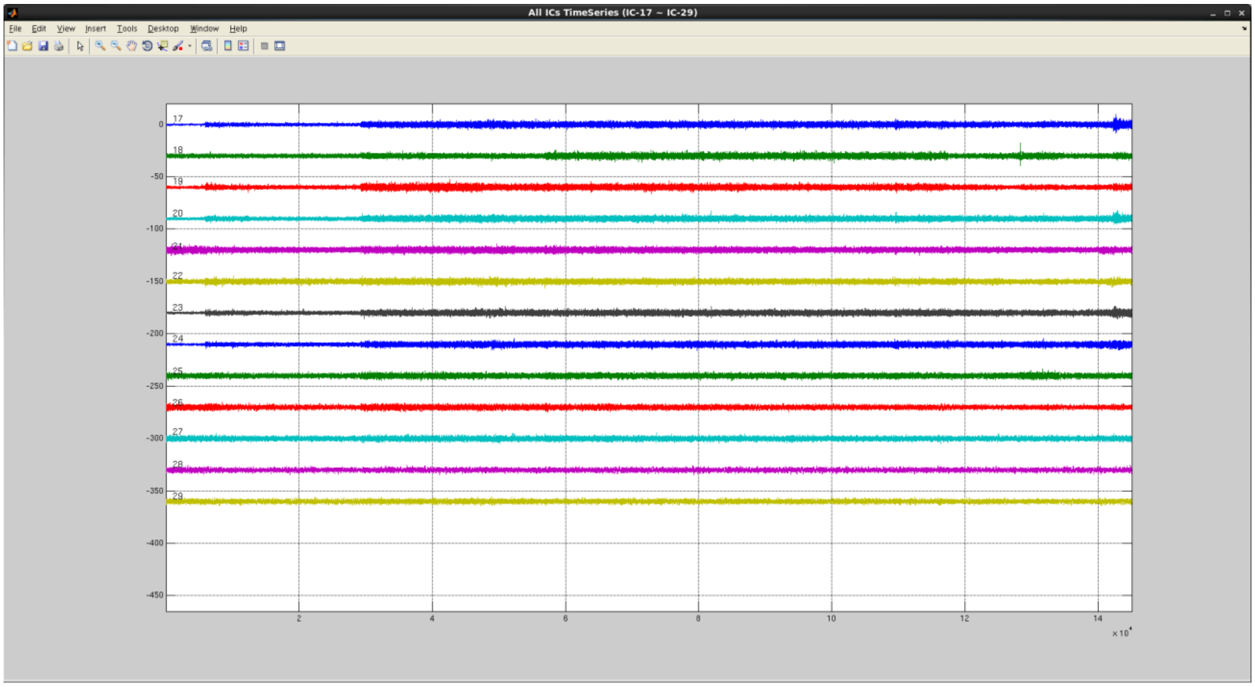
Summary displays for all ICs

By clicking **'Topo'** (i.e., "topographies") within the 'All ICs Display' box, a pop-up window appears displaying the topographies for all ICs.



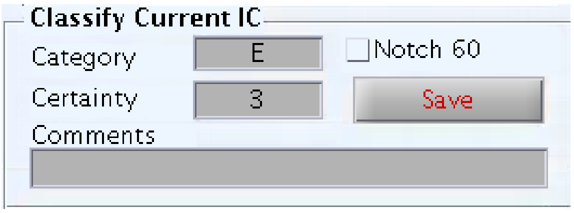
Colors within the topographical plots represent the polarity of signal (green = neutral, yellow or red = positive, cyan or blue = negative). Above each topographical plot are a series of characters. First is the IC number (e.g., “IC-3”). Next is the classification assigned to each IC (e.g., “B”). Finally, percentage of variance for each IC is listed (e.g., 23.4%). To the right of the topographies are the power density spectrums corresponding to each IC.

Furthermore, the ‘All ICs Display’ box lists a **‘TimeSeries’** button. This function generates a new pop-up window with a summary depiction of the time series across all the ICs. This can be used to assess amplitude patterns and further assist with the classification of ICs.


All ICs TimeSeries (IC-17 ~ IC-29)
File Edit View Insert Tools Desktop Window Help
17
18
19
20
21
22
23
24
25
26
27
28
29
0
-50
-100
-150
-200
-250
-300
-350
-400
-450
2
4
6
8
10
12
14
x 10^4

Classifying the currently selected IC

The ‘Classify Current IC’ box has fields to describe the IC category, the certainty of the classification, and any comments to be noted about the IC. Comments should be entered about any abnormalities in the IC. The ‘Notch 60’ filter for each IC is selected here, and should be used to filter out 60 Hz noise from brain dominant ICs.



Classify Current IC

Category E

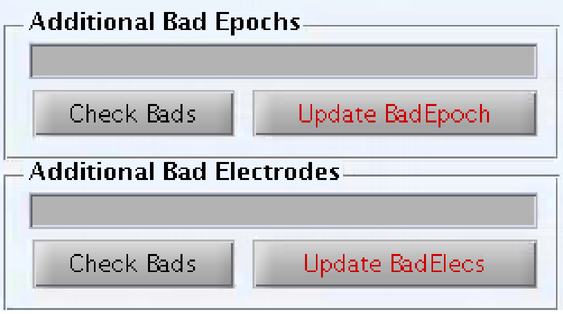
Certainty 3

□ Notch 60

Comments

Additional Bad Epochs and Electrodes

Any additional bad electrodes or bad epochs found during this step of preprocessing could be entered using the boxes displayed below.



In order for an electrode or epoch to be deleted in ICA inspection, the electrode/epoch in question needs to be bad in **more than 1 brain dominant IC.** CAB Lab members have found that deleting epochs in ICA inspection is quite rare, but deleting electrodes at this stage occurs quite frequently. Checking for bad electrodes in ICA inspection is best accomplished by viewing and inspecting the topographies of each IC by clicking the **'Topo'** button, and checking the topography of each brain-dominant IC. This is especially relevant for the brain-dominant ICs that were also labeled with an 'X' (e.g. BX, BMX, etc.).

If you identify **two or more** brain-dominant ICs with the same bad electrode(s), you can than delete that electrode using the 'Additional Bad Electrodes' field. In order to know which electrode to delete, you will need to access the 'EEG Analysis' folder within the 'Sponheim Lab Documents' folder on the Google drive. Within the EEG Analysis folder, there are image files that correspond to whichever EEG montage you are using. For BioSemi 128 elecN, open **'ElecNumbers_128.png.'** For BrainVision 125 elecN, open **'BrainVision_125elecN.png.'** These image files should look exactly the same as the montage you see in ICA Inspection, but instead of dots there are numbers that correspond to each electrode.

Applying low-pass filters to select ICs

Low-pass filtering can be an important tool for selectively reducing noise in ICs. Users can select a threshold under which lower frequencies are allowed to remain and higher frequencies are effectively removed. This can be used for removing higher frequency muscle noise contaminating ICs with prominent brain activity. 40 Hz is often an appropriate low-pass cut-off frequency for filtering. However, supplemental tools are available to assist. The **'lowpassDetermination.m*'** script can found under the following filepath:

/labs/srslab/static_files/shared_apps/matlab_toolboxes/ssk_eegtoolbox/ICAcleanEEG.v.1.3/

Enter the IC number in line 3 and run the script to generate two scalable figures. The first figure generated is a power spectrum plot and the second figure is a power spectrum by epoch plot. The 'data cursor' tool found on tool bar above the figures can be used to highlight individual points on both figures. Users can select points to generate coordinates (x,y) for the plot. Multiplying the y value by -2 and adding 128 to this will give you the value in Hz of the point you just selected. A reasonable cut-off point may be where some muscle noise is present, but no brain activity is easily observable.

Common ICs

Users should be using the topographies, power spectra, time series (both full time series and epoched views), and ERP plots simultaneously to make determinations about IC identities. The goal within 'ICA_inspection' is to label ICs with these designations:

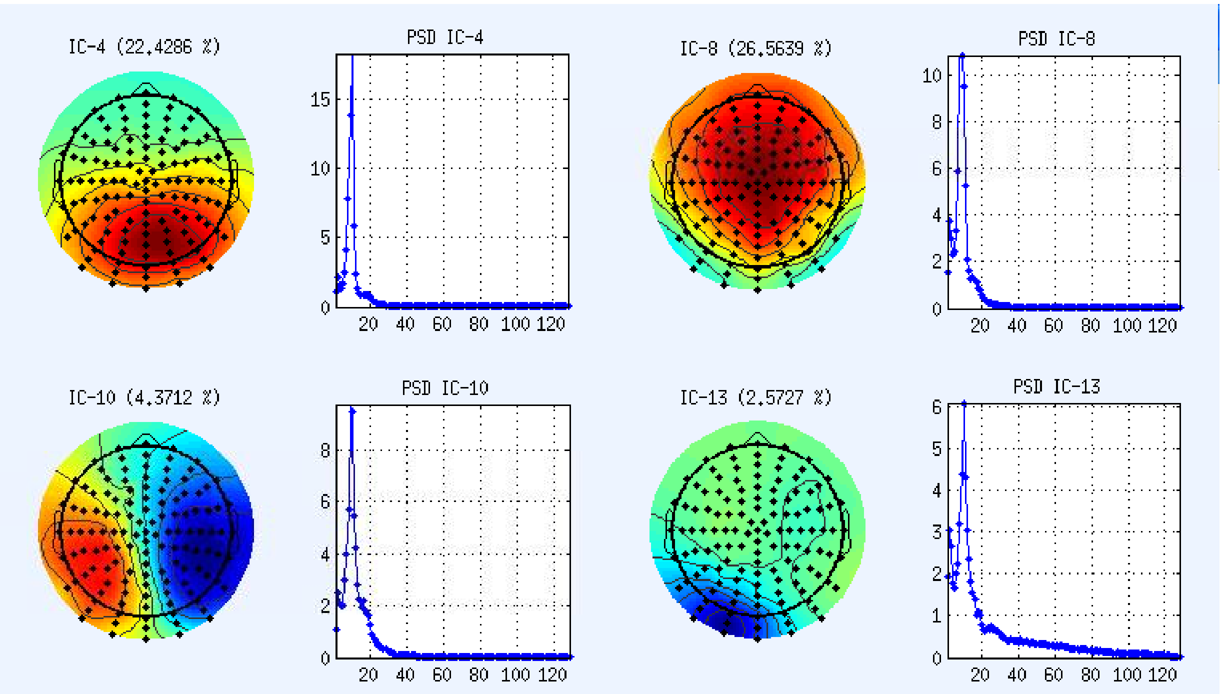
'B'	Brain
'M'	Muscle

‘E’	Eye Movement
‘K’	ECG/EKG
‘X’	Bad Electrode
‘S’	Slow Artifact
‘O’	Other Noise

The following section will go in-depth about each component type, and then elaborate upon best practices for handling mixed or ambiguous components. Readers are referred to **Appendix B** for additional examples of ICs and for CAB Lab-specific classification rules currently in use for mixed ICs.

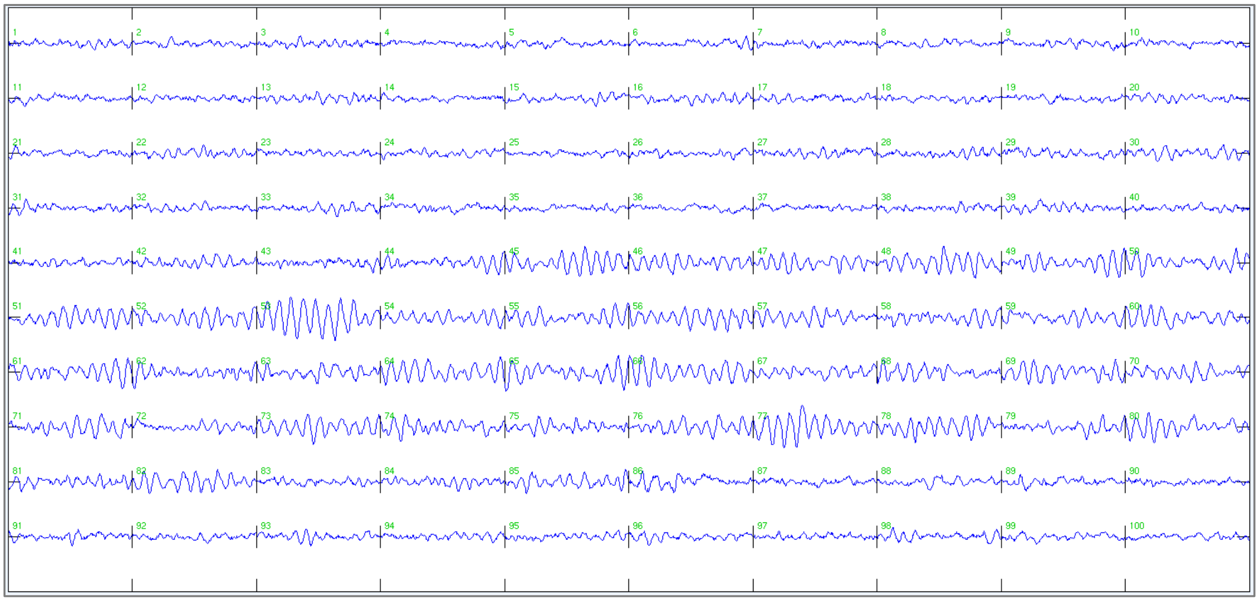
Brain (‘B’)

The topographies and power spectra from primary brain components can vary greatly in appearance, but usually possess several common characteristics for distinguishing them as neurophysiological in origin.



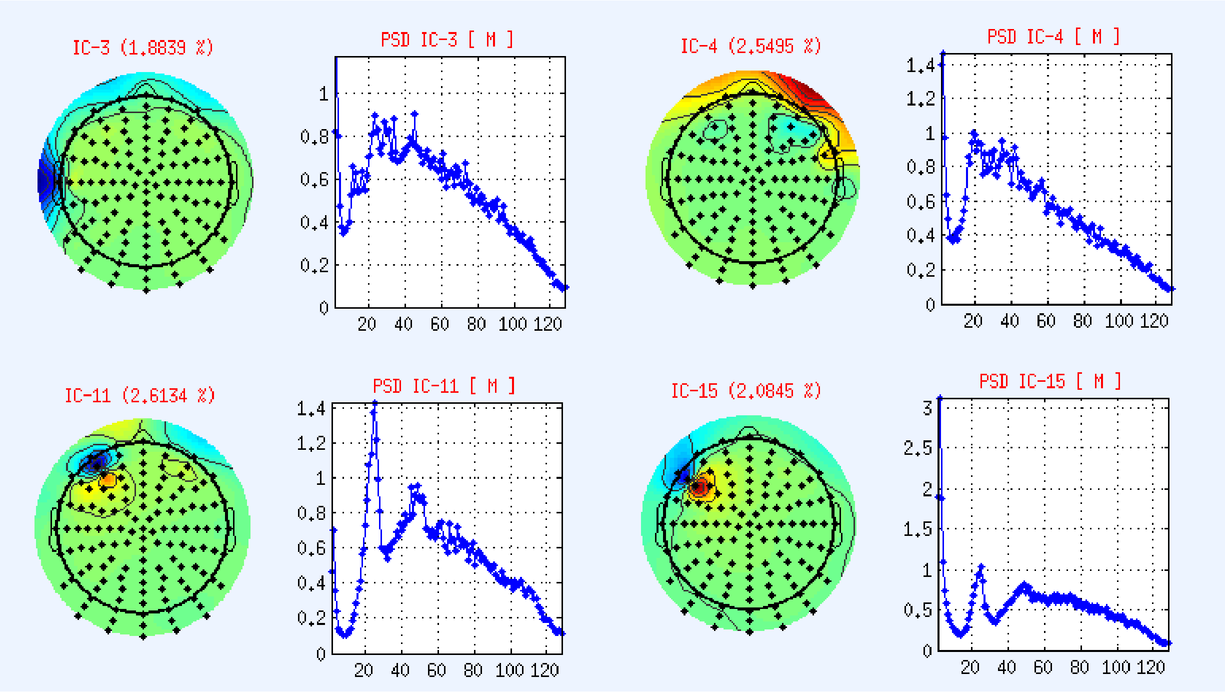
Brain topographies tend to be more widespread across the scalp, and include more of the central or “inner” electrodes. Topographies for brain components frequently display a “full rainbow” effect across the brain. Power spectra tend to show peaks between 8 Hz and 30 Hz, and few peaks elsewhere (depending on the quality of the component). 60 Hz noise is the exception to this, which is discussed below.

Time series of brain components are characterized by signal with regular oscillations in lower frequencies. Such oscillations should be readily apparent in the epoched time series display, particularly for alpha frequency band components.

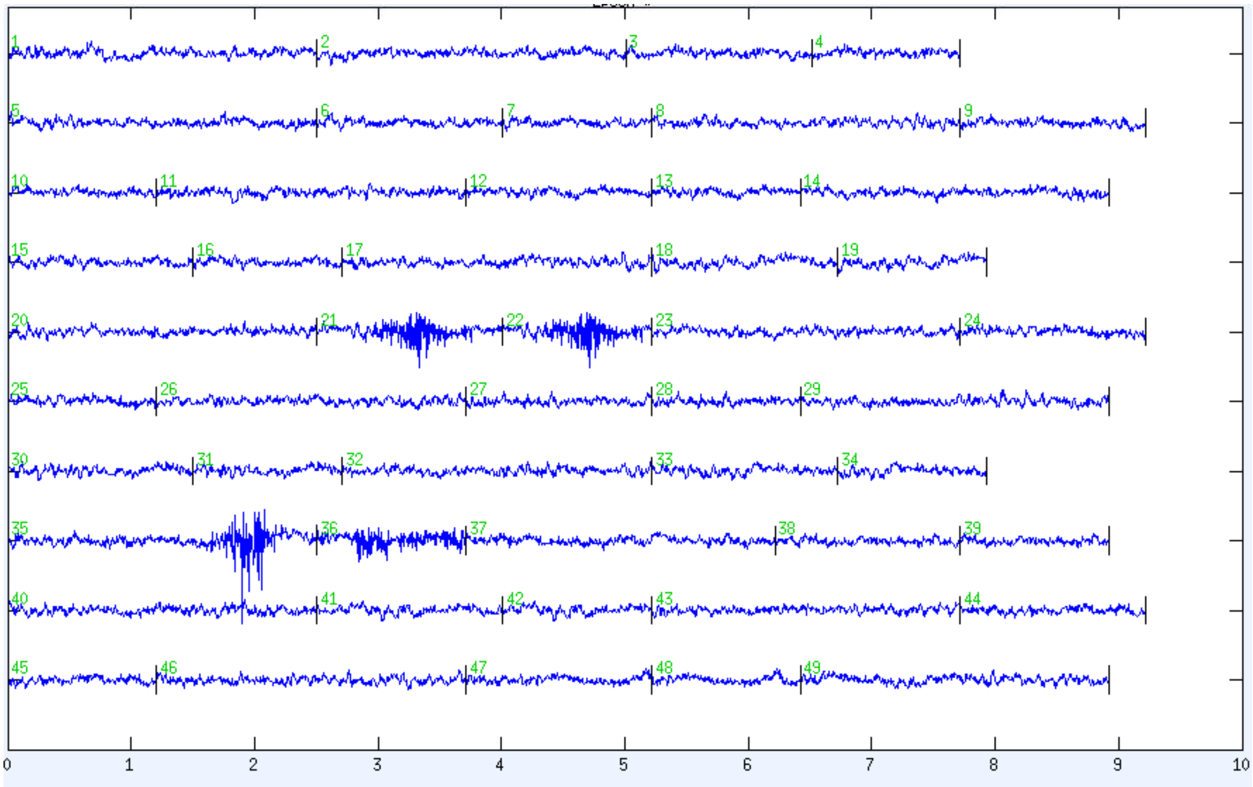


Muscle ('M')

In contrast to brain ICs, muscle ICs are more frequently maximal near marginal electrodes on the outside of the recording array. Furthermore, the activity is less widespread and often exhibits a "dipolar" configuration (i.e., focal blue and red color polarities closely positioned on the scalp). Muscle components are easiest to identify based on the presence of stronger peak activity in higher frequencies. Power spectra for muscle components are often more erratic and possess a greater integral for higher frequencies (i.e., area under the power spectrum curve).

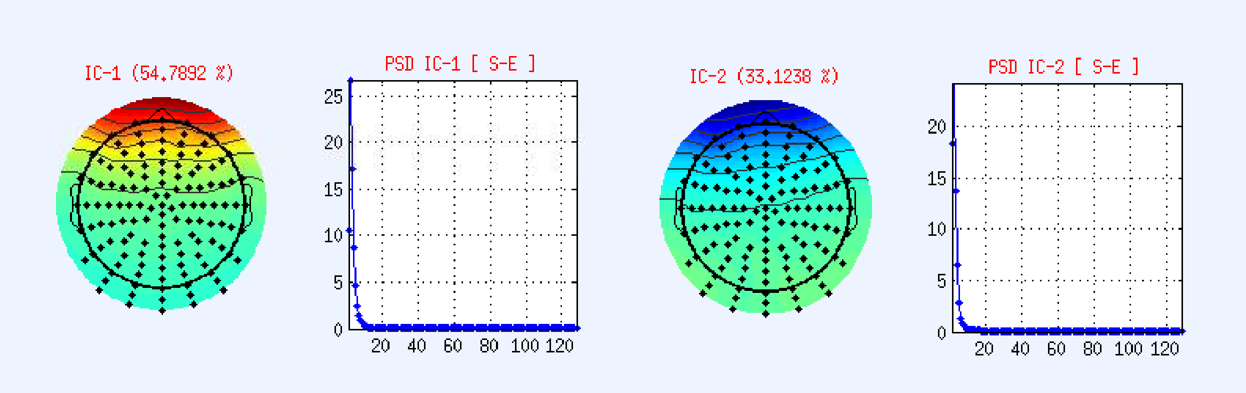


Muscle time series can be difficult to interpret conclusively. The decision to classify an IC as muscle should definitely be supported using the power spectrum and topographical information. However, also look for bursts of strong, dense signal within the time series as a sign of muscle artifact.

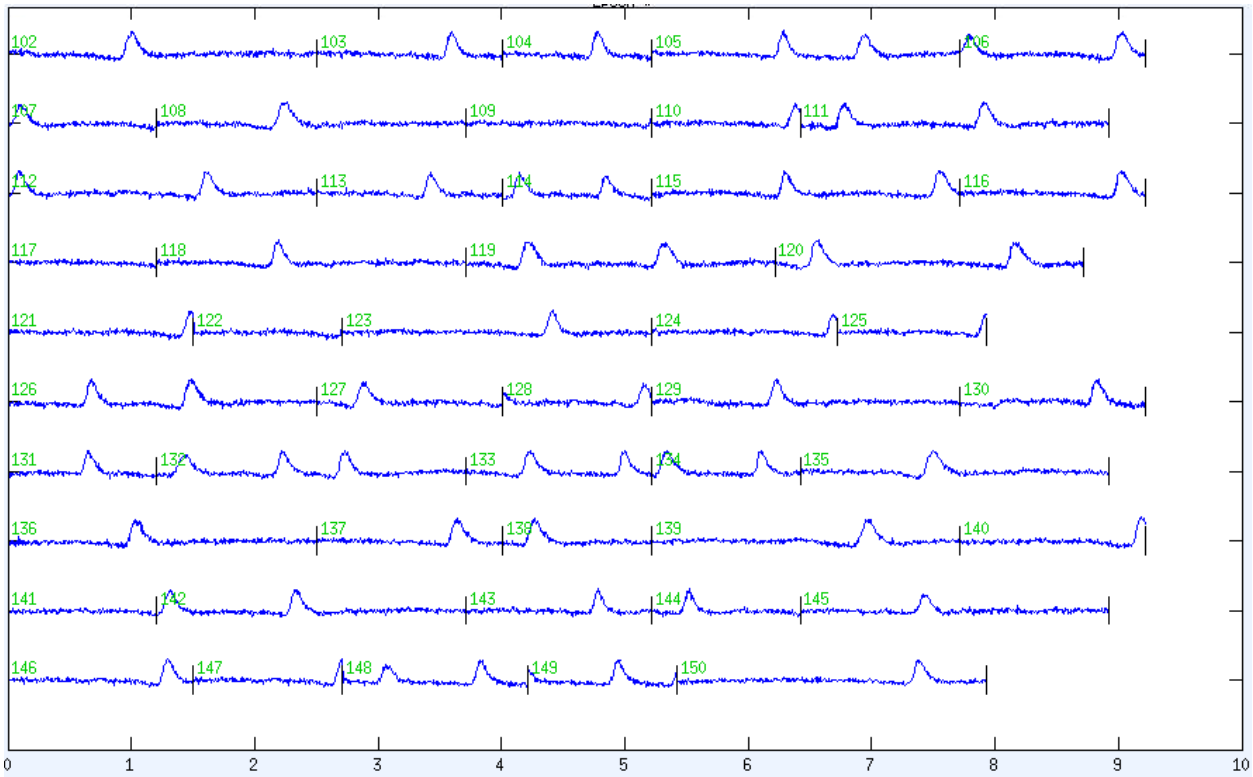


Vertical EOG ('E')

Vertical electro-oculographic ("VEOG") eye blink IC topographies are very recognizable with strong concentrations of activity around the participant's face and a stepwise decrease in amplitude to the posterior portion of the electrode array. Eye blink power spectrums are characterized by concentrations of activity within low frequencies. This visually creates a hyperbolic (i.e., L-shaped) curve across the frequency bands.

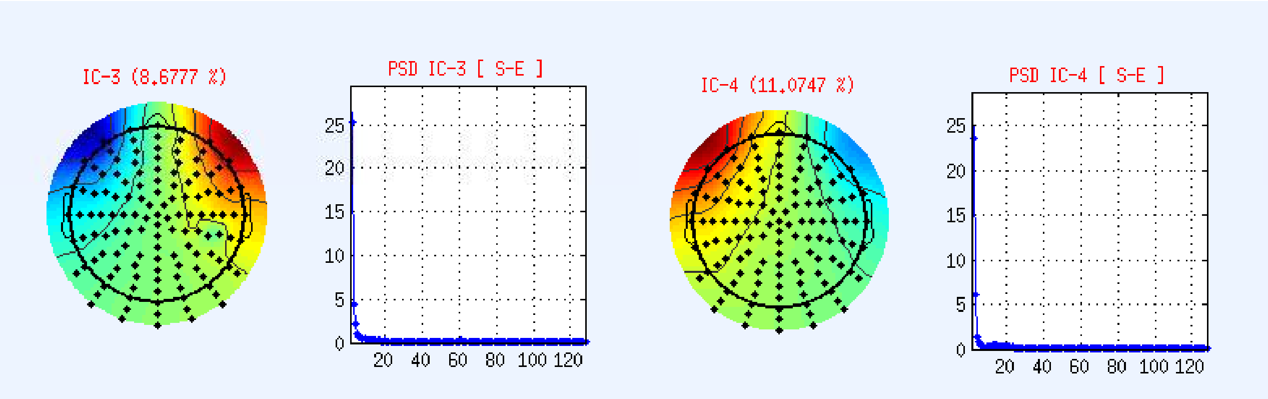


Eye blink ICs are even easier to recognize within the epoched time series plot. VEOG will appear as occasional triangular peaks surrounded by relatively flat signal. Blinks can sometimes be time-locked (e.g. eye blinks after something is flashed on the screen), but the time-locked activity will, in most cases, occur later than time-locked brain responses.

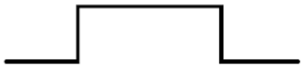


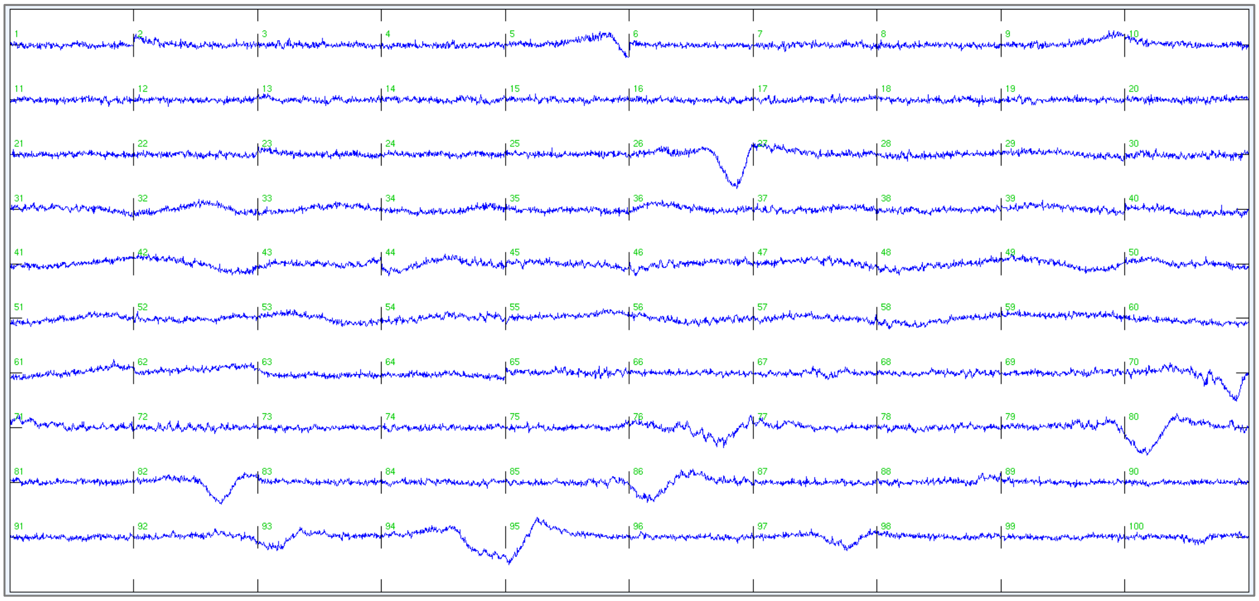
Horizontal EOG ('E')

Horizontal electro-oculographic ("HEOG") eye movement IC topographies are less commonly observed. HEOG IC topographies can present as two distinct, semi-circular signals that are roughly centered between the nose and each ear. Power spectra for HEOG ICs are similar to that of VEOG—exclusively low frequency activity appearing as a hyperbolic curve.



Time series signal for HEOG and VEOG are similar. However, HEOG artifacts often appear wider and take on a rectangular shape more akin to a “boxcar” than a triangle (e.g., time segments 94 and 95 below).

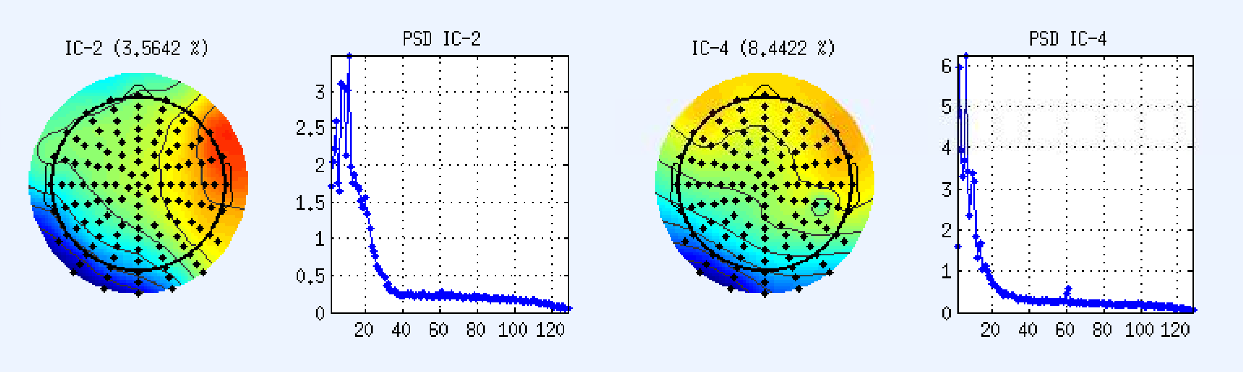




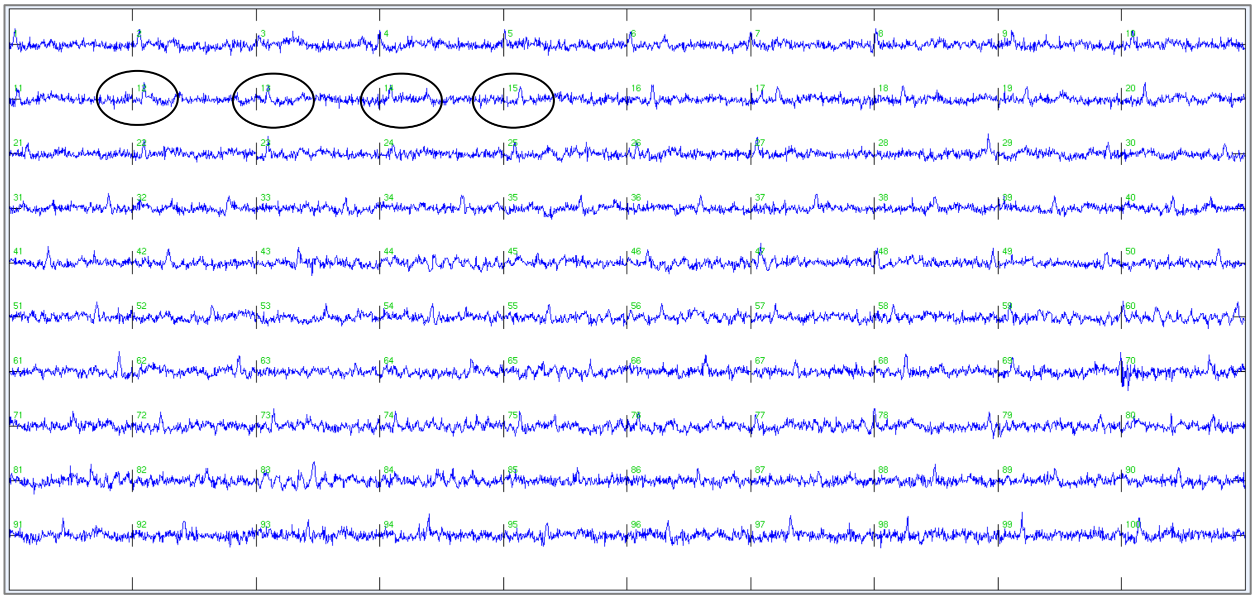
Many time series of ICs dominated by VEOG and HEOG can appear very similar. In practice, there can be some HEOG in VEOG ICs (and vice versa). Topographies will help to clarify which type of artifact is primarily present in each IC, but this is unnecessary effort in the grand scheme of preprocessing with ICAcleanEEG. Ultimately, both receive an ‘E’ label and are excluded.

ECG (‘K’)

Another less frequent IC, the electrocardiographic (“ECG” or “EKG”) artifact, has a topography that often appears as a diagonally tilted gradient between the front right and back left quadrants of the head. Some individuals will display a topography gradient from back right to front left. The ECG power spectrum is comprised of primarily low frequency activities. If other noise is absent, ‘K’ components should generally appear as hyperbolic curves with slow frequency activity.

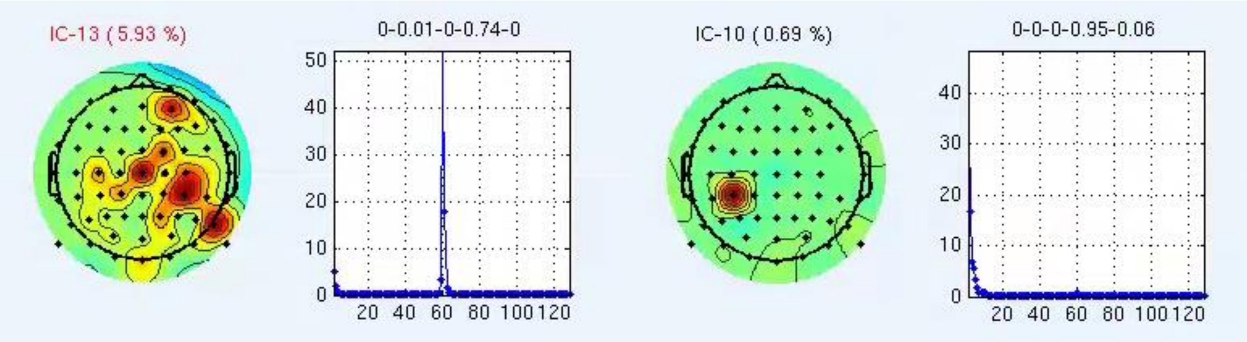


ECG signal may appear in subtle ways at the scalp. It can be difficult to distinguish between brain and heart rate signal because both have similar topographies and power spectra. Furthermore, ICA can decompose both types of signal together into the same mixed IC (see 'BBK' IC section below). Time series information can be extremely helpful for distinguishing between 'B' and 'K'. ECG signal will have regularly appearing, medium-sized peaks with "sharp" appearances. These may appear as either positive or negative deflections occurring roughly once every epoch.

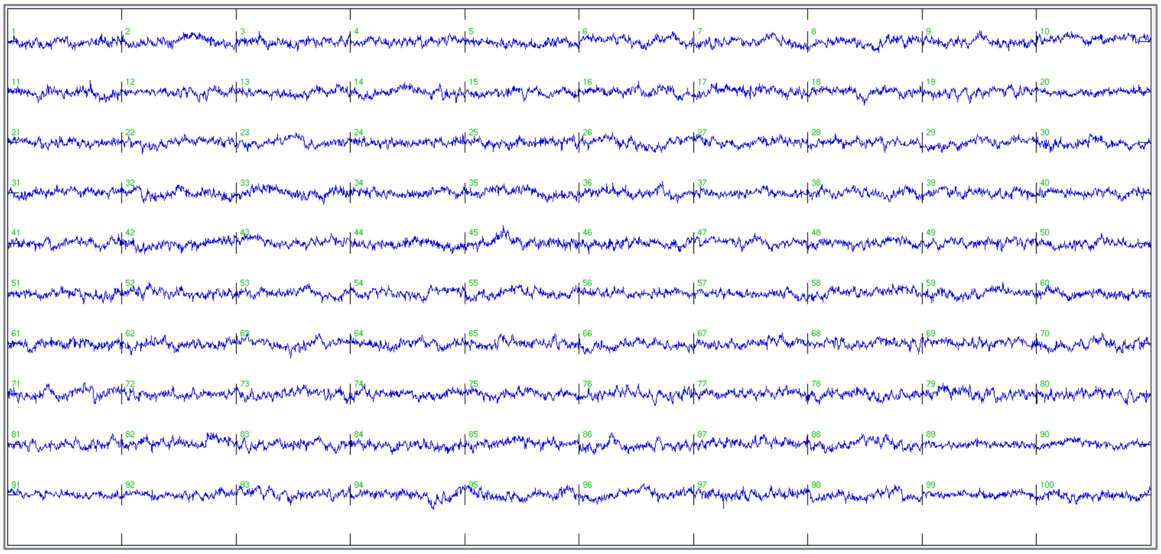


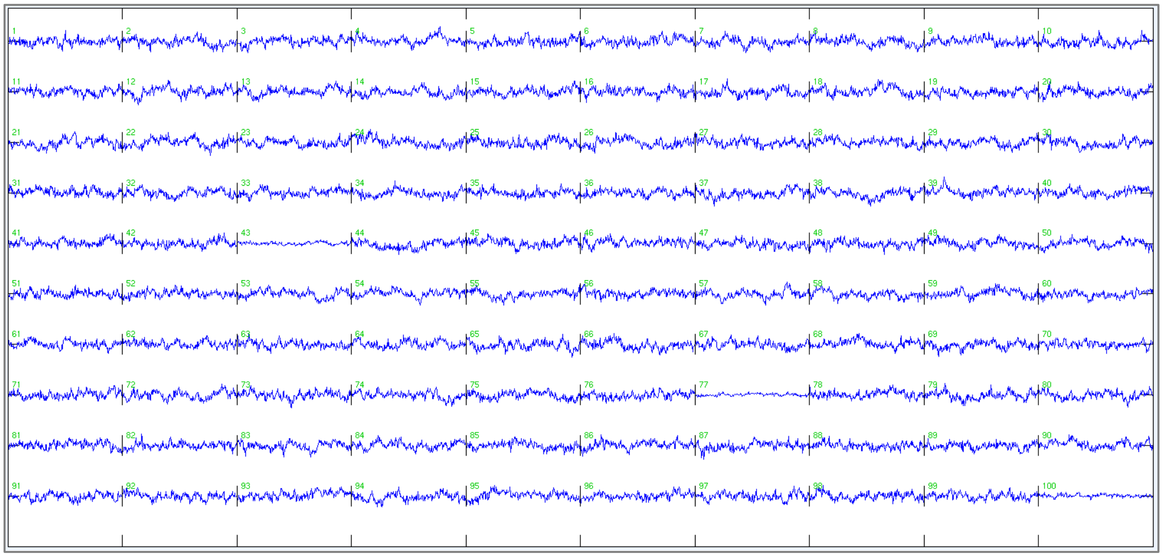
Bad Electrode ('X')

Bad electrode topographies exhibit "islands" of strong activity centered on specific electrode(s). There is generally little other electrical activity besides the signal generated from these isolated electrode(s). Isolated electrode activity is unlikely due to brain-based sources because actual neurophysiological signal is expected to propagate across the scalp. The power spectra will generally show little activity, but may show strong 60 Hz noise in some cases.



Time series for bad electrodes are widely variable. Users may mistake bad electrodes ICs for other types of components if the time series is considered in isolation of other information. Below are two epoched time series of bad electrode ICs:

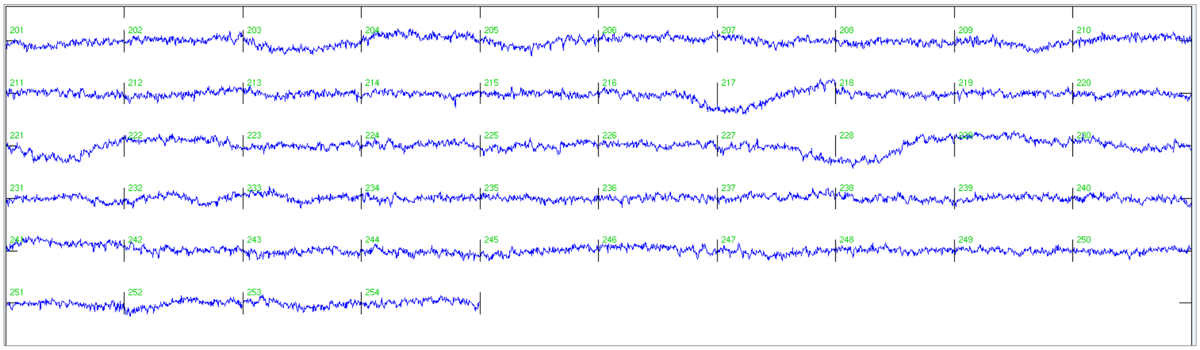




Perhaps the most informative feature of these two displays is the *lack* of clear structure. For bad electrodes, time series displays should be used primarily to support classifications derived from topographies or spectral displays. Time series plots may or may not be ultimately informative when examining ‘X’ ICs.

Slow Artifact (S)

Prominent slow waves ICs are commonly produced by sweat and should be marked as ‘S’.



Other Noise (‘O’)

The label ‘O’ is reserved for ICs that do not fit well into the other categories (but can be confidently determined to be noise). This category should be used sparingly.

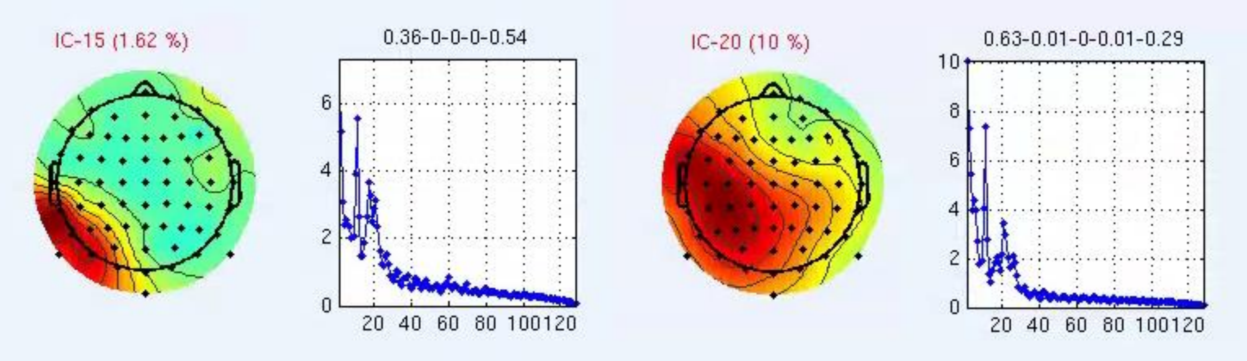
Mixed ICs

ICA is meant to decompose signal into meaningful and clear ICs, but this does not always happen as desired. It is common for two or more components to be lumped together as “mixed” ICs. However, all is not lost. Several principles can assist users with navigating these situations to ensure quality post-processed data. Although the number of possible mixed components is large, a review of several example ICs is sufficient to demonstrate the core guidelines. Readers should also refer to **Appendix B** for more detailed information about these judgments and CAB Lab-specific classification rules.

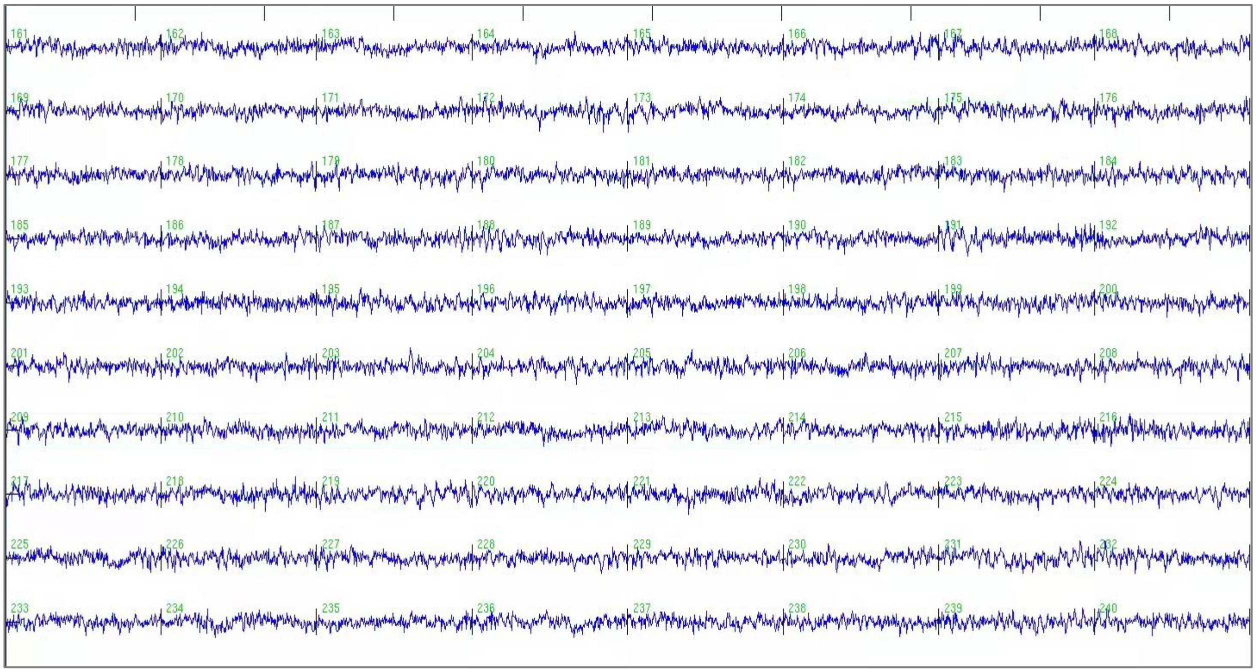
Brain and Muscle (‘BM’)

Mixed brain and muscle ICs contain neurophysiological activity with various amounts of muscle-related noise. The variability across these types of ICs can be significant. Users should carefully inspect each IC and evaluate the quality of the converging evidence for brain activity.

‘BM’ topographies may show characteristic brain IC distributions with peripheral muscle dipoles distorting the “rainbow”-like gradient. However, it is possible for noisier brain and muscle ICs to not resemble this topography. Often, the power spectrum shows a peak around 8-15Hz, which is a hallmark of neurophysiological signal. When the component is mixed, the lower frequency peak may be followed by a series of higher frequency muscle noise peaks and/or a greater “integral” over the higher frequency bands.

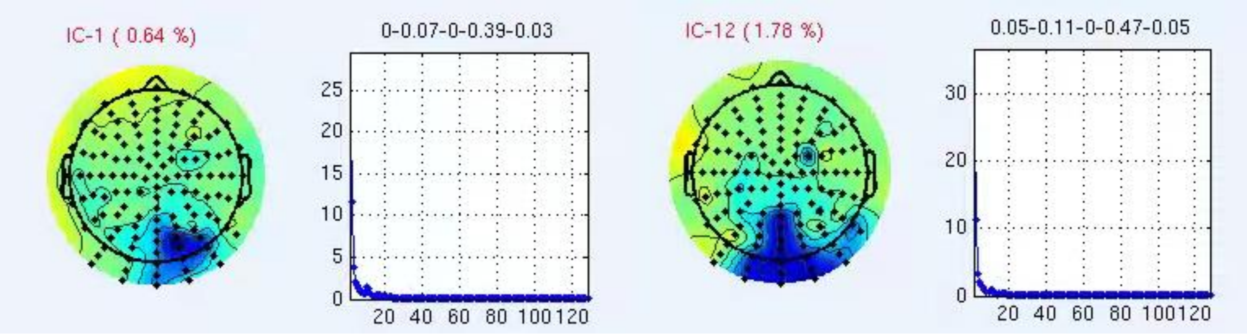


The time series of such ICs may display visible neural oscillations, but the signal may appear “thicker” because of high frequency muscle bursts. In some cases, the signal may be too noisy to make a confident identification of brain oscillations. When this occurs, ERP visualization can be informative.

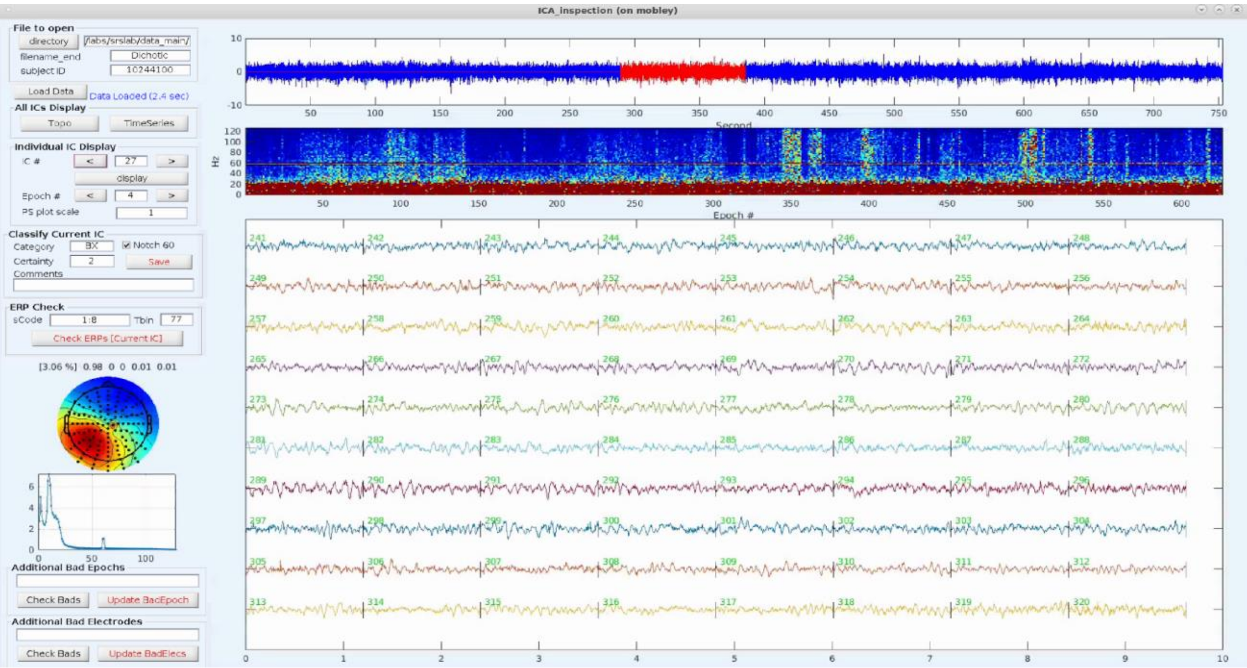


Brain and Bad Electrode (‘BX’)

Mixed brain and bad electrode ICs (i.e., ‘BX’) are similar to ‘BM’ ICs in their variability and complexity. ‘BX’ topographies often appear rather strange with bad electrode “islands” and more background activity than expected. Two challenging examples of this phenomenon are below.



Time series for these ICs are highly variable, which makes them difficult to characterize based on visual inspection alone. In the example below, note the electrode "island" near the center of the recording array and the presence of 60 Hz noise. Again, the 'ERP Check' function is frequently informative in these cases.


ICA_inspection (on mobley)

File to open
directory: /labs/srslab/data_main/
filename_end: Dichotic
subject ID: 10244100
Data Loaded (2.4 sec)

Individual IC Display
IC #: 27
Epoch #: 4
PS plot scale: 1

Classify Current IC
Category: BX ☑ Notch 60
Certainty: 2
Comments:

ERP Check
sCode: 1:8 Tbin: 77

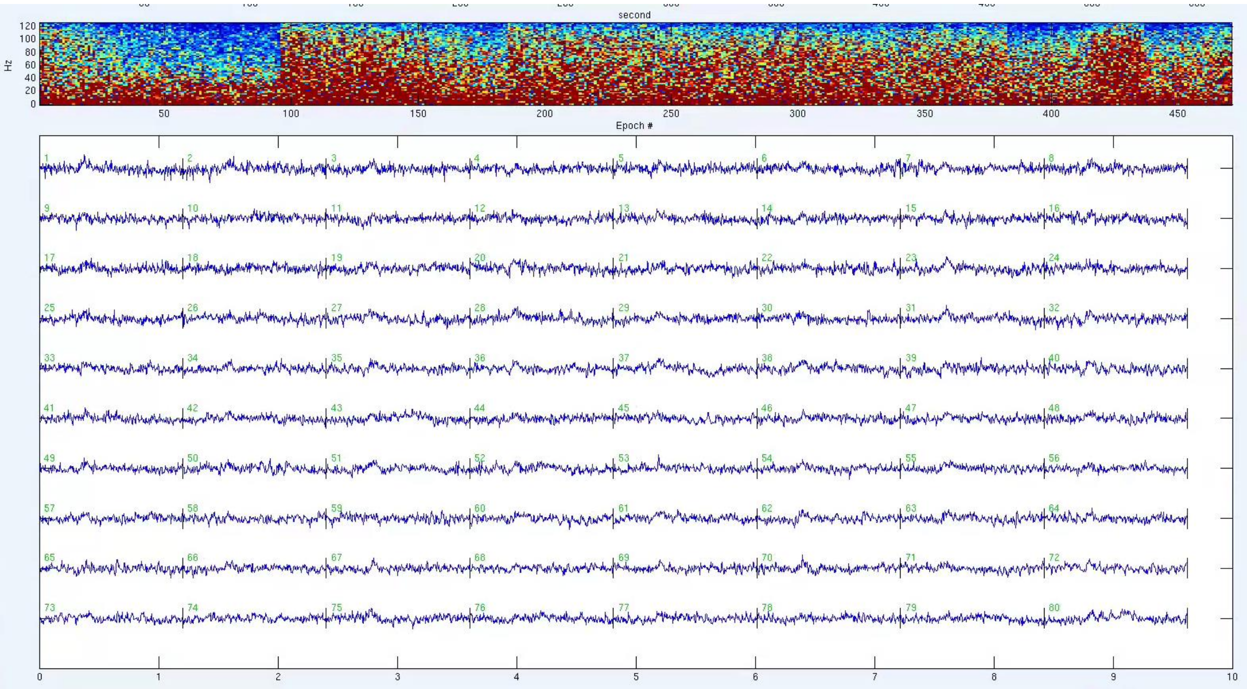
[3.06 %] 0.98 0 0 0.01 0.01

Additional Bad Epochs

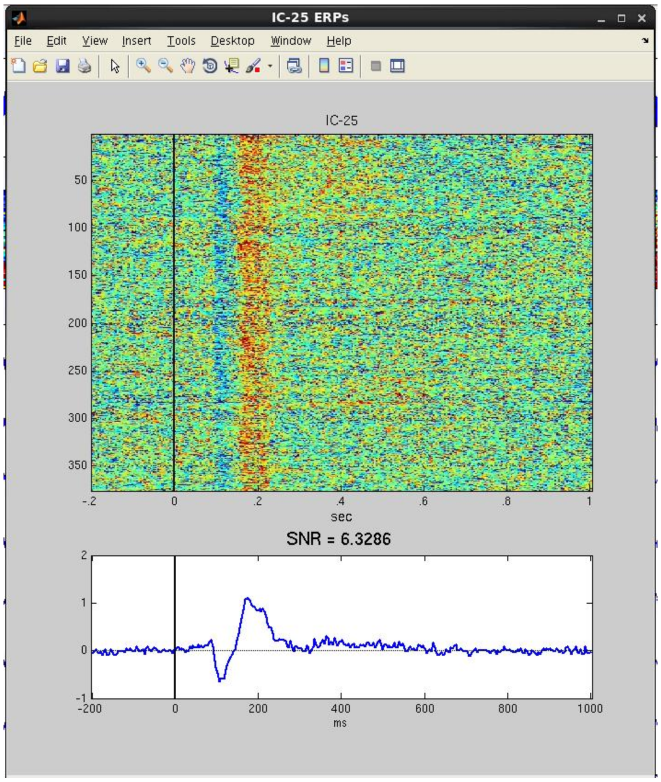
Additional Bad Electrodes

Mixed brain ICs for later ERP analysis only 'BB_'

The 'BB_' is a classification used to characterize ICs where brain activity is evident in the ERP function, but the power spectra is too contaminated by muscle/bad electrode/ECG/EOG/slow-wave activity to be useful for time-frequency analysis. These ICs are visually noisy and complex, and are oftentimes points of discussion at weekly CAB Lab ambiguous IC consensus meetings. Although these components are acceptable for post-processed ERP analyses, the additional noise makes 'BB_' ICs unusable for time-frequency analysis. The IC below was automatically classified as 'BM'. Noise above 40 Hz is readily observable across the recording.



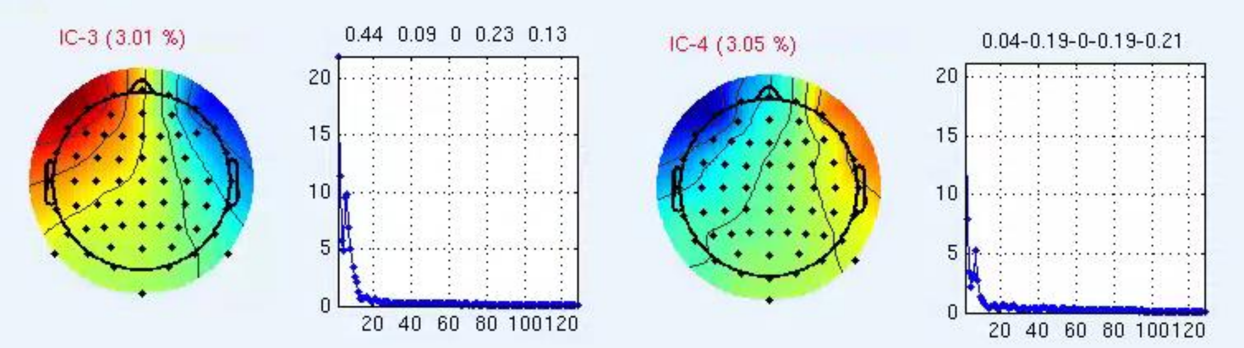
By only examining the time series and power spectrum, it would be possible to misjudge this IC as being primarily noise. For this example, the underlying neural signal can still be used for ERP analysis and should be retained. The ‘ERP Check’ function illustrates how strong and consistent the time-locked neural activity is across the epochs.



It may be difficult to decide whether an IC should be classified as ‘BM’ or ‘BBM’. It is not well established how much high frequency noise is acceptable for time-frequency analysis. A conservative approach would be to employ ‘BBM’ if there is ambiguity about the amount of high frequency noise present. Investigators in the CAB Lab review all ‘BBM’ ICs before preparing their denoised datasets for time-frequency analysis.

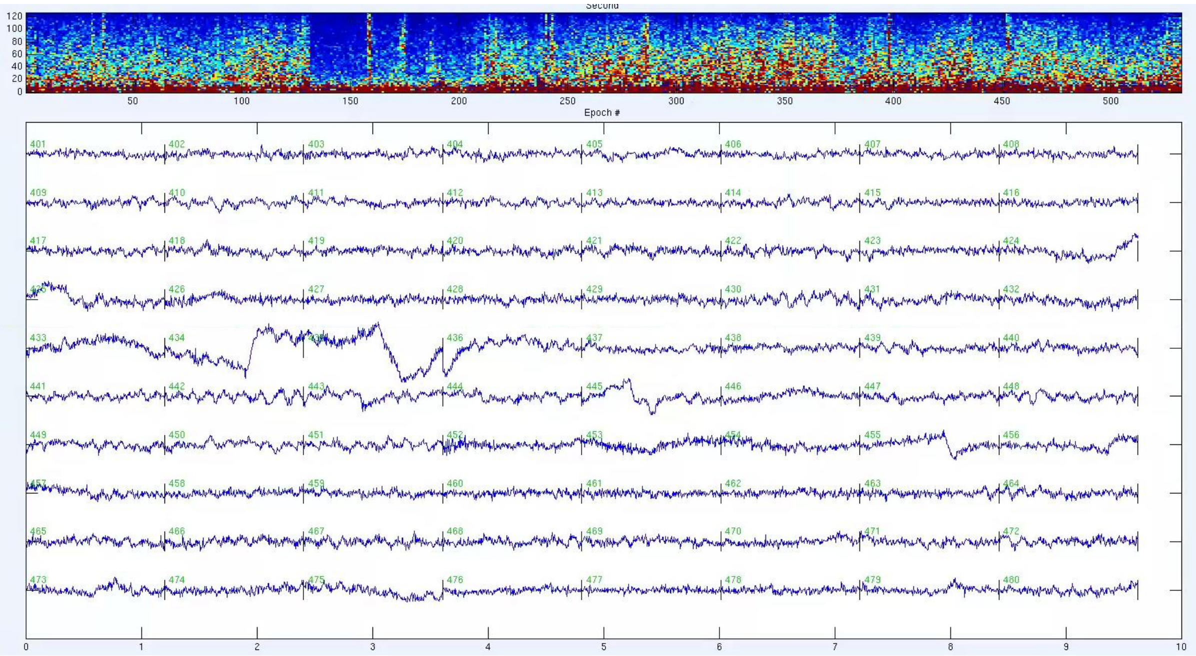
Brain and Eye (‘BBE’)

Mixed ICs with brain and eye signal have topographies that appear very similar to those of regular eye movement ICs. However, frontal artifacts may be less prominent and there may be overlap with legitimate brain signal.



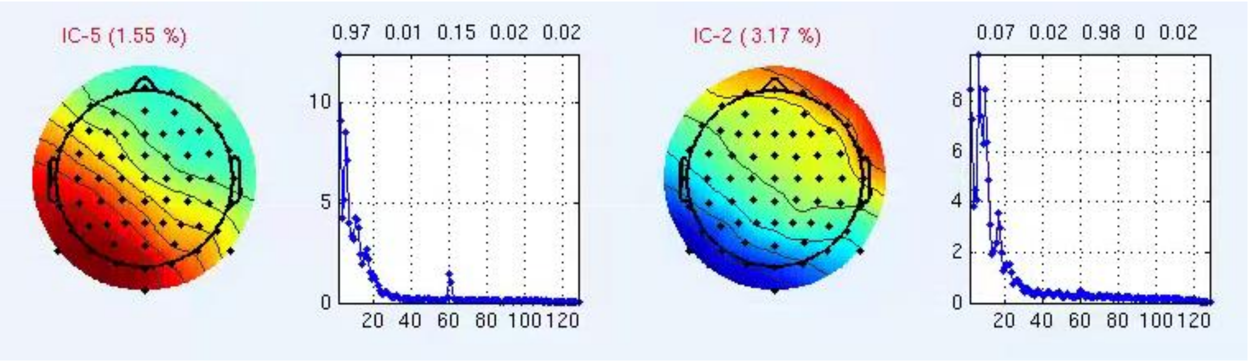
Time series of ‘BBE’ ICs often contain evidence of brain signal with eye deflections appearing during isolated sections of the recording. Task characteristics may contribute to the complexities involved with evaluating these ICs during preprocessing. Some visually stimulating tasks can evoke time-locked eye blinks resemble some aspects of primary visual cortex activations. The ERP functionality should be used in these cases to make a classification.

Below is a mixed brain-HEOG IC. The power spectrum shows strong, but short duration power bursts when blinks or eye movements occur. This is most clearly corroborated in the time series between segments 434 and 435— a large “boxcar”-like HEOG deflection can be observed contemporaneously with ongoing neural oscillations.

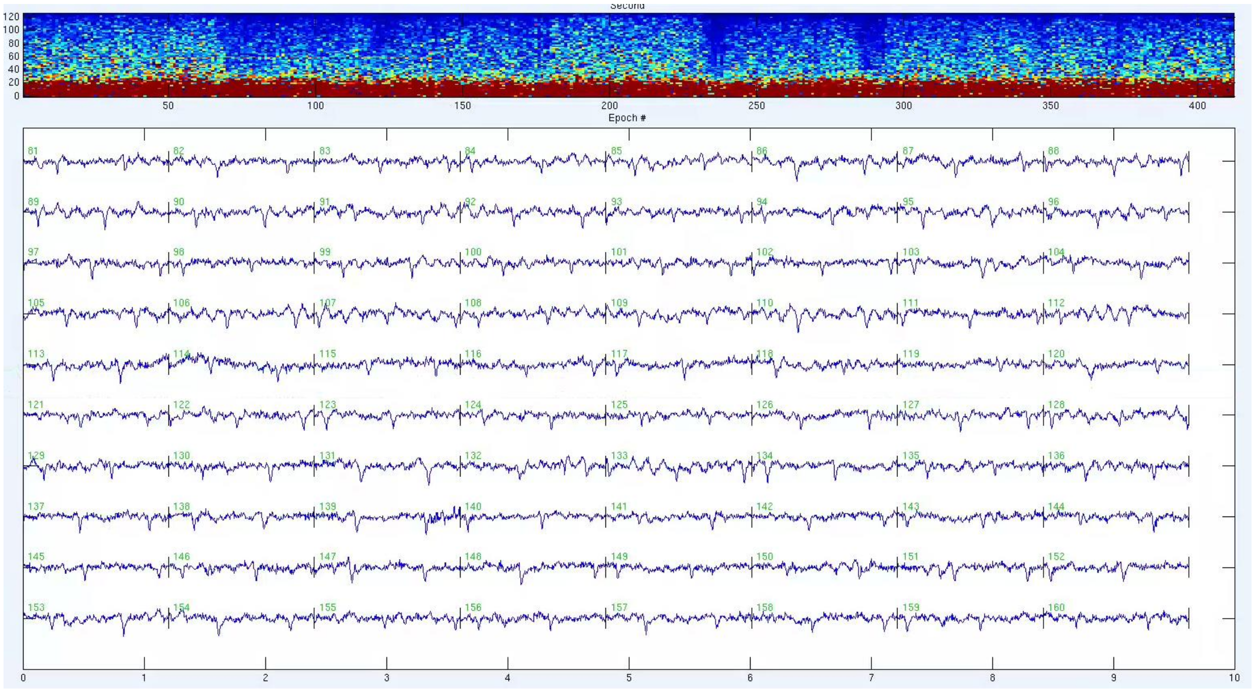


Brain and EKG ('BBK')

Mixed brain and ECG ICs are commonly combined during IC decomposition due to their similar topographies and power spectra. Users should closely examine potential ECG ICs to determine if a meaningful amount of brain is also included. Again, the ERP function is valuable for classification in these cases because ECG will not be time-locked enough to appear in the ERP waveform. For both example ICs below, the topographies appear like typical 'K' gradients.



However, the time series for these ICs show both brain oscillations and ECG-related deflections. Also, the power spectrums show most of the activation below 20 Hz. Examination of ERP waveforms with the 'ERP Check' function would likely confirm the presence of evoked brain signal.



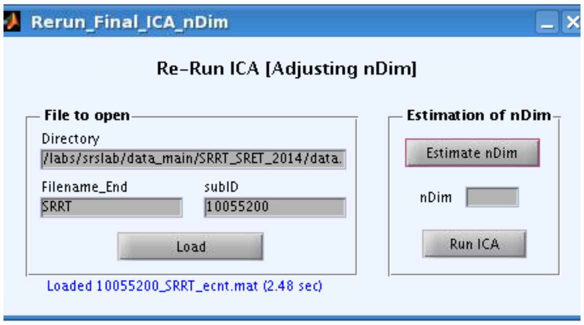
Unimportant mixed ICs

There are many combinations of ICs that do not require substantial precision during classification (e.g., ‘MX’, ‘EX’, ‘KM’, ‘MKX’, etc.). These ICs *do not* include brain and will not be incorporated into post-processed data for any type of analysis. If IC labels like these are automatically generated, it is still important to check the classification and search for usable brain signal.

Conversely, there are many mixed ICs that *do* contain some trivial amount of brain (e.g., ‘MB’, ‘EB’, ‘XB’, etc.). Users must make decisions after considering the quality of the neural signal represented across the plots. If the amount of brain is minimal, the components should be left as is. Distinguishing when a mixed brain IC is worthy of further attention is will be discussed in-depth later in the manual.

Rerunning ICA using ICA2 (optional)

On occasion, one or more bad electrodes or epochs may be having an undue influence on multiple mixed brain ICs. It may be necessary to remove these electrodes and rerun the ICA using the ‘ICA2’ tools. First, double-check the ‘Additional Bad Electrodes’ and ‘Additional Bad Epochs’ box fields for accuracy within the ‘ICA_inspection’ window. Click **‘Re-Run ICA (nDim)’** to open the ‘Rerun_Final_ICA_nDim’ window.



Directory: /labs/srslab/data_main/SRRT_SRET_2014/data.

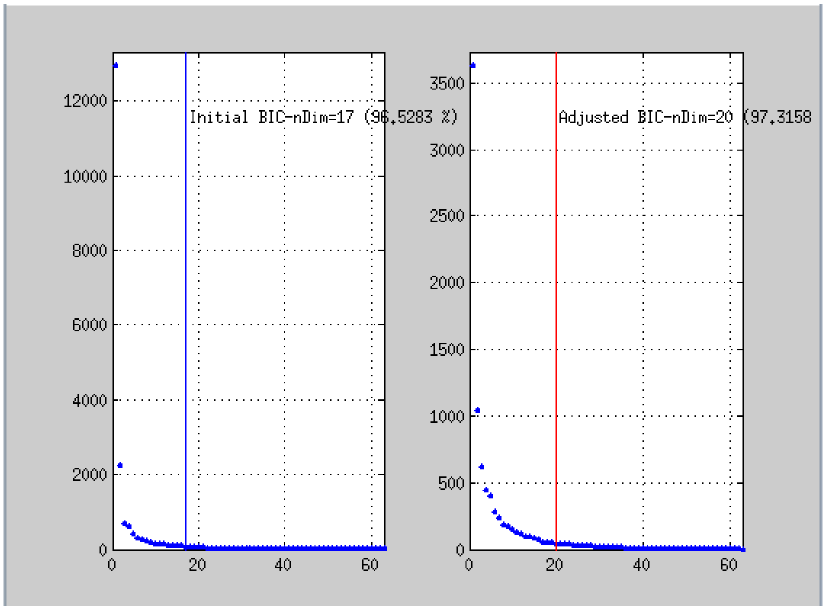
Filename_End: SRRT

subID: 10055200

nDim:

Loaded 10055200_SRRT_ecnt.mat (2.48 sec)

Click **'Load'** and then **'Estimate nDim'**. This generates a new fit estimate of the number of dimensions to be decomposed for the new ICA. A brief automated process will run followed by the appearance of the 'Eigenvalue Scree Plot & Estimated nDim' window.



Within the left plot, a blue vertical line indicates the original IC dimensionality estimation. For this example file, 17 ICs were originally produced. On the right, there is a second plot showing an alternative dimensionality estimation of 20. While the initial estimation accounted for 96% of the variance in the data, the adjusted estimation will account for 97%. Increasing dimensionality with the adjusted value will produce ICs that account for a slightly higher percentage of the data while also giving the users a greater total number of ICs to evaluate.

ICA2 inspection

Once 'Run ICA' is complete, click **'ICA2 inspection'**. This GUI operates in exactly the same way as the first 'ICA inspection' screen, except that it will display the ICs using the updated dimensionality estimation. This step is useful for visualizing the impact that deleting electrodes and epochs. If you do not see significant changes relative to the first ICA decomposition, you can simply go back to the first ICA classification step and discard the results from ICA2. If you are not satisfied with either the first or second ICA decompositions, it is possible run ICA2 again with a different 'ndim' value.

ICA2 caveat

Anecdotally, current CAB Lab ICAcleanEEG users do not have strong evidence that increasing dimensionality (i.e., changing the 'ndim' value) without performing any other changes to a recording (e.g., deleting epochs or electrodes) can more reliably separate various sources of noise from brain activity. Determining the conditions in which dimensionality changes can improve ICA performance may warrant further exploration. However, expanding dimensionality for the sole sake of improving decomposition quality is not advised at this time.

Special Considerations for Resting (i.e., non-event-related) Recordings

In the CAB Lab, every participant completes a "resting" EEG recording without any cognitively or perceptually demanding tasks. Participants are asked to sit quietly and alternate every 45 seconds between keeping their eyes open or closed. This is done to assess activity across brain networks in a way that is independent of any task requirements. Kang et al., 2018 is an example of the frequency-based analyses that can be performed using resting EEG recordings. In general, resting recordings can be preprocessed in a similar way as task-based recordings. Nevertheless, there are several special considerations to be aware of. The threshold for including ICs will be *more conservative* as a consequence of having less information about IC quality.

Typically, the 'ERP Check' visualization tool is a great additional source of information about event-related brain signal within an EEG recording. However, resting recordings do not contain time-locked events suitable for identifying signal and noise ICs. Users must rely solely on the other available ICA visualization tools to make determinations about IC quality. In the absence of ERPs, it can be useful to place increased weight upon the quality of the waveforms depicted in the time-series. Furthermore, one should be vigilant for evidence of alpha-band activity within the power spectrum density plot. When in doubt, use the 'BB_' designation and seek consultation. It is common for users to more frequently rely on group consensus decisions when preprocessing resting recordings.

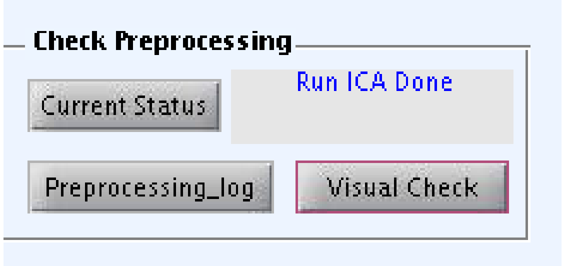
Resting recordings are almost always analyzed in terms of signal power in the frequency domain. Therefore, many artifacts averaged out during ERP analyses will still have effects within various frequency spectra. For example, heart rate signal is of particular concern. Any mixed 'B' and 'K' ICs should be labeled 'KB.' Similarly, any ICs with pervasive slow-wave artifacts throughout the time-series should be labeled 'SB.' When unsure, users can label ICs as 'BBS' and seek consultation about the pervasiveness of the slow waves. Most mixed 'B' and 'E' ICs should be labeled 'EB.' Finally, mixed 'B' and 'M' (i.e., strong and pervasive 'M') that would normally be low-pass filtered (e.g., 40 Hz) should be labeled 'MB.' As a rule, ICs should not be low-pass filtered because both low and high frequency information can serve as dependent variables.

8. Create Denoised Data

The final step of preprocessing involves reconstructing or “reconstituting” the epoched EEG recording using only brain ICs identified via ‘ICA_inspection’ with any requested filters. Enter a ‘Threshold for tBadElec’ value for temporary bad electrodes and click ‘Run’ (5 standard deviations is typical across tasks). This is the variability threshold an electrode must exceed to be considered temporarily bad at this step.

9. Check Preprocessing

ICAcleanEEG generates a log of the preprocessing steps taken for each recording file. It is sometimes helpful to access this log to check the status of a recording. In the main ICAcleanEEG GUI, several options are available within the 'Check Preprocessing' box.



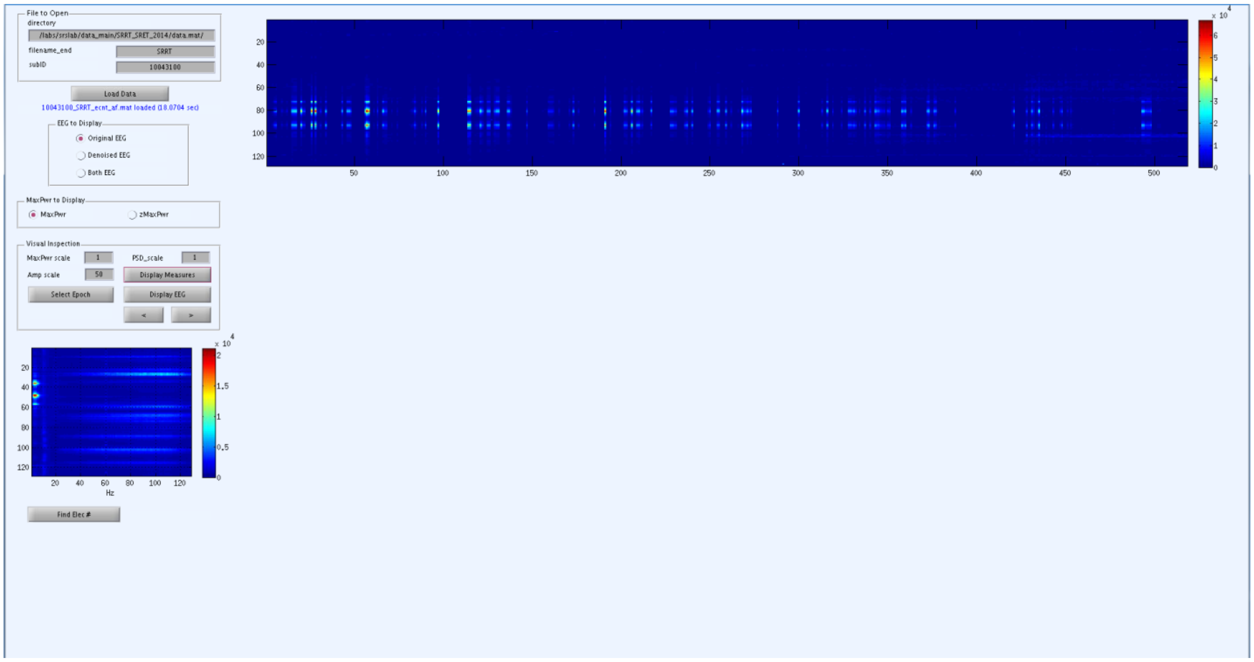
Click **'Current Status'** and the last step completed will be displayed in blue to the right. In the picture above, 'Run ICA' has been completed, meaning that the file still has not been run through ICA inspection.

Clicking **'Preprocessing_log'** will open the text file associated with the preprocessing pipeline. This file records more detailed information than simply what step is needed next. The log includes a record of when each step was completed, how many electrodes and epochs were deleted or filtered, which codes were accessed in order to run the pipeline, and how many ICs were decomposed from the data, etc.

Pre-post visualizations

The ICA preprocessing approach can appear like a "black box" methodology. This is particularly true when it is implemented in a fully automated fashion. Users may not fully appreciate the impact ICA is having on their recordings, which can result in either a too liberal or too conservative application of the technique. ICAcleanEEG was designed to combat this problem. We believe preprocessing quality can be improved by providing users with multiple visualization tools at every step of preprocessing to visualize the impact various decisions are having on the recording. ICA risks appearing like an easy fix for all EEG noise sources. In ICAcleanEEG, users are provided opportunities to appreciate when ICA performs effectively and adjust the process when additional improvements can be made.

ICAcleanEEG includes one final collection of features embodying this theme of visualization for the purpose of appreciating how recordings have been modified through preprocessing. Click the **'Visual Check'** button for a depiction of changes in the recording before and after ICAcleanEEG preprocessing ('VisualCheck_denoised_data' GUI). This provides a sense of the impact ICAcleanEEG prcedures can have on the data. Click **'Load Data'**. Once the file has loaded, change the ‘EEG to Display’ options by selecting **‘Original EEG’** and change the ‘MaxPwr to Display’ field by selecting **‘MaxPwr’**. Finally, click **‘Display Measures’**.



File to Open
directory: /labs/srslab/data_main/SRRT_SRET_2014/data.mat/
filename_end: SRRT
subID: 10043100

Load Data
10043100_SRRT_ecnt_af.mat loaded (18.0704 sec)

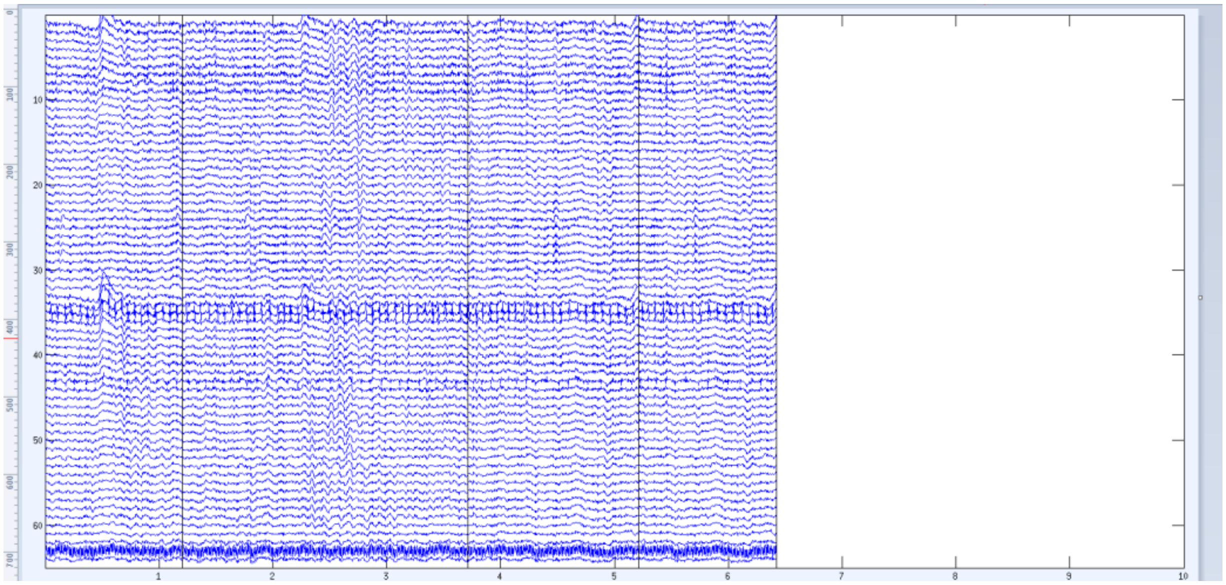
EEG to Display
● Original EEG
○ Denoised EEG
○ Both EEG

MaxPwr to Display
● MaxPwr ○ zMaxPwr

Visual Inspection
MaxPwr scale: 1
PSD_scale: 1
Amp scale: 50

On the bottom left of the GUI, a power spectrum density plot is depicted with frequency power on the x-axis and electrode number on the y-axis. At the top of the screen, max power is plotted. As expected, this plot shows time in seconds on the x-axis and electrode numbers on the y-axis. In contrast, if you select **‘zMaxPwr’** and click **‘Display Measures’**, you will see slightly different information in the max power plot. Here, the data are z-scored (i.e., statistically standardized) across epochs and within electrodes. In this case for a given electrode, a visible spike in power at an epoch will be depicted *relative* to the mean power of that same electrode across time.

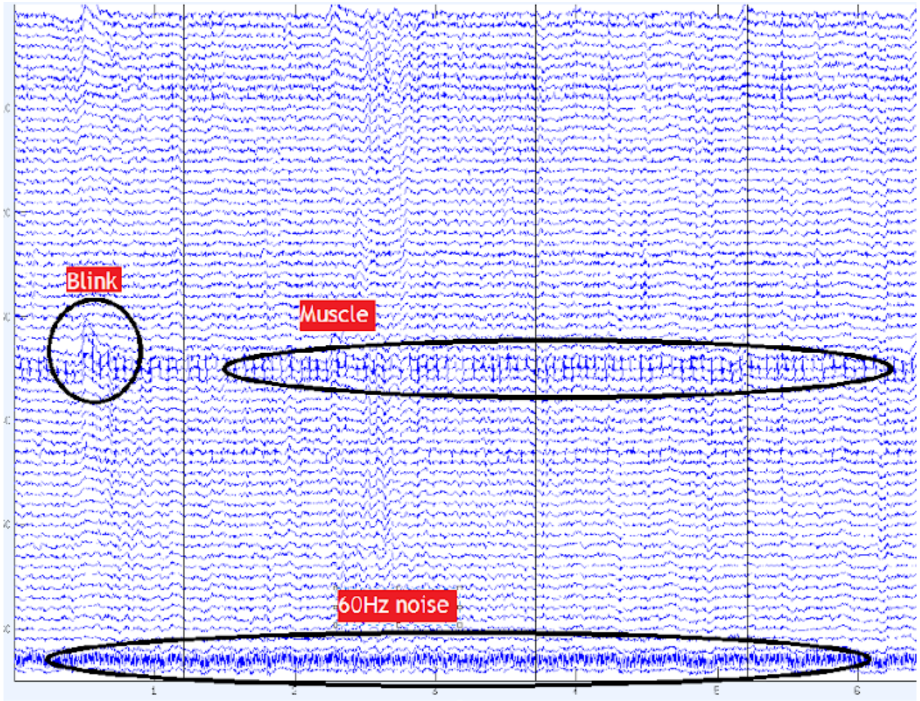
Switch between **‘Original EEG’** and **‘Denoised EEG’** to appreciate changes in the recording that occurred as a function of the preprocessing. Several types of alterations are often obvious. On the power spectrum density plot, notice how 60 Hz and/or higher frequency noise has been diminished or removed. Also, examine the lower frequencies to make sure plausible brain signal has not inadvertently been discarded. On the max power plot, look to see that isolated segments of noise have been removed. Bursts of noise visible on the ‘Original EEG’ plot are often the result of eye blinks. To take a closer look, click **‘Select Epoch’** in the ‘Visual Inspection’ box and use the crosshairs to choose a time point with a burst of high power. This will bring up an epoched time series, such as the one depicted data below.



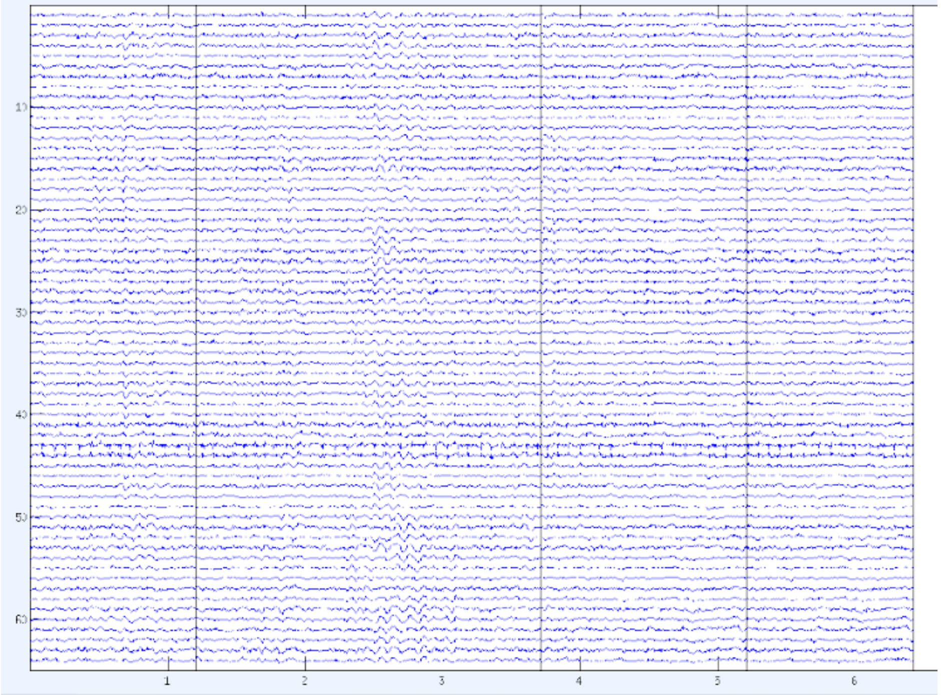
Notice that the entire 10 second plot is not filled with data. Epochs are not always defined as 1,000 ms duration time segments. This plot will only display whole epochs and a set of epochs will not always add up to exactly 10 seconds. You can scroll through the epochs using the arrow buttons in the 'Visual Inspection' box. Feel free to adjust the scale of the time series **('Amp scale'**), max power plot **('MaxPwr'**), and power spectrum density plot (**'PSD scale'**) using the respective fields in the 'Visual Inspection' box.

The max power and power spectrum density plots can only show original or denoised EEG data at a time. The time series, however, can show both sets of data superimposed onto each other. Select **'Both EEG'**. Now, the original EEG data appears in green and the denoised data is blue.

Below is an example of how to use this information:



The time series above shows a problematic period in the original EEG time series, which was identified through examination of the power spectrum plot. By examining the time series, several non-brain elements are apparent including a vertical blink, some muscle signal, and an electrode contaminated with 60 Hz noise. Now, select **'Denoised EEG'** and **'Display EEG'**. It is now apparent that ICAcleanEEG was implemented effectively for correcting these noise sources in the post-processed data.

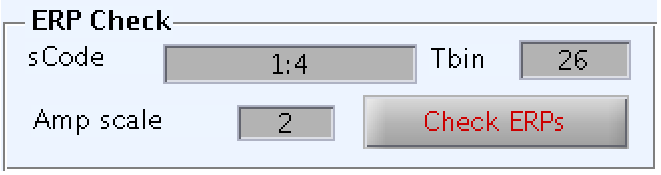


Important note about pre-post visual checks

These ‘VisualCheck_denoised_data’ plots *only* show epochs that exist in both the original and post-processed denoised data. Therefore, deleted epochs are not included in the display. Electrode numbers that have been deleted *do* appear due to interpolation.

ERP Check (within Visual Check)

The final visualization tool is the **‘ERP Check’** within the Visual Check GUI.



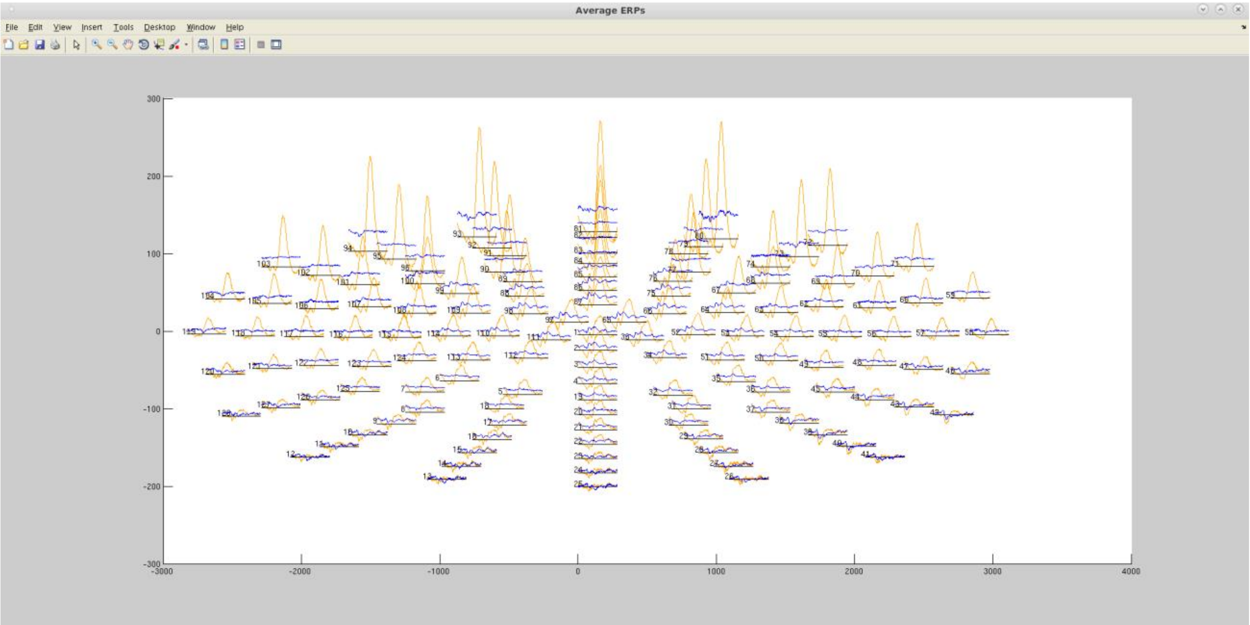
ERP Check

sCode: 1:4

Tbin: 26

Amp scale: 2

Using this tool, we can see the average waveform at each electrode on the scalp after the data has been preprocessed. In order to check the denoised ERPs, enter the sCode and Tbin information, and change the Amp scale to your liking (2 or 3 is recommended). sCode and Tbin information can be found within the sweep variable, and should be the same values used during ICA inspection. After entering the information, click **‘Check ERPs’.** A window titled **‘Average ERPs’** will appear.



The electrode montage is displayed in this window, and each electrode will have two overlapping average waveforms next to it. The orange-colored waveform is the original data and the blue-colored waveform is the denoised data. In the example, notice the large spike in the original data that is most-prominent at frontal sites, but is absent in the denoised data. This corresponds to the eye blink variance identified and removed via ICA.

Preprocessing completed

At this stage, preprocessing is now considered “complete.” The recordings for each subject are now ready for quantification. Post-processed EEG analysis is beyond the scope of this tutorial. In the CAB Lab, quantification of individual recordings is most frequently done using MATLAB scripts. However, additional scripts are also available to export post-processed data to other software environments such as BESA and EEGLAB/ERPLAB. As a research assistant, seek guidance before attempting to quantify post-processed data for analysis. Best practices for these latter steps deserve separate tutorials in their own right (e.g., “An Introduction to the Event-Related Potential Technique” by Steven Luck [2014; MIT Press]).

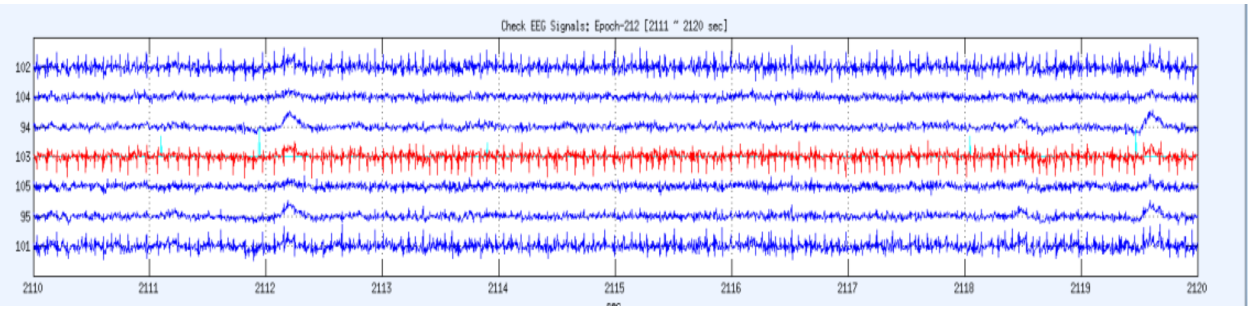
Appendix A: Various Artifact Examples

Recording artifacts can often appear very different across participants. Below are some additional time series examples users may encounter, which are accompanied with recommended courses of action.

Muscle

Example 1

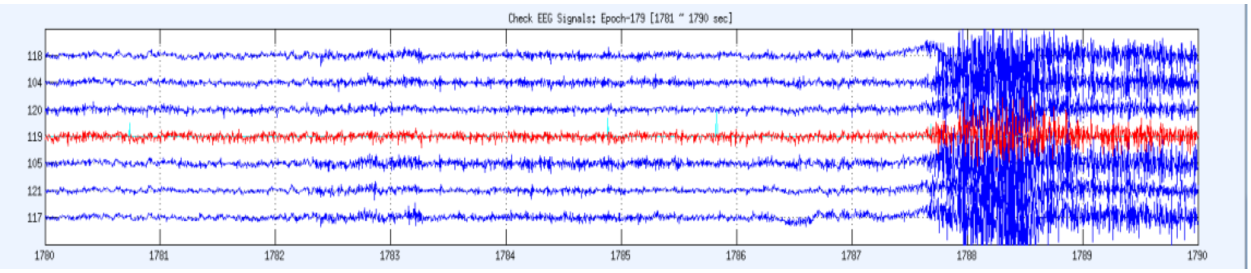
Muscle signal can manifest in many ways. Below is an example of small amplitude, but temporally consistent muscle signal observed in electrode 103 and several neighboring electrodes (e.g., 101, 102). The signal does have frequent jagged peaks, but also reflects the signal patterns of less-contaminated nearby electrodes.



This type of muscle artifact is not a problem for ICA. No epochs or electrodes need to be deleted during the ‘badElec’ step. ICA decomposition should be able to isolate this muscle noise from underlying neural signal.

Example 2

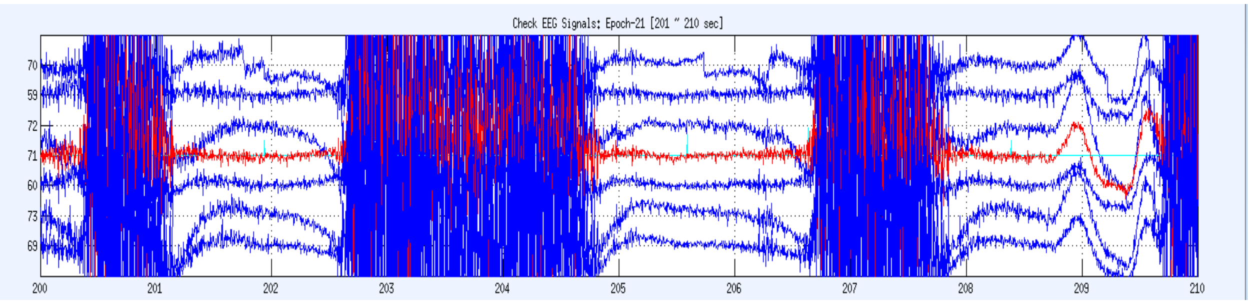
Muscle can also appear as bursts of strong noise. This typically happens if a participant momentarily clenches their jaw, moves their head to the side, or coughs.



If the artifacts are strong enough that it is impossible to distinguish nearby electrode signals from one another (using 100 microvolt amplitude scaling), these bursts should be treated as strong bad epochs to be deleted. The case above is not prominent enough for the time segment to be deleted. ICA decomposition should be able to remove the muscle noise from the underlying neural signal.

Example 3

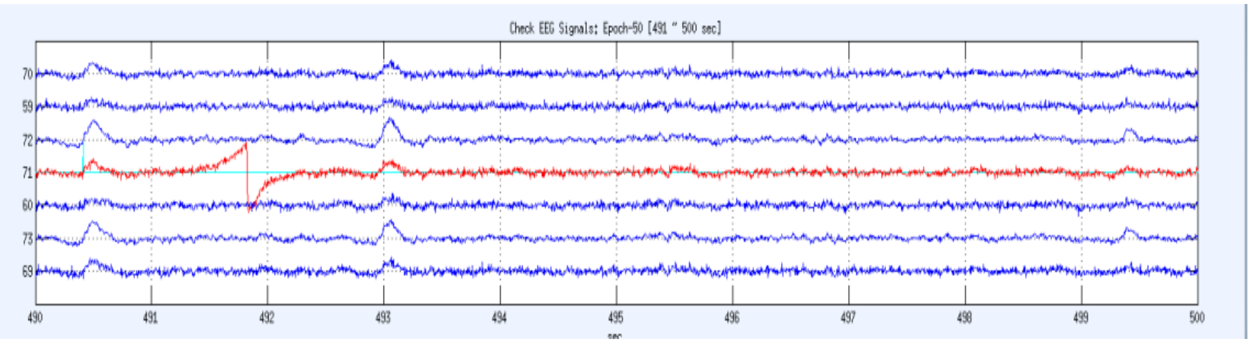
Below is a more extreme example of a possible muscle-related artifact:



This time series shows very frequent, extreme bursts. The signal in between the bursts also seems to be noisy. Because of the strong muscle activity and the failure of the signal to recover in between the bursts, all the time segments shown should be deleted. If you see a case like this with the extreme muscle bursts contaminating large portions of the file, please make a note. We may have to investigate if this file is usable at all.

Small deflections

In the example below, electrode 71 exhibits a brief deflection near time segment 492. This deflection is not large enough to be concerning for the purpose of ICA decomposition. For this case, do not delete any epochs. However, if electrode 71 shows these deflections frequently throughout the recording, it may be best to delete it before moving to ICA.

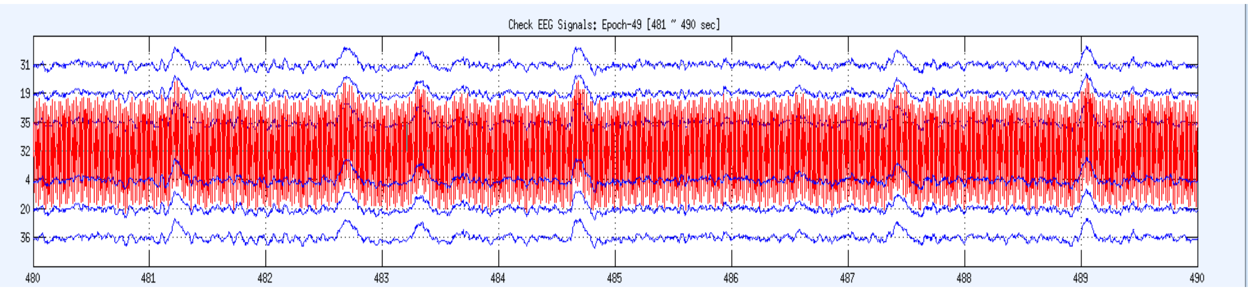


60 Hz noise

Recognizing 60 Hz noise contamination can be difficult for beginning users, but it can be easier to identify after reviewing examples and noticing several key features.

Example 1

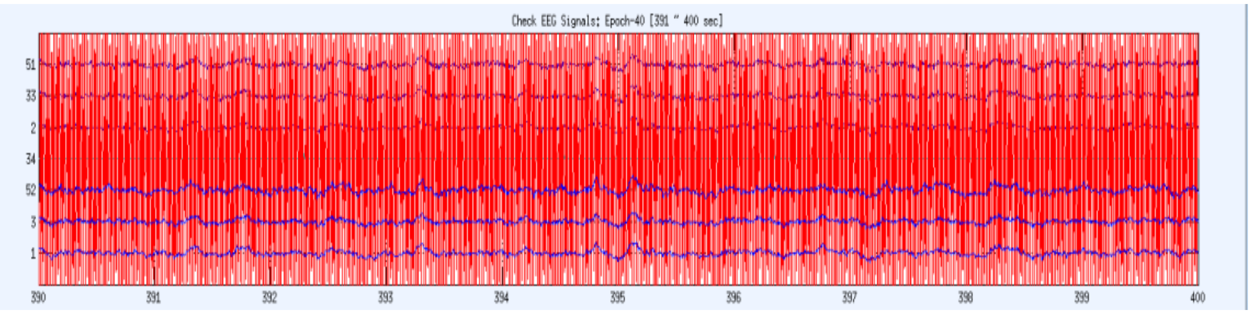
Below is an example of an electrode contaminated with 60 Hz noise. Electrode 32 (in red) appears much denser and more saturated than its neighbors on the scalp.



However, the signal in this electrode also follows the peaks and valleys of the neighboring electrodes. In this case, electrode 32 should be filtered for 60 Hz noise rather than deleted and interpolated.

Example 2

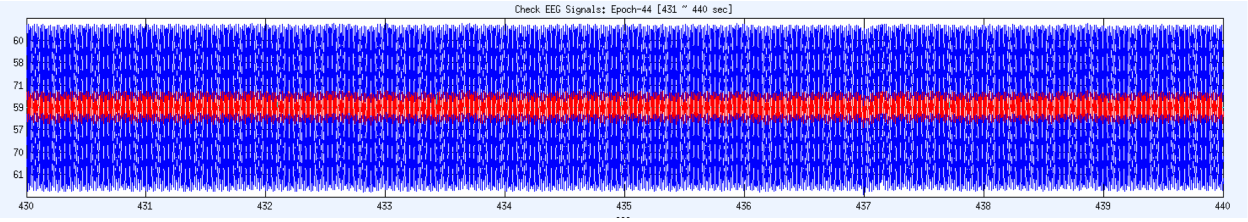
The signal at this electrode is also contaminated with 60 Hz noise, but it is impossible to detect any underlying signal that may be covarying with neighboring electrodes.



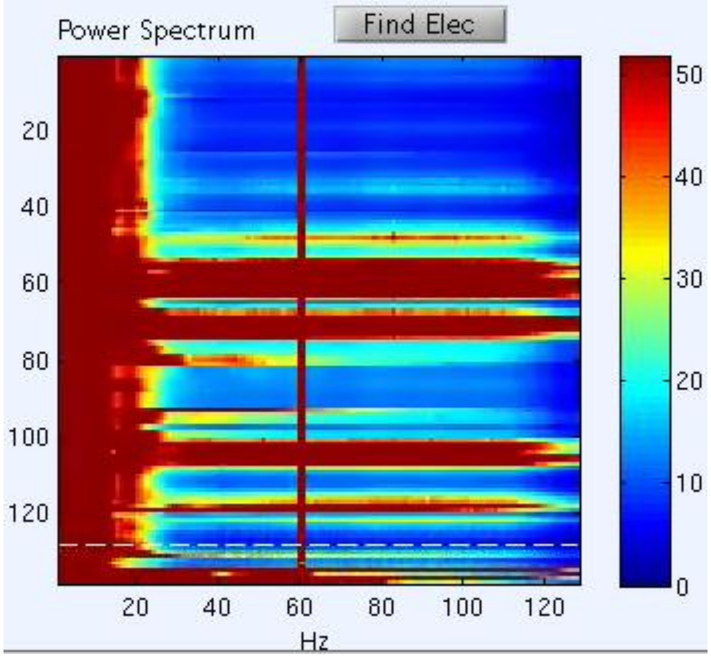
First, users should filter electrode 34 for 60 Hz noise during ‘badElec’. During the later ‘ICA_inspection’ step, this electrode could also be deleted if the notch filtering does not reveal any meaningful neural signal.

Example 3

Finally, all electrodes in some recording files may be contaminated with 60 Hz noise to some degree.



Filter all electrodes for 60 Hz noise and proceed cautiously to the ‘ICA_inspection’ step to assess the impact of the filtering. The most useful tool for confirming 60 Hz noise is the power spectrum plot:

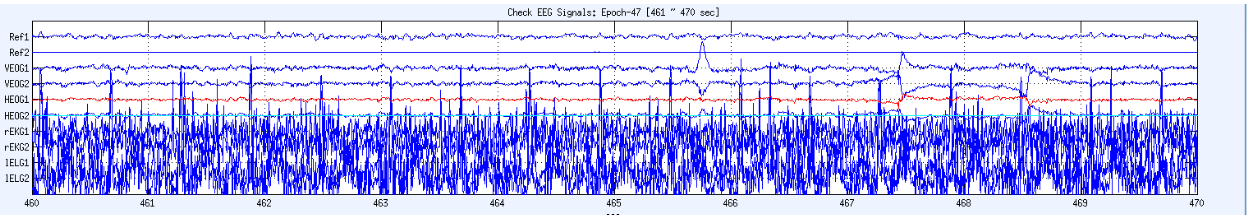


In this example, the vertical red stripe at 60 Hz tells us there is strong 60 Hz noise in all channels.

Referencing problems

As explained at the beginning of this manual, reference electrodes are critically important because EEG signal cannot be measured “in a vacuum”; it must be quantified relative to another signal of some sort. Reference electrodes provide this context with which to interpret the measured electrical signal.

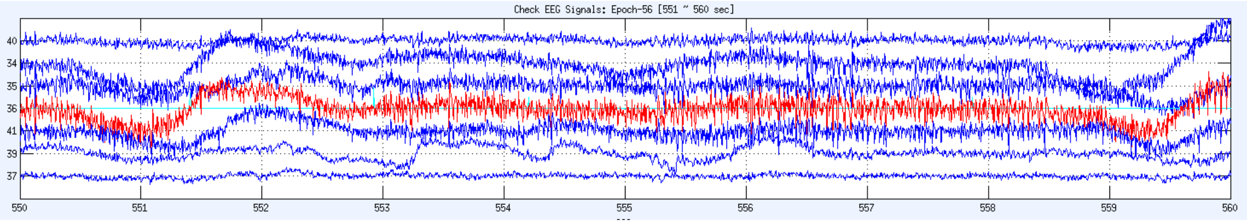
If there is something wrong with the reference electrodes, the remaining EEG channels may appear altered due to incorrect referencing. Make sure to check the peripherally recorded reference electrodes during ‘BadElec’ to avoid this problem. For example, below is a depiction of the Ref2 channel when it has been completely corrupted. There is absolutely no usable signal recorded by Ref2. This is may be due to the electrode being improperly attached, falling off, or being broken.



This is an issue that requires documentation, but can be fixed. Do not delete this electrode. Make a clear record of the problem so the EEG signal can be referenced to only Ref1 for this file during preprocessing.

Slow waves

Here is an example of slow waves. Slow waves are commonly caused by sweat. This case is a bit trickier. If you see slow oscillations like this, scroll through the file to assess if the waves are temporary or continue throughout a large portion of the file.



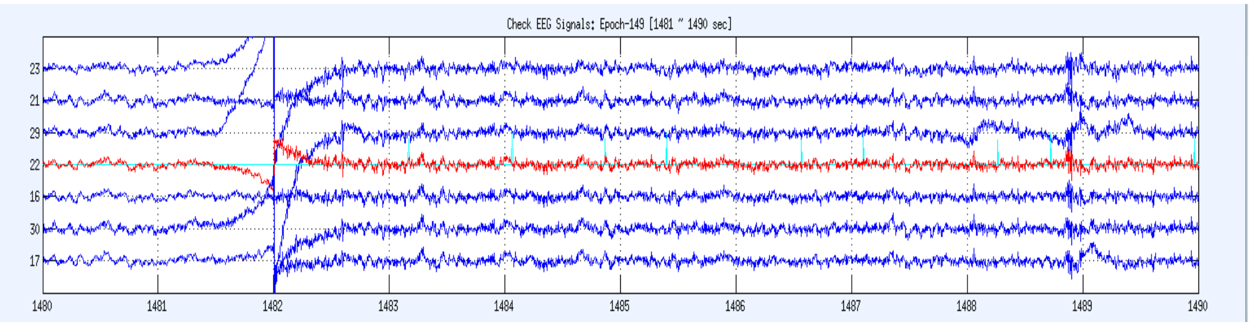
If the slow waves are short-lived, delete the time segments where these artifacts are discernible. If slow waves are pervasive across most of the file, make a note and leave the file unaltered with regards to this artifact. It may be possible for the slow wave signal to be isolated during ICA inspection. If that is not the case, there may be a more complicated problem present requiring some sort of unique solution.

EEG Recording discontinuities

Occasionally, there may be other problems with the recording hardware while the actual EEG electrodes are functioning appropriately.

Example 1

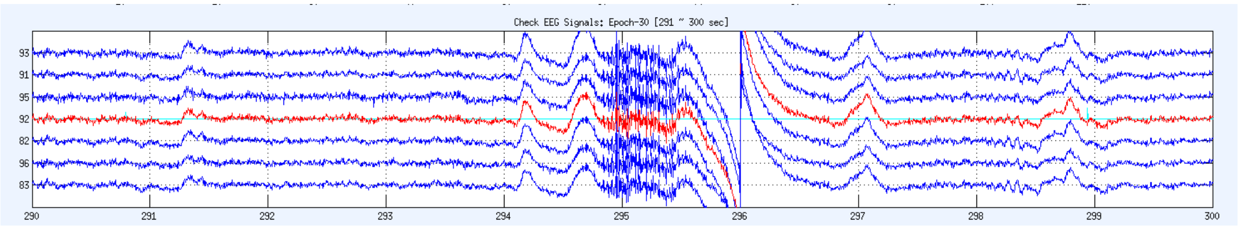
Within time segment 1482 and 1483 depicted below, there is a marked interruption in the signal across all electrodes. This can be caused by problems a momentary electrical interruption occurs. Also, discontinuity like this might occur when technicians administering the tasks manually stop and restart a recording session.



Delete these affected time segments (1482 1483) because of the relative prominence of the high amplitude noise deflections.

Example 2

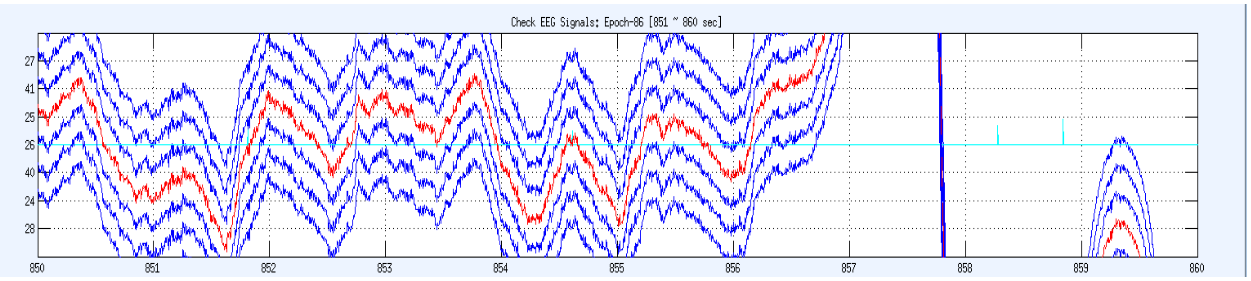
Below is a slightly different example of the same type of discontinuity.



In this example, time segments 296 & 297 do not necessarily need to be deleted because there are no task events occurring within 1 second of each time segment. The closest task event occurs right before time segment 299 (small, cyan-colored vertical line).

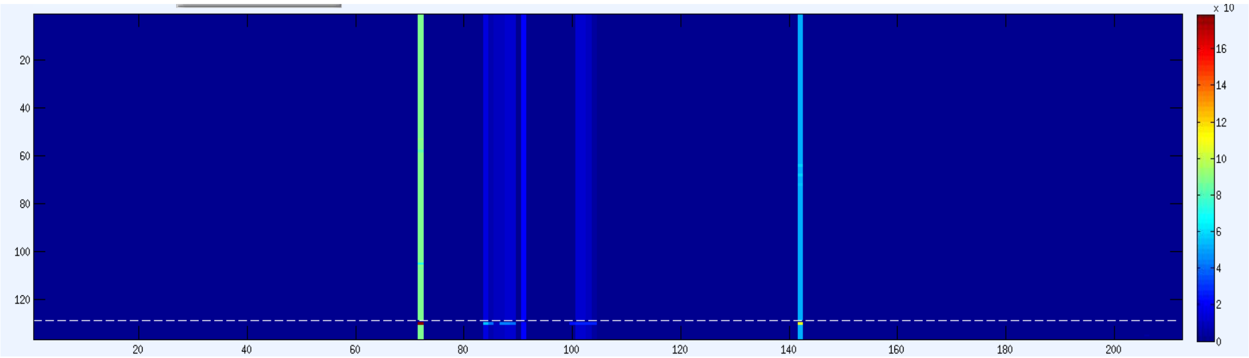
Example 3

High-density EEG systems often organize electrodes into grouping "braids." For both the BioSemi and Brain Products systems in the CAB Lab, each braid contains 32 electrodes. Sometimes, there are problems isolated to specific braids, which leads to recording noise and discontinuities that are more extreme.



In these cases, delete all time segments that look like the time segments above. None of the signal was collected properly. Typically, a significant number of time segments need to be deleted when this occurs. Be sure note why a significant amount of the file needed to be deleted.

The max power plot in 'BadElec' can provide clues as to when recording artifacts may be occurring in the time series (below).



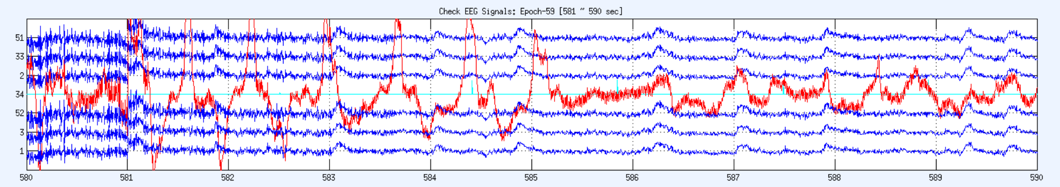
The plot above shows vertical bands of extreme noise while the rest of the plot appears essentially monochromatic. This is not a definitive hallmark of hardware problems, but it may be a helpful indicator of where a recording could be particularly problematic.

Bad electrodes

Electrodes can be "bad" for any number of reasons and may be identifiable based several possible signal patterns. Delete an electrode if the signal appears consistently erroneous and does not consistently mimic the surrounding electrodes.

Example 1

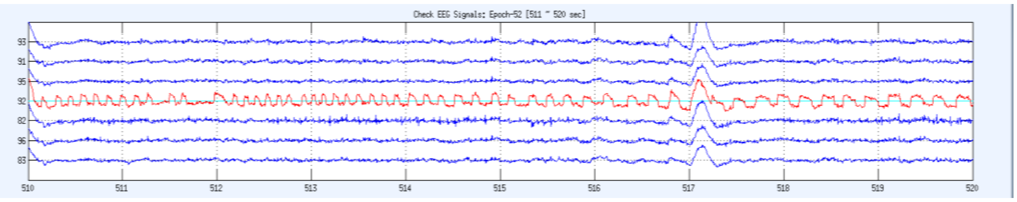
For some subjects, bad electrodes may appear like electrode 34 in the time series below.



Delete this electrode.

Example 2

For the example below, electrode 92 is quite odd. The signal is consistent with itself across time, but is inconsistent with the neighboring electrodes.



Delete this electrode.

Appendix B: 'ICA_inspection' classification examples

IC classification takes practice. Checking the automatically generated ICs in 'ICA_inspection' is a skillset based on expertise that is developed over time. Referencing the manual while preprocessing will allow you to internalize the recommended decision-making steps and become quicker at identifying components.

Take time to review several additional ICs in this Appendix B before beginning preprocessing with real recordings. These examples are by no means exhaustive, but do provide concrete examples of EEG phenomena that are frequently observed during ICA. As always, consult with other staff if you run into strange ICs with unusual phenomena.

For the screenshots below, the 'Category' listed in the 'Classify Current IC' box is the automatically generated classification. As such, they may not reflect the actual correct classification of the ICs and should be ignored. Furthermore, note the GUI layout with 'ICA_inspection' on the left and 'ERP Check' on the right. This is the recommended way to use these plots.

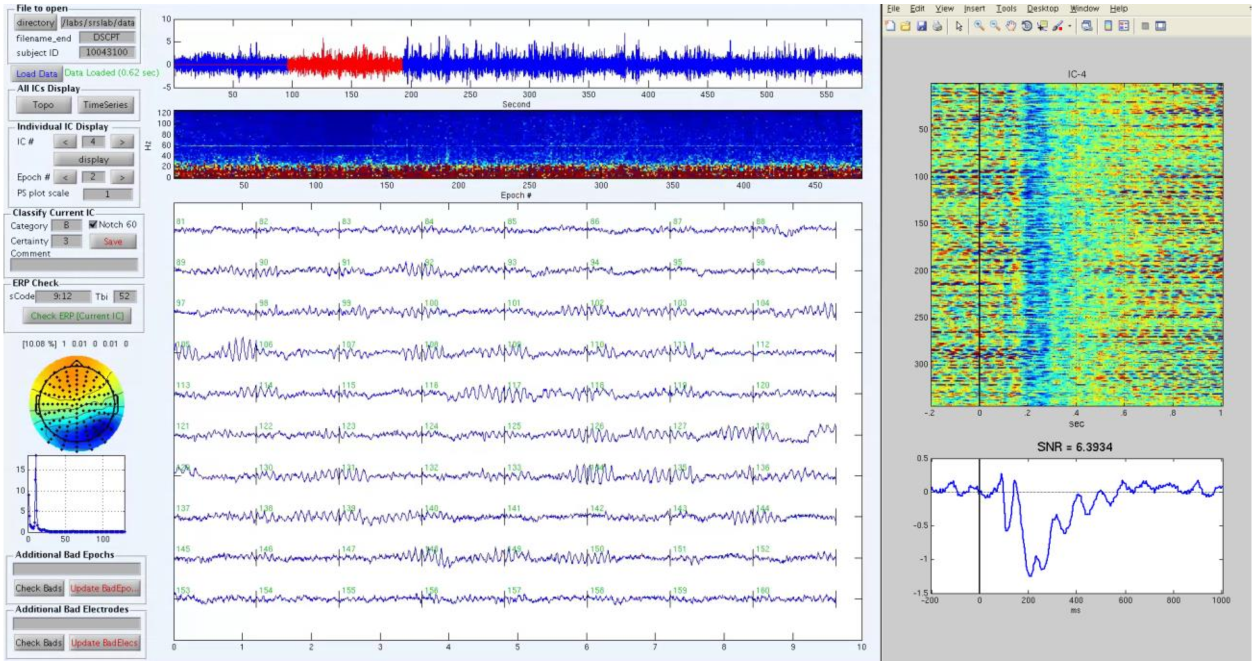
Brain ('B')

Current CAB Lab criteria for brain ICs are met when at least one of the following is true:

ERPs are evident: A grand average ERP can be observed using the 'ERP Check' function and there are consistent deflections across trials congruent with the averaged ERP shape (and incongruent with blinks or ECG signal, if the blinks or ECG deflections are time locked). 'SNR' values are not currently used to make this decision.

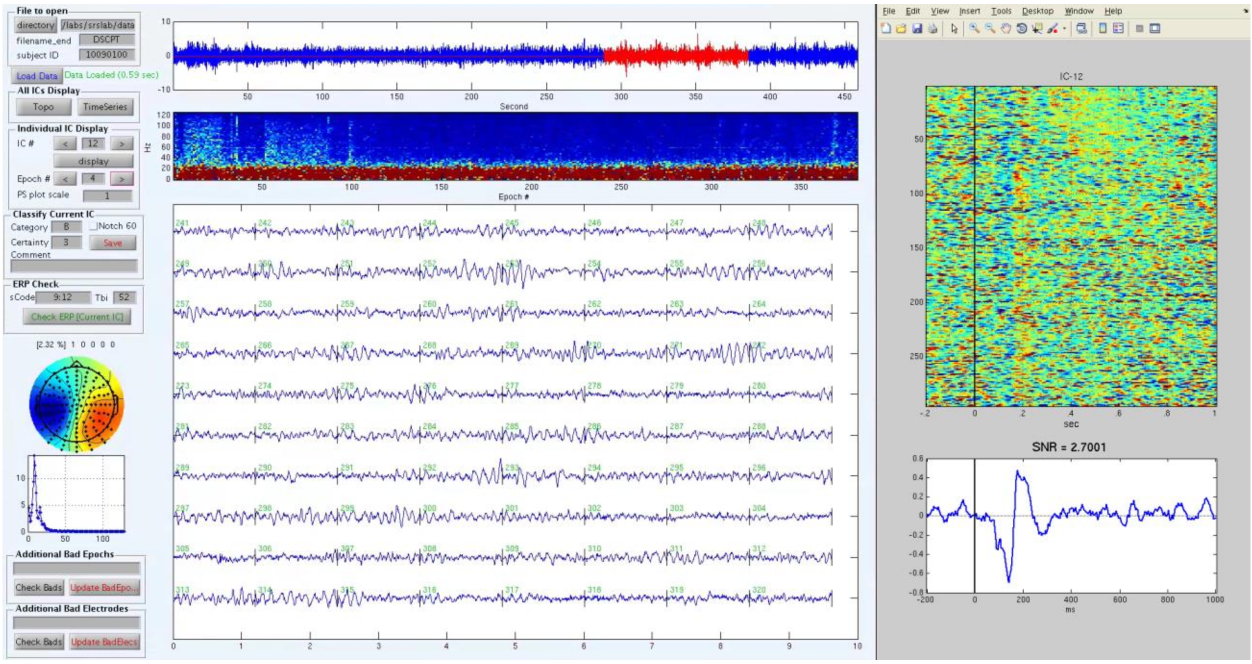
Time series oscillations: There are oscillations in the time series resembling brain activity (e.g., alpha band [8 – 15 Hz] activity). These oscillations may or may not be time-locked to the onset of the task stimuli as visualized using the 'ERP Check' function.

Example 1



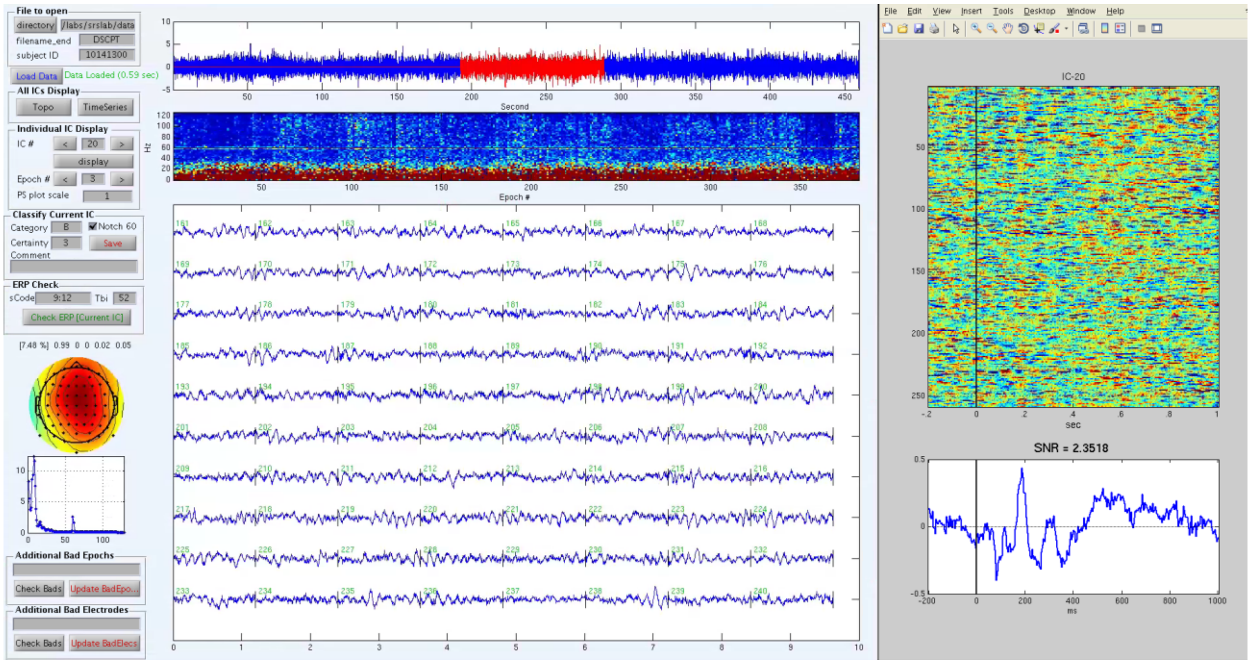
File to open
directory /labs/srslab/data
filename_end DSCPT
subject ID 10043100
Data Loaded (0.62 sec)
Individual IC Display
IC # 4
Epoch # 2
PS plot scale 1
Classify Current IC
Category 8 ☑ Notch 60
Certainty 3
Comment
ERP Check
sCode 9:12 Tbi 52
[10.08 %] 1 0.01 0 0.01 0
IC-4
SNR = 6.3934

Example 2



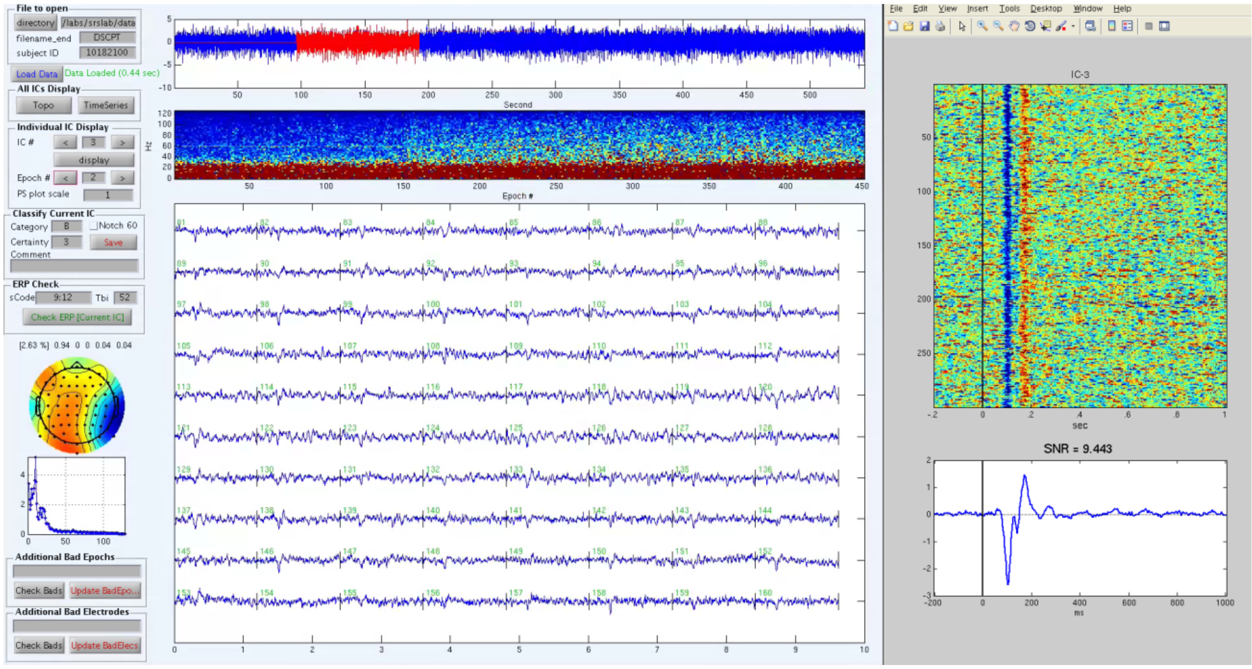
File to open
directory /labs/srslab/data
filename_end DSCPT
subject ID 10090100
Data Loaded (0.59 sec)
Individual IC Display
IC # 12
Epoch # 4
PS plot scale 1
Classify Current IC
Category 8 □ Notch 60
Certainty 3
Comment
ERP Check
sCode 9:12 Tbi 52
[2.32 %] 1 0 0 0 0
IC-12
SNR = 2.7001

Example 3



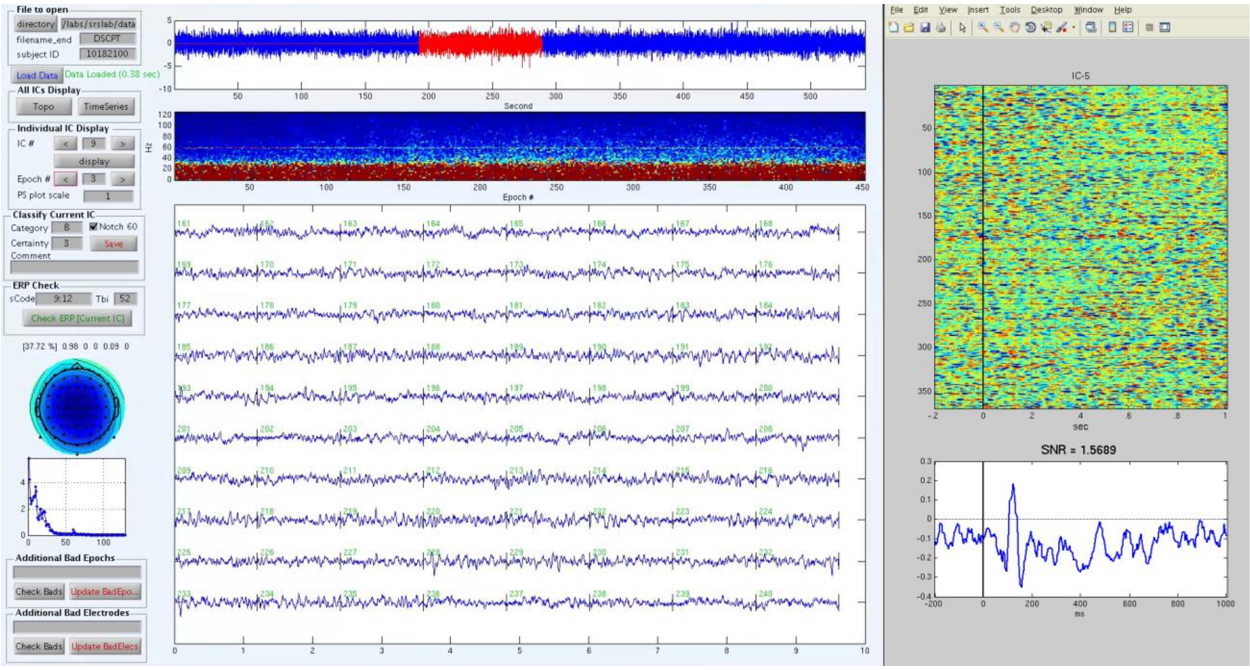
File to open
directory /labs/srslab/data
filename_end DSCPT
subject ID 10141300
Data Loaded (0.59 sec)
Individual IC Display
IC # 20
Epoch # 3
PS plot scale 1
Classify Current IC
Category B ☑ Notch 60
Certainty 3
Comment
ERP Check
sCode 9:12 Tbi 52
[7.48 %] 0.99 0 0 0.02 0.05

Example 4



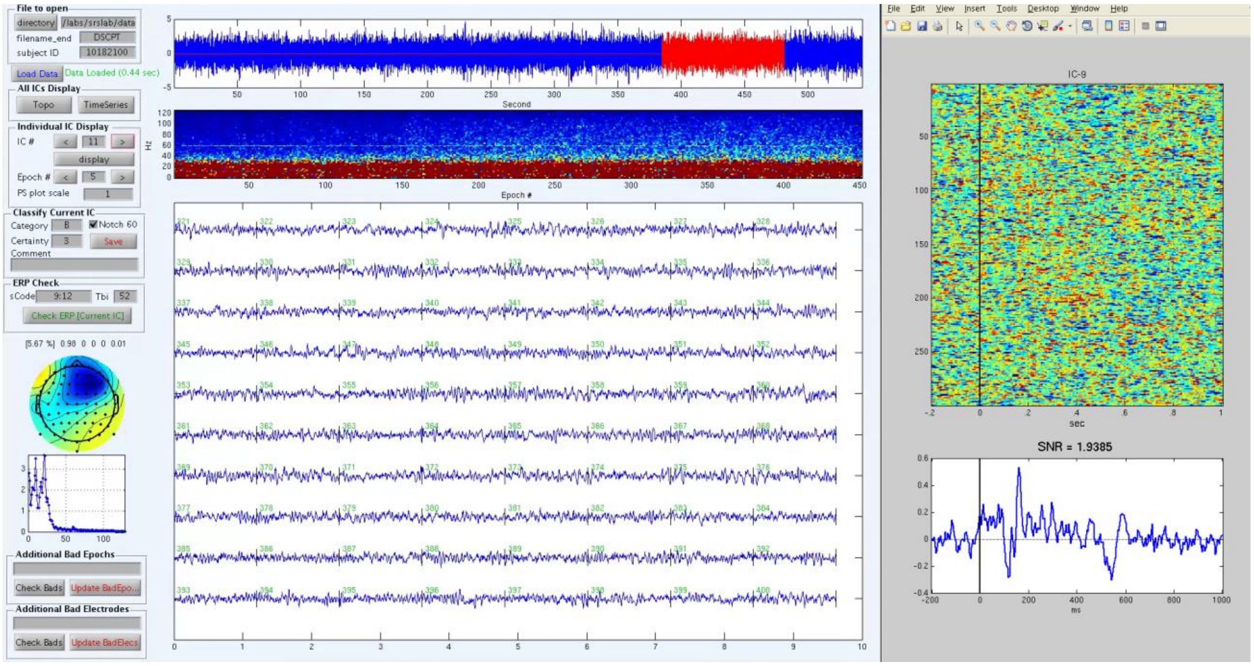
File to open
directory /labs/srslab/data
filename_end DSCPT
subject ID 10182100
Data Loaded (0.44 sec)
Individual IC Display
IC # 3
Epoch # 2
PS plot scale 1
Classify Current IC
Category B ☐ Notch 60
Certainty 3
Comment
ERP Check
sCode 9:12 Tbi 52
[2.63 %] 0.94 0 0 0.04 0.04

Example 5



File to open
directory: /labs/srslab/data
filename_end: DSCPT
subject ID: 10182100
Data Loaded (0.38 sec)
IC #: 9
Epoch #: 3
PS plot scale: 1
Classify Current IC
Category: B
Certainty: 3
☑ Notch 60
ERP Check
sCode: 9:12
Tbi: 52
[37.72 %] 0.98 0 0 0.09 0

Example 6

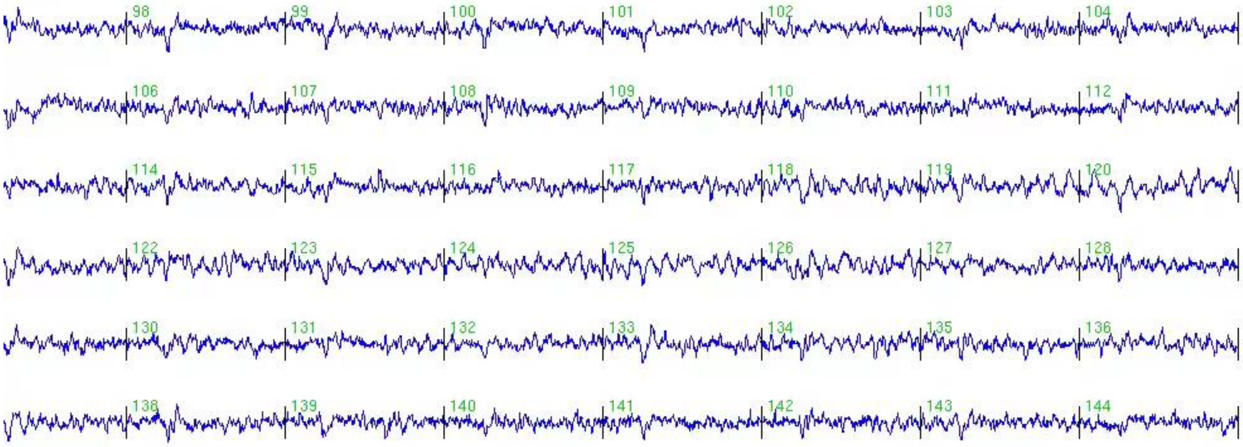


File to open
directory: /labs/srslab/data
filename_end: DSCPT
subject ID: 10182100
Data Loaded (0.44 sec)
IC #: 11
Epoch #: 5
PS plot scale: 1
Classify Current IC
Category: B
Certainty: 3
☑ Notch 60
ERP Check
sCode: 9:12
Tbi: 52
[5.67 %] 0.98 0 0 0 0.01

Commentary

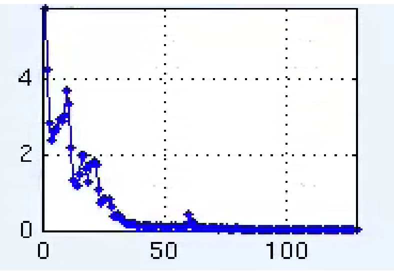
ICs examples 1 – 3 represent some of the most common and easily identifiable types of brain ICs (‘B’) in EEG recordings. The power spectrum, power spectrum density, and time series displays are all very similar across ICs. The power spectrum and PSD plots both show the strongest power in lower frequencies. The time series plots show periods of readily apparent neurophysiological oscillations. However, topographies for Examples 1 – 3 are all different from one another. Despite this difference, each IC is clearly depicting neural processes. Brain topographies often show a diffuse pattern across the entire scalp rather than being isolated to scalp edge electrodes (e.g., muscle ICs) or single electrodes (e.g., bad electrodes). The color of the topographies often covers the entire "rainbow" color spectrum and is not limited to only "hot" or "cold" colors.

Example 4 is slightly more complex because it depicts a brain IC composed of at least two different neurophysiological phenomena. It is clearly brain and should be classified as 'B'. An obvious ERP is evident within the 'ERP Check' window. However, the time series appears more complicated.



A deflection consistently appears following the onset of each epoch. Since this task involved visual stimuli, this activity likely represents the initial processing of the visual stimuli in the primary visual cortex. The Example 4 IC is unique in that there are also regular alpha oscillations co-occurring with this visual response seen clearly in epochs 118-128. It is common to observe at least some alpha oscillations mixed with primary cortex activation ICs because their topographies and frequency compositions are similar.

Example 5 is also a brain IC ('B'). Similar ICs often take up a significant portion of the file variance, and usually appear topographically as either all blue/green or all red/orange. This IC depicts brain activity perhaps derived from deeper or more widespread sources in the brain. Notice the higher frequency composition than the first few brain ICs (powers spectrum reproduced below). There are peaks appearing at approximately 15-30 Hz, which are still lower than typical muscle frequencies.



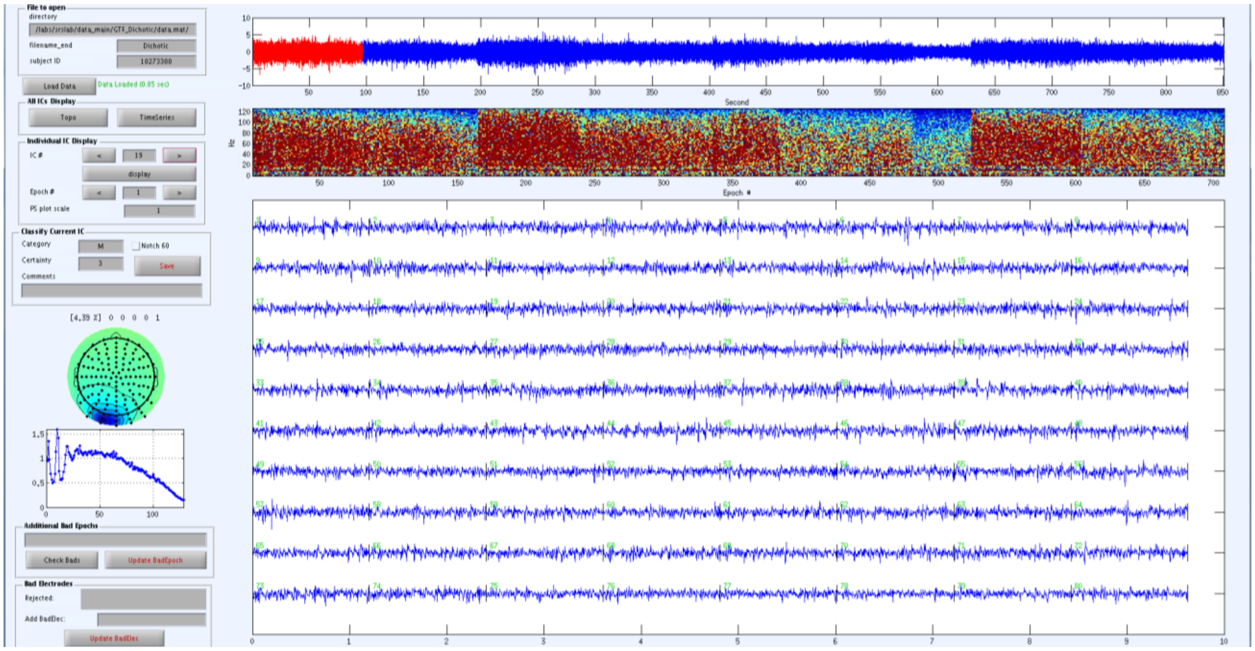
Example 6 is another brain IC with a mix of two types of brain activity. In this case, it contains alpha oscillations with even higher frequency brain activity. This is illustrated by the second peak in the power spectrum density plot around 25-30 Hz. In addition, quickly oscillating neural signals can be seen in the time series plot and the topography is generally centered on more anterior/frontal regions. It can be hard to distinguish high frequency brain activity from muscle noise. It is important to use the 'ERP Check' function and see if the high frequency activity extends beyond 25-50 Hz. ICs with even higher frequency neural activity are hard to identify, so it is best to ask questions if you encounter them. Unless you are specifically preprocessing the data to be used for an analysis of beta and gamma activity, you shouldn't spend too much time determining if you are looking at high frequency neural or muscle activity.

Muscle ('M')

Muscle ICs are the most common type of components in EEG files. ICA attempts to identify independent sources of both brain and noise within individual recordings. Jaw clenching creates a different pattern in EEG deflections than coughing, which means the muscle activity distinct to each can potentially be identified as different ICs during preprocessing.

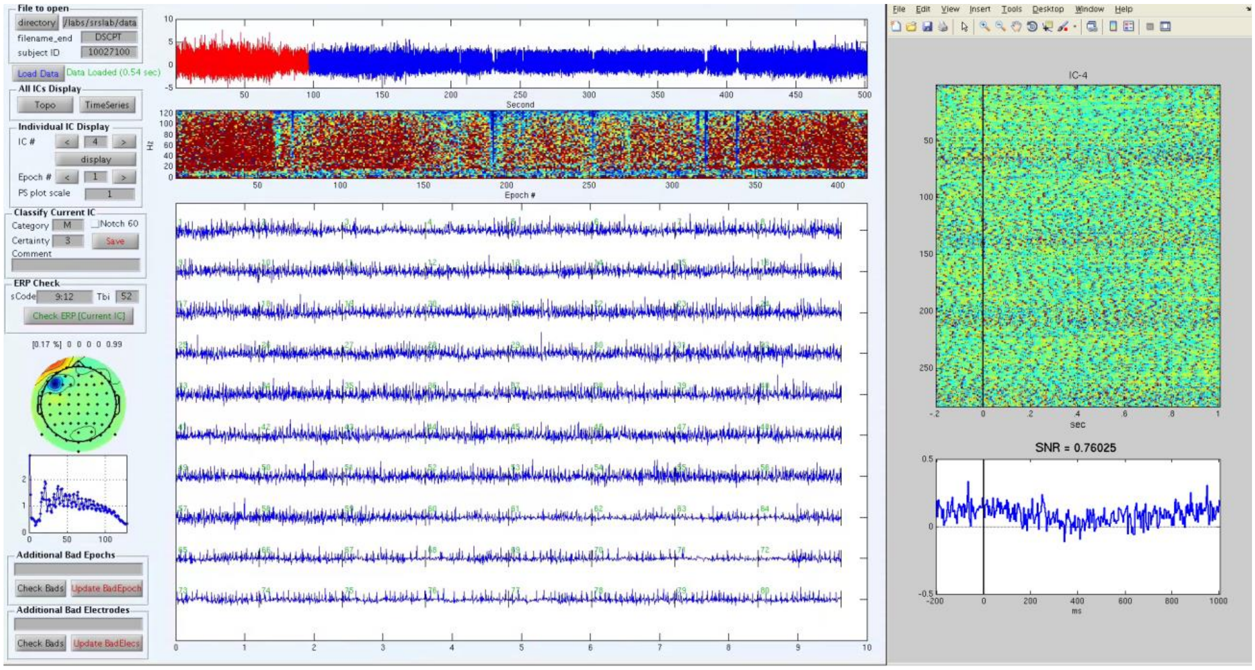
Muscle activity is identifiable primarily using the power spectrum plots. Solid red vertical lines may indicate that muscle activity is present for that given epoch. Any apparent brain activity during those pronounced epochs are likely unusable for either ERP or time frequency analysis. This is particularly true when there may be some observable brain activity in the time series, but there is reason to believe it not be separable from muscle activity even after applying a low pass filter.

Example 1



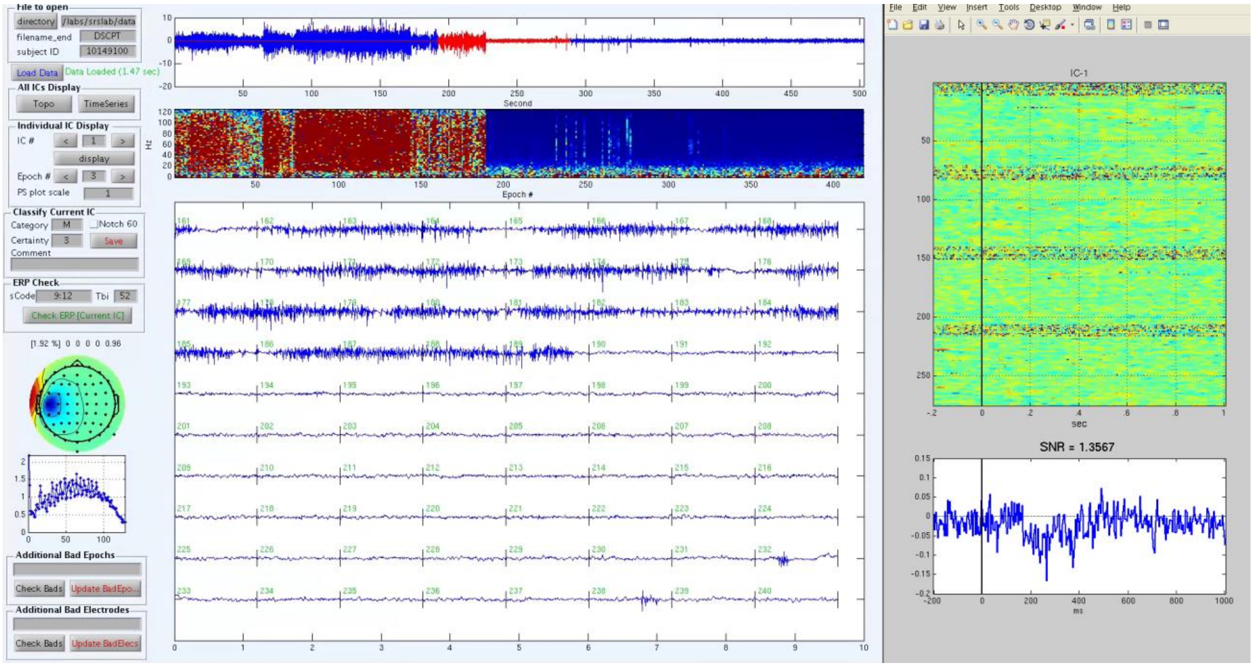
File to open
directory: /labs/srslab/data_main/GTF_Dichotic/data.mat/
filename_end: Dichotic
subject ID: 10273300
Data Loaded (0.85 sec)
IC #: 19
Epoch #: 1
PS plot scale: 1
Category: M
Certainty: 3
□ Notch 60
[4.39 %] 0 0 0 0 1

Example 2



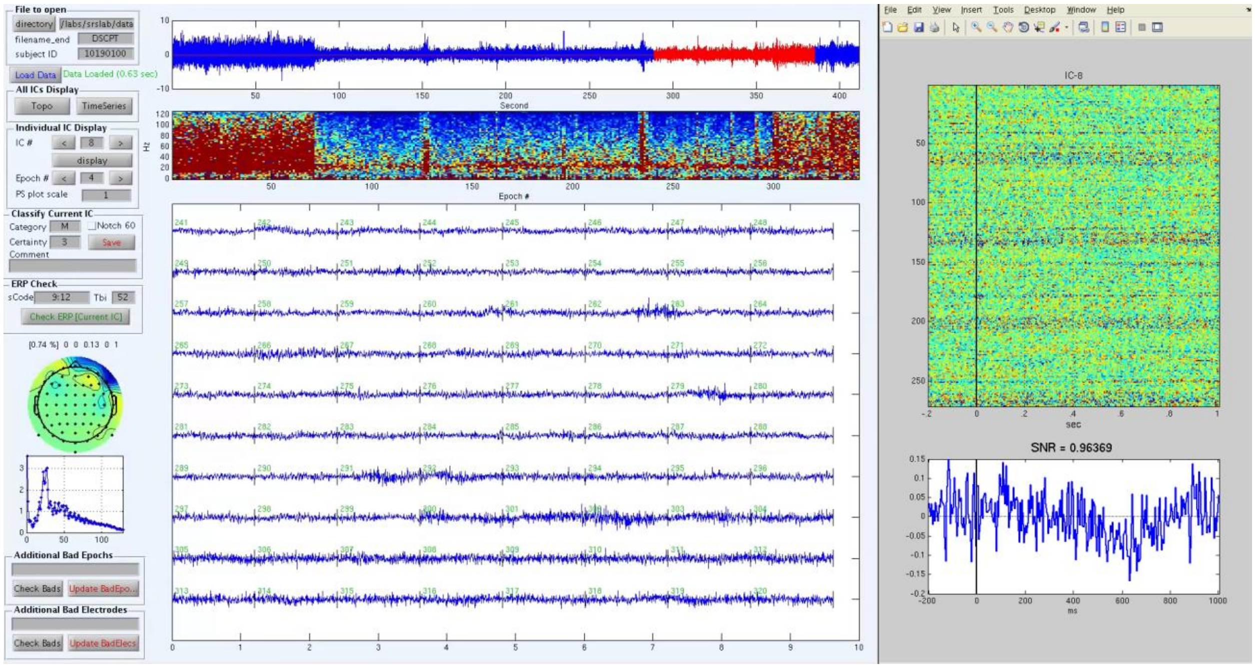
File to open
directory: /labs/srslab/data
filename_end: DSCPT
subject ID: 10027100
Data Loaded (0.54 sec)
IC #: 4
Epoch #: 1
PS plot scale: 1
Category: M
Certainty: 3
□ Notch 60
ERP Check
sCode: 9:12 Tbi: 52
[0.17 %] 0 0 0 0 0.99
IC-4
SNR = 0.76025

Example 3



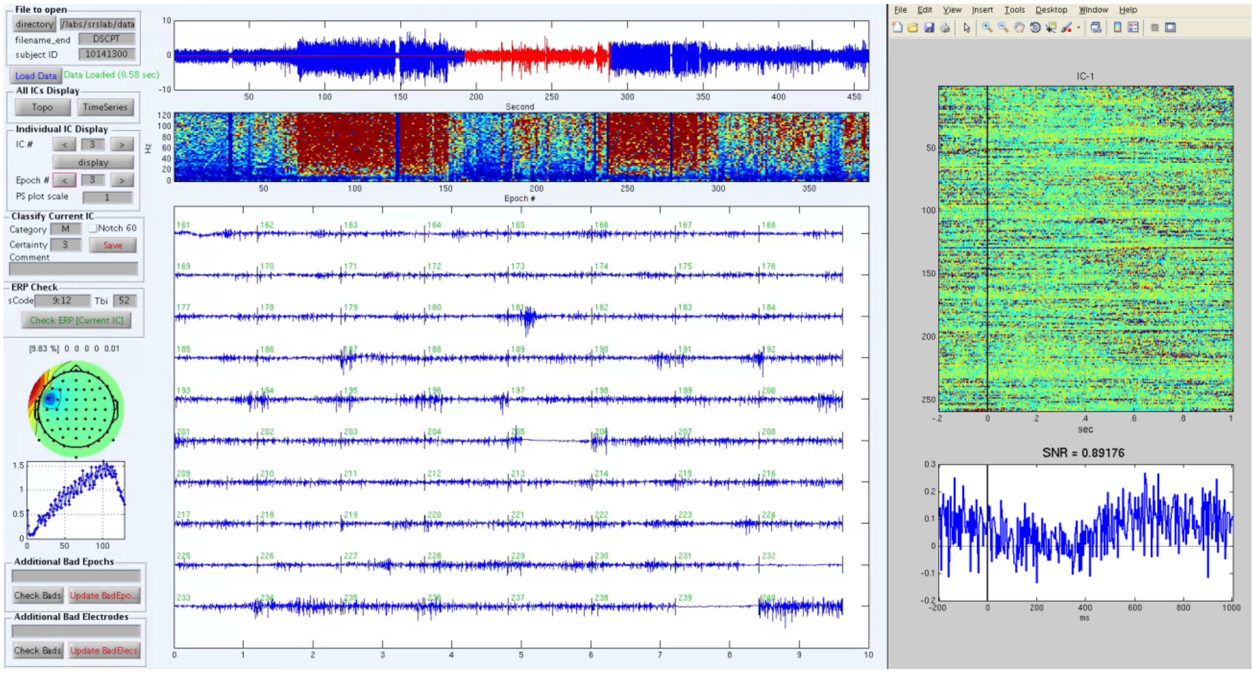
File to open
directory /labs/srslab/data
filename_end DSCPT
subject ID 10149100
Load Data Data Loaded (1.47 sec)
IC # 1
Epoch # 3
PS plot scale 1
Category M
Certainty 3
sCode 9:12 Tbi 52
[1.92 %] 0 0 0 0 0.96
SNR = 1.3567

Example 4



File to open
directory /labs/srslab/data
filename_end DSCPT
subject ID 10190100
Load Data Data Loaded (0.63 sec)
IC # 8
Epoch # 4
PS plot scale 1
Category M
Certainty 3
sCode 9:12 Tbi 52
[0.74 %] 0 0 0.13 0 1
SNR = 0.96369

Example 5



File to open
directory: /labs/srslab/data
filename_end: DSCPT
subject ID: 10141300
Data Loaded (0.58 sec)
IC #: 3
Epoch #: 3
PS plot scale: 1
Category: M
Certainty: 3
sCode: 9:12
Tbi: 52
[9.83 %] 0 0 0 0 0.01
SNR = 0.89176

Commentary

ICA is capable of isolating muscle activity that is continuously present across the recording (Examples 1 – 2) or isolated to specific epochs (Examples 3 – 5). Across all examples, notice the strong signal in higher frequency bands and the topographies that are isolated to only part of the head. Often, topographies exhibit a pronounced and spatially concentrated dipole (i.e., strong "warm" colors near strong "cool" colors) such as that observed in Example 2. Muscle time series typically appear very jagged without the smooth rise and fall that is characteristic of neural activity. Brain waves should appear like sine waves. In contrast, muscle produces erratic and high amplitude deflections that do not exhibit a gradual transition from peak to trough. The power spectrum density plots vary across muscle ICs, but generally show a jagged "saw tooth"-like pattern within the higher frequency range. Examples 3 – 5 are exemplars of this phenomenon.

Mixed brain and muscle ICs

Brain and muscle signal often co-occurs within ICs. It is important to know whether a particular IC is brain-dominant or muscle-dominant; the former should be retained and the latter should be removed from further processing. If you can quickly recognize brain activity in a recording, classify the IC as brain dominant in some form.

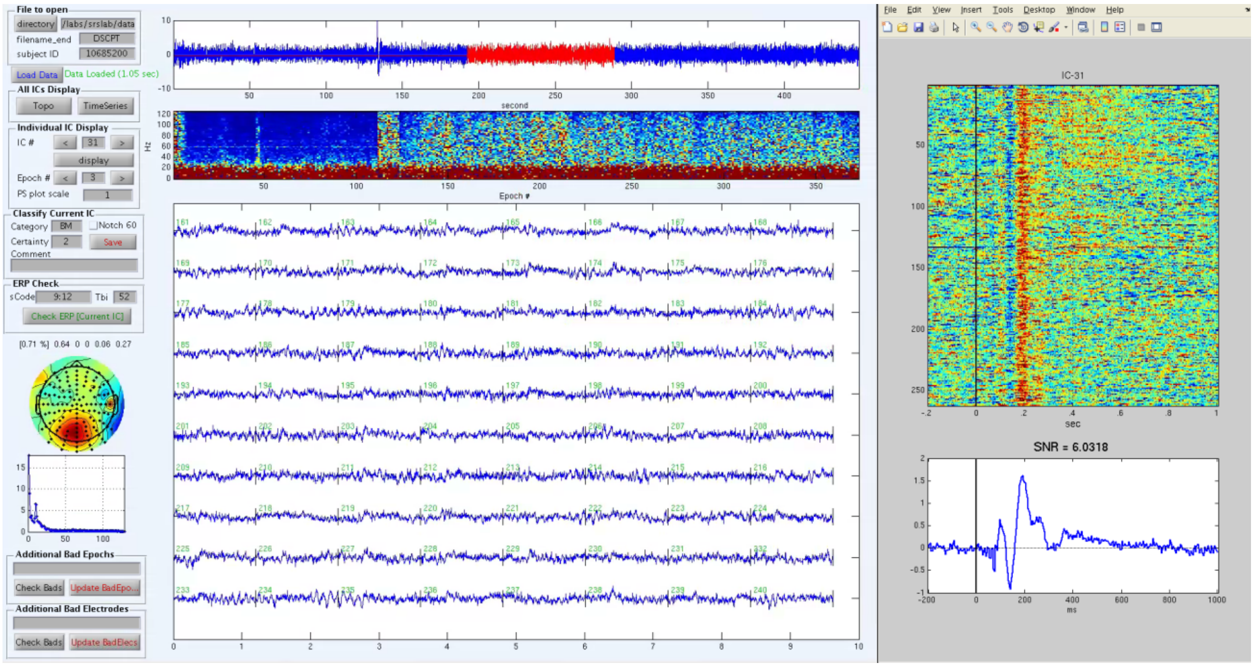
The amount of acceptable muscle contamination depends heavily on the type of analysis to be performed after preprocessing. ERP analysis is more lenient and can withstand more muscle artifact in the final dataset. In contrast, time-frequency analysis is quite sensitive to muscle contamination and it is more important to have clean neural data than to have the entirety of the available neural signals represented. CAB Lab personnel currently have several guidelines for making this distinction and place particular emphasis on using the power spectrum plot.

'BM': There is apparent brain activity and only a small amount of muscle activity (i.e., **10% or less** of the IC).

'BBM': There is brain activity in an IC, but there may be too much muscle for time-frequency analysis. "Too much muscle" would be cases where **more than 10% of all epochs** have significant activity across all frequencies. This would be evident as vertical stripes of near red color in the power spectrum plot for significant periods of the recording. It is often appropriate to apply a low pass filter to 'BBM' ICs.

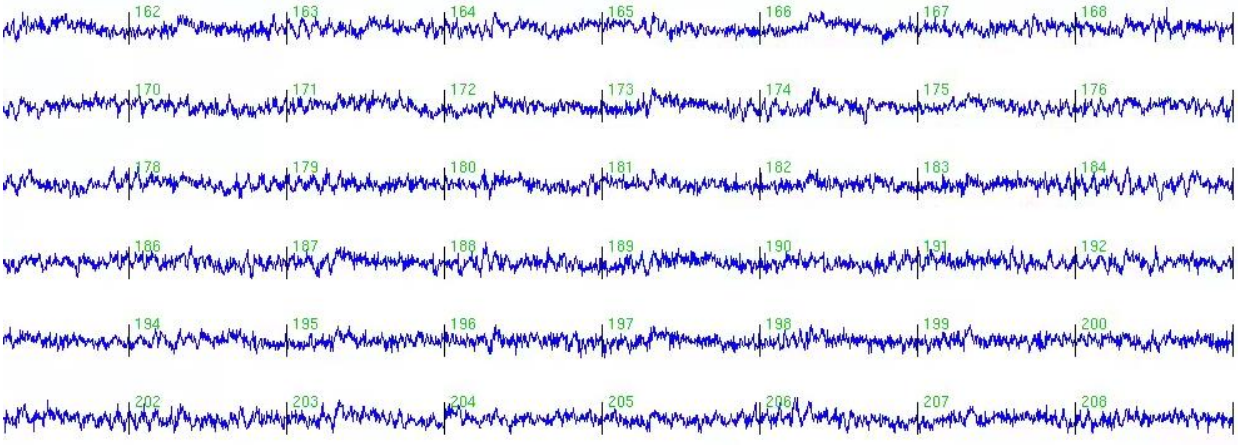
'BM' ICs

For 'BM' ICs, the topography and the power spectrum density plots should indicate the presence of brain activity. Users should look for signs such as a widespread topographical distribution, peaks in lower frequency bands, and clear ERP activity.



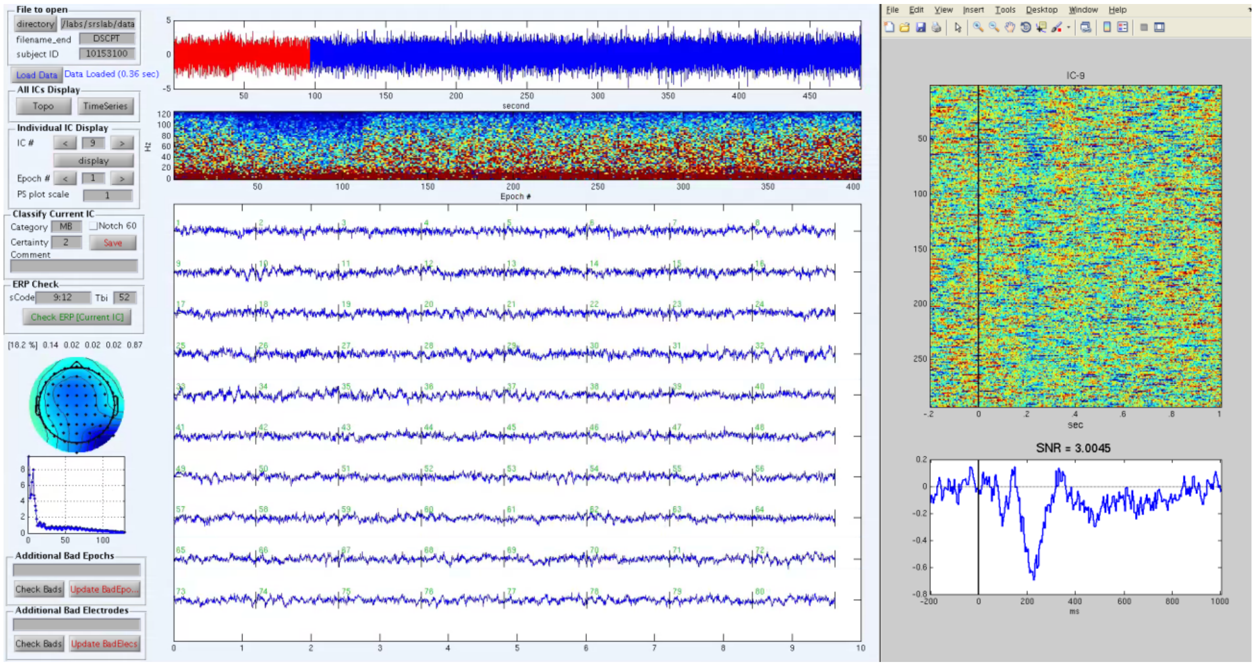
File to open
directory: /labs/srslab/data
filename_end: DSCPT
subject ID: 10685200
Data Loaded (1.05 sec)
Individual IC Display
IC #: 31
Epoch #: 3
PS plot scale: 1
Classify Current IC
Category: BM
Certainty: 2
Comment:
ERP Check
sCode: 9:12
Tbi: 52
[0.71 %] 0.64 0 0 0.06 0.27
SNR = 6.0318

The majority of the time series should support this classification as well. In the example (enlarged below), the signal is consistent with neurophysiological activity, but appears somewhat dense with jagged spikes between the wider waveforms. There is also a bad electrode in the topography which could be noted using the classification ('BMX').

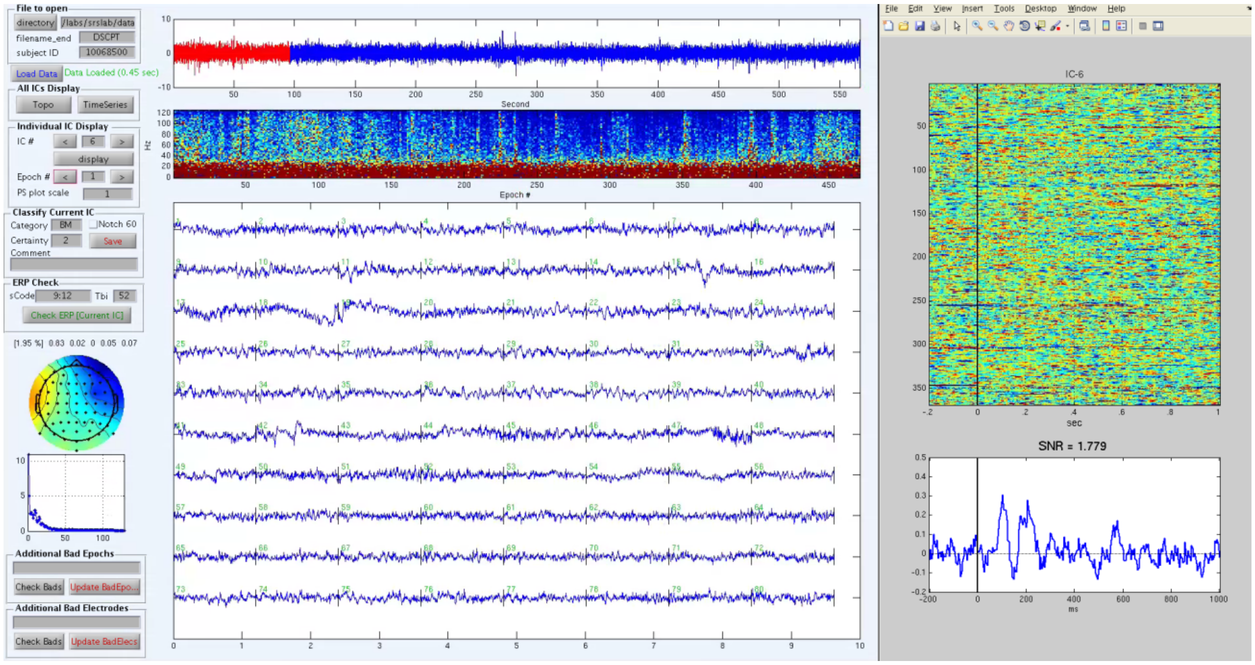


The fact that brain signal can be discerned with little effort after examination of the time series most likely means there is significant neural signal present. Several examples of 'BM' ICs are presented below.

Example 1


File to open
directory /labs/srslab/data
filename_end DSCPT
subject ID 10153100
Data Loaded (0.36 sec)
Individual IC Display
IC # 9
Epoch # 1
PS plot scale 1
Classify Current IC
Category MB
Certainty 2
☐Notch 60
Comment
ERP Check
sCode 9:12 Tbi 52
[18.2 %] 0.14 0.02 0.02 0.02 0.87
Additional Bad Epochs
Additional Bad Electrodes
IC-9
SNR = 3.0045

Example 2



File to open
directory: /labs/srslab/data
filename_end: DSCPT
subject ID: 10068500
Data Loaded (0.45 sec)

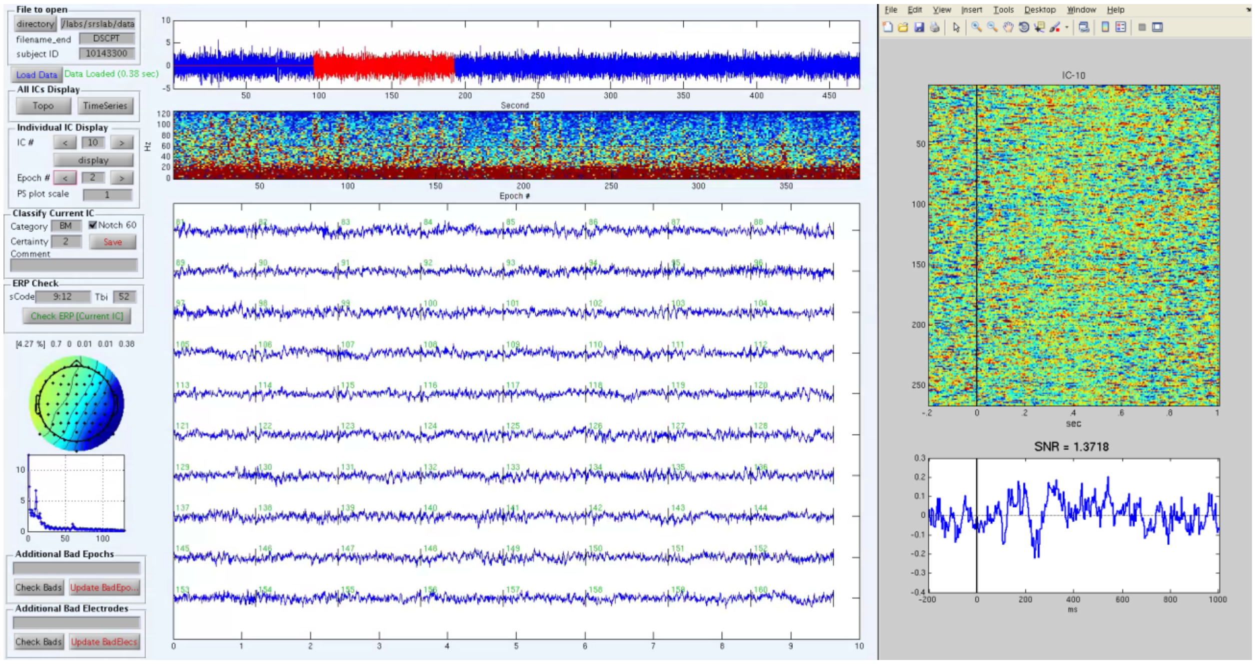
Individual IC Display
IC #: 6
Epoch #: 1
PS plot scale: 1

Classify Current IC
Category: BM
□ Notch 60
Certainty: 2
Comment:

ERP Check
sCode: 9:12
Tbi: 52

[1.95 %] 0.83 0.02 0 0.05 0.07

Example 3



File to open
directory: /labs/srslab/data
filename_end: DSCPT
subject ID: 10143300
Data Loaded (0.38 sec)

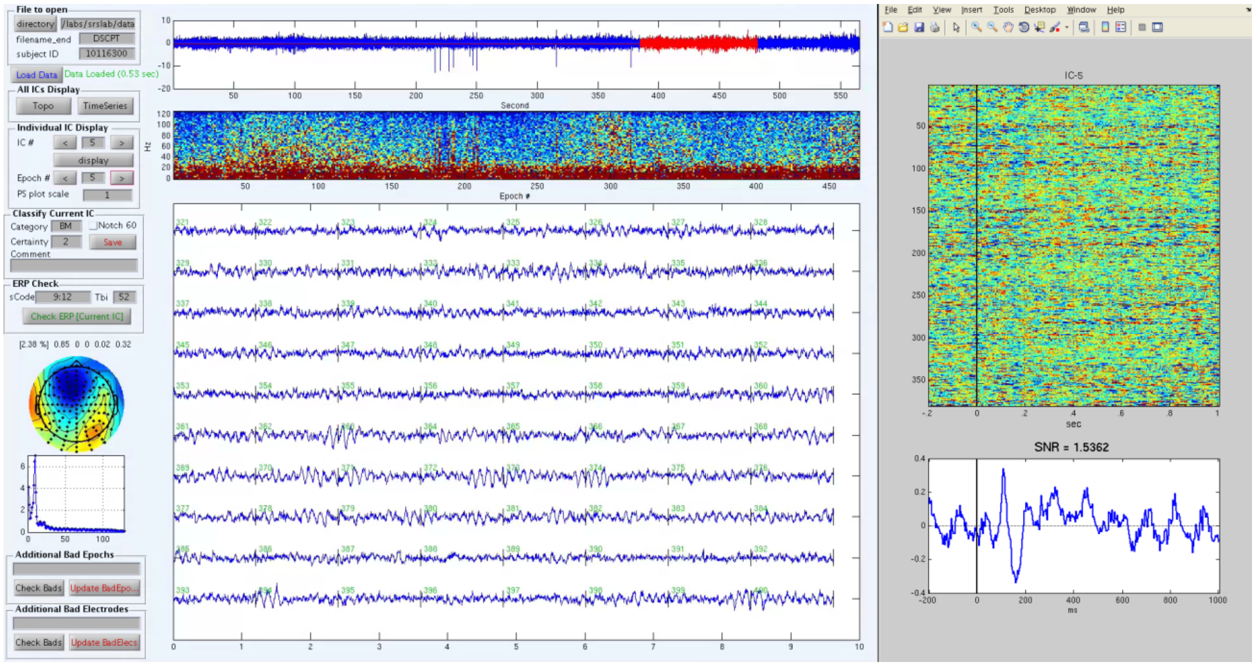
Individual IC Display
IC #: 10
Epoch #: 2
PS plot scale: 1

Classify Current IC
Category: BM
☑ Notch 60
Certainty: 2
Comment:

ERP Check
sCode: 9:12
Tbi: 52

[4.27 %] 0.7 0 0.01 0.01 0.38

Example 4



File to open
directory /labs/srslab/data
filename_end DSCPT
subject ID 10116300
Data Loaded (0.53 sec)

Individual IC Display
IC # 5
Epoch # 5
PS plot scale 1

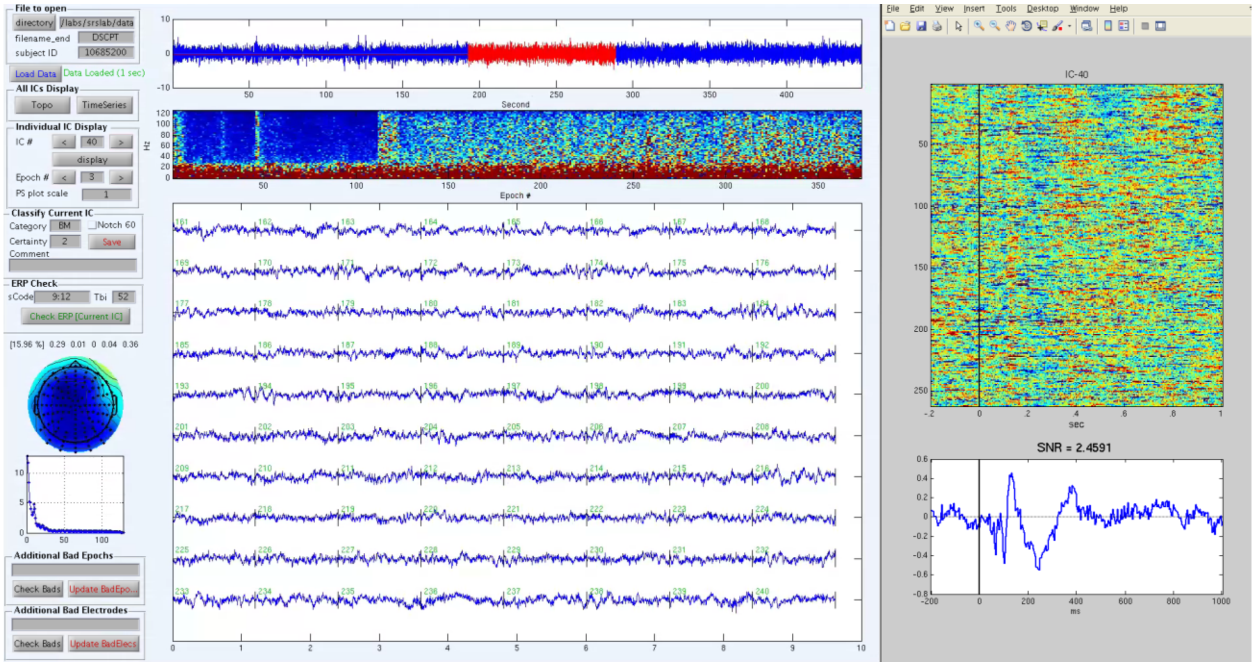
Classify Current IC
Category BM
□ Notch 60
Certainty 2
Comment

ERP Check
sCode 9:12 Tbi 52

[2.38 %] 0.85 0 0 0.02 0.32

SNR = 1.5362

Example 5



File to open
directory /labs/srslab/data
filename_end DSCPT
subject ID 10685200
Data Loaded (1 sec)

Individual IC Display
IC # 40
Epoch # 3
PS plot scale 1

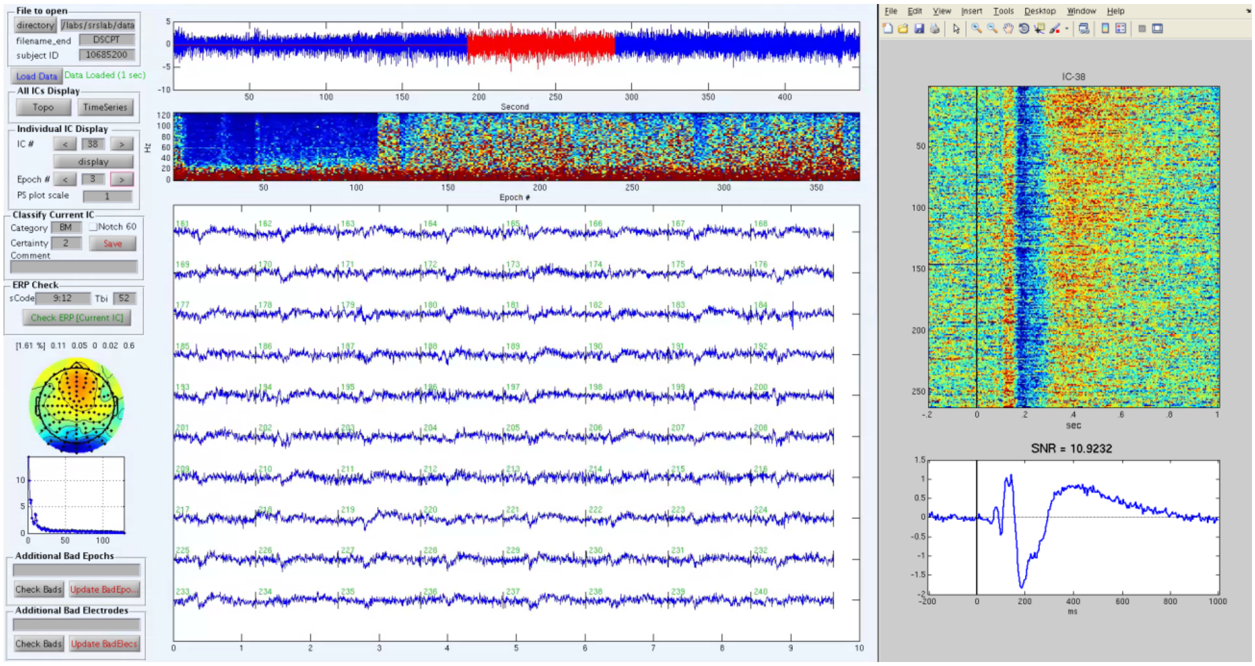
Classify Current IC
Category BM
□ Notch 60
Certainty 2
Comment

ERP Check
sCode 9:12 Tbi 52

[15.96 %] 0.29 0.01 0 0.04 0.36

SNR = 2.4591

Example 6



File to open
directory: /labs/srslab/data
filename_end: DSCPT
subject ID: 10685200
Data Loaded (1 sec)
Individual IC Display
IC #: 38
Epoch #: 3
PS plot scale: 1
Classify Current IC
Category: BM
□ Notch 60
Certainty: 2
Comment:
ERP Check
sCode: 9:12
Tbi: 52
[1.61 %] 0.11 0.05 0 0.02 0.6

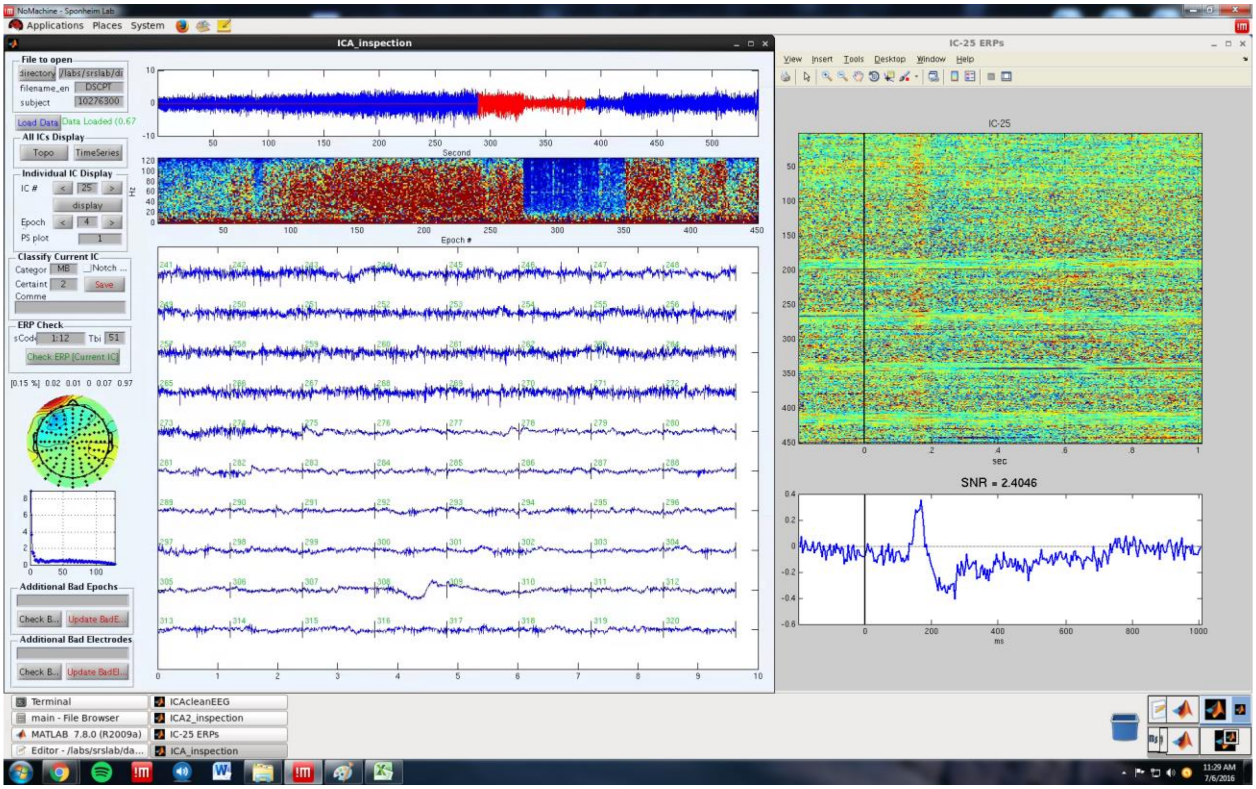
IC-38

SNR = 10.9232

‘BBM’ ICs

Below are four examples of mixed brain and muscle ICs that are too noisy for time-frequency analysis. Each would be classified as ‘BBM’.

Example 1



File to open
directory /labs/srslab/di
filename_en DSCPT
subject 10276300
Data Loaded (0.67

Individual IC Display
IC # 25
Epoch 4
PS plot 1

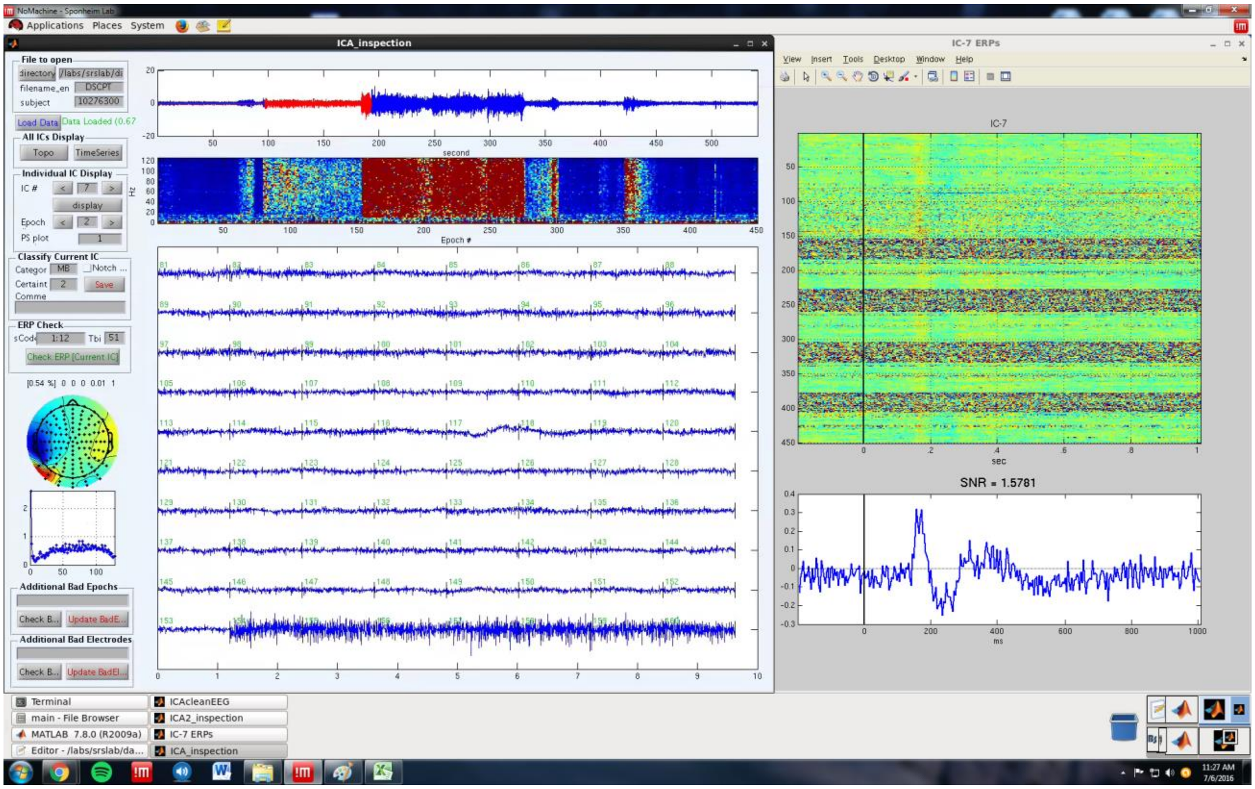
Classify Current IC
Categor MB
Certaint 2

ERP Check
sCod 1:12 Tbi 51

[0.15 %] 0.02 0.01 0 0.07 0.97

SNR = 2.4046

Example 2



File to open
directory /labs/srslab/di
filename_en DSCPT
subject 10276300
Data Loaded (0.67

Individual IC Display
IC # 7
Epoch 2
PS plot 1

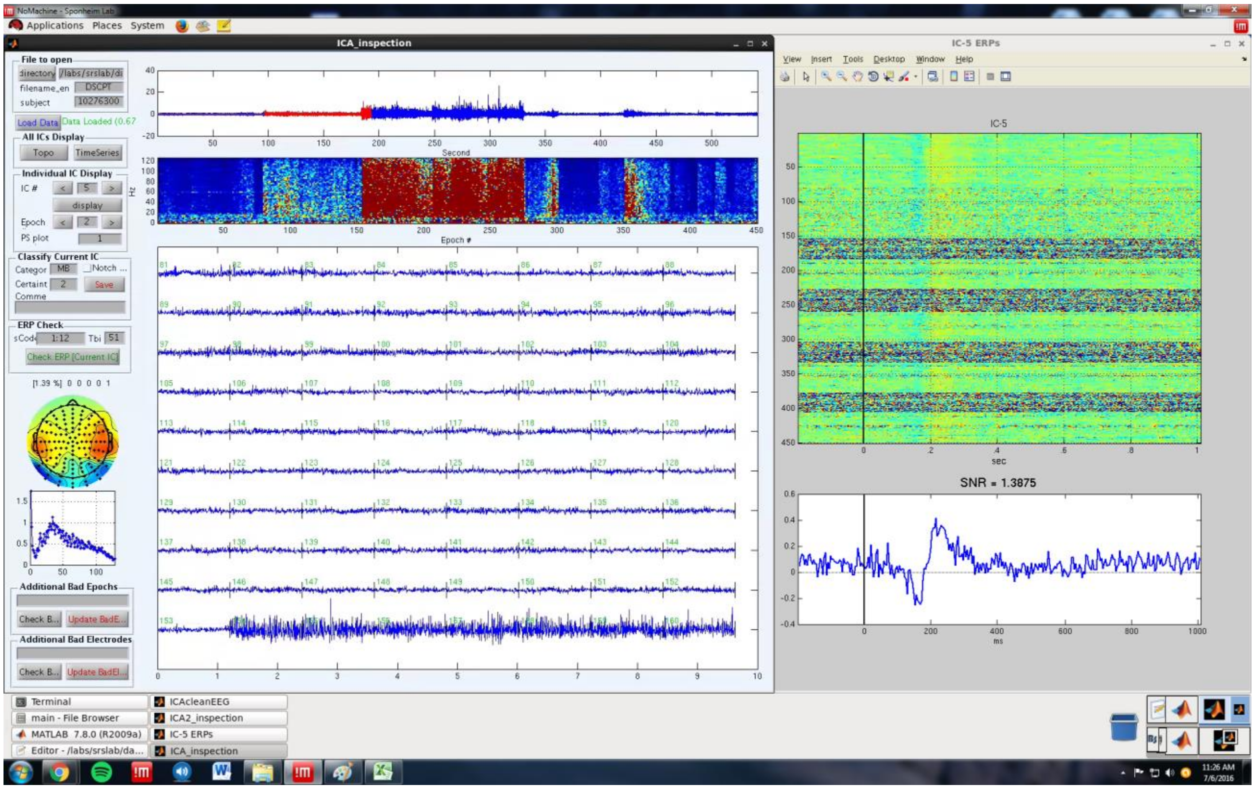
Classify Current IC
Categor MB
Certaint 2

ERP Check
sCod 1:12 Tbi 51

[0.54 %] 0 0 0 0.01 1

SNR = 1.5781

Example 3



File to open
directory /labs/srslab/di
filename_en DSCPT
subject 10276300

Data Loaded (0.67

IC # 5
Epoch 2
PS plot 1

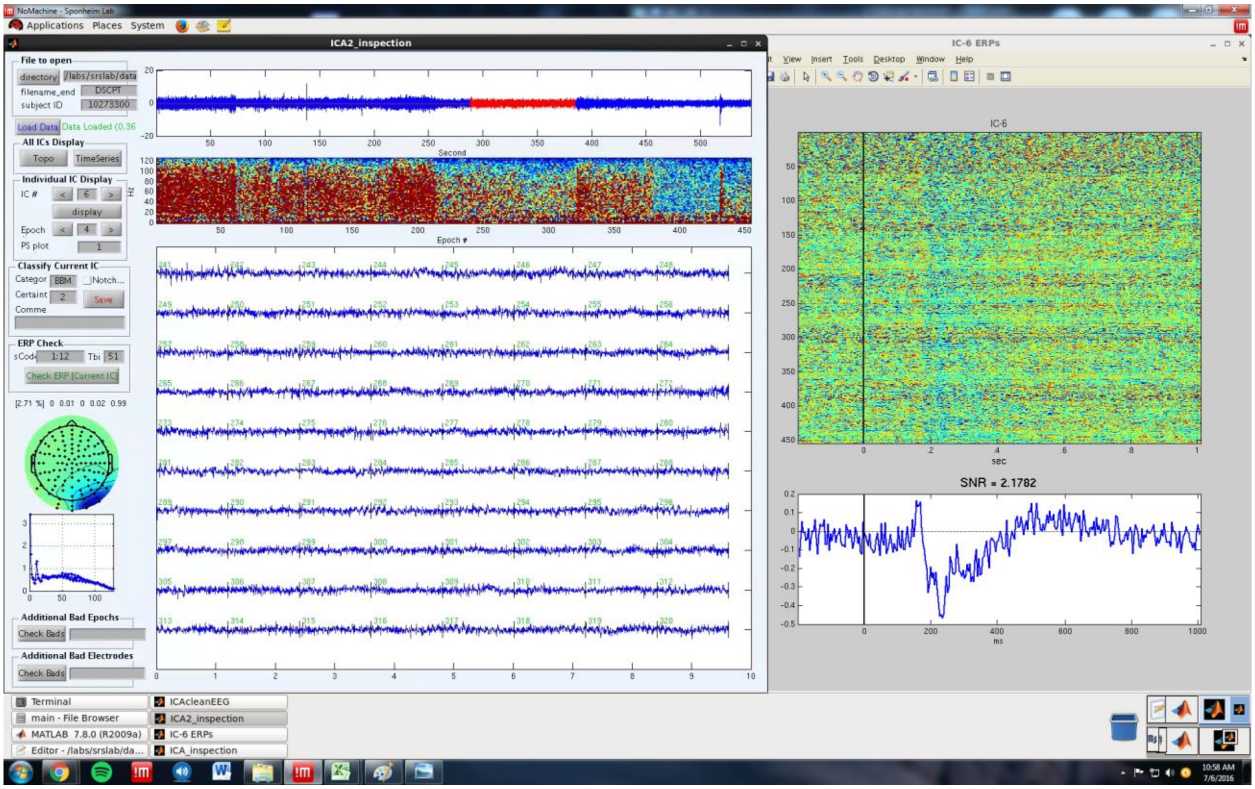
Classify Current IC
Categor MB
Certaint 2

ERP Check
sCod 1:12 Tbi 51

[1.39 %] 0 0 0 0 1

SNR = 1.3875

Example 4



File to open
directory /labs/srslab/data
filename_end DSCPT
subject ID 10273300

Data Loaded (0.36

IC # 6
Epoch 4
PS plot 1

Classify Current IC
Categor BBM
Certaint 2

ERP Check
sCod 1:12 Tbi 51

[2.71 %] 0 0.01 0 0.02 0.99

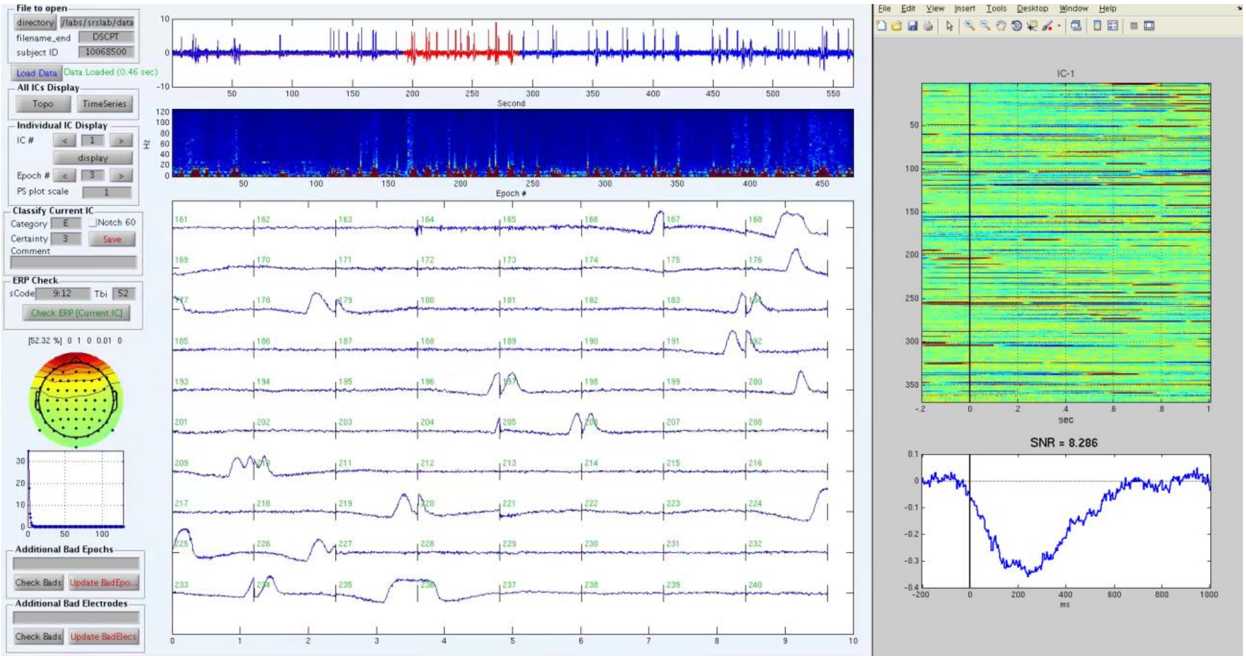
SNR = 2.1782

Each of these four ICs have prolonged periods of high frequency muscle activity, but also contain evidence of neural activity within the ‘ERP Check’ displays.

Eye movement ICs

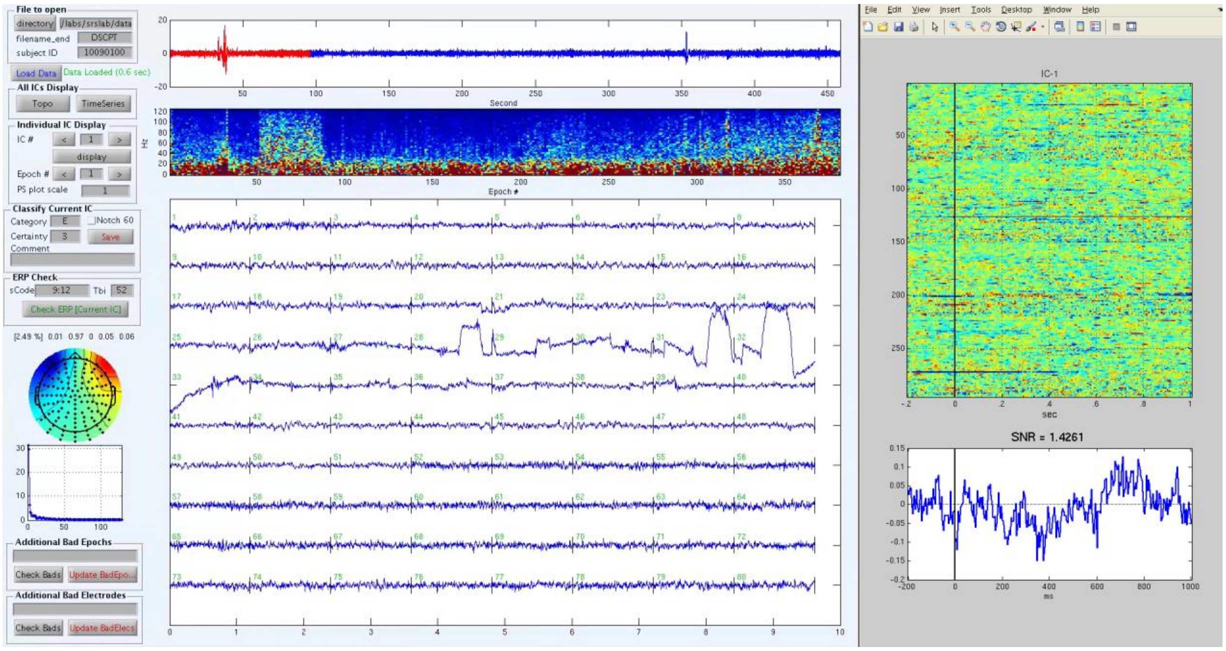
The topographies of eye movements should show strong frontal activation. This should be combined with low frequency power clearly depicted in the power spectrum density plot. Within the time series, the triangular deflections of VEOG or “box car” deflections of HEOG should be visible. When the time series is ambiguous, refer to the topography for clarification about whether the eye movement deflections originate from VEOG or HEOG.

Example 1



directory /labs/srslab/data
filename_end DSCPT
subject ID 10068500
Data Loaded (0.46 sec)
IC # 1
Epoch # 3
PS plot scale 1
Category E
Certainty 3
sCode 9:12 Tbi 52
[52.32 %] 0 1 0 0.01 0
SNR = 8.286

Example 2



File to open
directory /labs/srslab/data
filename_end DSCPT
subject ID 10090100
Data Loaded (0.6 sec)

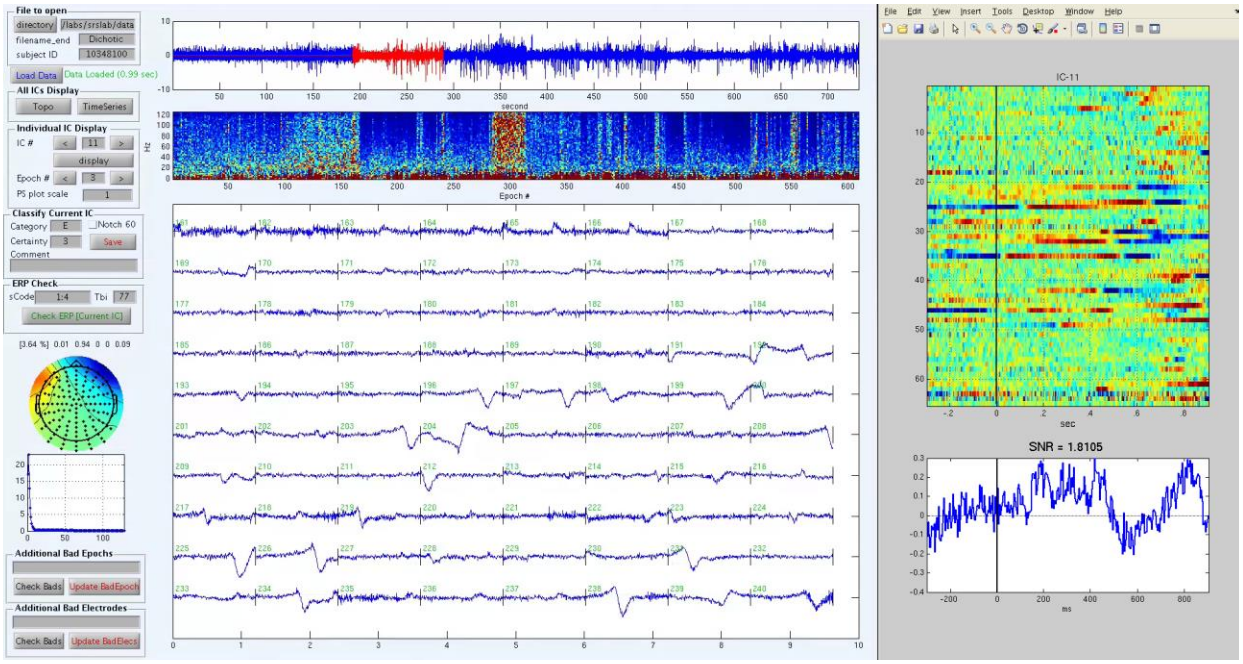
Individual IC Display
IC # 1
Epoch # 1
PS plot scale 1

Classify Current IC
Category E
□ Notch 60
Certainty 3
Comment

ERP Check
sCode 9:12 Tbi 52

[2.49 %] 0.01 0.97 0 0.05 0.06

Example 3



File to open
directory /labs/srslab/data
filename_end Dichotic
subject ID 10348100
Data Loaded (0.99 sec)

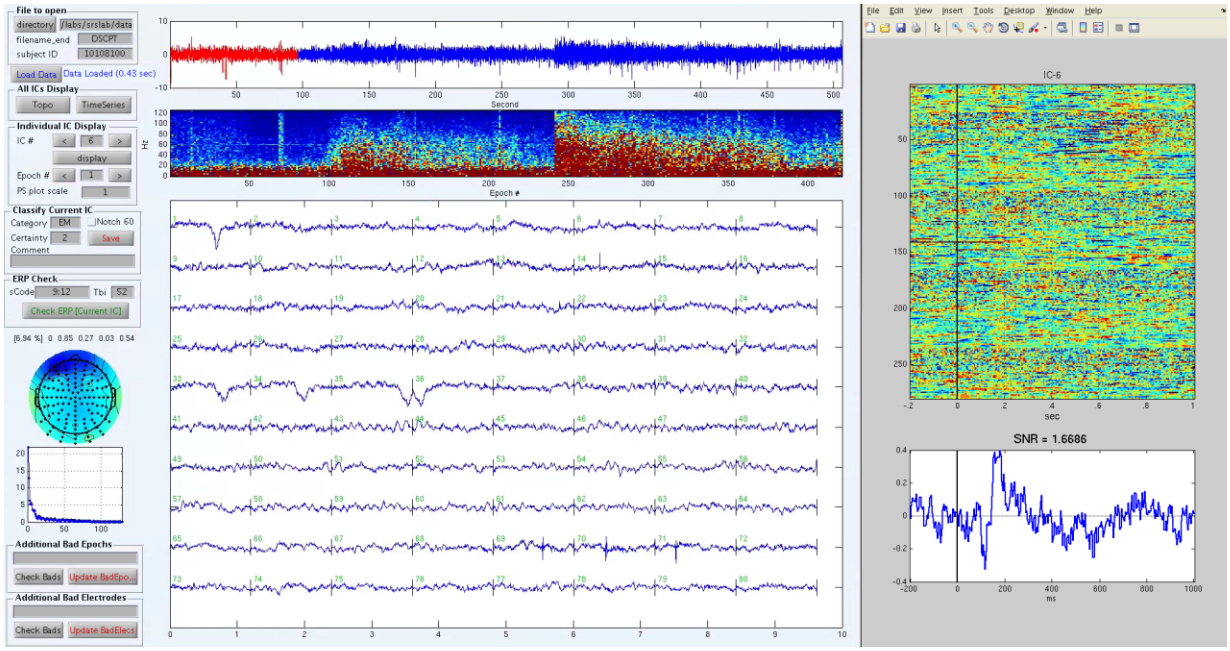
Individual IC Display
IC # 11
Epoch # 3
PS plot scale 1

Classify Current IC
Category E
□ Notch 60
Certainty 3
Comment

ERP Check
sCode 1:4 Tbi 77

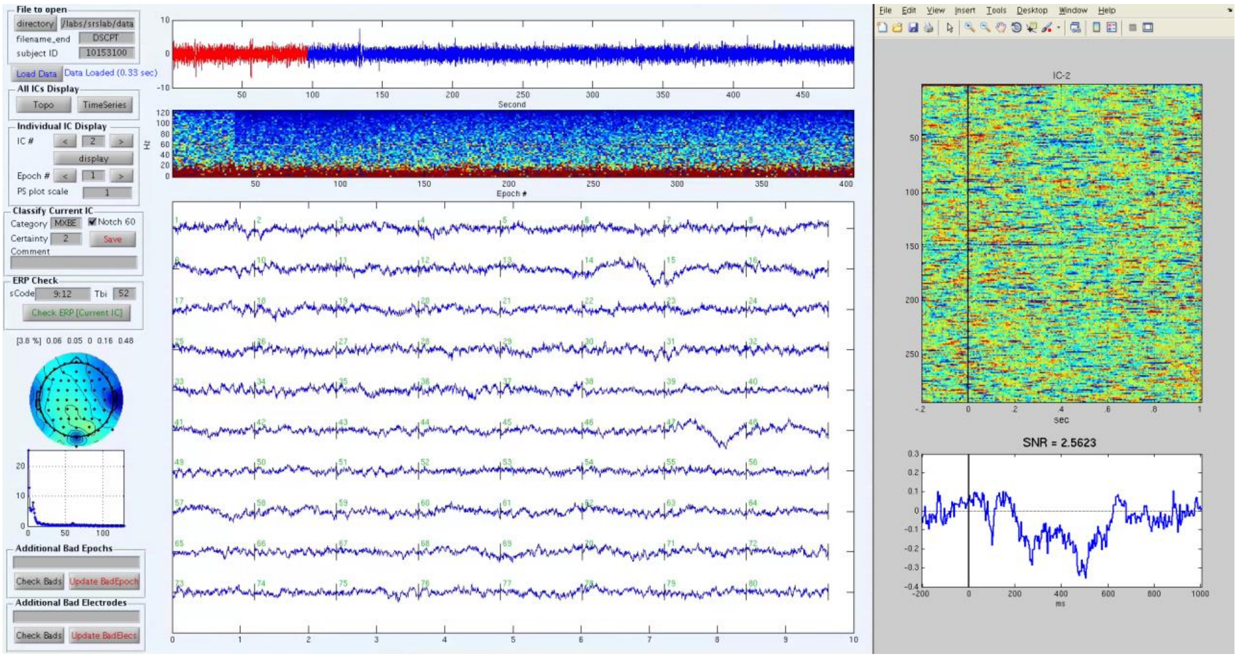
[3.64 %] 0.01 0.94 0 0 0.09

Example 4



File to open
directory /labs/srslab/data
filename_end DSCPT
subject ID 10108100
Data Loaded (0.43 sec)
Individual IC Display
IC # 6
Epoch # 1
PS plot scale 1
Classify Current IC
Category EM □Notch 60
Certainty 2
Comment
ERP Check
sCode 9:12 Tbi 52
[6.94 %] 0 0.85 0.27 0.03 0.54

Example 5



File to open
directory /labs/srslab/data
filename_end DSCPT
subject ID 10153100
Data Loaded (0.33 sec)
Individual IC Display
IC # 2
Epoch # 1
PS plot scale 1
Classify Current IC
Category MXBE ☑Notch 60
Certainty 2
Comment
ERP Check
sCode 9:12 Tbi 52
[3.8 %] 0.06 0.05 0 0.16 0.48

Commentary

Given the strength of eye movement signal, it is often simple for ICA to identify the artifacts and isolate them as separate ICs. Example 1 depicts an exemplar VEOG IC with clear triangular deflections in the time series, low frequency power activity, and a frontally dominated topographical distribution. HEOG can be more challenging to identify in IC time series in part because it generally appears less frequently in typical EEG tasks. For Example 2, the frontal dipole topography is helpful for confirming the identity of HEOG signal within this IC.

Example 3 demonstrates an IC with both VEOG and HEOG present simultaneously. For the purpose of ICA preprocessing, the distinction between the two is unnecessary. Example 3 can be discarded from the post-processed data. Accurate categorization of mixed signal sources becomes more critical when brain may be present. Example 4 depicts an IC with VEOG activity co-occurring with higher frequency muscle activity and some possible brain activity as seen in the ‘ERP Check’ display. Based on the relative dominance of the eye signal, Example 4 would be classified as ‘EMB’ and subsequently discarded.

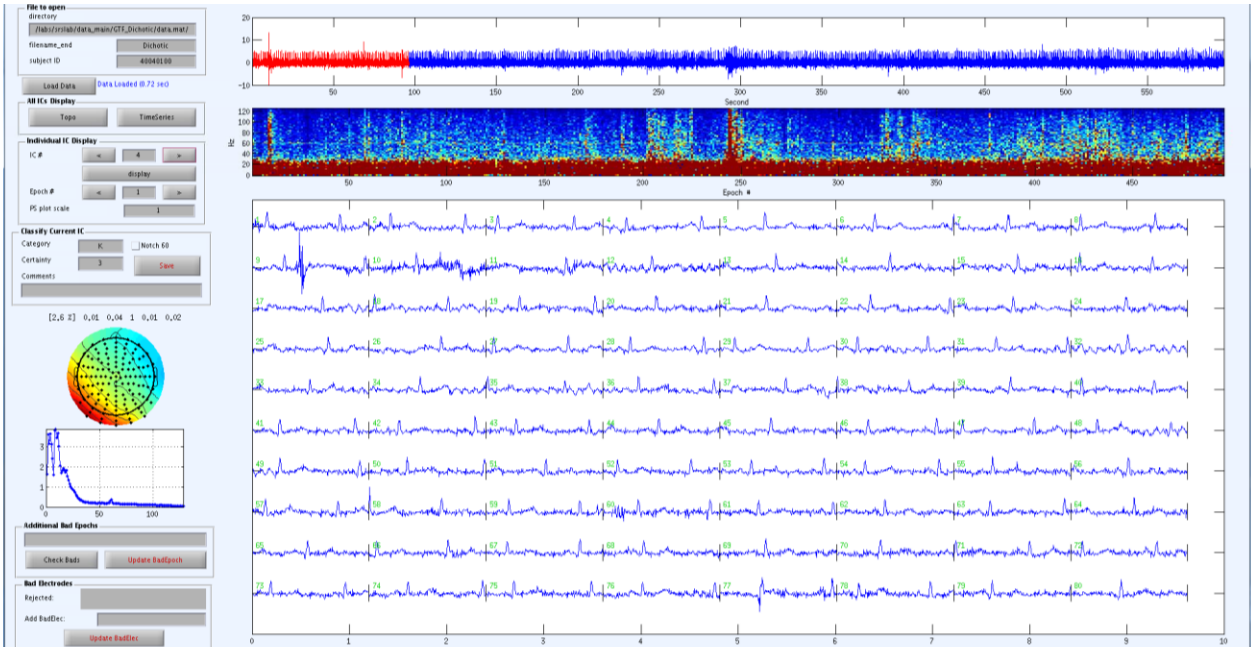
Although rare, users may come across ICs where eye signal is present and also secondary to other sources. In such cases, it is useful to note the type of eye signal observed in the notes. Example 5 depicts a possible ‘MBE’ or ‘MBXE’ IC where muscle, brain, and bad electrodes are contaminated with VEOG signal (i.e., an overall very messy IC). Regardless, this would be discarded.

ECG ICs

Make sure it is not brain and vice versa

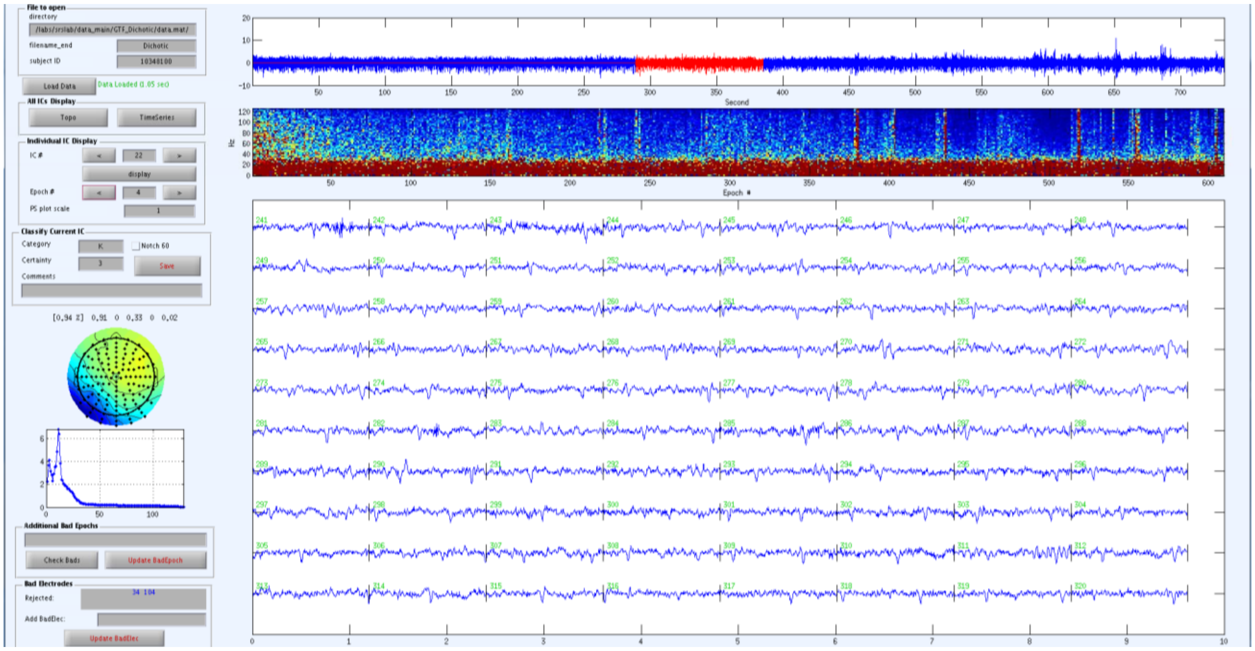
‘K’ (i.e., ECG or “EKG”) ICs tend to be recognizable during visual inspection of the time series and topographies. That being said, it is very difficult to mathematically differentiate ‘K’ from ‘B’ using ICA. Both are characterized by low frequency signal, exhibit wide topographical distributions, and contain regular bursts of activity. Given this difficulty, it is very important to inspect ICs when ‘K’ or ‘B’ is identified to double check the ‘ICA_inspection’ algorithm. ‘K’ ICs have rhythmic, but not time-locked, moderately sharp peaks in the time series. The topography is always diagonal from the back left to the front right of the scalp. The power spectrum density usually has several peaks in the low frequency range.

Example 1



File to open
directory: /labs/srslab/data_main/GTF_Dichotic/data.mat/
filename_end: Dichotic
subject ID: 40040100
Data Loaded (0.72 sec)
IC #: 4
Epoch #: 1
PS plot scale: 1
Category: K
Certainty: 3
[2.6 %] 0.01 0.04 1 0.01 0.02

Example 2



File to open
directory: /labs/srslab/data_main/GTF_Dichotic/data.mat/
filename_end: Dichotic
subject ID: 10348100
Data Loaded (1.05 sec)
IC #: 22
Epoch #: 4
PS plot scale: 1
Category: K
Certainty: 3
[0.94 %] 0.91 0 0.33 0 0.02
Rejected: 34 104

Commentary

'K' time series should be identified based on the presence of regularly appearing, relatively sharp peaks. Although these peaks occur recurrently throughout the recording, they are not time-locked to the onset on the task stimuli. This is often key for differentiating brain from ECG activity. Also, look for a diagonal pattern in the topography that does not show a full color spectrum (e.g., red to blue or blue to yellow). ‘BBK’ is often an appropriate designation, but do not employ ‘BK’. This will ensure these ICs are not included in time-frequency analysis. Use the ‘BBK’ classification liberally for similar cases unless ‘KB’ is more appropriate.

Bad Electrodes

Bad electrode ICs (‘X’) are often identifiable by an erratic time series. There may be occasional strong deflections, or consistently small and jagged shapes throughout. “Wayward noise” has also been occasionally observed in CAB Lab recordings induced by the active shielding features of some EEG system amplifiers. This may appear as noise varying in frequency over the course of a recording. Other times, broken electrodes or inconsistent electrode signal quality at the scalp can be observed the data files.

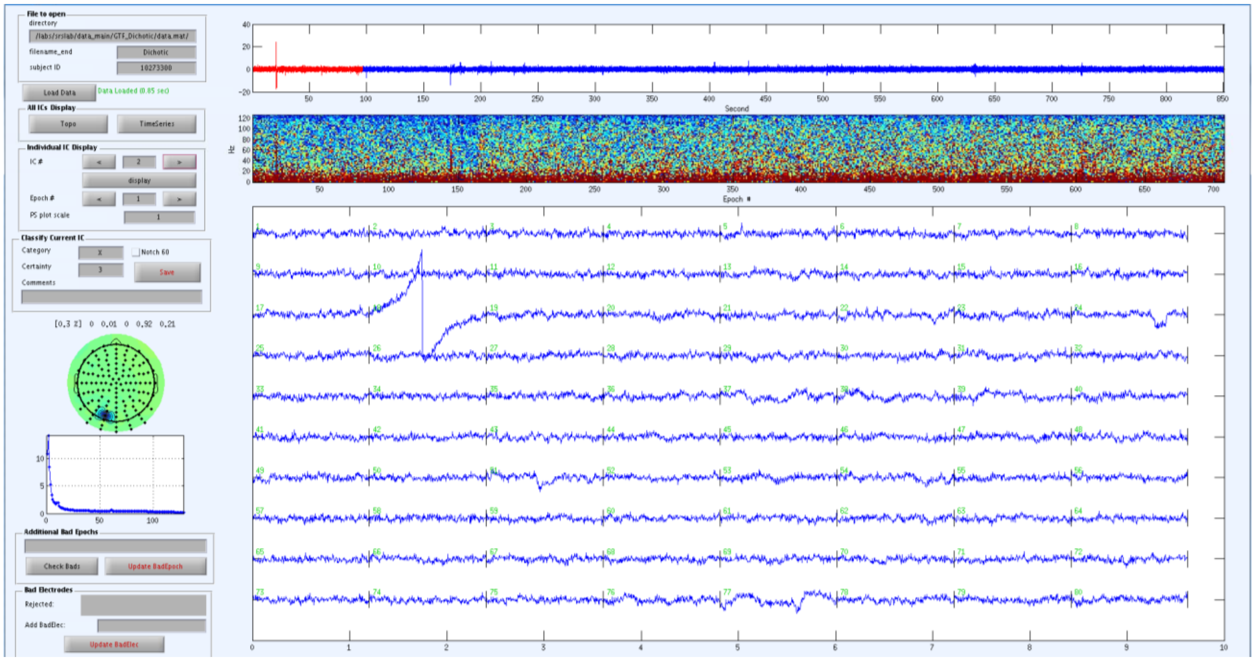
The best way to identify a bad electrode is to check the power spectrum for signs of wayward noise or the topographies for clear electrode “islands”. Deletion of an apparent bad electrode at the ‘ICA_inspection’ step is appropriate when a bad electrode is visible in more than one brain dominant IC or if the frequency of a “wayward noise” containing electrode partially overlaps with brain activity in a brain dominant IC. After deleting an electrode, conduct ‘ICA2_inspection’ to ensure that bad electrode interference is no longer present.

BX: Used for a brain dominant IC and an electrode “island” in the topography that does not show up in any other brain-dominant ICs.

BBX: This is appropriate for ICs where there is apparent brain activity and evidence of “wayward noise” *NOT* associated with a clear electrode “island” in the topography. Comment **“lowpassX, wayward noise”** to identify this phenomena. In this case, it may be helpful to plot the A matrix to determine which electrode(s) may be causing the wayward noise. First delete the electrodes with the strongest wayward noise, then low pass the IC as low as possible, getting rid of as much wayward noise without low passing below 30 Hz. If the topography shows activity exclusively on one or a few electrodes, but the ERP shows brain, review other ‘B’ ICs carefully to determine if these electrodes are bad in other ‘B’ ICs. If so, examine in ‘BadElec_inspection’ and consider removing them, or exclude in ‘ICA_inspection’ and move to ICA2. If these electrodes do not show up as “islands” in the other ‘B’ ICs, classify as ‘BBX’ and comment **“bad electrode hiding brain”**. ‘BBX’ can also be used for a mixed B and X IC in which the time-series is too messy for time-frequency analysis.

XB: Use this classification if there is no prominent ERP, but there is reason to believe there is some non-time locked brain activity.

Example



File to open
directory: /labs/srslab/data_main/GTF_Dichotic/data.mat/
filename_end: Dichotic
subject ID: 10273300
Data Loaded (0.85 sec)
IC #: 2
Epoch #: 1
PS plot scale: 1
Category: X
Certainty: 3
[0.3 Z] 0 0.01 0 0.92 0.21

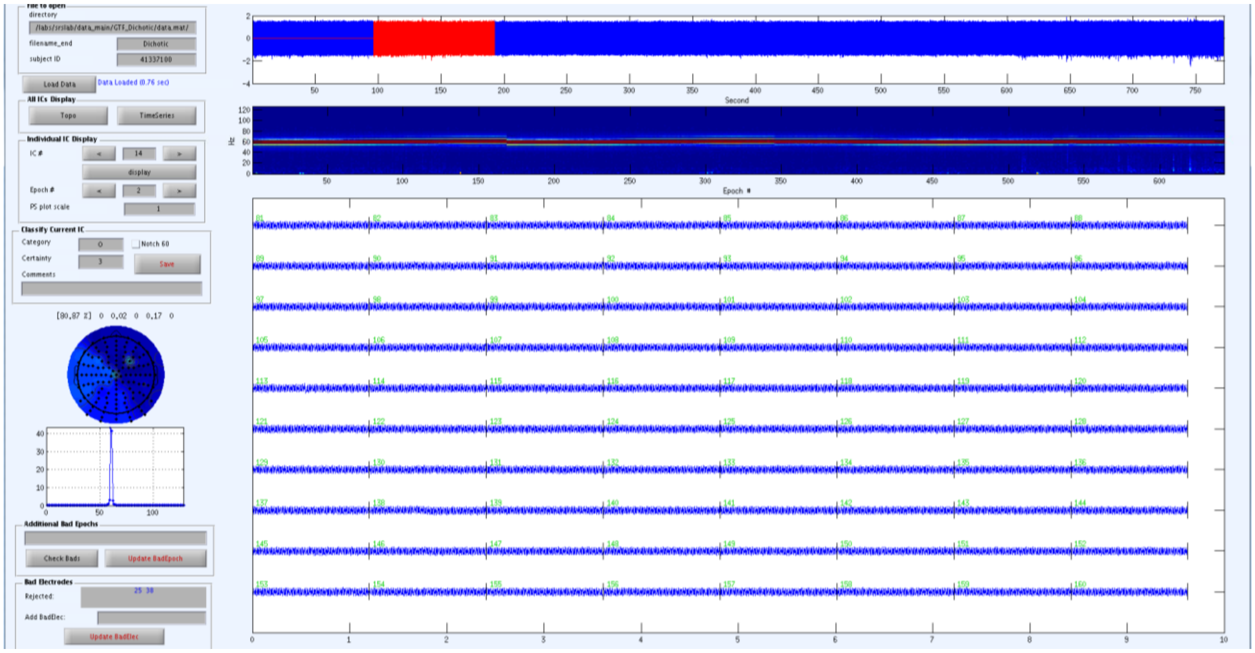
Commentary

The power spectrum and power spectrum densities can vary widely across different ‘X’ ICs. The topography will usually show signal centered on one or two electrodes. Since neurophysiological signal is unlikely to occur at only a single electrode site, this is a sign you may be looking at a bad electrode.

60 Hz noise

60 Hz noise can appear either as its own IC or as contamination in other ICs. Like ‘X’ ICs, 60 Hz ICs have topographies that show activations that do not have plausible neurophysiological distributions. These types of ICs would technically be labeled ‘O’ for other noise if they could not be successfully filtered.

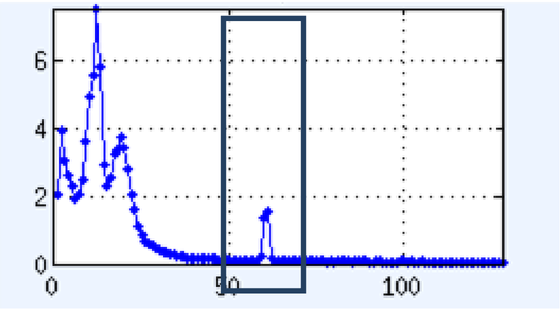
Example

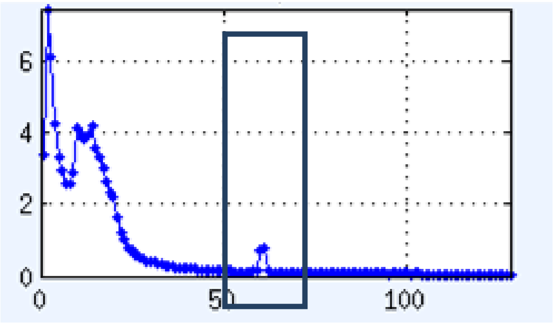


File to open
directory /labs/srslab/data_main/GTF_Dichotic/data.mat/
filename_end Dichotic
subject ID 41337100
Data Loaded (0.76 sec)
IC # 14
Epoch # 2
PS plot scale 1
Category 0
Certainty 3
□ Notch 60
Comments
[80.87 %] 0 0.02 0 0.17 0
Bad Electrodes
Rejected: 25 38
Add BadElec:

The odd topographical distribution, the band of 60 Hz signal in the power spectrum density plot, the power spectrum plot, and the compact signal in the time series all point to 60 Hz noise. Users would mark the IC as ‘O’ and note **“60 Hz noise”** in the comments.

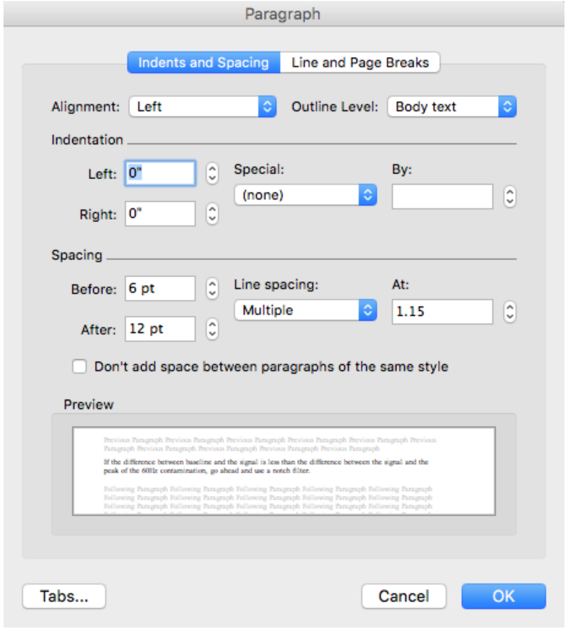
Only mark 60 Hz noise and apply a ‘Notch60’ filter to an IC if it is brain or brain-dominant. If 60 Hz noise is contaminating a brain-dominant IC, we do not necessarily filter for 60 Hz noise. The filter will remove all the signal from 58 to 62 Hz in the IC. If the 60 Hz noise is not strong, filtering will remove more signal than noise. To determine whether or not we should filter, look at the power spectrum density plot. Below are two example ICs for which it would be appropriate to apply a 60 Hz notch filter.




6
4
2
0
0
50
100

Assorted Notes

Formatting of this document



Alignment: Left

Outline Level: Body text

Indentation

Left: 0"

Right: 0"

Special: (none)

By:

Spacing

Before: 6 pt

After: 12 pt

Line spacing: Multiple

At: 1.15

□ Don't add space between paragraphs of the same style

Preview

Previous Paragraph Previous Paragraph Previous Paragraph Previous Paragraph Previous Paragraph Previous Paragraph Previous Paragraph Previous Paragraph Previous Paragraph Previous Paragraph

If the difference between baseline and the signal is less than the difference between the signal and the peak of the 60Hz contamination, go ahead and use a notch filter.

Following Paragraph Following Paragraph Following Paragraph Following Paragraph Following Paragraph Following Paragraph Following Paragraph Following Paragraph Following Paragraph Following Paragraph Following Paragraph Following Paragraph Following Paragraph Following Paragraph Following Paragraph