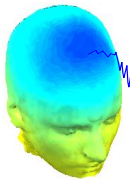


EEG Preprocessing

Importing, rejecting, and preparing data for ICA decomposition

EEGLAB Workshop XXIV

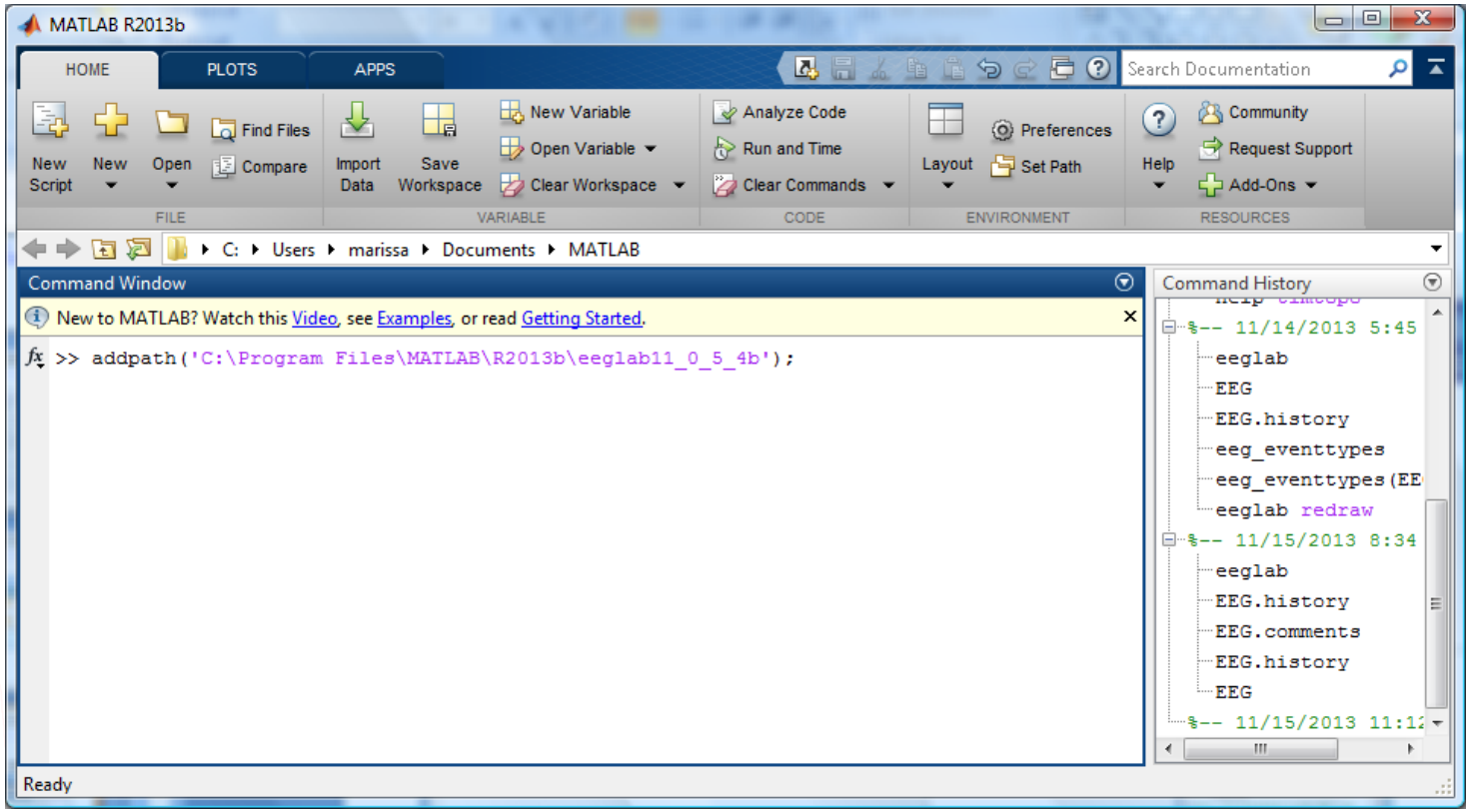
Aspet, France

July 4, 2017

Day 1, 11:30-12:30

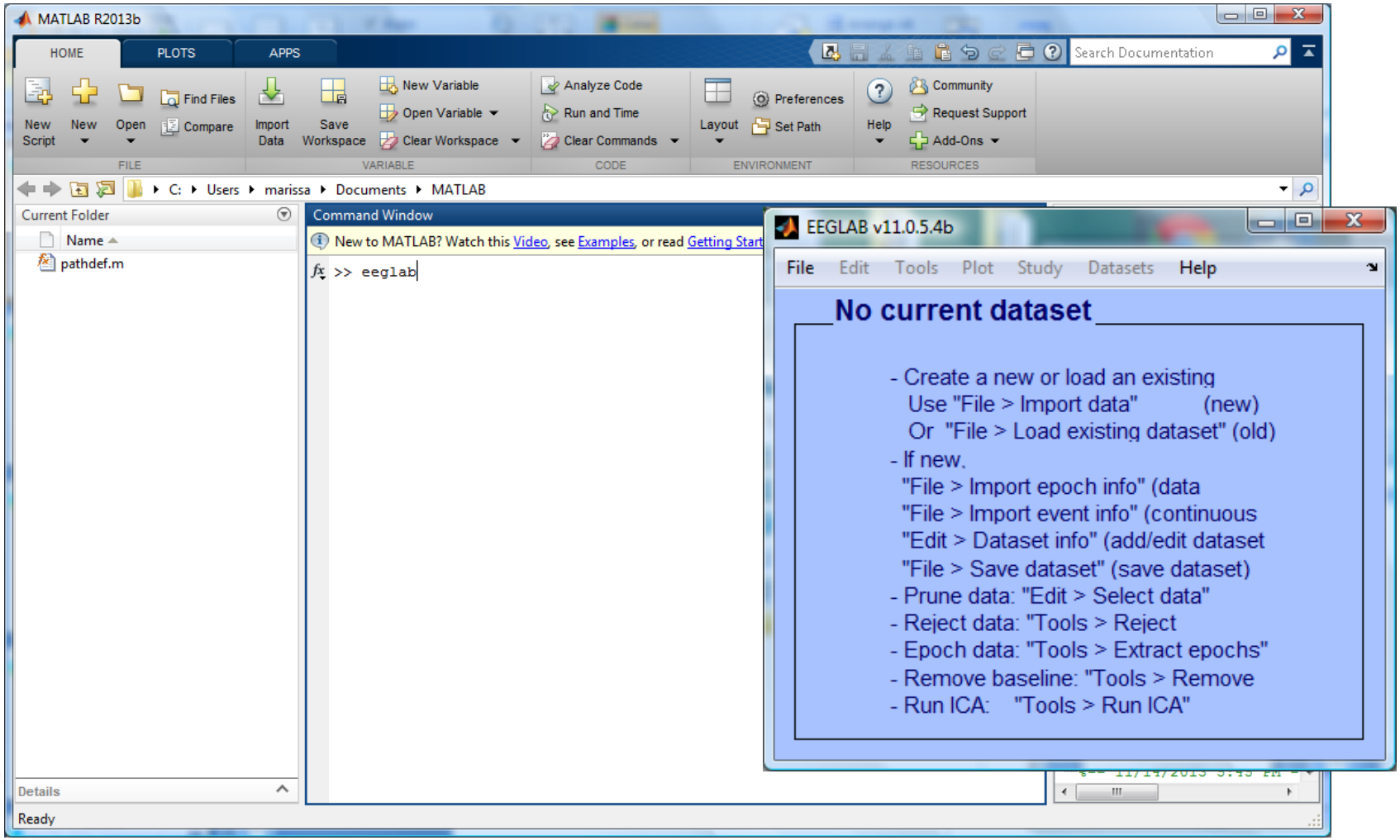
Installing EEGLAB and data folder

- Start Matlab
- Add the EEGLAB folder to your Matlab path:

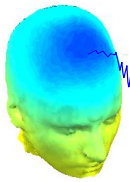


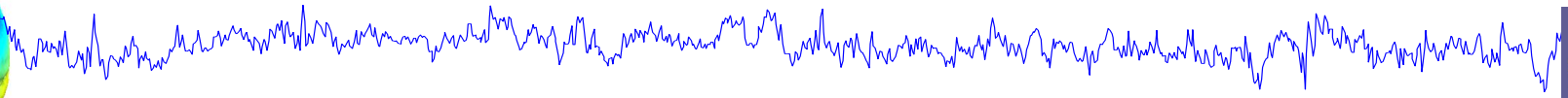
The EEGLAB Matlab software

main graphic interface



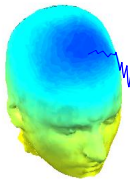
“Secrets” to a good ICA decomposition

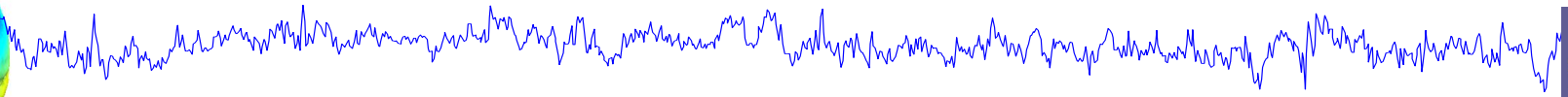




- Garbage in… garbage out (it’s not magic)
- Remove large, non-stereotyped artifacts
- Do you have enough data? (based mostly on time, not frames)
- High-pass filter to remove slow drifts (no low-pass filter needed)
- Remove bad channels

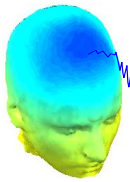
The Goal of Preprocessing

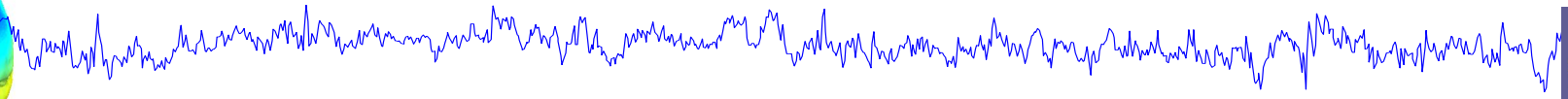




- Create a complete EEGLAB data set with
  - EEG
  - Channel Locations
  - Events
- Preprocess the EEG data to yield optimal ICA decompositions
  - Re-referencing
  - High Pass Filtering
  - Remove bad channels, noisy segments of data
- Run ICA decomposition

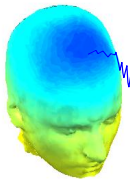
Many Preprocessing Variants

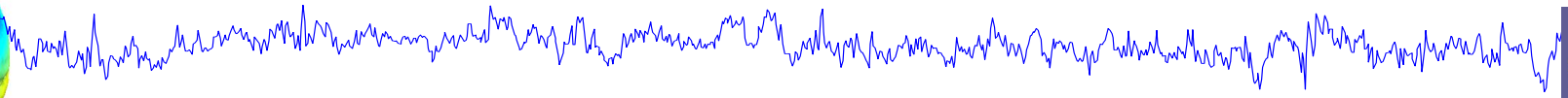




- Resources
  - EEGLAB wiki “Quick Tutorial on Rejection”
    - http://sccn.ucsd.edu/wiki/Quick_Rejection_Tutorial
  - Makoto’s Preprocessing Pipeline
    - http://sccn.ucsd.edu/wiki/Makoto%27s_preprocessing_pipeline
  - Bigdely-Shamlo et al (2015): PREP Pipeline
    - http://dx.doi.org/10.3389/fninf.2015.00016
  - EEGLAB Plugins
    - Kothe’s clean_rawdata plugin
    - Miyakoshi’s trimOutlier plugin

Pre-processing pipeline





Swartz Center for Computational Neuroscience

Collect high-density EEG data (>30 chan)

Import into EEGLAB

Import event markers and channel locations

Re-reference/ down-sample (if necessary)

High pass filter (~.5 – 1 Hz)

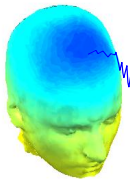
Examine raw data

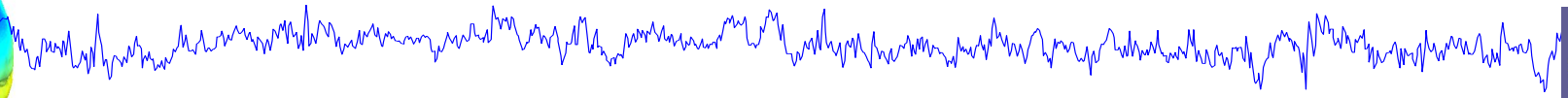
Data Cleaning: Identify/reject bad channels

Data Cleaning: Reject large artifact time points

Run ICA

Pre-processing pipeline





Swartz Center for Computational Neuroscience

Collect high-density EEG data (>30 chan)

Import into EEGLAB

Import event markers and channel locations

Re-reference/ down-sample (if necessary)

High pass filter (~.5 – 1 Hz)

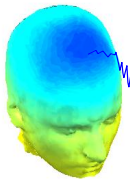
Examine raw data

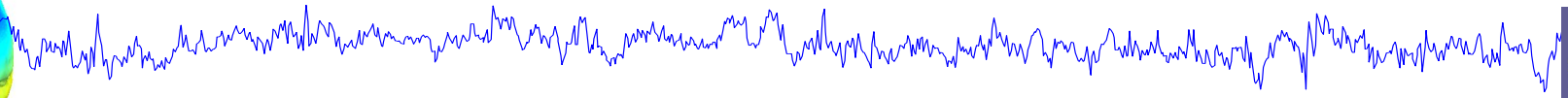
Data Cleaning: Identify/reject bad channels

Data Cleaning: Reject large artifact time points

Run ICA

Pre-processing pipeline





Swartz Center for Computational Neuroscience

Collect high-density EEG data (>30 chan)

Import into EEGLAB

Import event markers and channel locations

Re-reference/ down-sample (if necessary)

High pass filter (~.5 – 1 Hz)

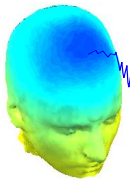
Examine raw data

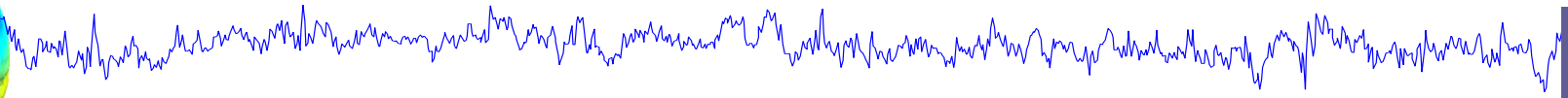
Data Cleaning: Identify/reject bad channels

Data Cleaning: Reject large artifact time points

Run ICA

Sample data: basic P300 paradigm





File

SimpleOddball.set

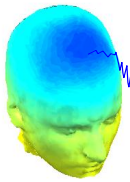
Data

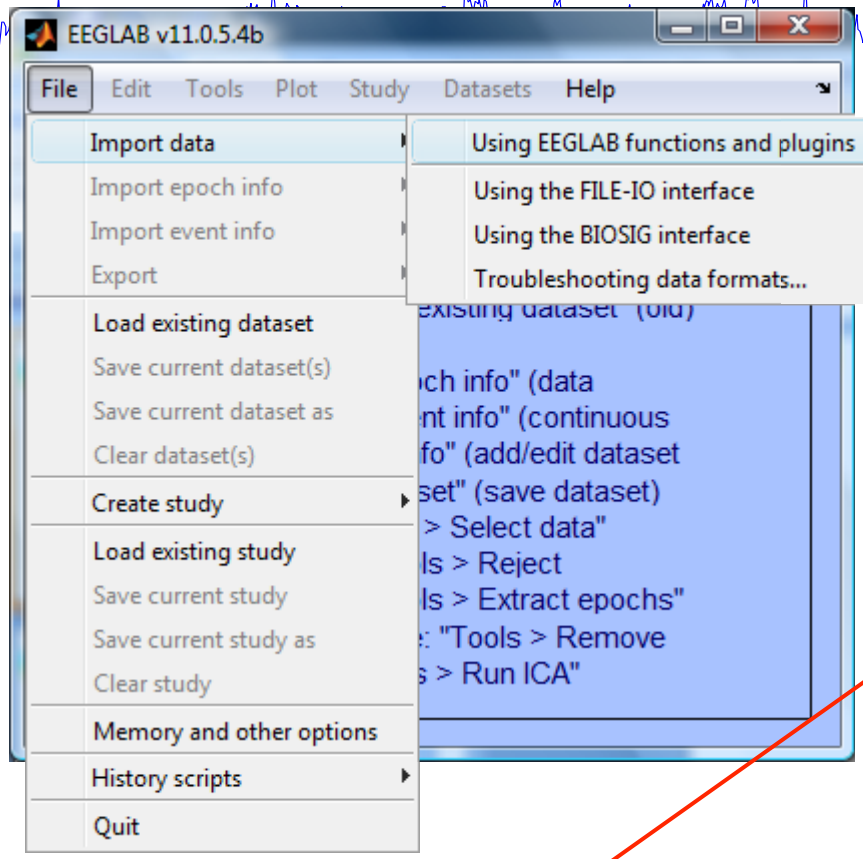
64 channel EEG, 256 Hz sampling rate, Biosemi system, re-referenced during import to averaged left and right mastoid electrodes

Task

speeded button press response to star shape (no response to circle shape), 100 ms presentation duration, 200 trials

Importing a dataset



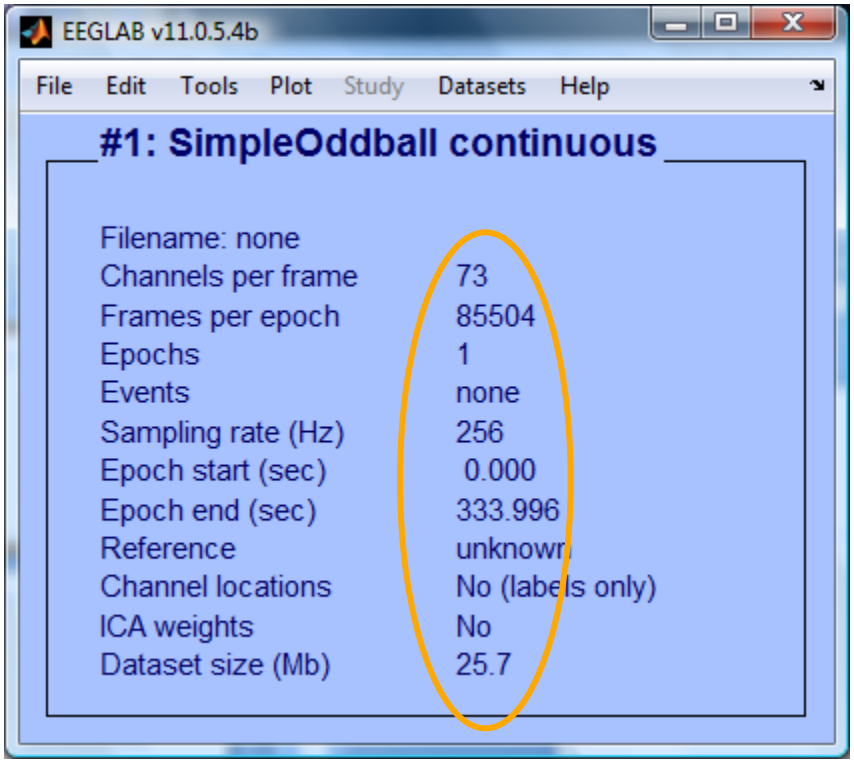


Tip for Biosemi users:

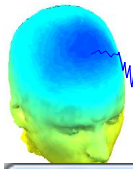
Use the 'BDF plugin' version

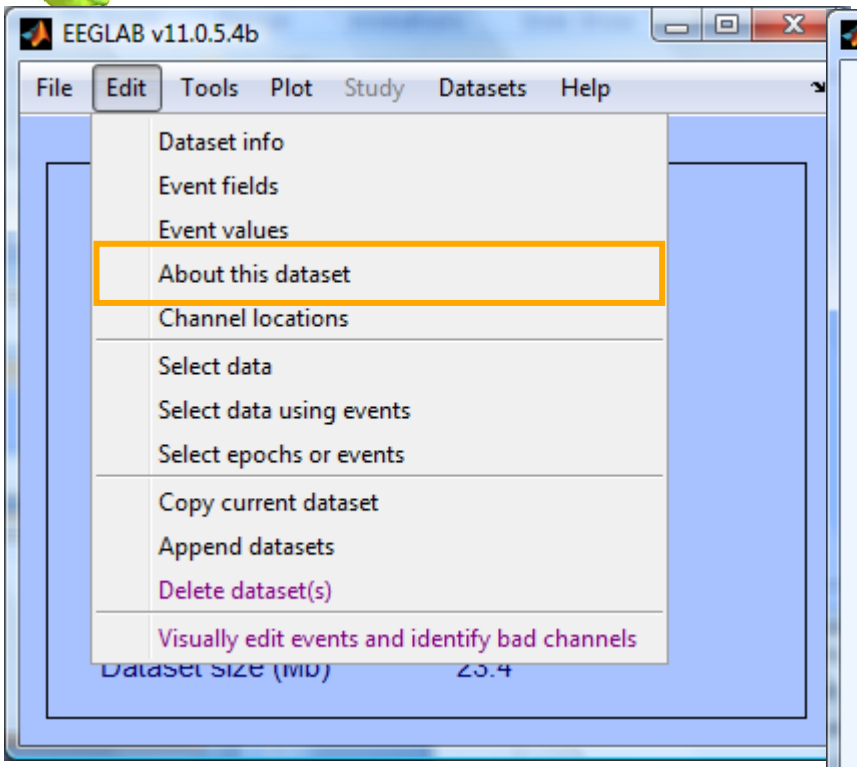
of the Biosemi BDF/EDF importer

Imported EEG data



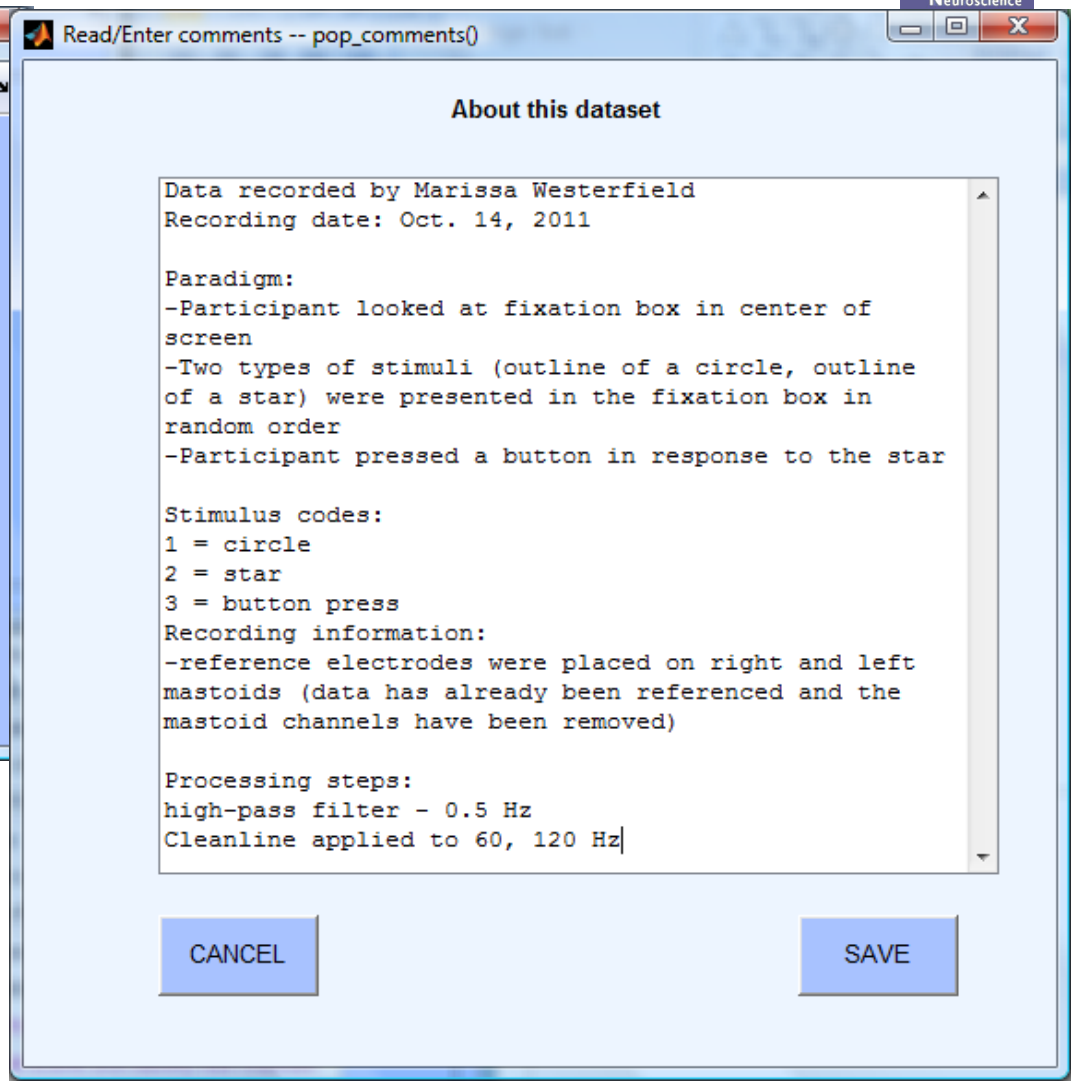
Comments and dataset history



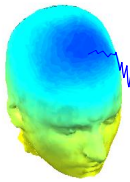


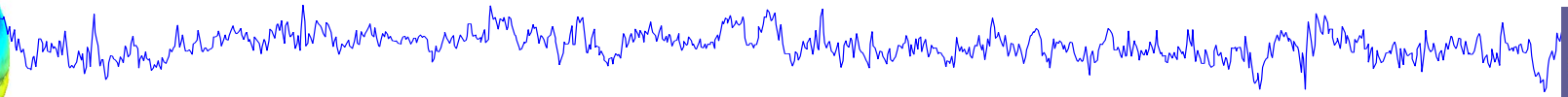
Also:
>> EEG.comments

and
>> EEG.history



Pre-processing pipeline





Swartz Center for Computational Neuroscience

Collect high-density EEG data (>30 chan)

Import into EEGLAB

Import event markers and channel locations

Re-reference/ down-sample (if necessary)

High pass filter (~.5 – 1 Hz)

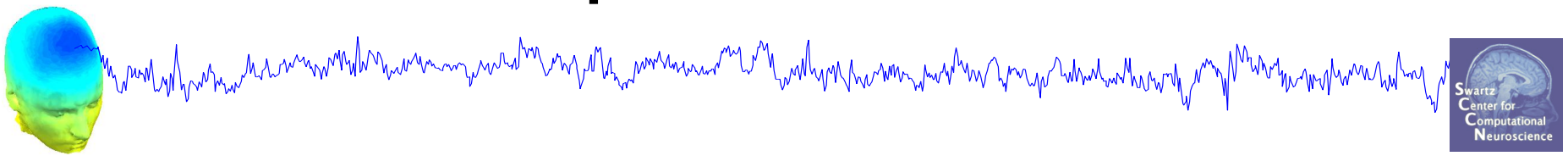
Examine raw data

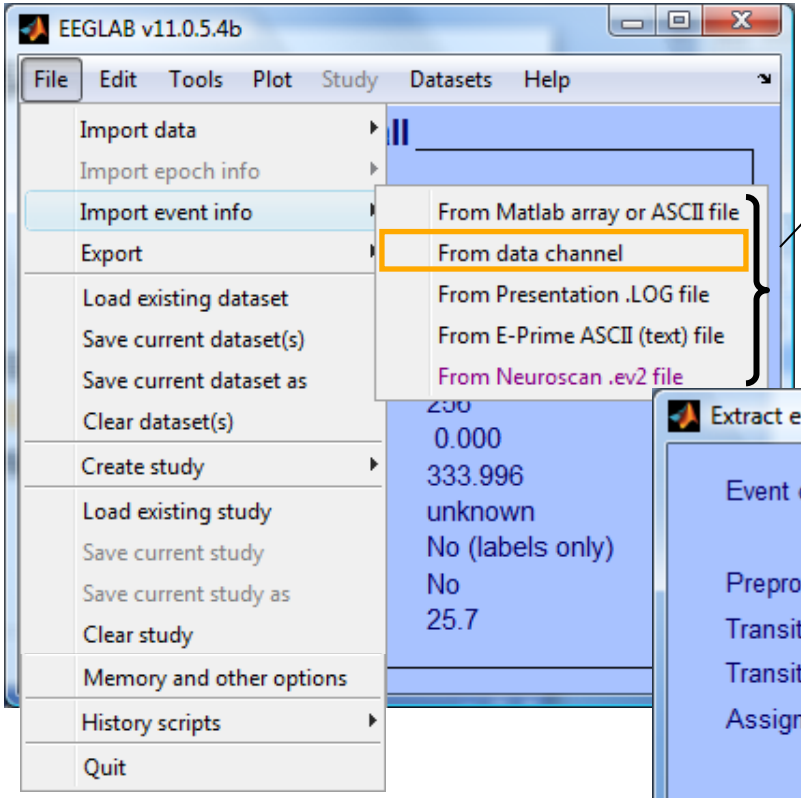
Data Cleaning: Identify/reject bad channels

Data Cleaning: Reject large artifact time points

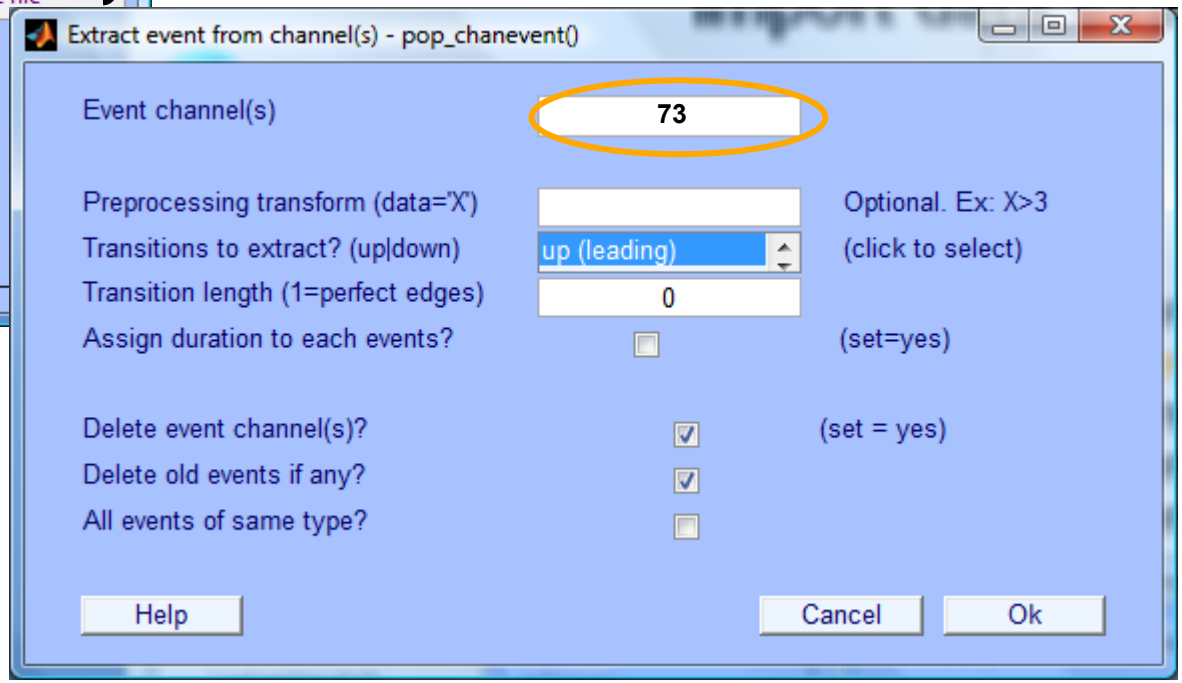
Run ICA

Import data events



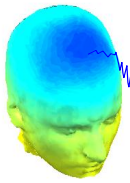


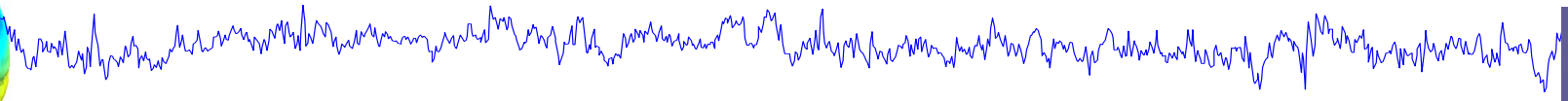
- Import events from Matlab array or ASCII file
- Import events from data channel
- Import from Presentation event file
- Import events from E-Prime event file
- Import events from Neuroscan event file

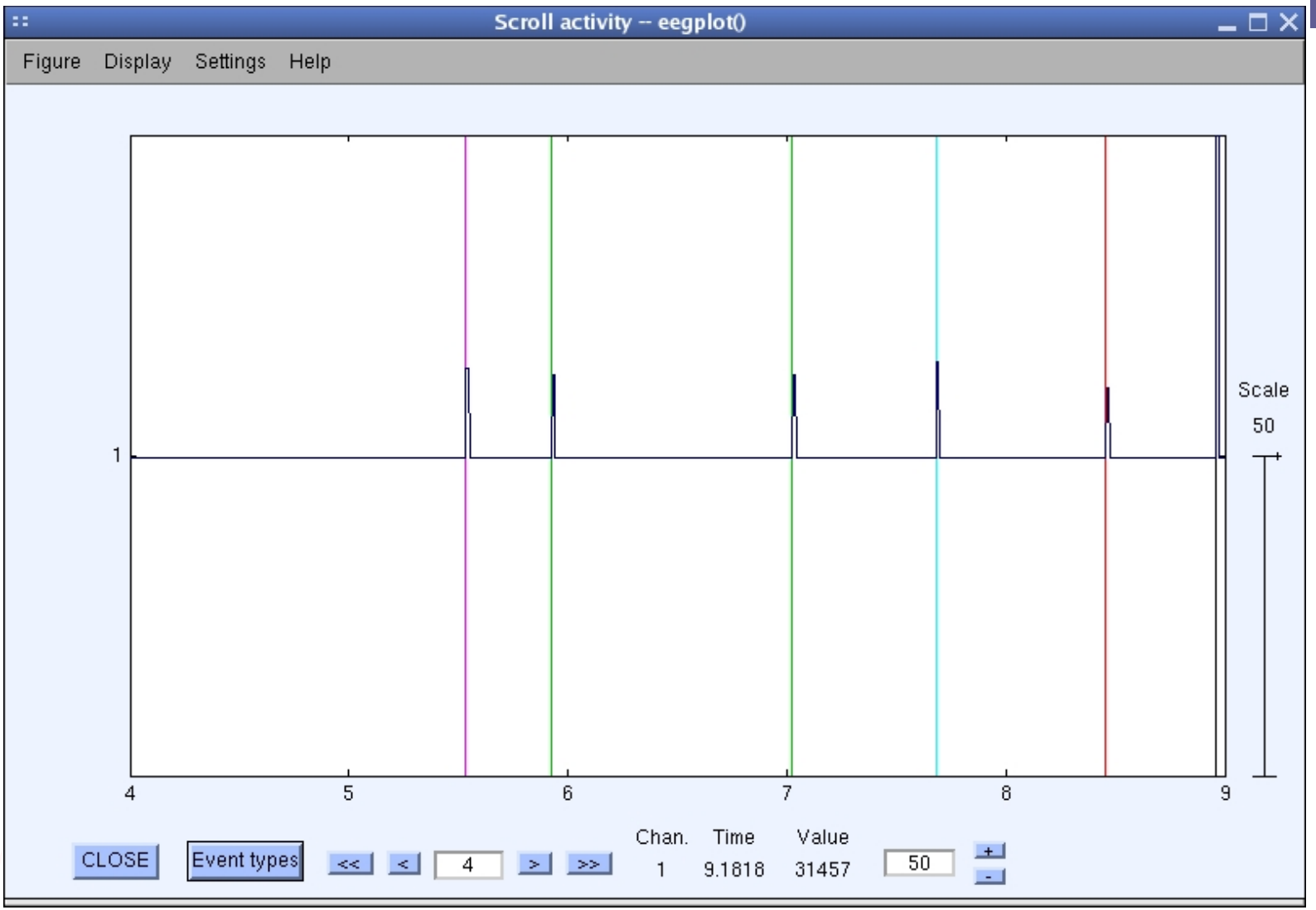


(Often imported automatically during data import)

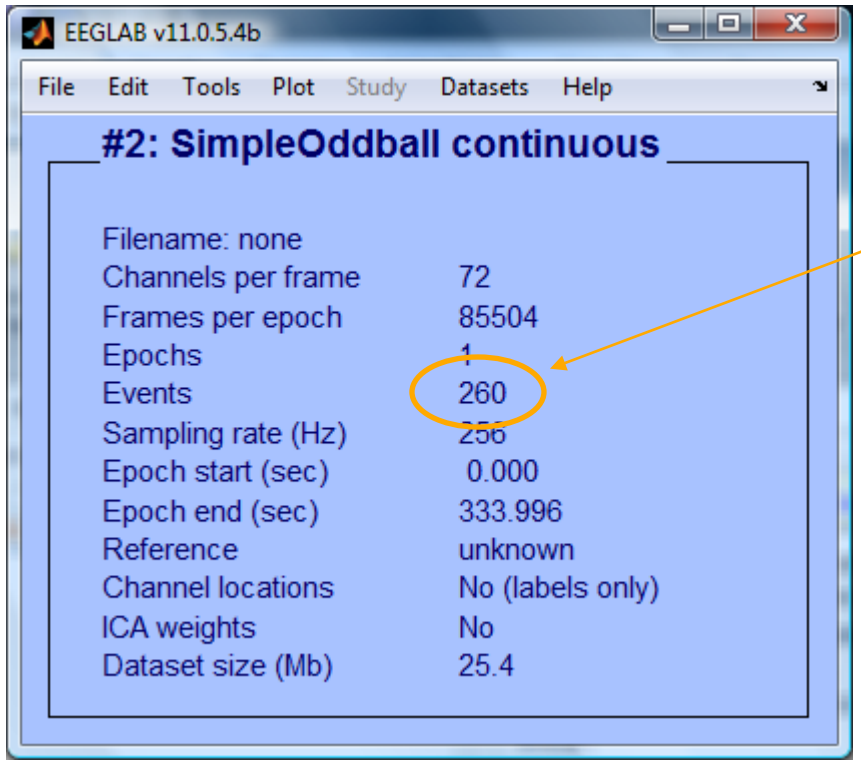
Appearance of an event channel in raw data





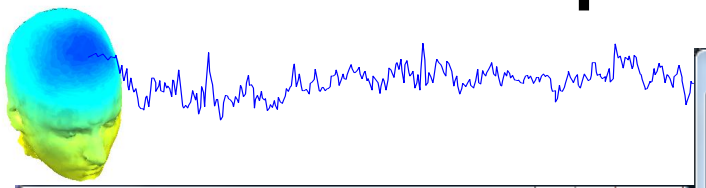


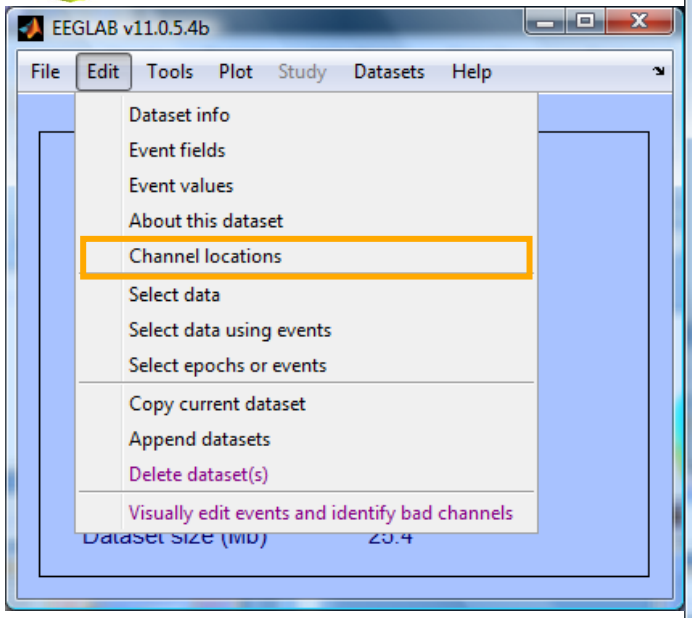
Imported data events

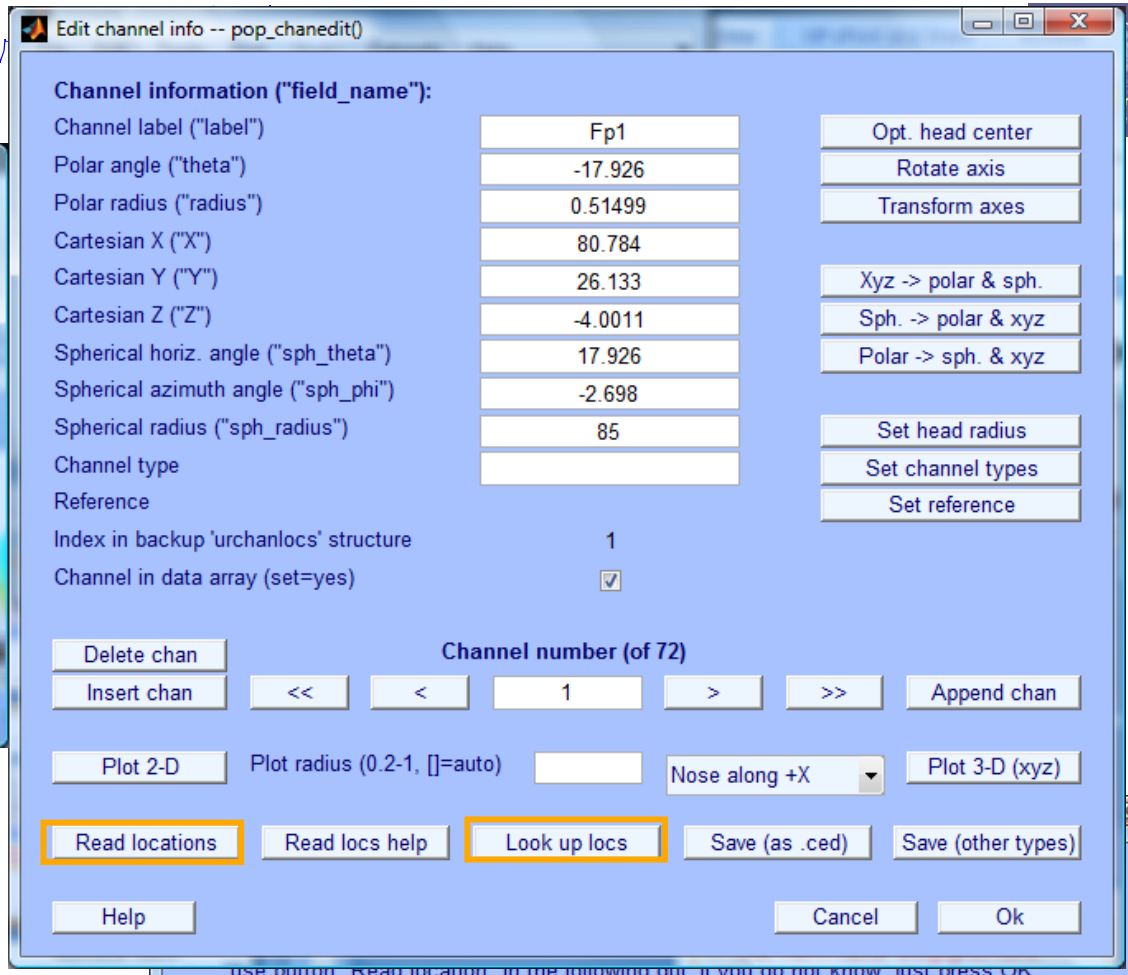


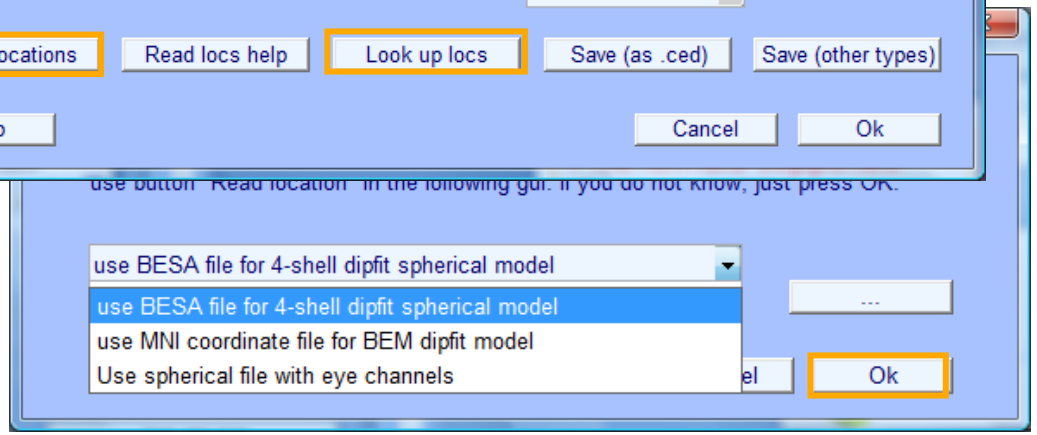
If event import was successful, you will see an appropriate number here

Import channel locations

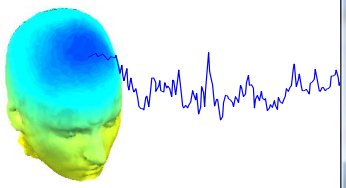








7 file formats supported (Polhemus, BESA, …)



Edit channel info -- pop_chanedit()

Channel information ("field_name"):

Channel label ("label")	HEOG	Opt. head center
Polar angle ("theta")	-42	Rotate axis
Polar radius ("radius")	0.65556	Transform axes
Cartesian X ("X")	55.7734	
Cartesian Y ("Y")	50.2186	Xyz -> polar & sph.
Cartesian Z ("Z")	-39.9051	Sph. -> polar & xyz
Spherical horiz. angle ("sph_theta")	42	Polar -> sph. & xyz
Spherical azimuth angle ("sph_phi")	-28	
Spherical radius ("sph_radius")	85	Set head radius
Channel type		Set channel types
Reference		Set reference
Index in backup 'urchanlocs' structure	68	
Channel in data array (set=yes)	☑	

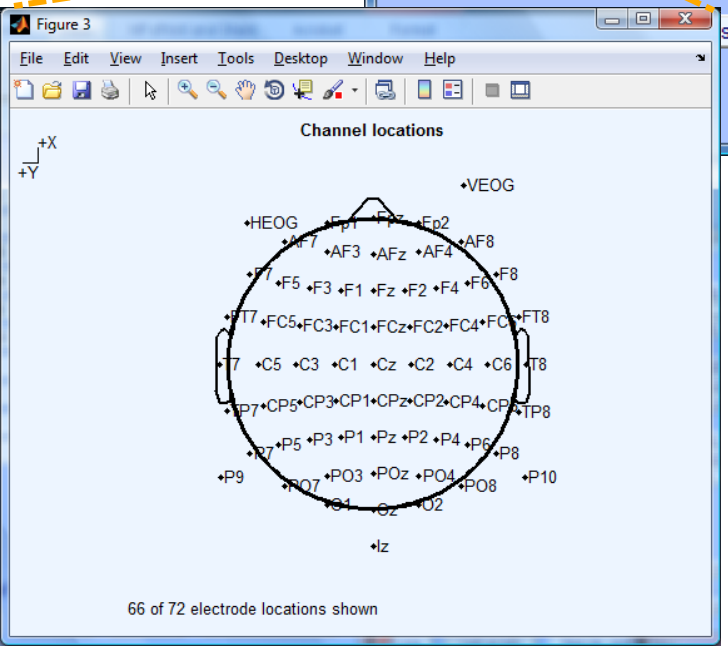
Delete chan | Insert chan | << | < | 68 | > | >> | Append chan

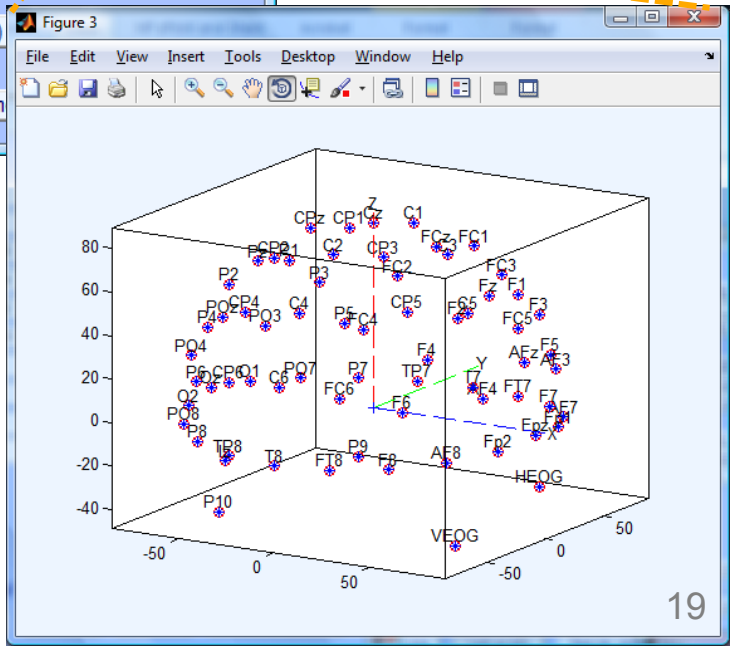
Channel number (of 72)

Plot 2-D | Plot radius (0.2-1, []=auto) | | Nose along +X | Plot 3-D (xyz)

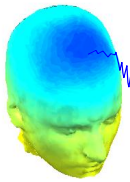
s help | Look up locs | Save (as .ced)

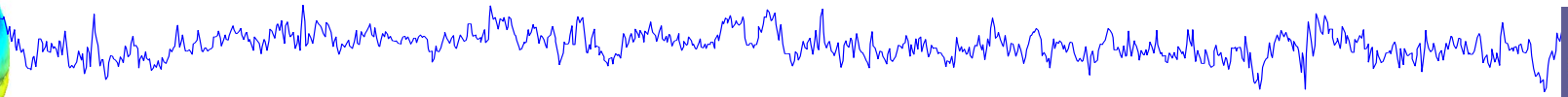
Can





Import channel locations





Edit channel info -- pop_chanedit()

Channel information ("field_name"):

Channel label ("label")	LEYE
Polar angle ("theta")	-45.1543
Polar radius ("radius")	0.54374
Cartesian X ("X")	0.79487
Cartesian Y ("Y")	0.79917
Cartesian Z ("Z")	-0.15585
Spherical horiz. angle ("sph_theta")	45.1543
Spherical azimuth angle ("sph_phi")	-7.8725
Spherical radius ("sph_radius")	1.1379
Channel type	EEG
Reference	
Index in backup 'urchanlocs' structure	
Channel in data array (set=yes)	☑

Opt. head center | Rotate axis | Transform axes | Xyz -> polar & sph. | Sph. -> polar & xyz | Polar -> sph. & xyz | Set head radius | Set channel types | Set reference

Delete chan | Insert chan | **Channel number (of 71)** << < 1 > >> | Append chan

Plot 2-D | Plot radius (0.2-1, []=auto) | Nose along +X | Plot 3-D (xyz)

Read locations | Read locs help | Look up locs | Save (as .ced) | Save (other types)

Help | Cancel | Ok

Convert channel locations -- pop_chancenter()

Optimize center location ☑ or specify center 0 0 0

Channel indices to ignore for best-sphere matching

Browse | Help | Cancel | Ok

Force electrode location -- forcelocs()

X/Y value: 0 | Coordinate: X (rotate X-Z plane) | Electrode list: Cz | Pick

Help | Cancel | Ok

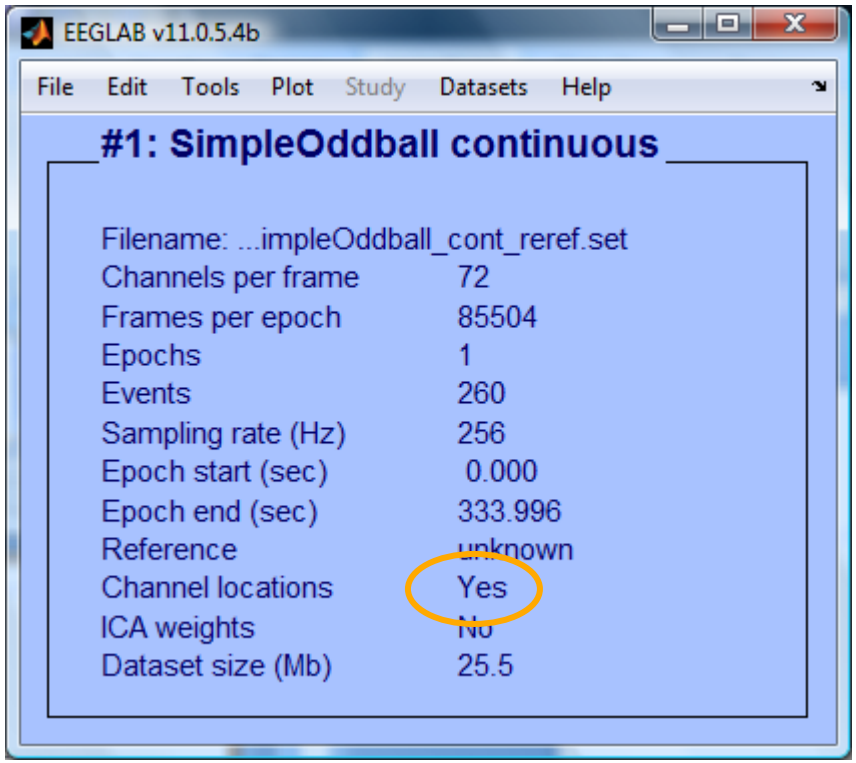
Set channel ...

Channel indices: 1:71

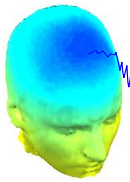
Type (e.g. EEG): EEG

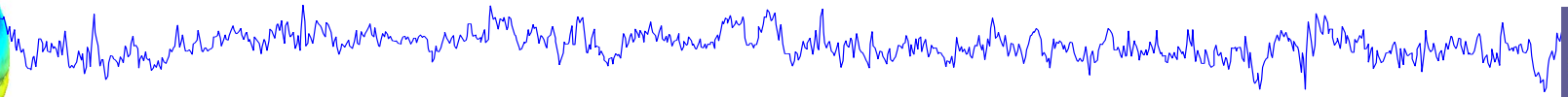
Help | Cancel | Ok

Imported channel locations



Pre-processing pipeline





Swartz Center for Computational Neuroscience

Collect high-density EEG data (>30 chan)

Import into EEGLAB

Import event markers and channel locations

Re-reference/ down-sample (if necessary)

High pass filter (~.5 – 1 Hz)

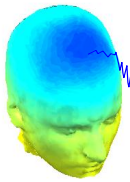
Examine raw data

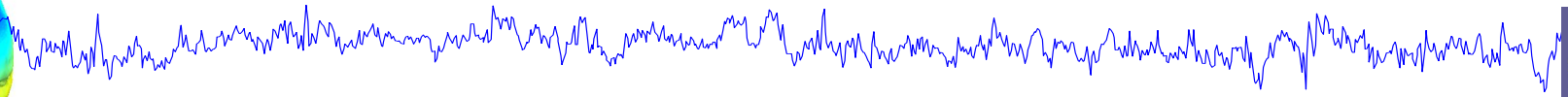
Data Cleaning: Identify/reject bad channels

Data Cleaning: Reject large artifact time points

Run ICA

Re-reference data (if necessary/desired)



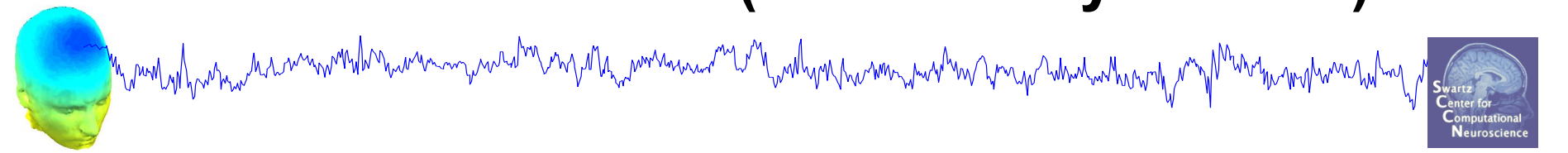


Swartz Center for Computational Neuroscience

Re-reference to

(e.g.) 'linked mastoids'

Re-reference data (if necessary/desired)



EEGLAB v10.2.2.1b

File Edit Tools Plot Study Datasets Help

Change sampling rate
Filter the data
Re-reference
Interpolate electrodes
Reject continuous
Extract epochs
Remove baseline
Run ICA
Remove componer
Automatic channel
Automatic epoch r
Reject data epochs
Reject data using IC
NFT plugin
SIFT
Locate dipoles usin
Peak detection usin
FMRIB Tools
Locate dipoles using LORETA

pop_reref - average reference or re-reference data

Current data reference state is: unknown

Compute average reference

Re-reference data to channel(s):

Retain old reference channels in data

Exclude channel indices (EMG, EOG) optional — LEYE REYE

Add current reference channel back to the data

Help Cancel Ok

Or,

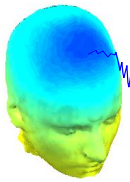
average reference

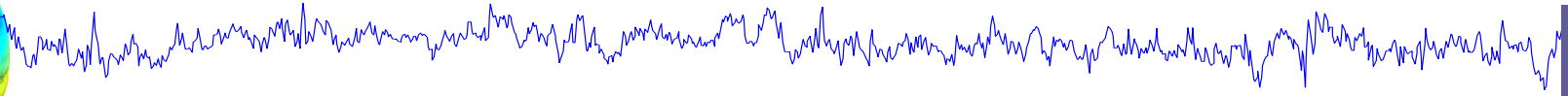
(use shift|ctrl to select several)

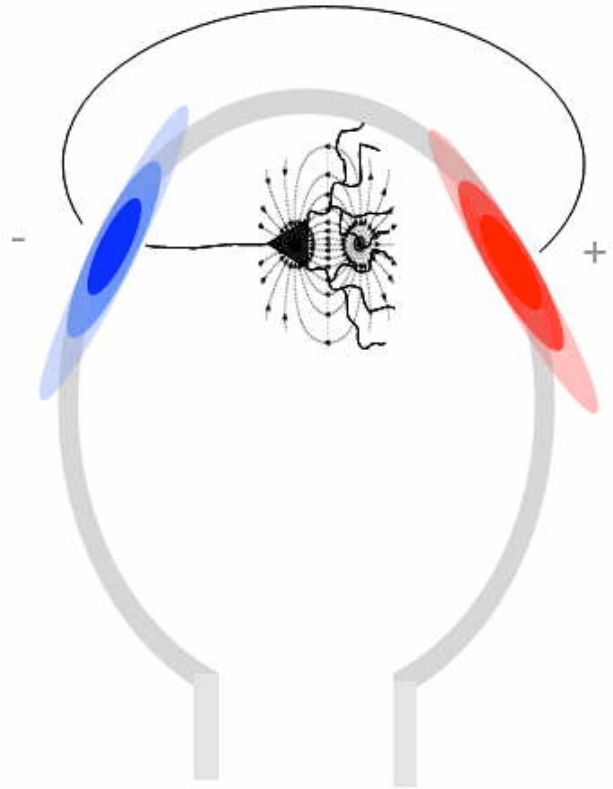
1 - LEYE
2 - REYE
3 - OZ
4 - O2
5 - FP1
6 - FPZ
7 - FP2
8 - AF7
9 - AF3
10 - AFZ
11 - AF4
12 - AF8
13 - F9
14 - F7
15 - F5
16 - F3
17 - F1
18 - FZ
19 - F2
20 - F4
21 - F6
22 - F8
23 - F10
24 - FT9
25 - FT7
26 - FC5

Cancel Ok

On Average Referencing



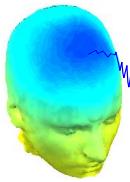


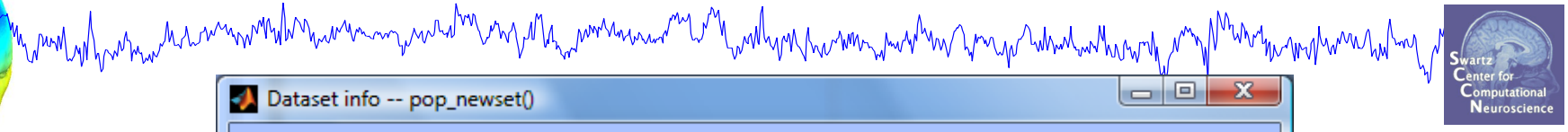


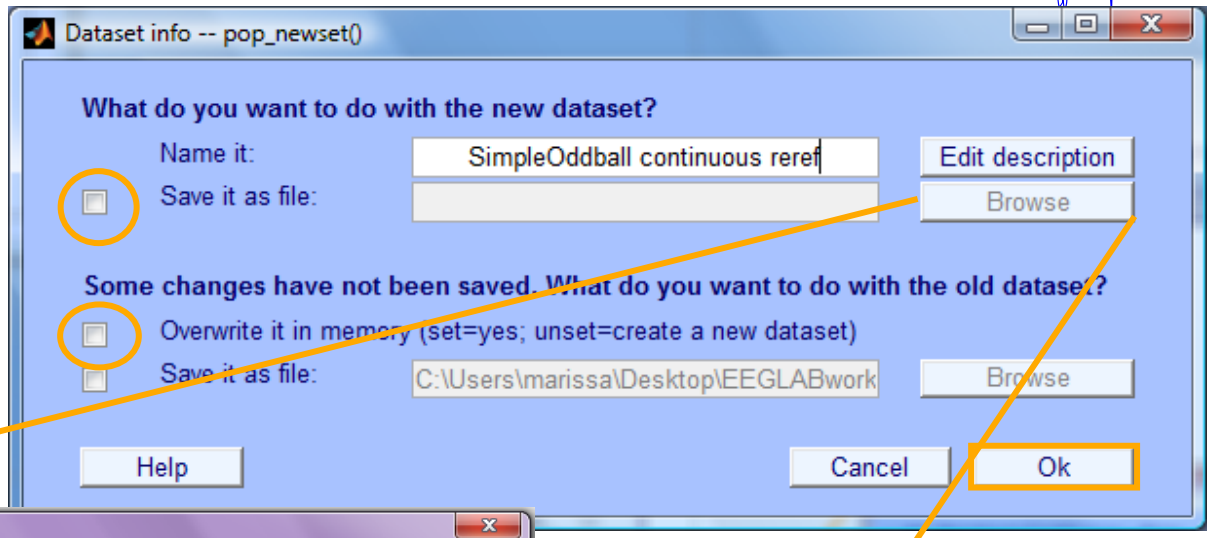
In theory, positive and negative current across entire head should balance—no net current source or sink: Average referencing enforces this.

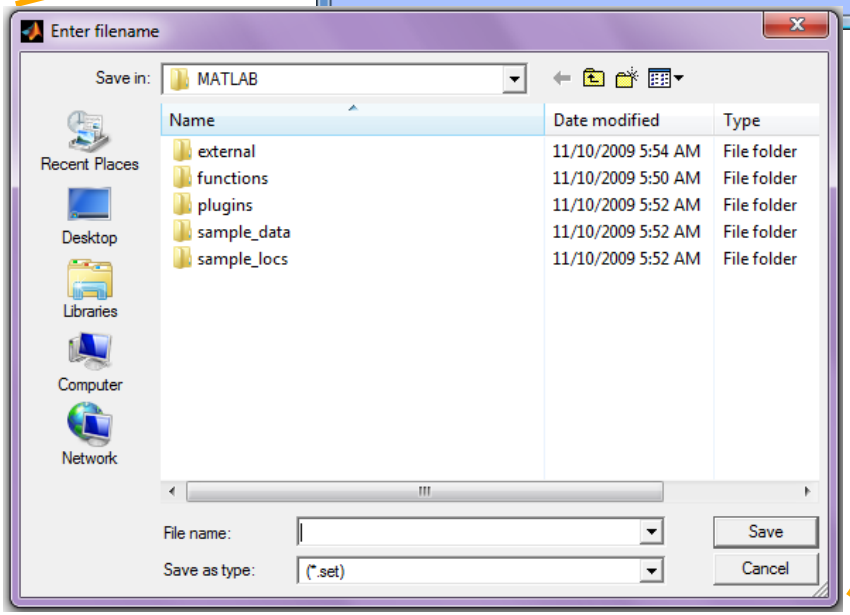
In practice, depends on distribution of electrodes.

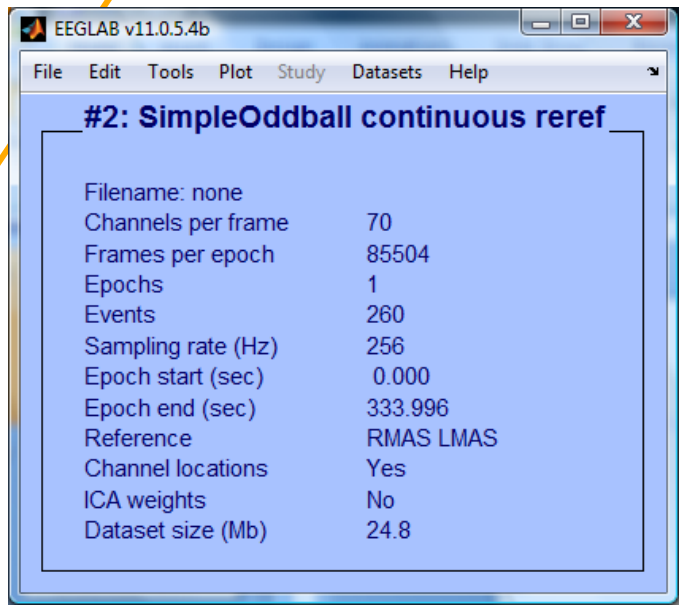
Save new dataset, keep old one



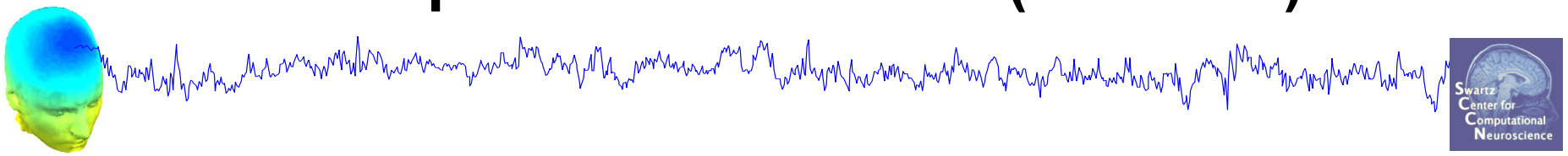


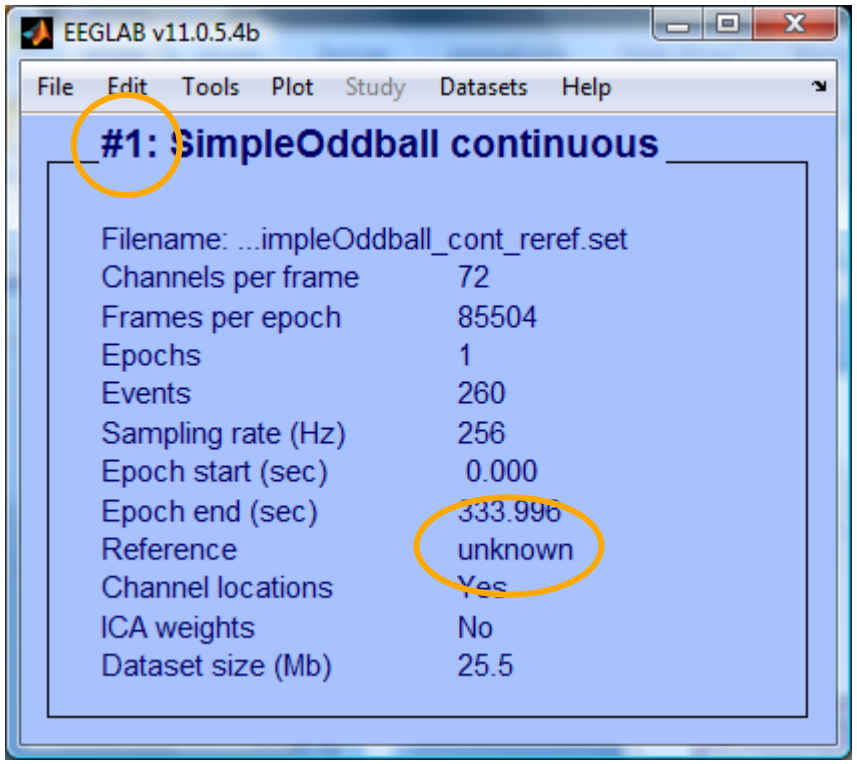


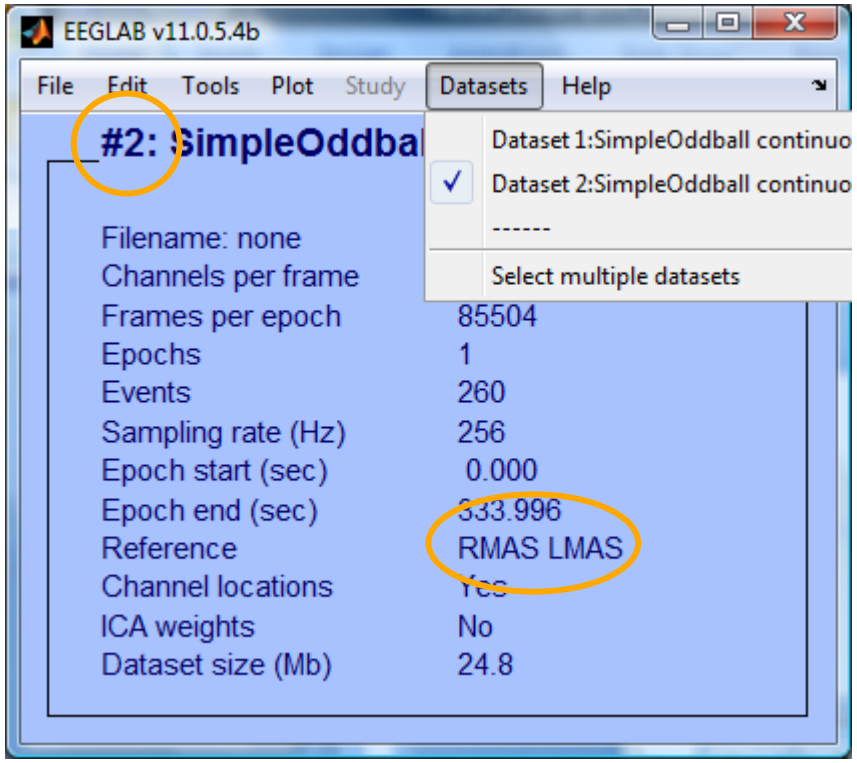




Multiple active datasets (ALLEEG)

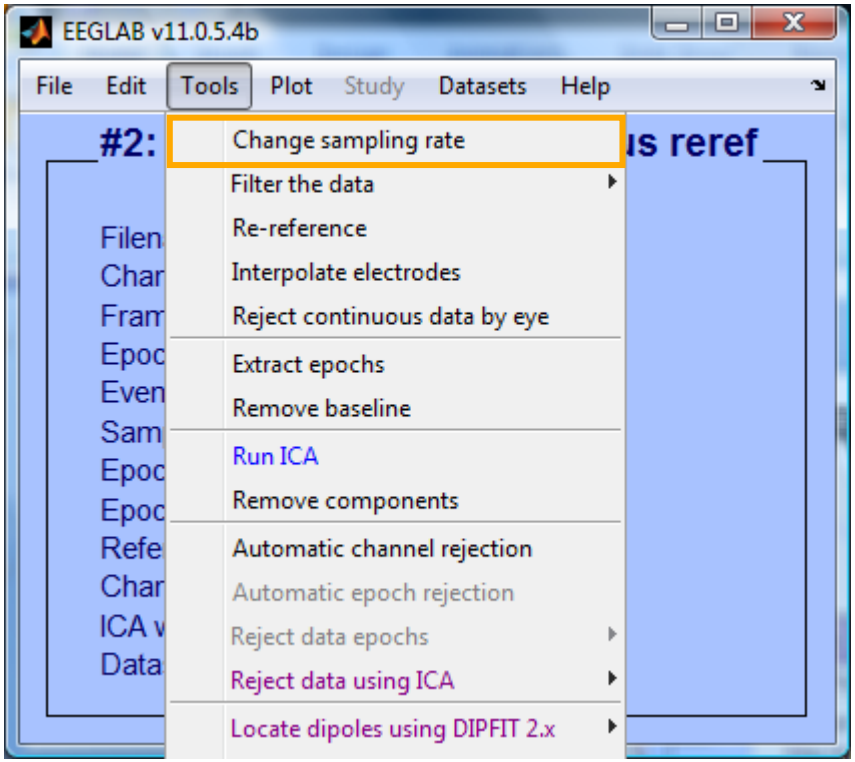


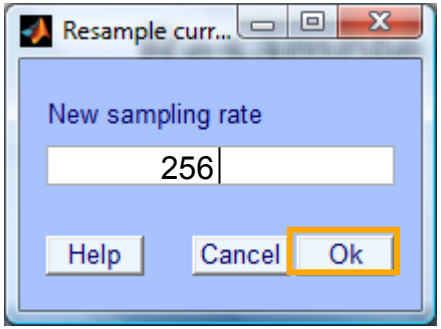




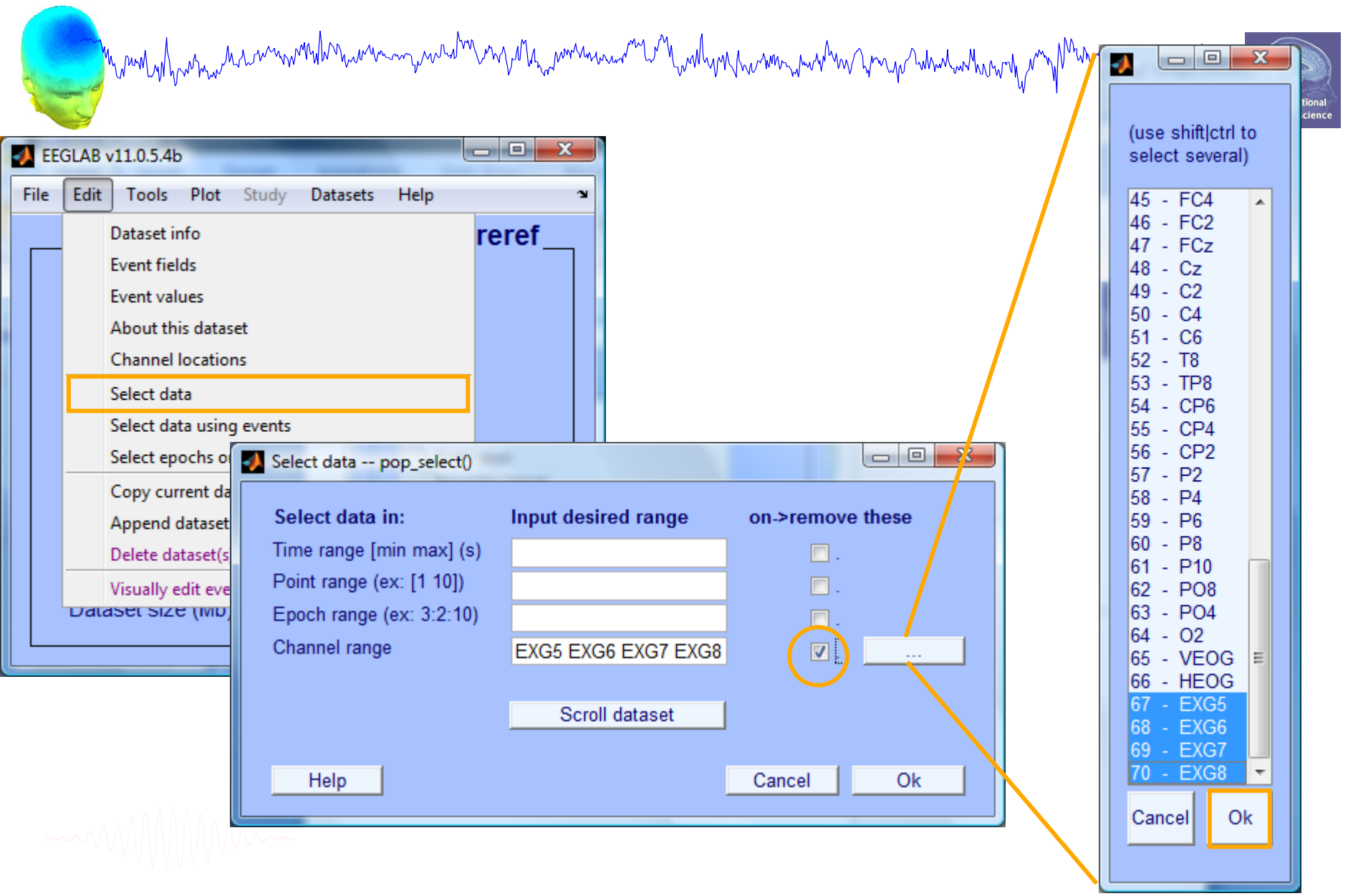
Resample data (if desired)

Reason: Reduce space, time. But keep nyquist and ICA data length requirements in mind…

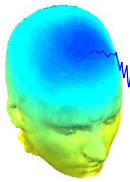


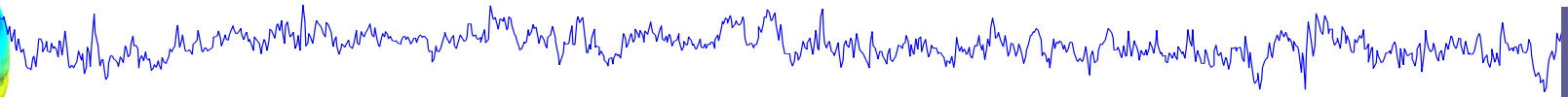


Remove unwanted channels



Pre-processing pipeline





Swartz Center for Computational Neuroscience

Collect high-density EEG data (>30 chan)

Import into EEGLAB

Import event markers and channel locations

Re-reference/ down-sample (if necessary)

STOP!

Save your continuous dataset here

(at least!)

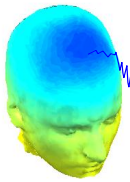
Examine raw data

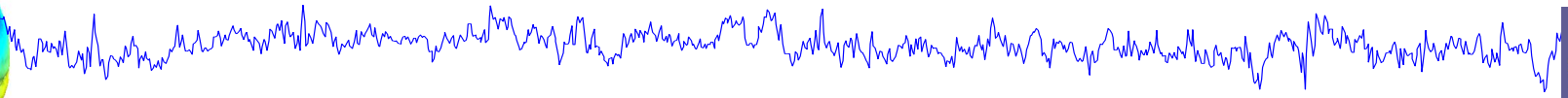
Data Cleaning: Identify/reject bad channels

Data Cleaning: Reject large artifact time points

Run ICA

Pre-processing pipeline





Swartz Center for Computational Neuroscience

Collect high-density EEG data (>30 chan)

Import into EEGLAB

Import event markers and channel locations

Re-reference/ down-sample (if necessary)

High pass filter (~.5 – 1 Hz)

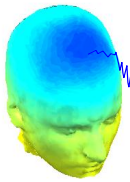
Examine raw data

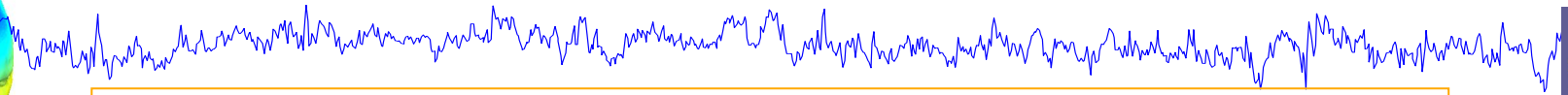
Data Cleaning: Identify/reject bad channels

Data Cleaning: Reject large artifact time points

Run ICA

High-Pass Filter the data





Swartz Center for Computational Neuroscience

Reason: remove slow, possibly large amplitude, drift

EEGLAB v11.0.5.4b

File Edit Tools Plot Study Datasets Help

#1:

Change sampling rate
Filter the data
Re-reference
Interpolate electrodes
Reject continuous data by eye
Extract epochs
Remove baseline
Run ICA
Remove components
Automatic channel rejection
Automatic epoch rejection
Reject data epochs
Reject data using ICA
Locate dipoles using DIPFIT 2.x
Peak detection using EEG toolbox
FMRIB Tools
Locate dipoles using LORETA
CleanLine

Basic FIR filter
ERPLAB Butterwort
ERPLAB Polynomia
Short non-linear IIR

Filter the data -- pop_eegfilt()

Lower edge of the frequency pass band (Hz) 0.5
Higher edge of the frequency pass band (Hz)
FIR Filter order (default is automatic)
Notch filter the data instead of pass band
Use (sharper) FFT linear filter instead of FIR filtering
(Use the option above if you do not have the Signal Pro
Use causal filter (useful when performing causal ana
Plot the filter frequency response
Use fir1 (check, recommended) or firls (uncheck, le

High-pass needed for ICA

Help Cancel Ok

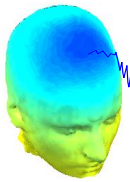
Dataset info -- pop_newset()

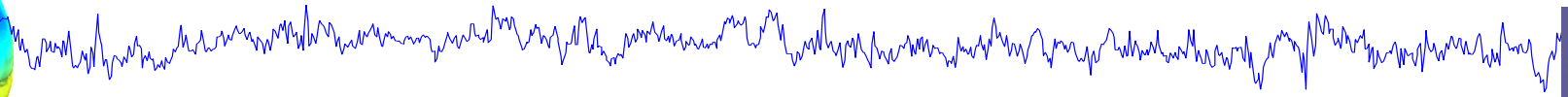
What do you want to do with the new dataset?
Name it: SimpleOddball hipass0.5 Edit description
Save it as file: Browse

What do you want to do with the old dataset (not modified since last saved)?
Overwrite it in memory (set=yes; unset=create a new dataset)

Help Cancel Ok

Pre-processing pipeline





Swartz Center for Computational Neuroscience

Collect high-density EEG data (>30 chan)

Import into EEGLAB

Import event markers and channel locations

Re-reference/ down-sample (if necessary)

High pass filter (~.5 – 1 Hz)

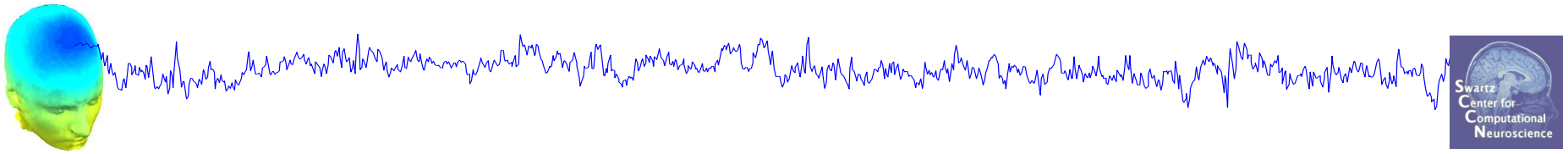
Examine raw data

Data Cleaning: Identify/reject bad channels

Data Cleaning: Reject large artifact time points

Run ICA

Scroll channel data



EEGLAB v11.0.5.4b

File Edit Tools Plot Study Datasets Help

Channel locations
Channel data (scroll)
Channel spectra and maps
Channel properties
Channel ERP image
Channel ERPs

EEGLAB v11.0.5.4b

File Edit Tools Plot Study Datasets Help

Change sampling rate
Filter the data
Re-reference
Interpolate electrodes
Reject continuous data by eye
Extract epochs
Remove baseline

Alternate GUI option, same function

Scroll channel activities -- eegplot()

Figure Display Settings Help

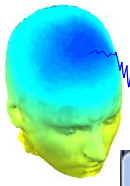
Stack
Norm

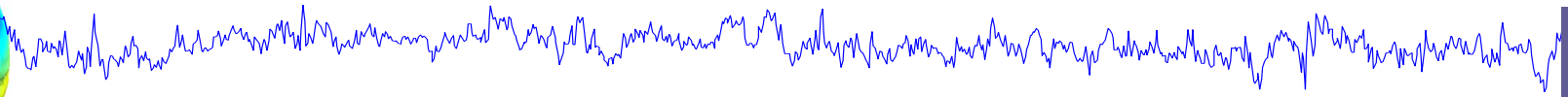
Scale 45

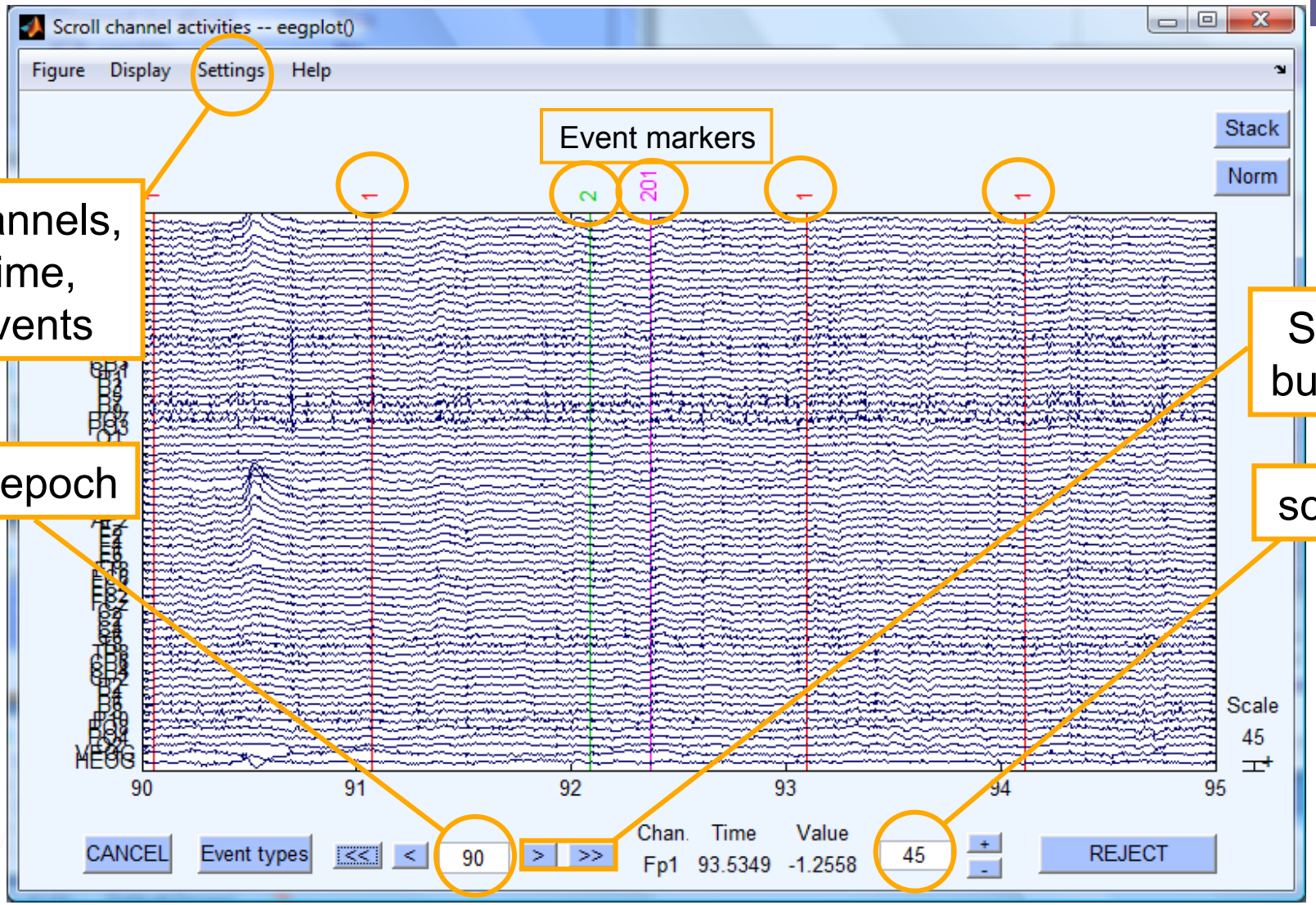
90 91 92 93 94 95

CANCEL Event types << < 90 > >> Chan. Time Value Fp1 93.5349 -1.2558 45 REJECT

Scroll channel data







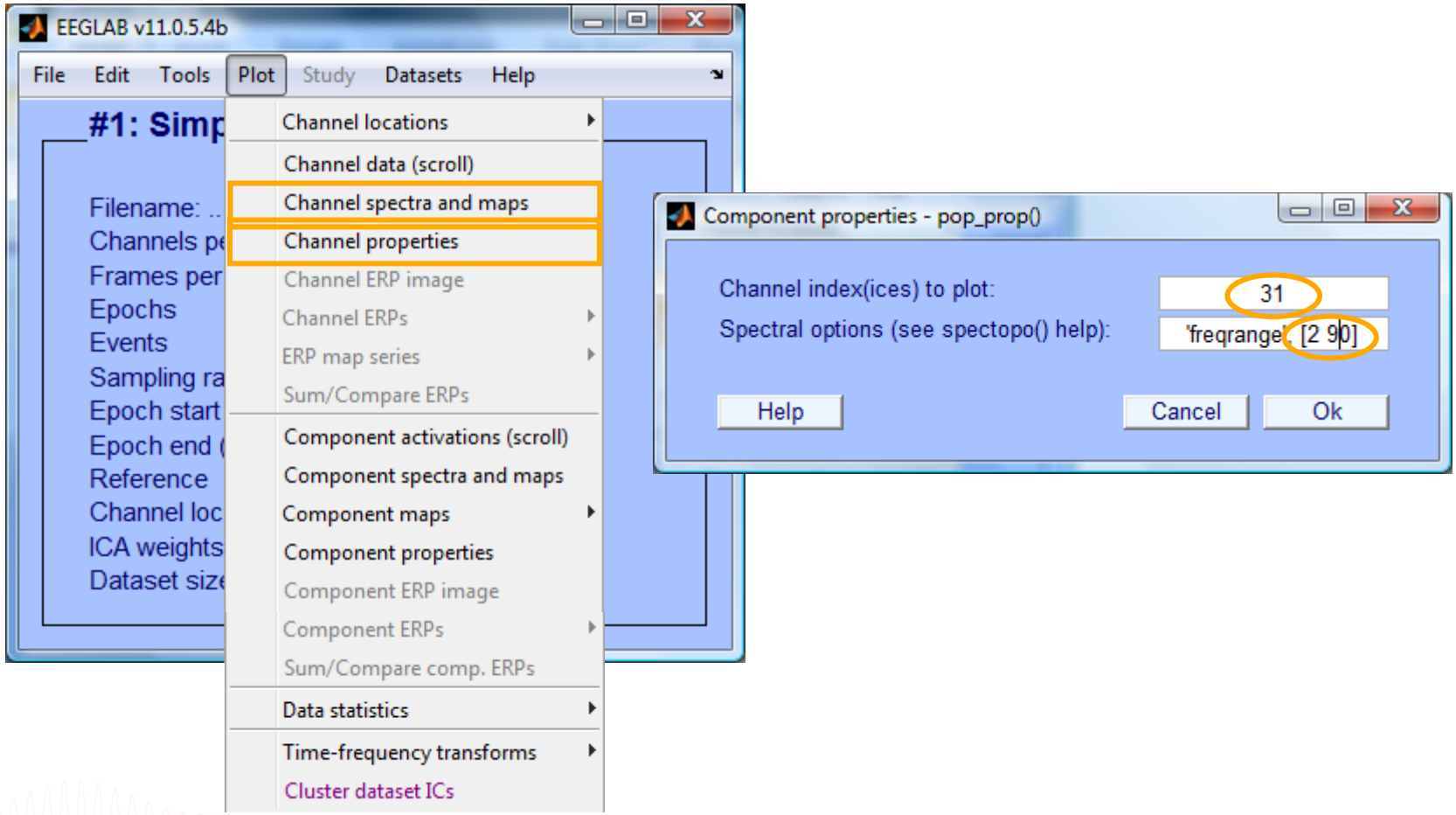
channels, time, events

sec/epoch

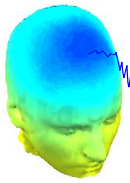
Scroll buttons

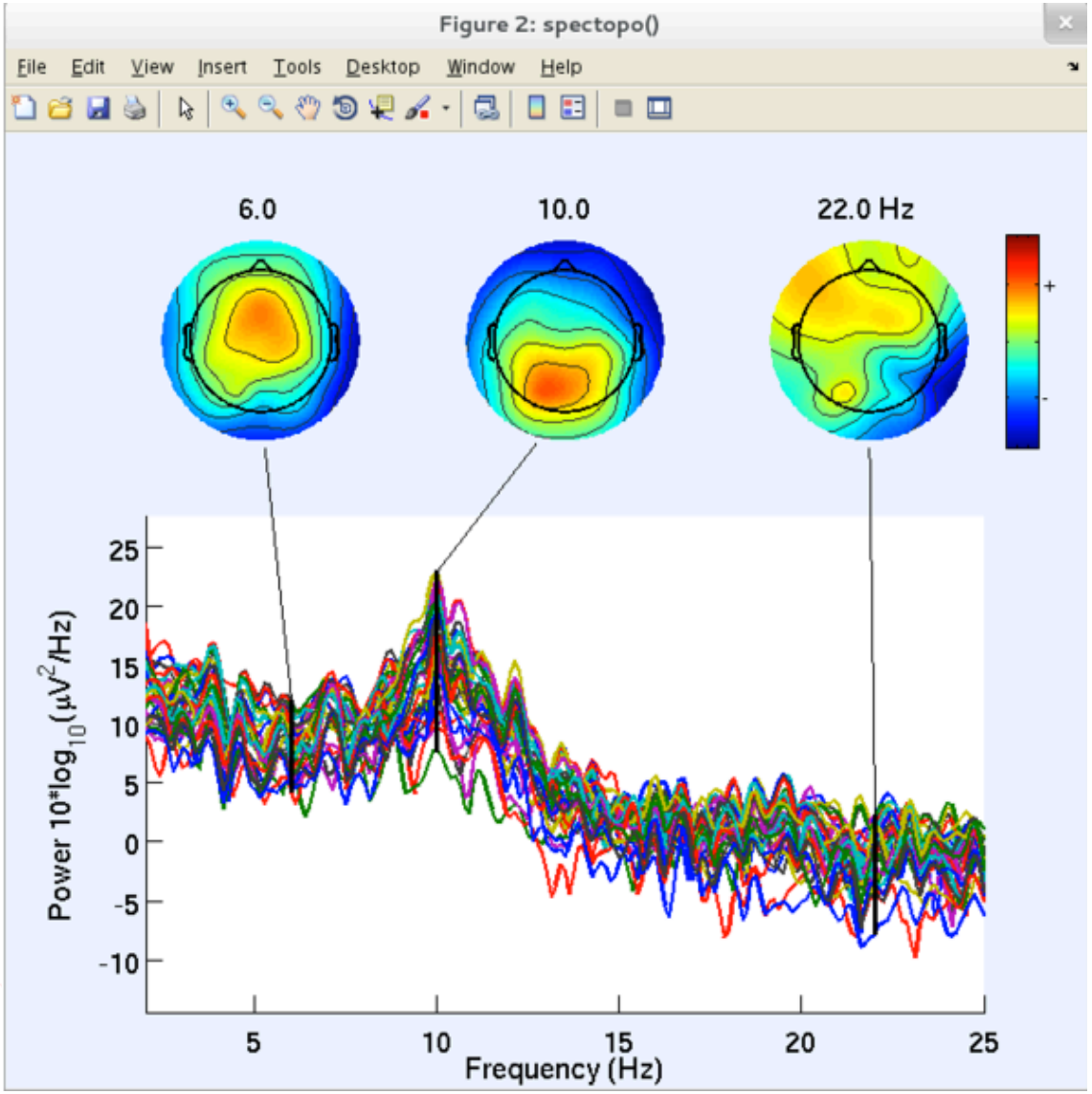
scaling

Plot channel spectra

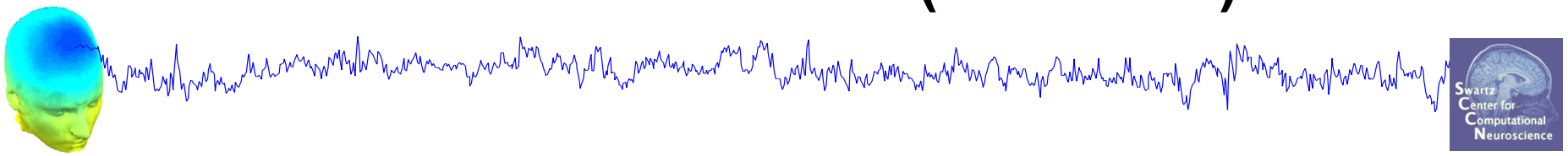


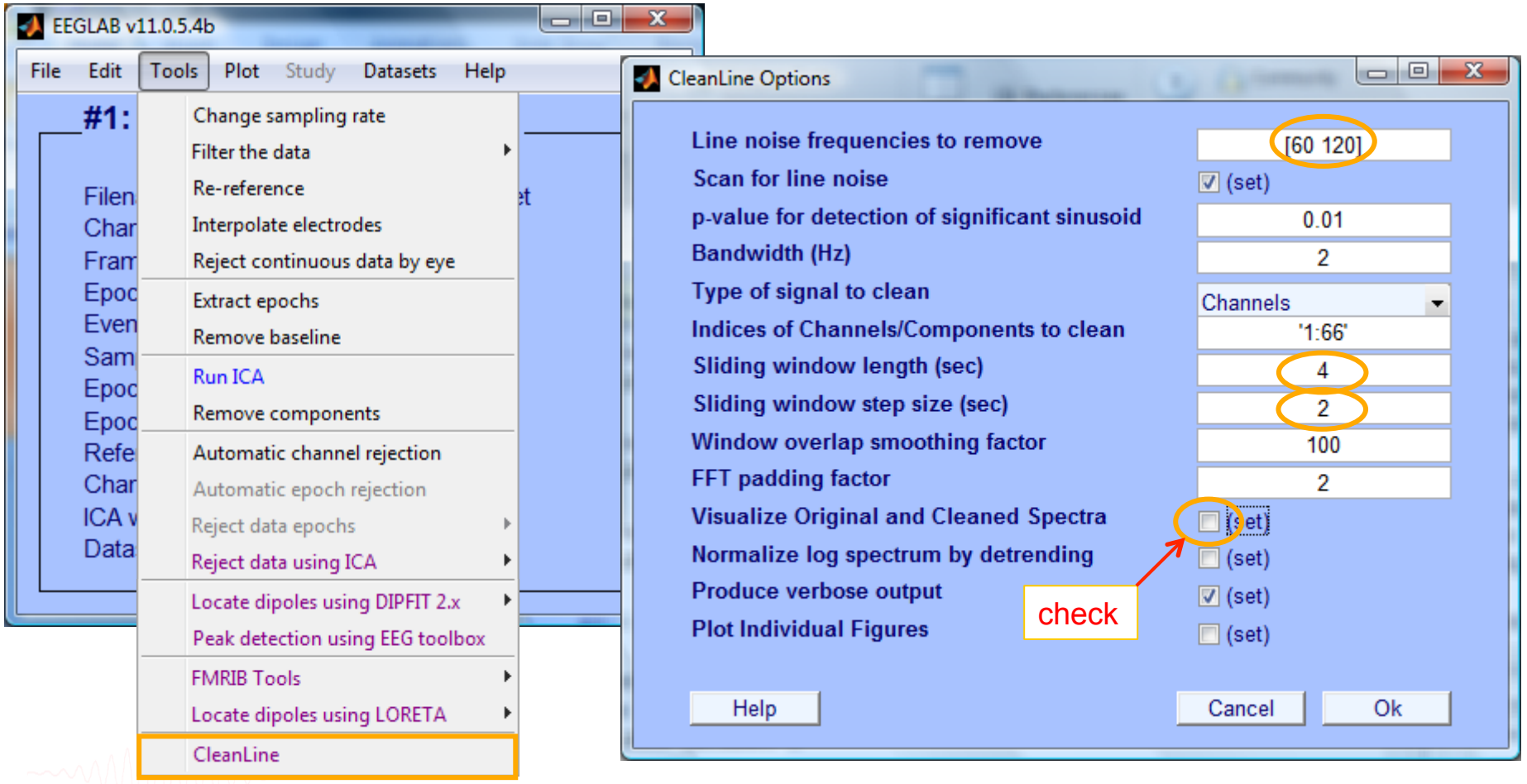
Plot channel spectra



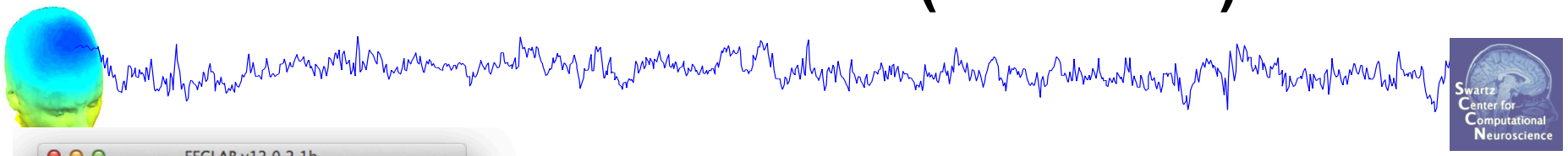


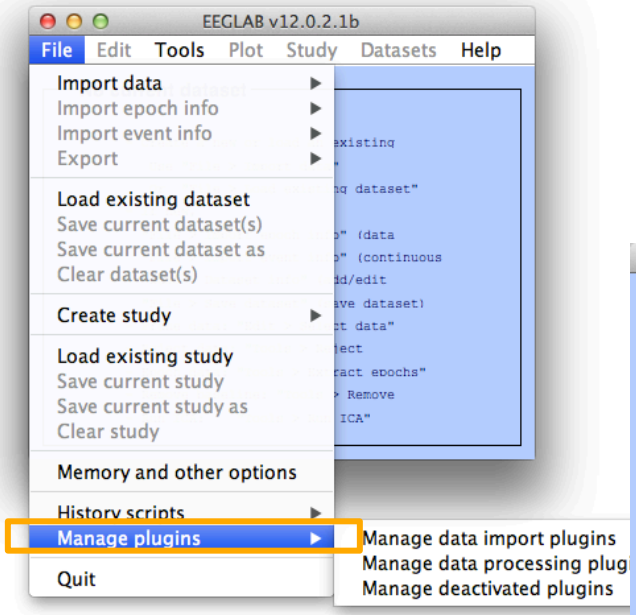
Remove line noise (Cleanline)

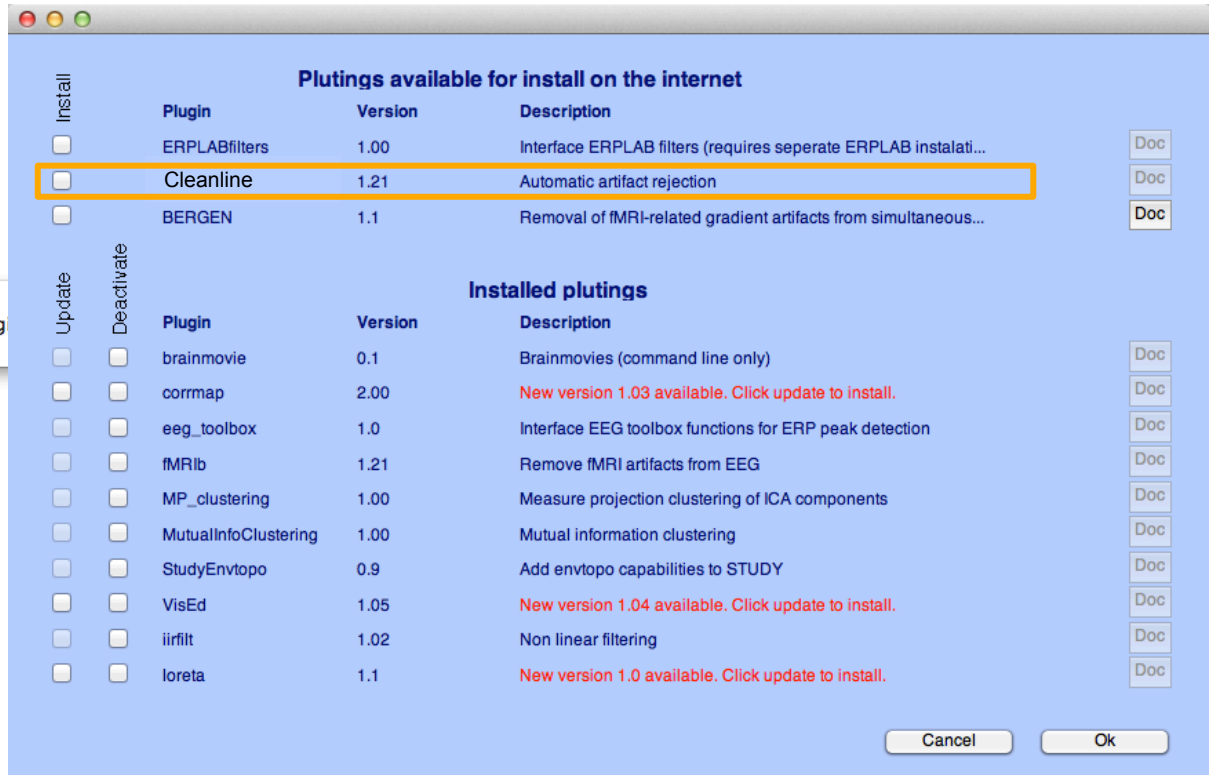




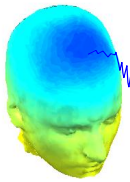
Remove line noise (Cleanline)

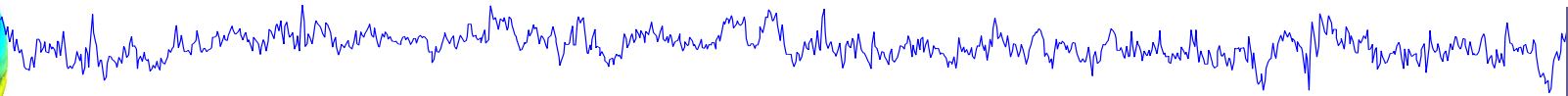






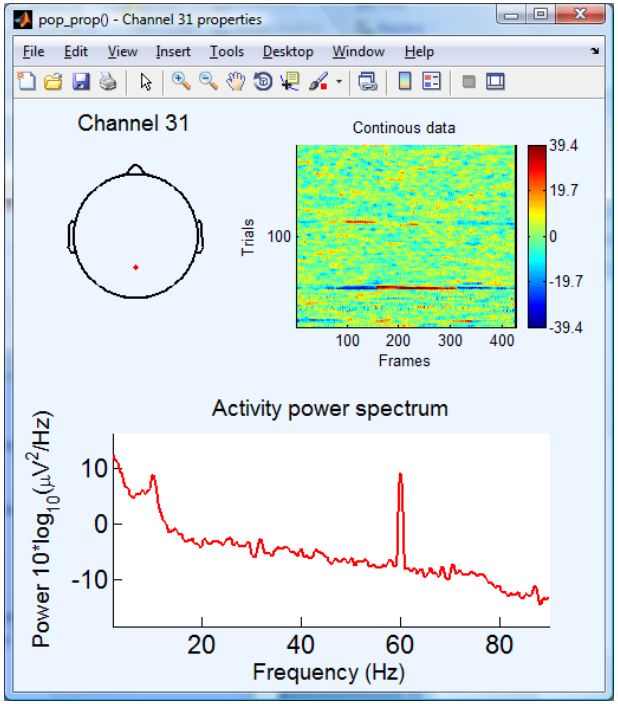
Filter comparisons



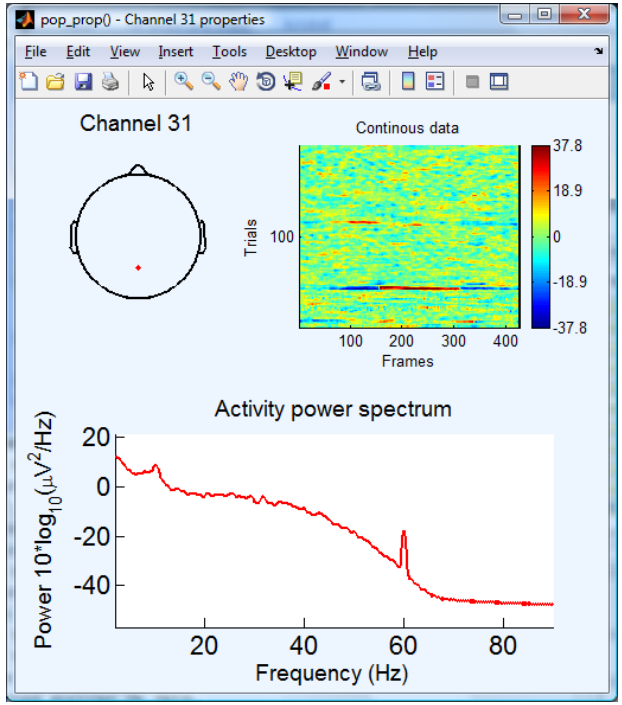


Swartz Center for Computational Neuroscience

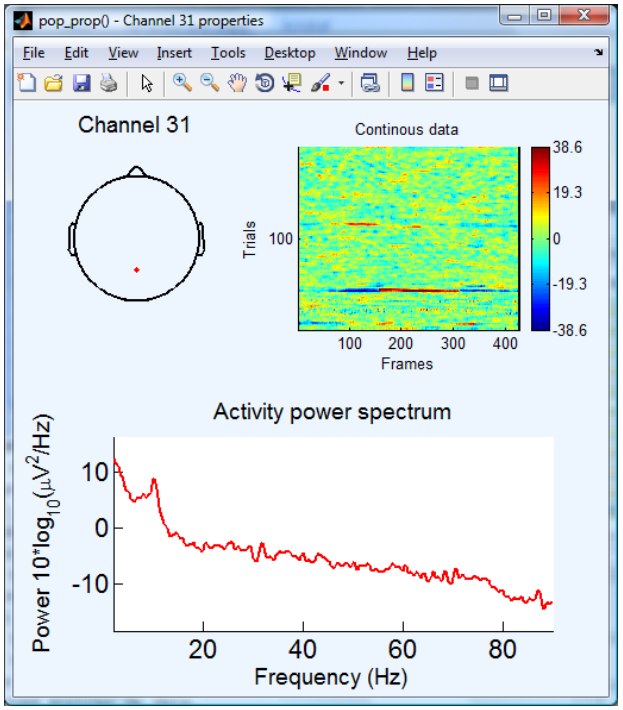
0.5 Hz high-pass filter



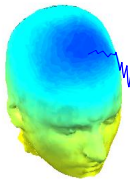
0.5 Hz high-pass filter
50 Hz low-pass filter

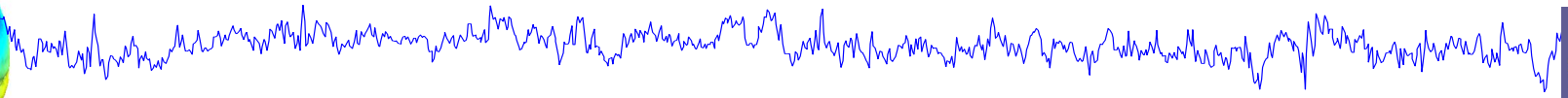


0.5 Hz high-pass filter
Cleanline



Pre-processing pipeline





Swartz Center for Computational Neuroscience

Collect high-density EEG data (>30 chan)

Import into EEGLAB

Import event markers and channel locations

Re-reference/ down-sample (if necessary)

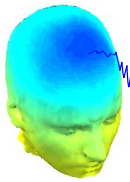
High pass filter (~.5 – 1 Hz)

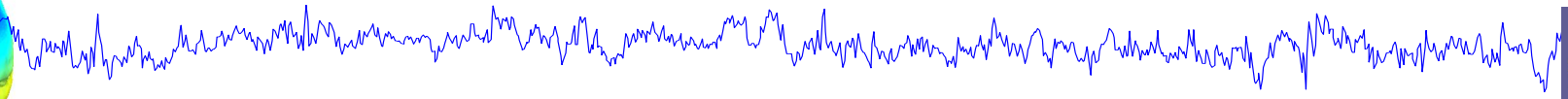
Examine raw data

Data Cleaning: Identify/reject bad channels

Data Cleaning: Reject large artifact time points

Run ICA

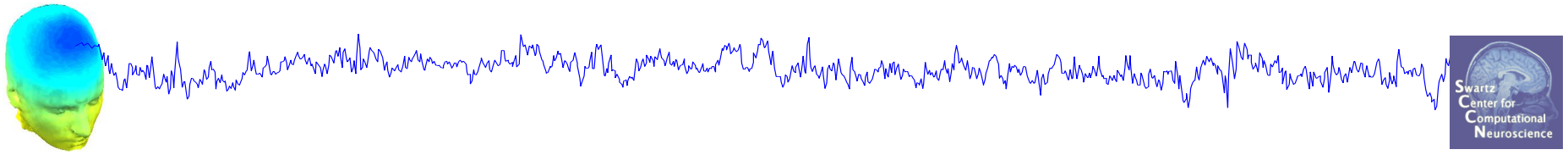


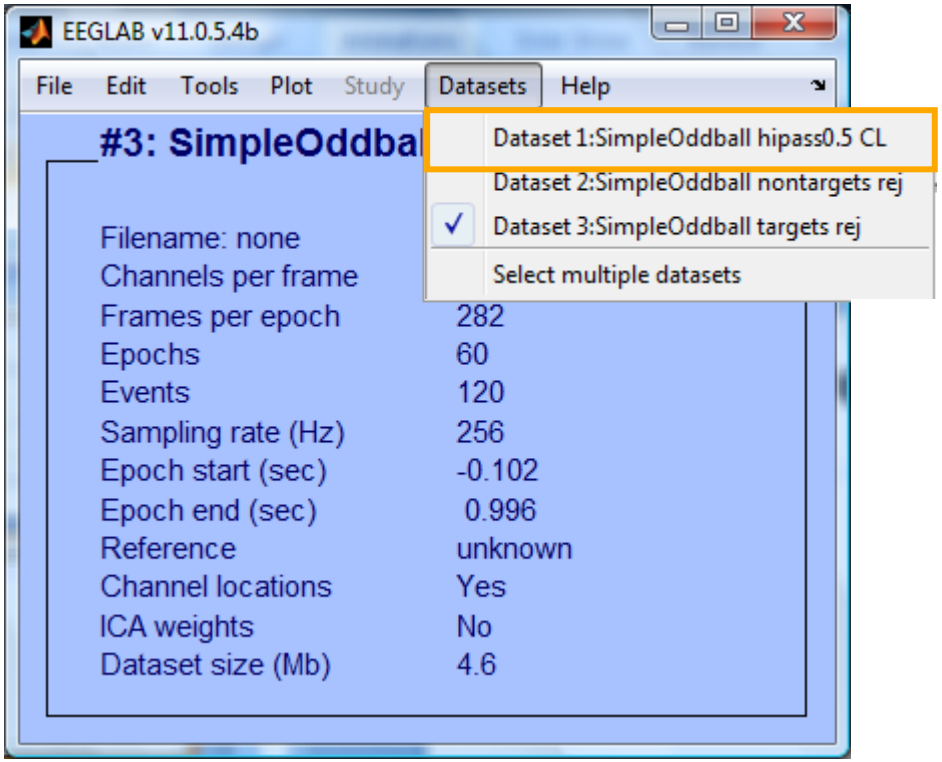


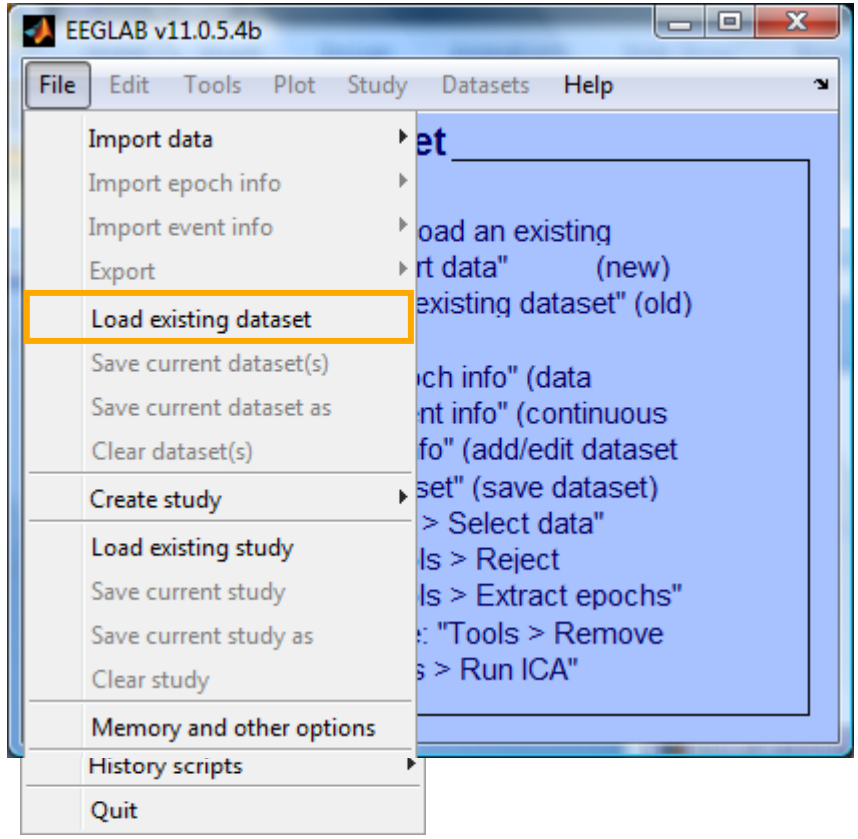
Data Cleaning for ICA

Variant 1: Continuous Data

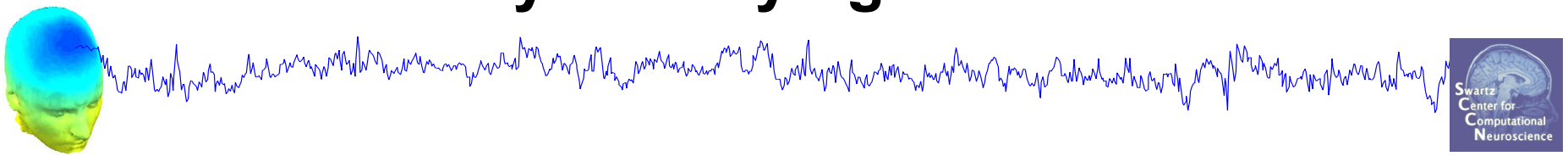
Continuous EEG dataset







Manually identifying bad channels



Scroll channel activities -- eegplot()

Figure Display Settings Help

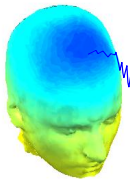
1) Identify bad channel

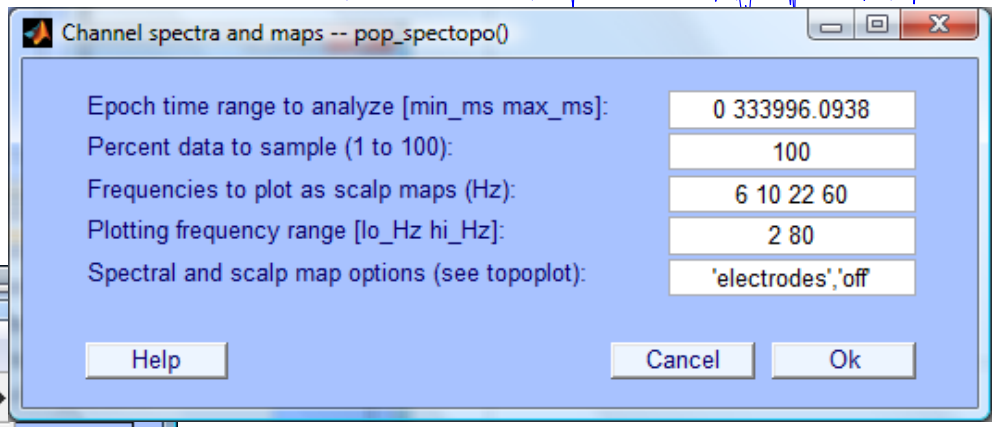
449 450 451 452 453 454

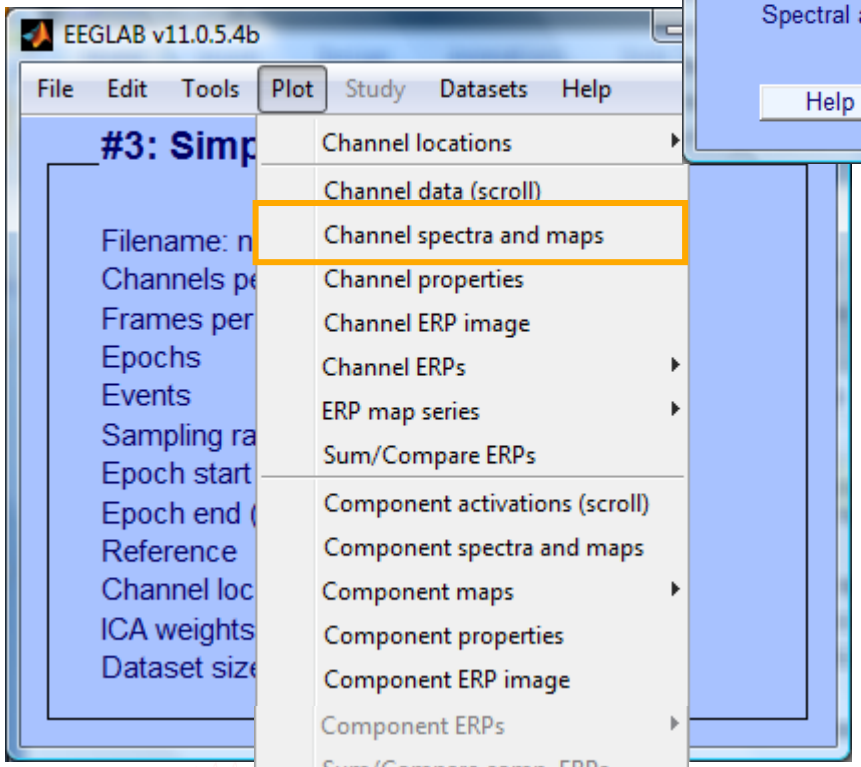
Scale 35

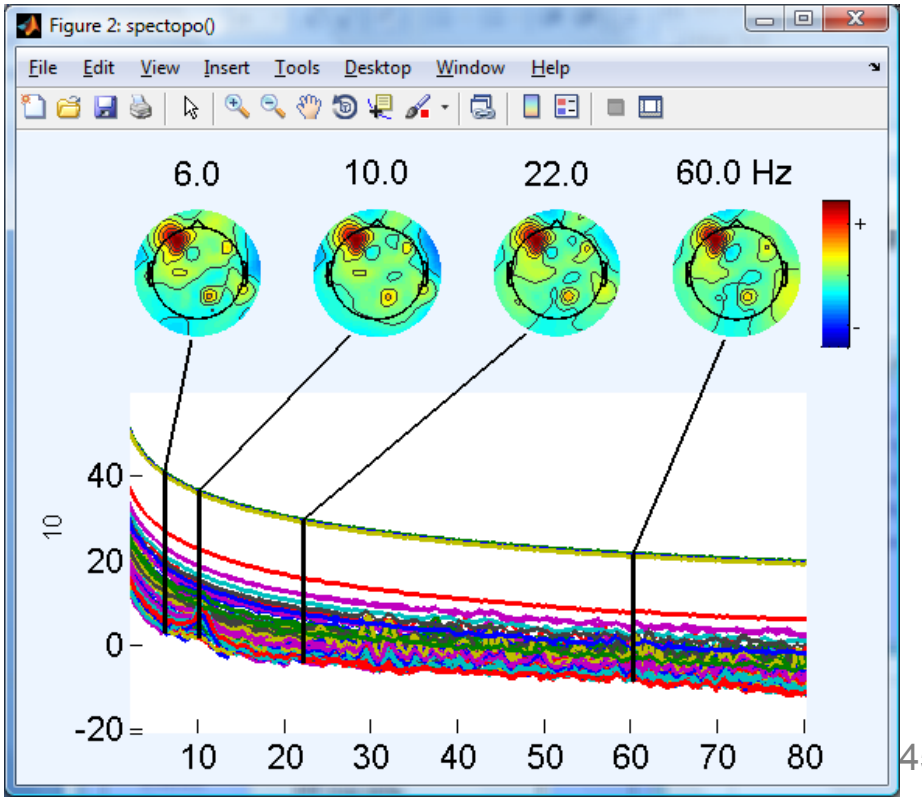
CANCEL Event types << < 449 > >> Chan. Time Value O1 451.0988 3.6619 35 + - REJECT

Manually identifying bad channels

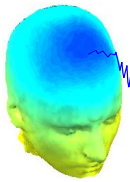


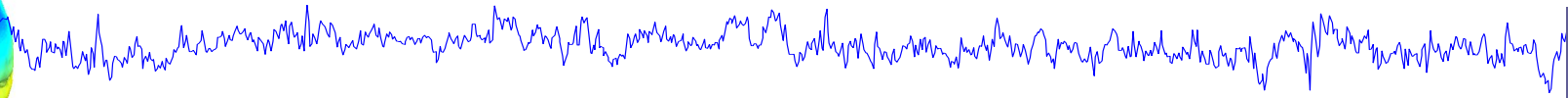


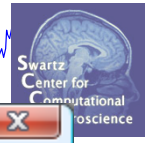


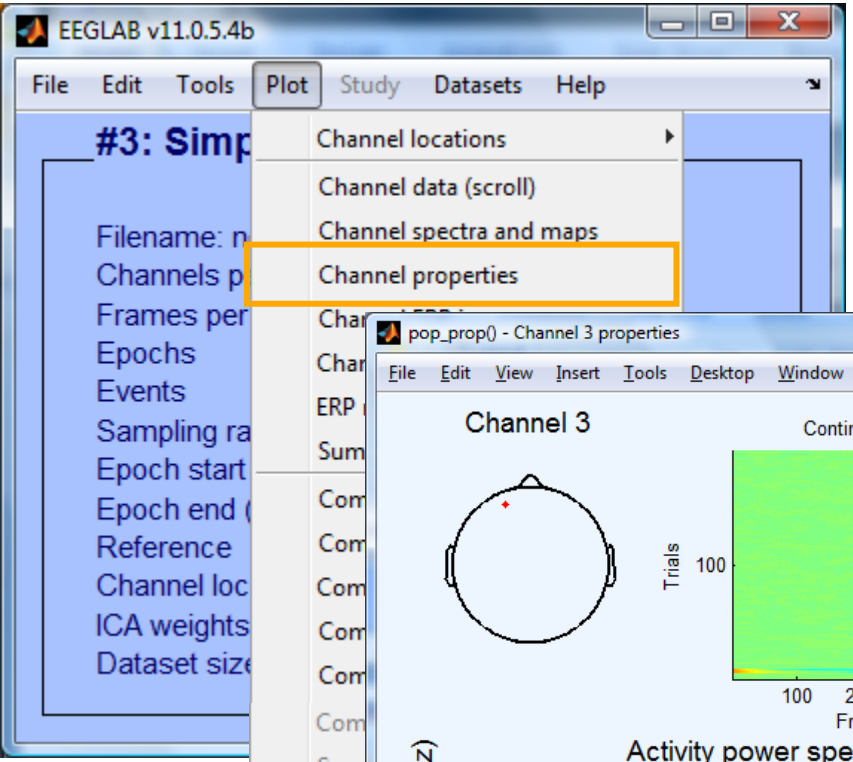


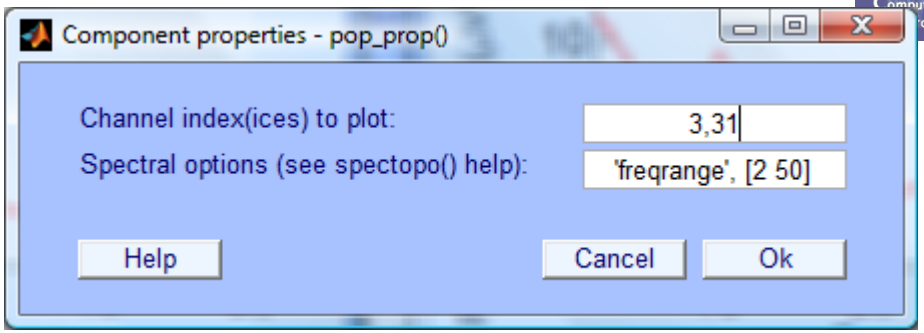
Manually identifying bad channels

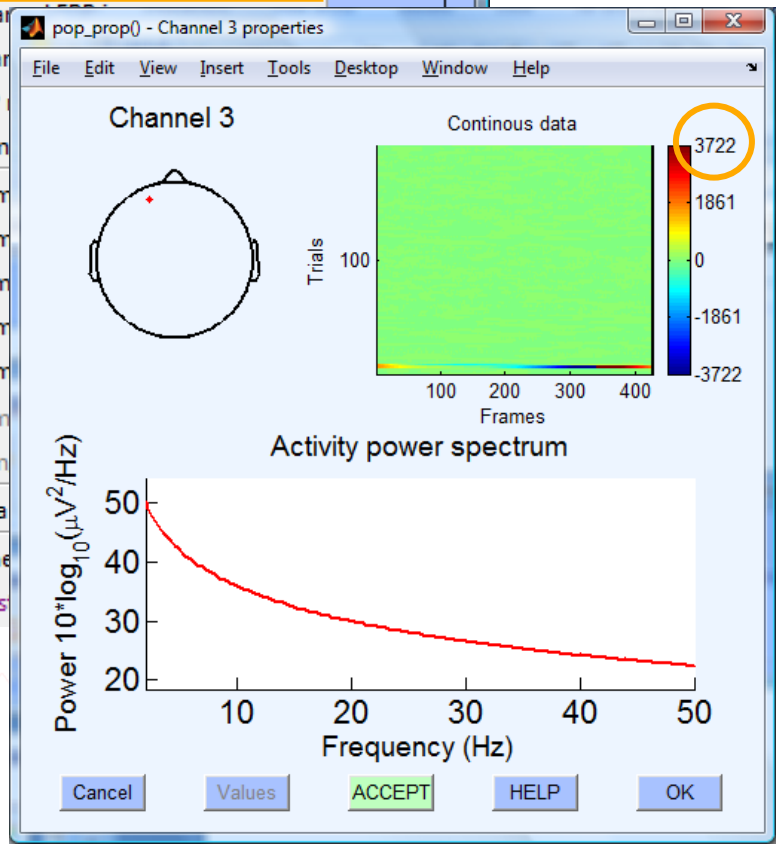


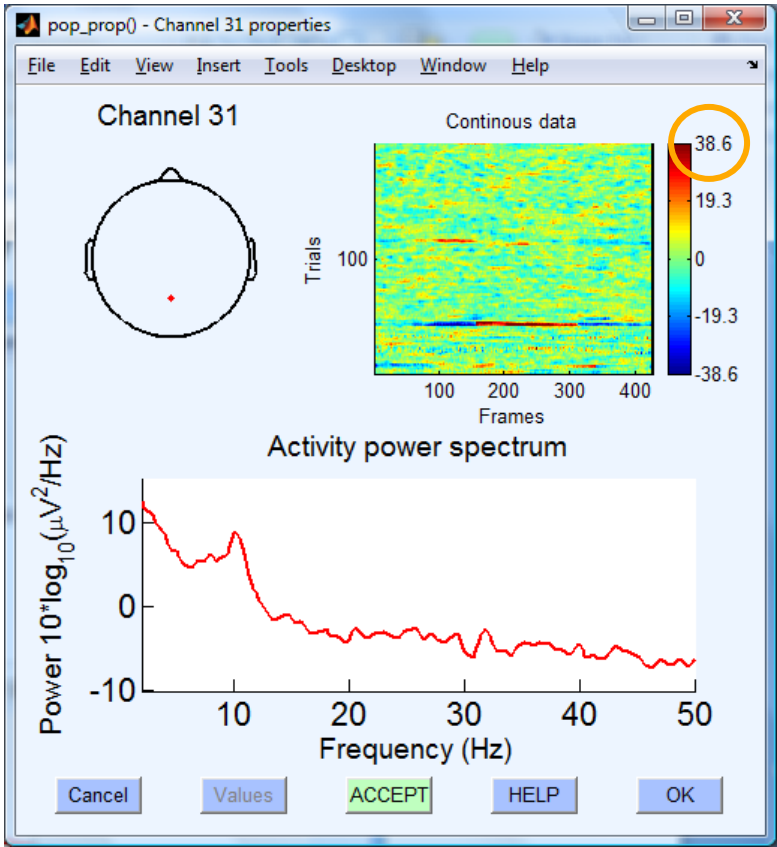




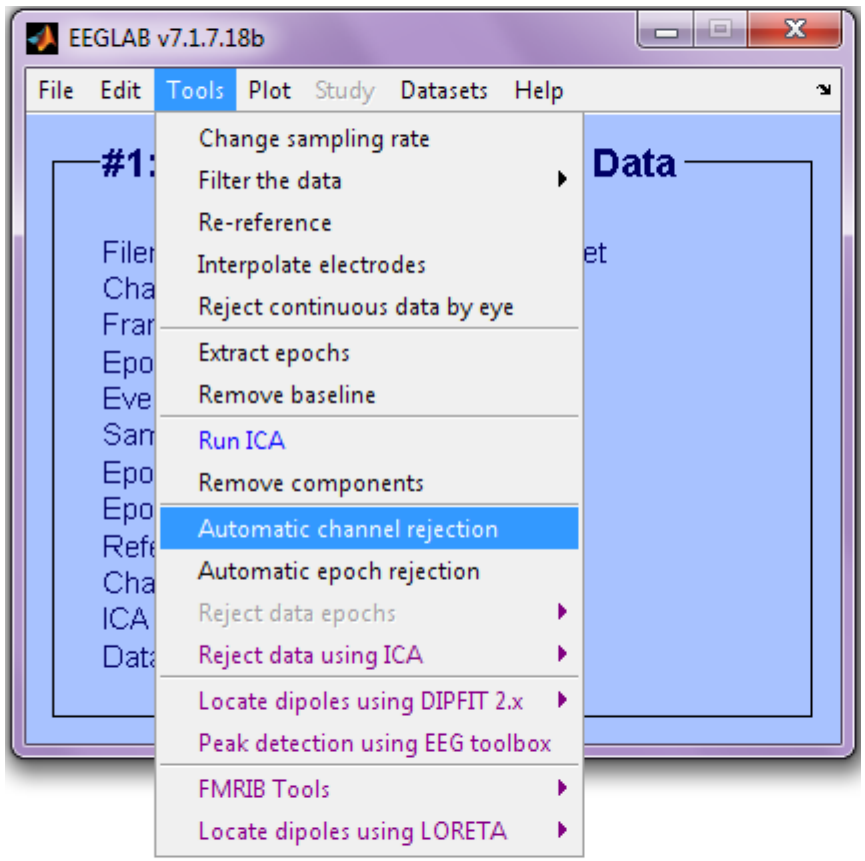


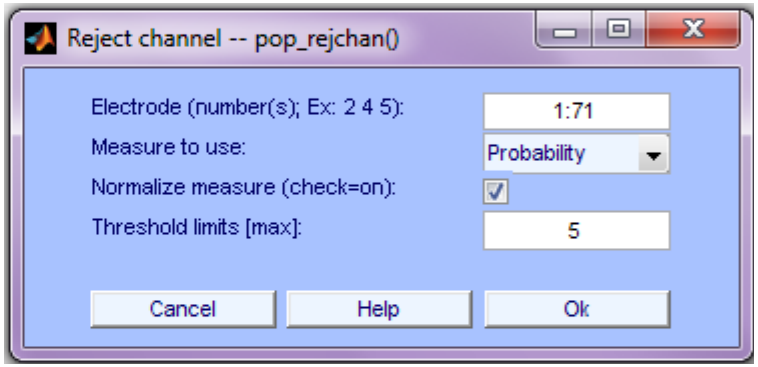




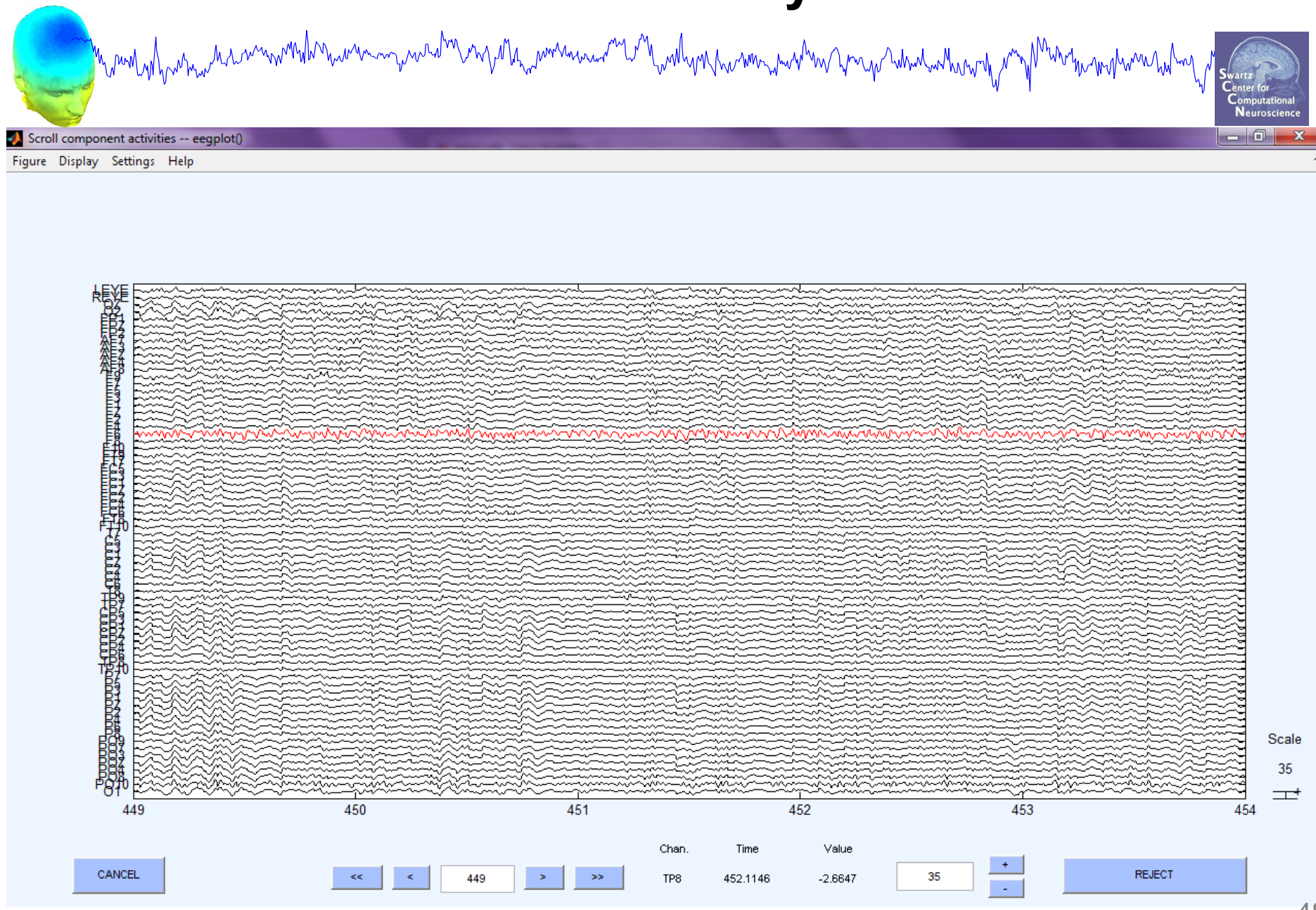


Auto-detection of noisy channels

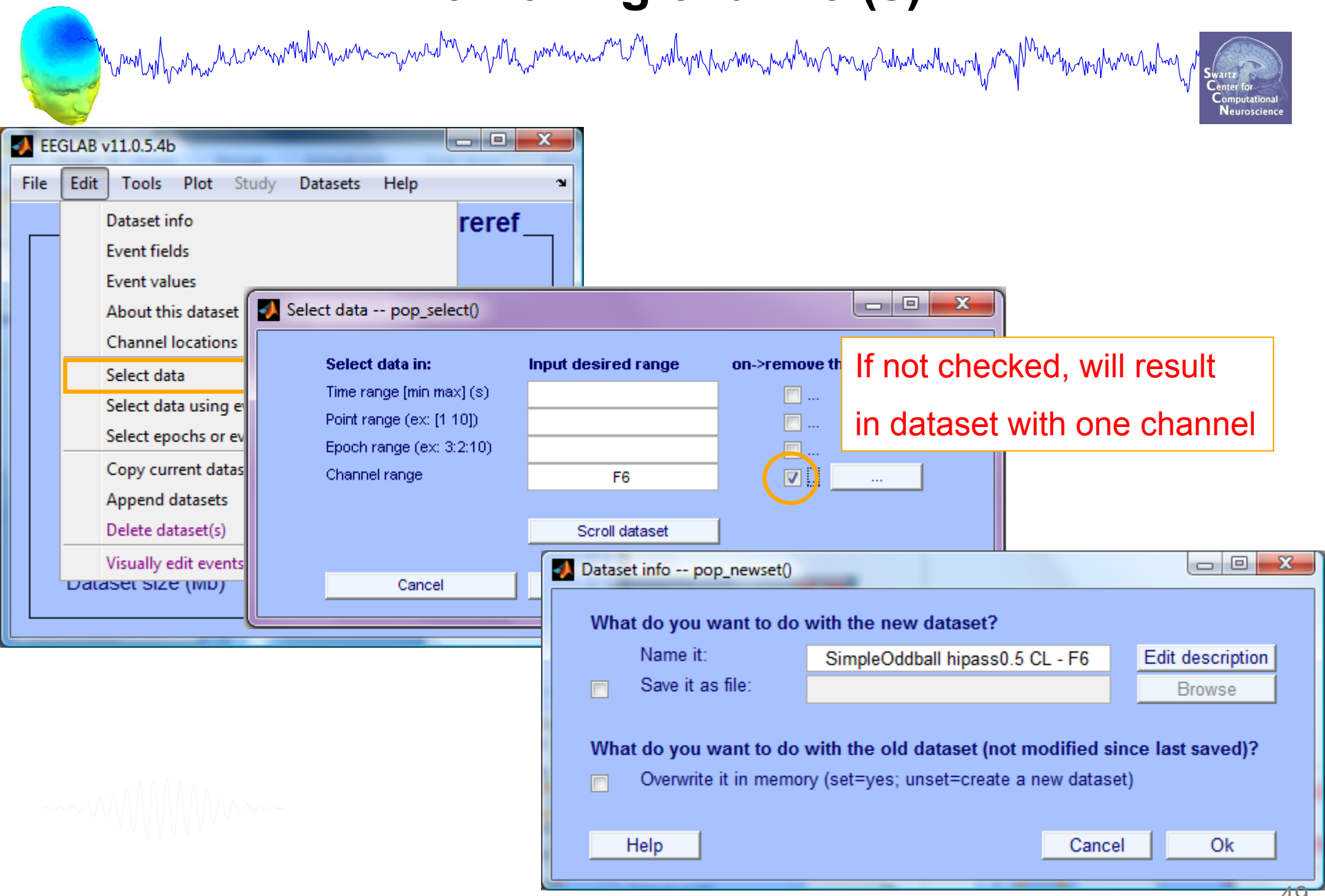




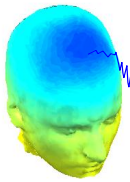
Auto-detected noisy channel

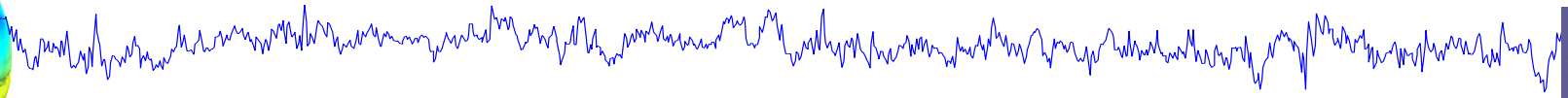


Removing channel(s)



Removing channel(s)

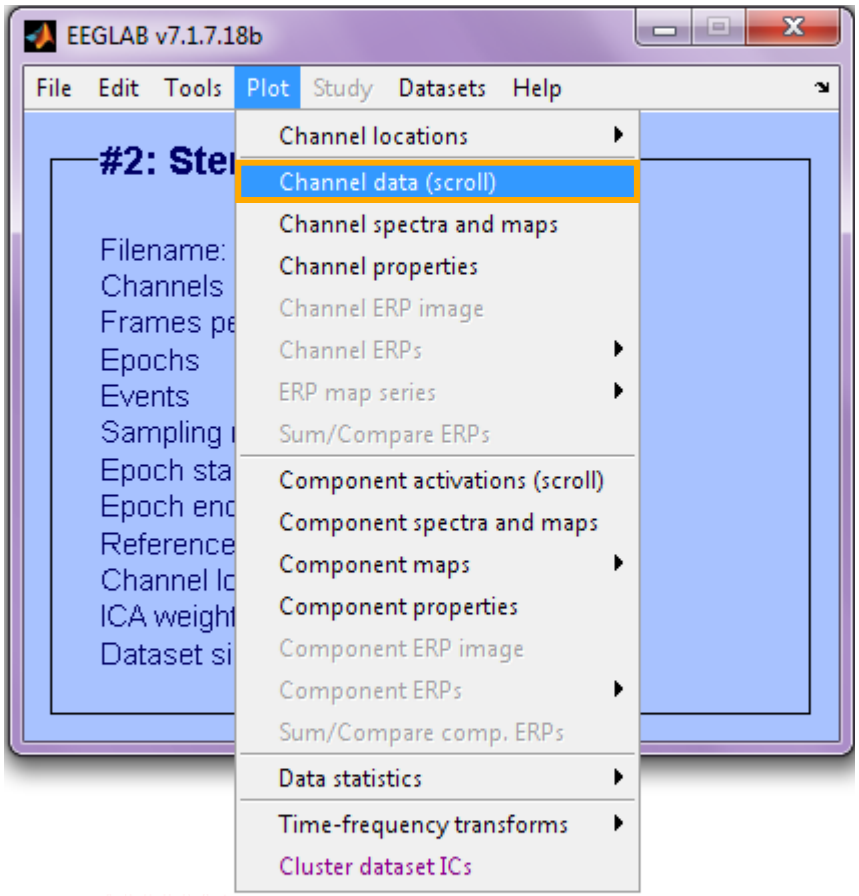


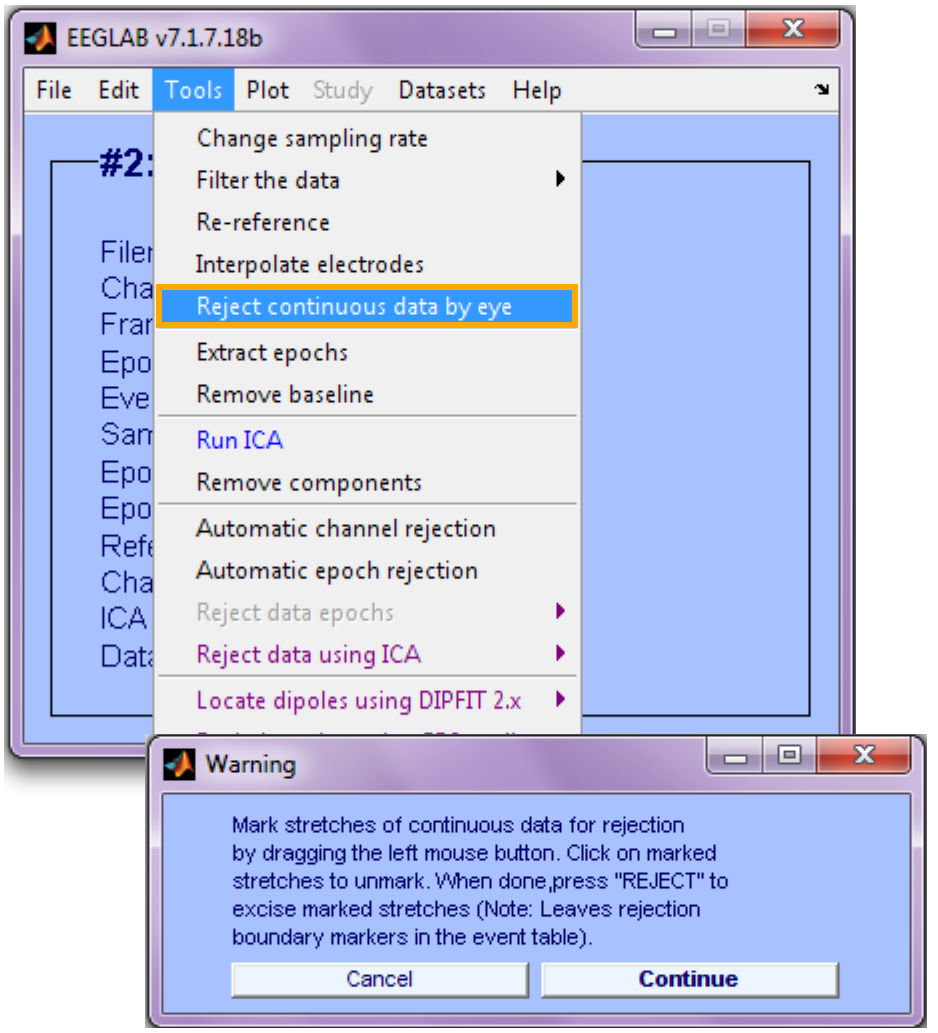


- You may want to interpolate bad channels rather than remove them altogether. Don't do this!
- The loss in dimensionality will affect the ICA decomposition
- Preferred solution:
  - **Delete** the bad channels before running ICA
  - STUDY tools will interpolate missing channels automatically

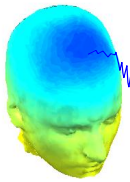
Reject continuous data

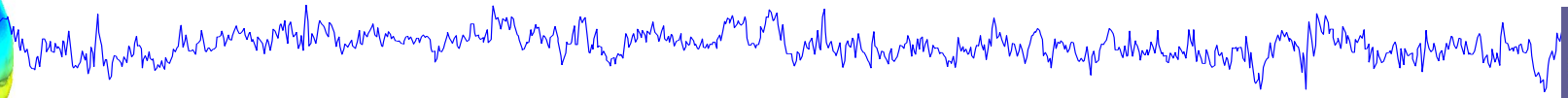
Equivalent





Pre-processing pipeline





Swartz Center for Computational Neuroscience

Collect high-density EEG data (>30 chan)

Import into EEGLAB

Import event markers and channel locations

Re-reference/ down-sample (if necessary)

High pass filter (~.5 – 1 Hz)

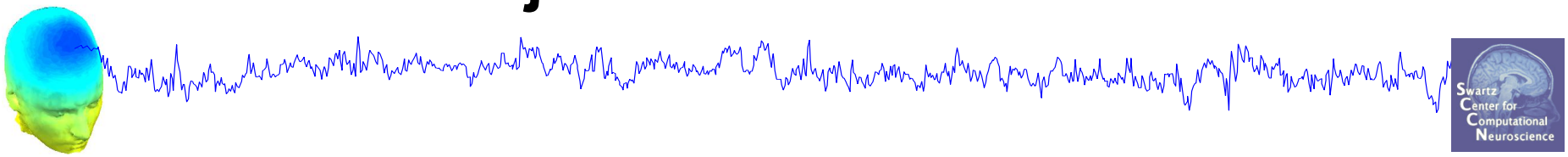
Examine raw data

Data Cleaning: Identify/reject bad channels

Data Cleaning: Reject large artifact time points

Run ICA

Reject continuous data



Scroll channel activities -- eegplot()

Figure Display Settings Help

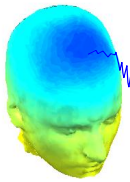
Click and drag with mouse over noisy data to reject

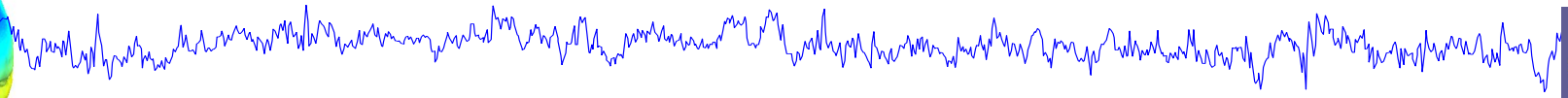
536 537 538 539 540 541

Scale 35

CANCEL | Event types | << < 536 > >> | Chan. FC6 | Time 539.9355 | Value 4.8773 | 35 + - | REJECT

Rejecting data for ICA



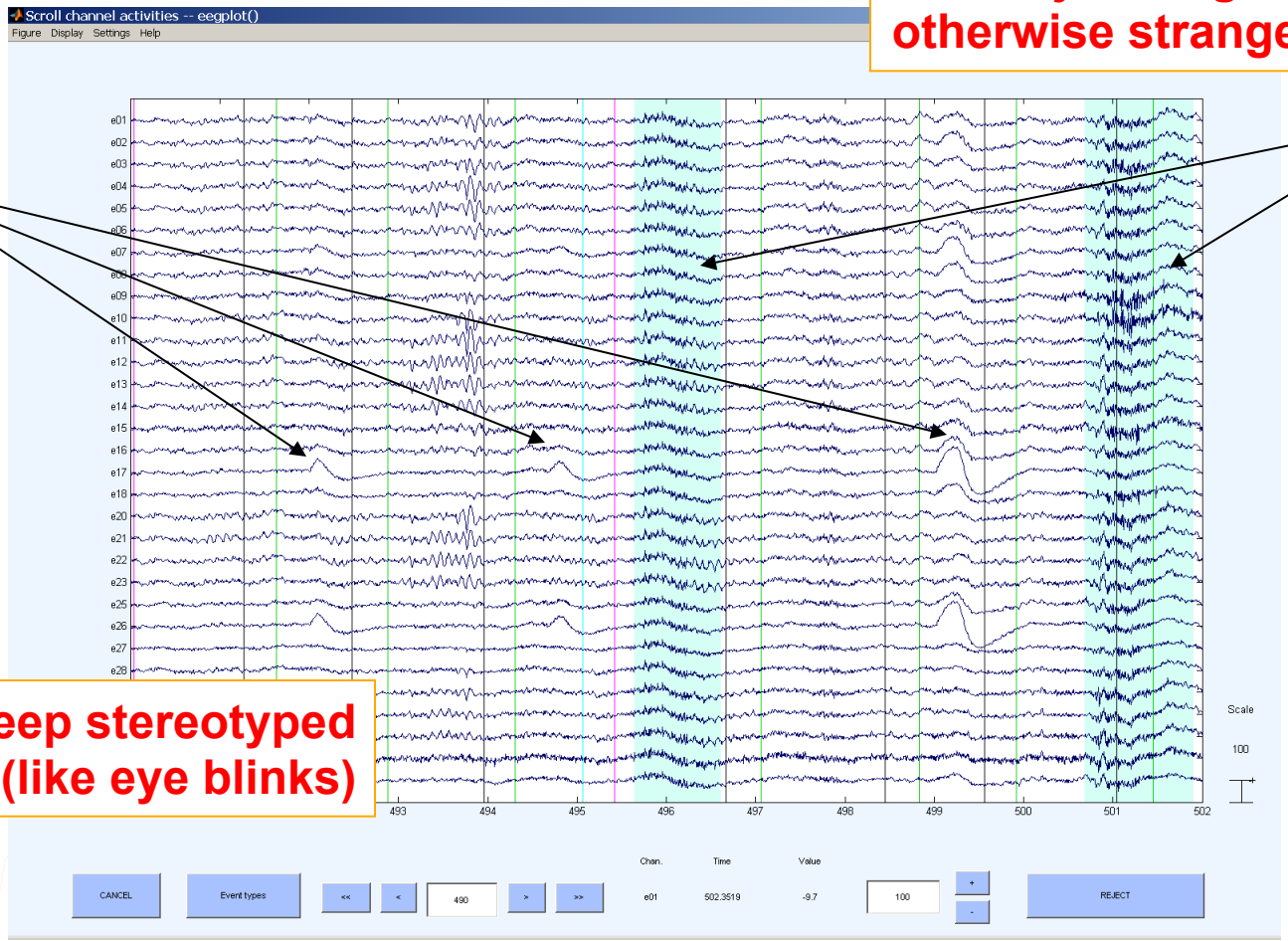


To prepare data for ICA:

Reject large muscle or otherwise strange events...

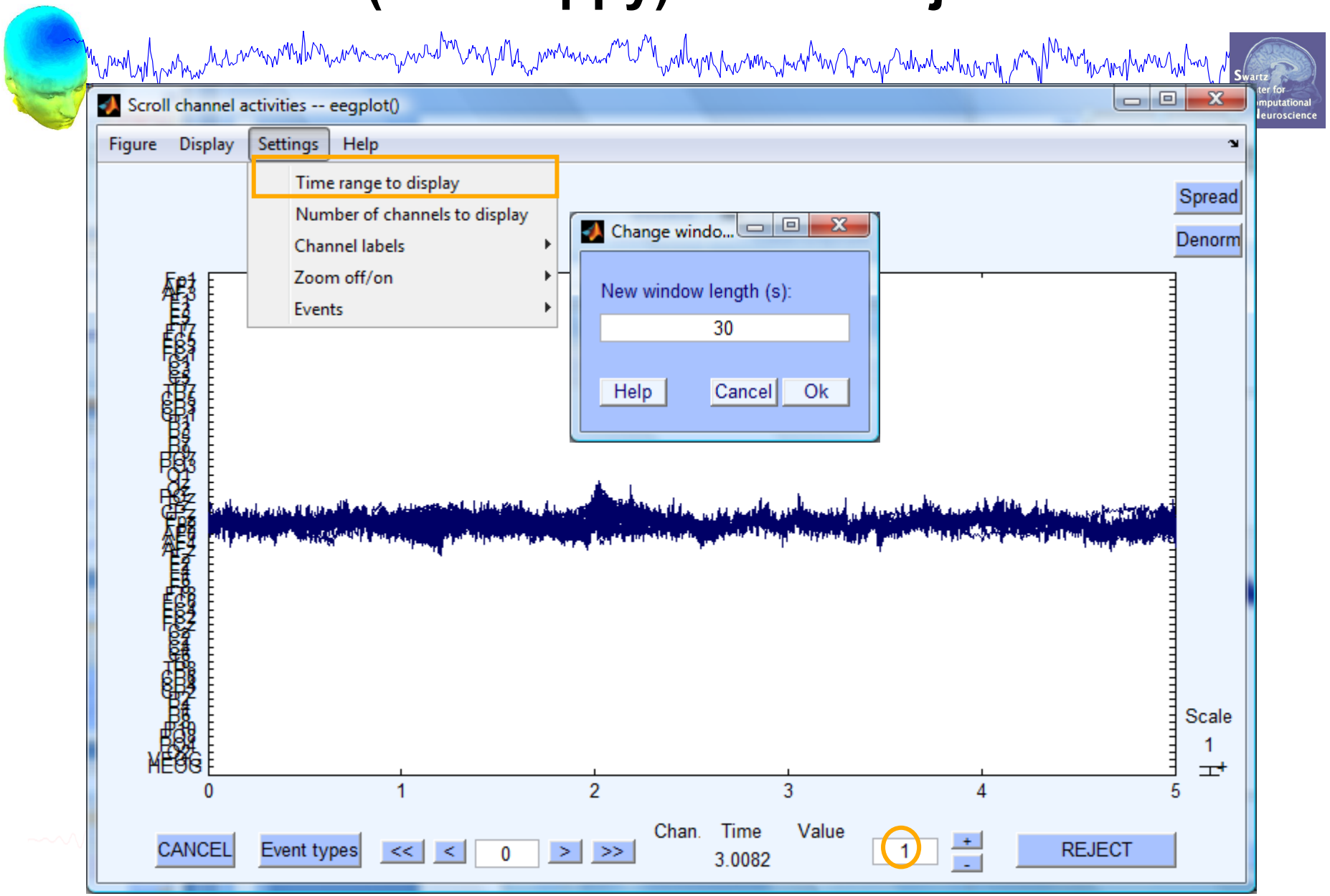
Keep

Reject

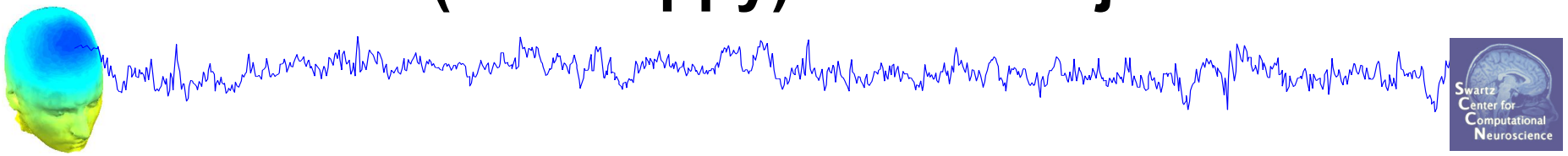


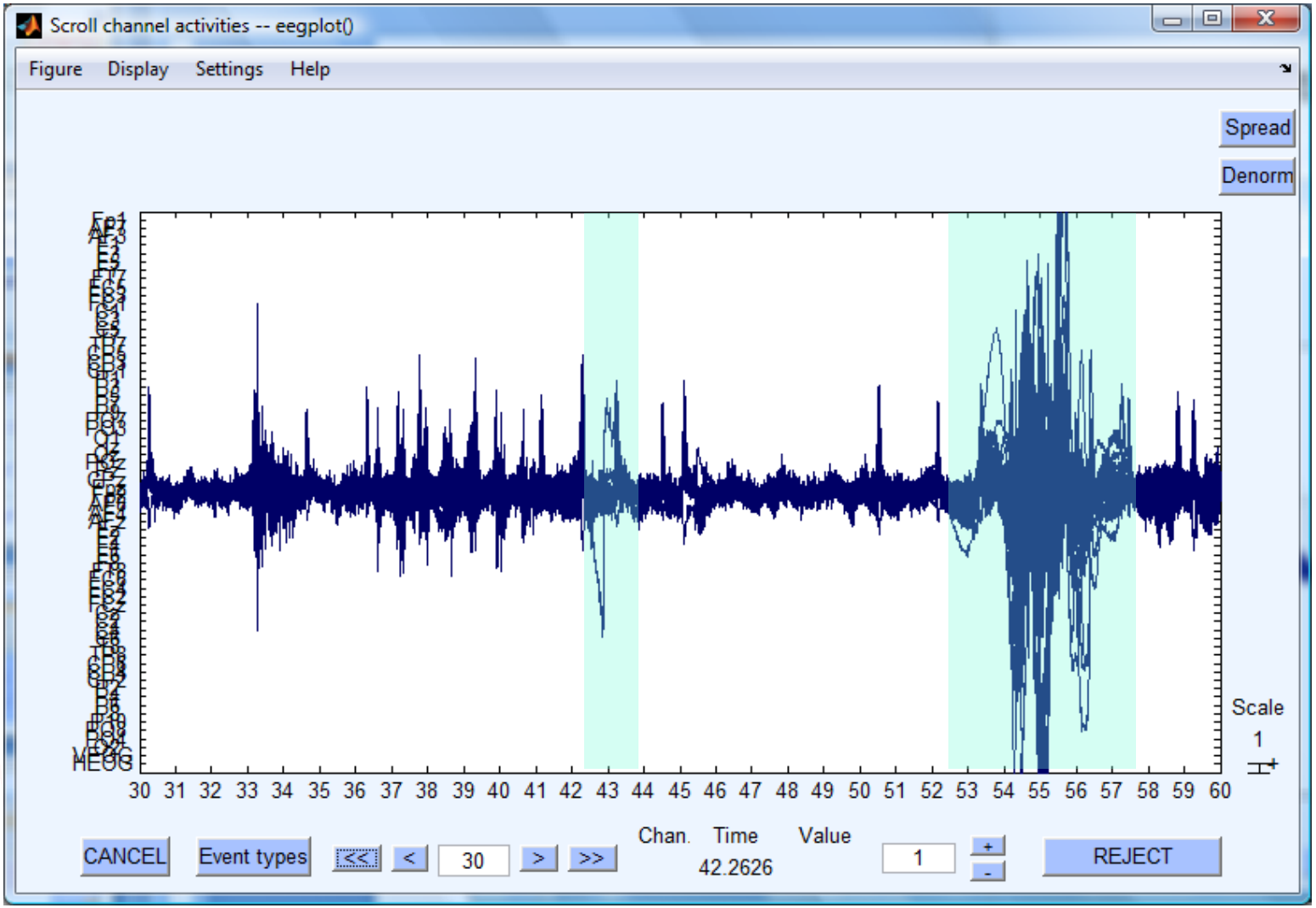
... but keep stereotyped artifacts (like eye blinks)

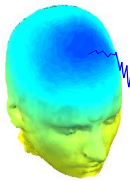
Fast (but sloppy) artifact rejection

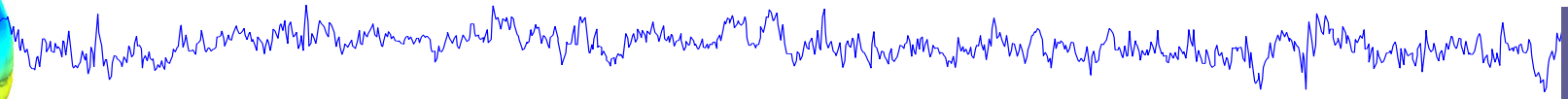


Fast (but sloppy) artifact rejection





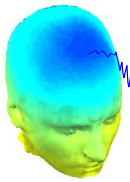


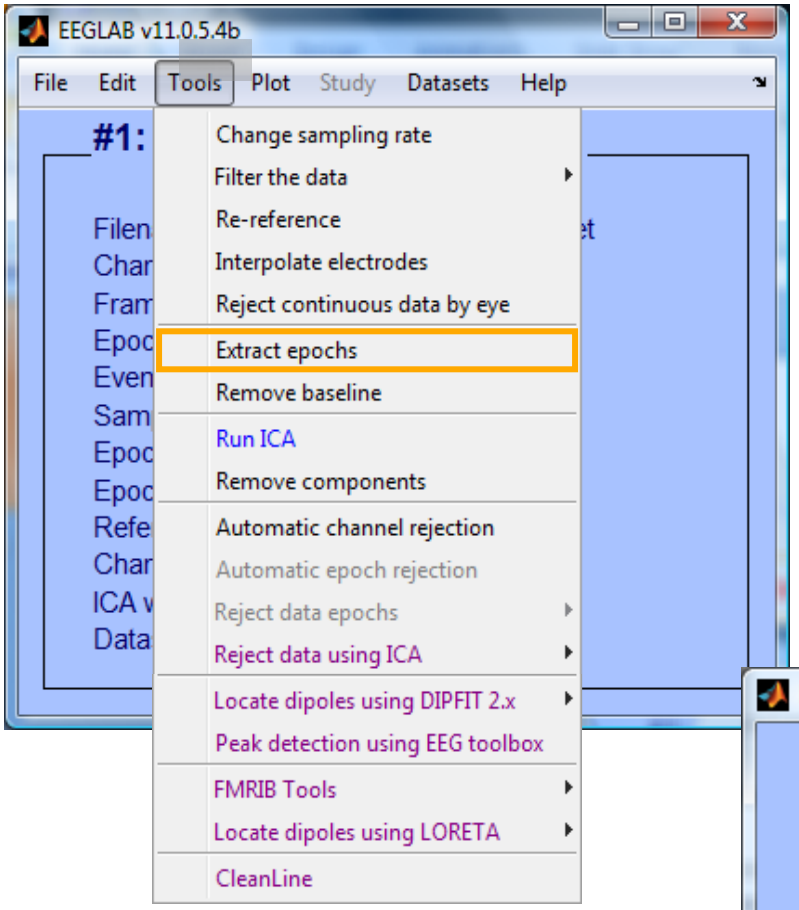


Data Cleaning for ICA

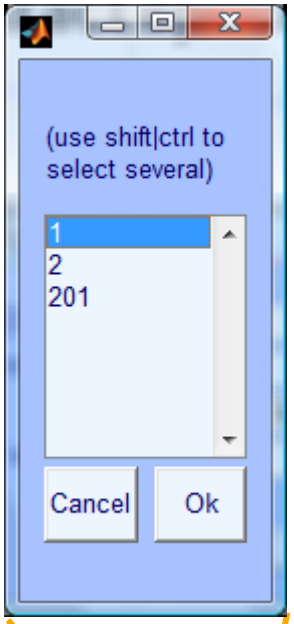
Variant 2: Epoched Data

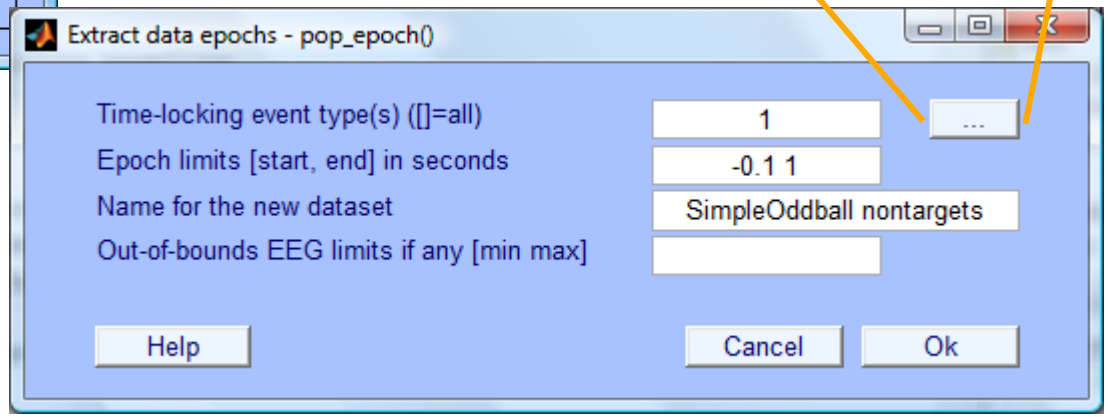
Extract epochs



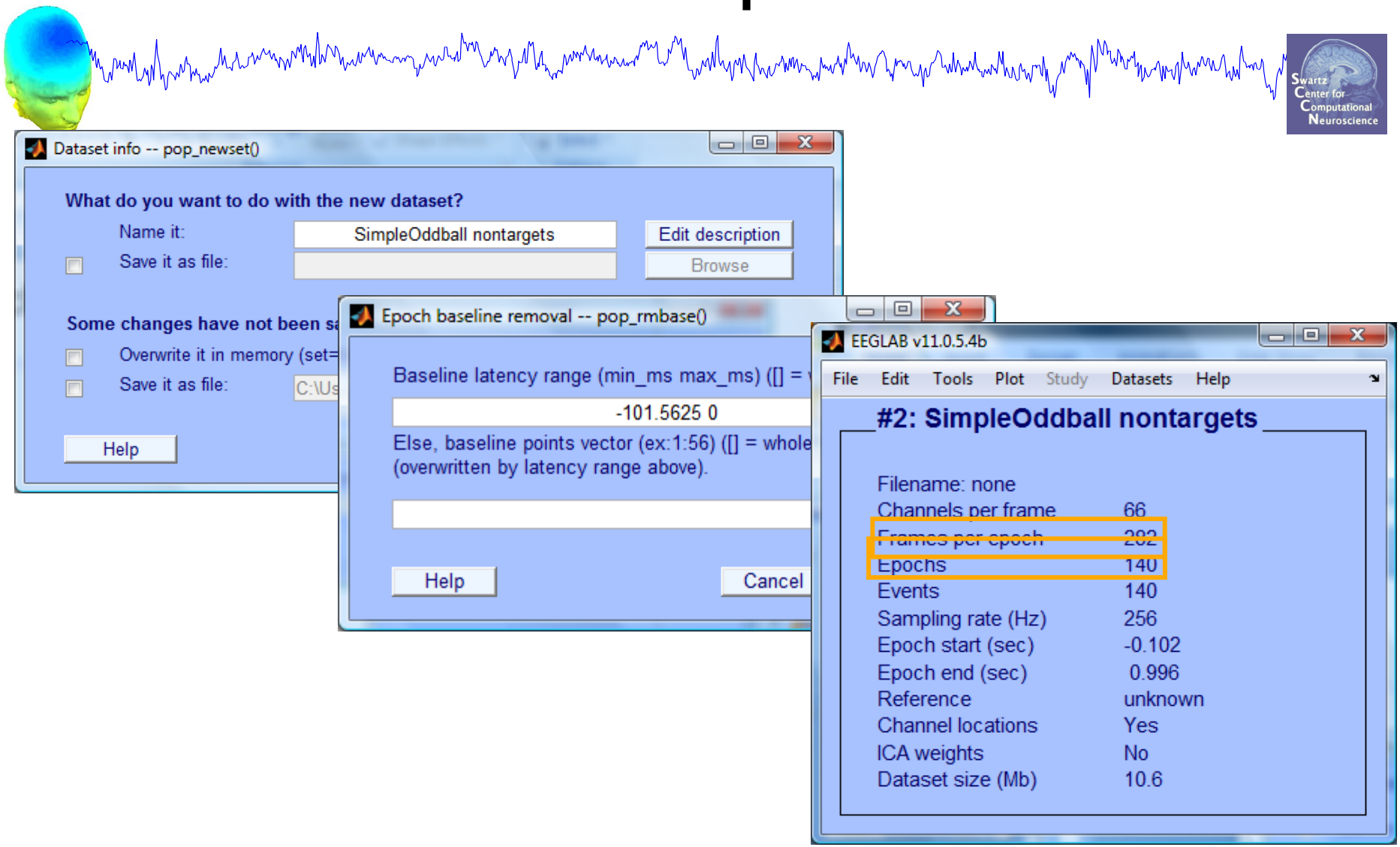



```
>> eeg_eventtypes(EEG)

   1       140

   2       60

   201     60
```

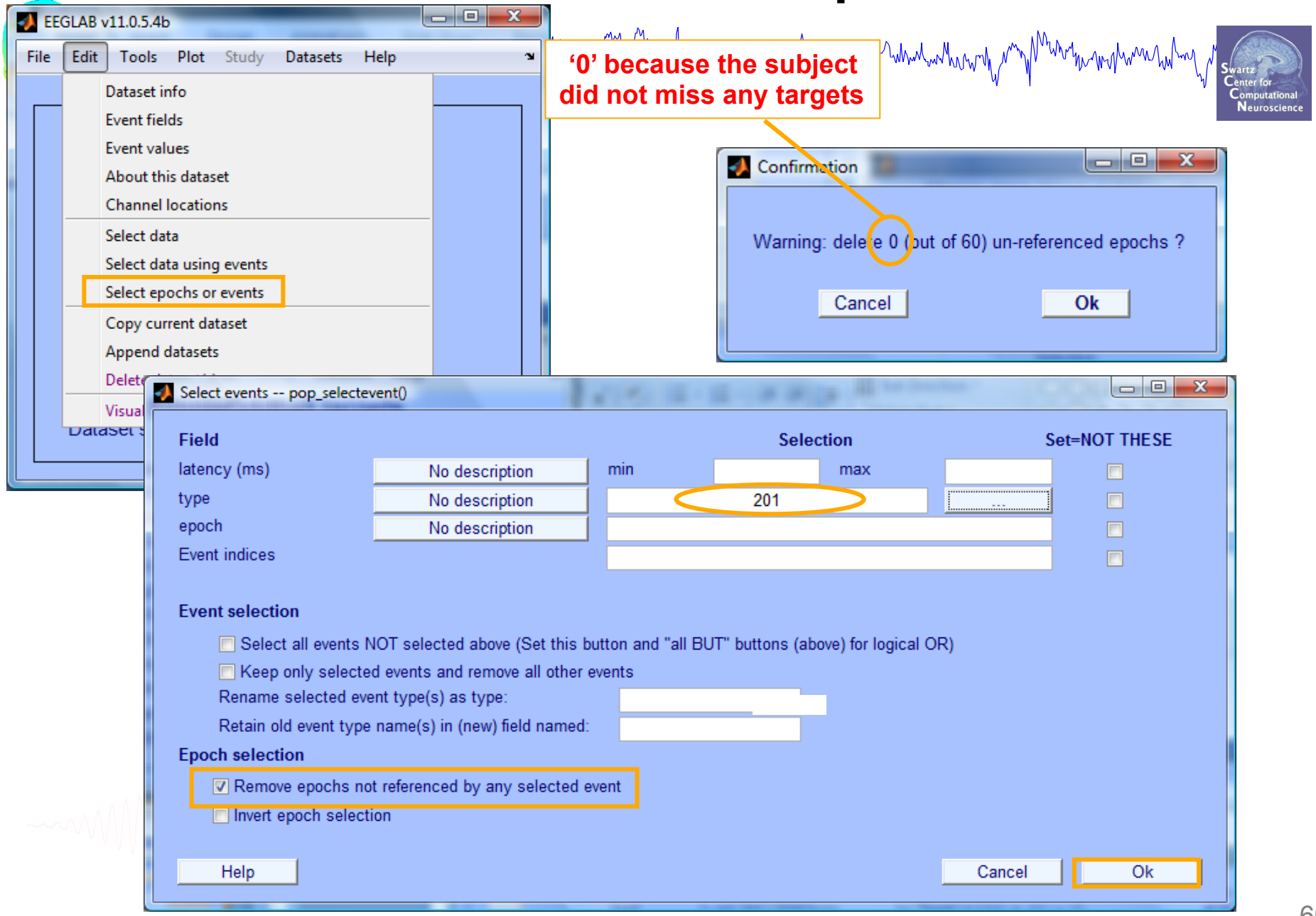


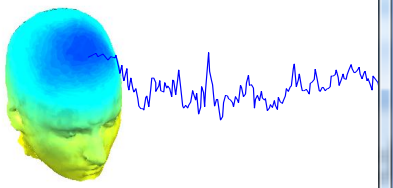


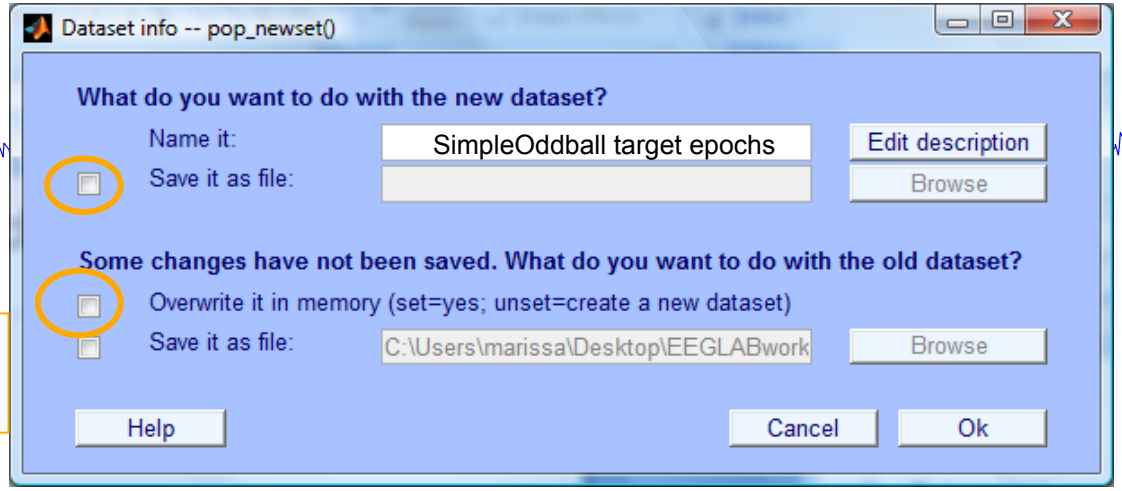
Extract epochs



Select a subset of epochs

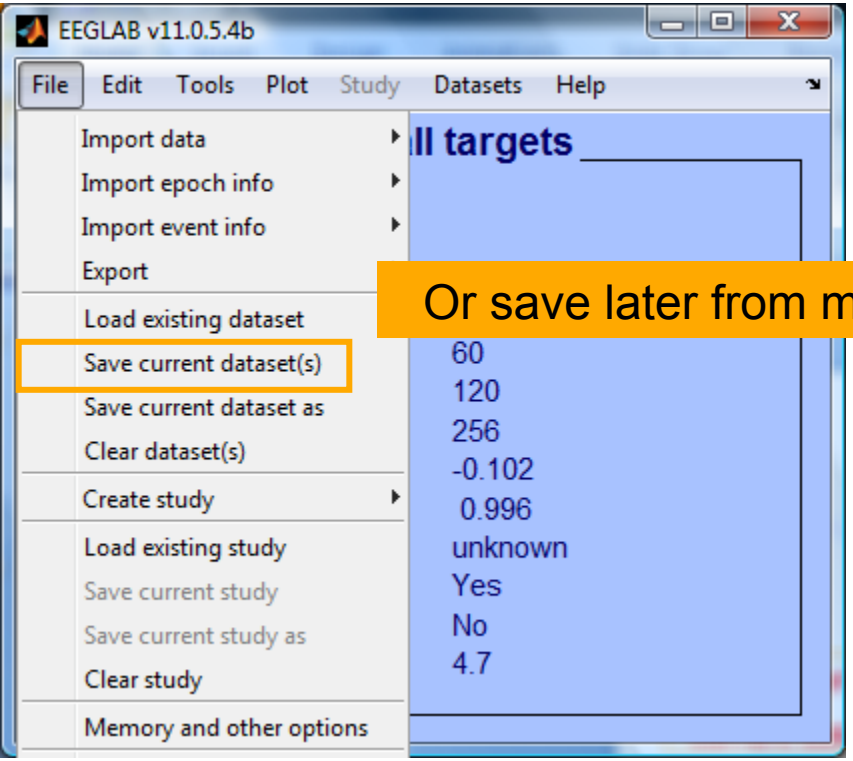




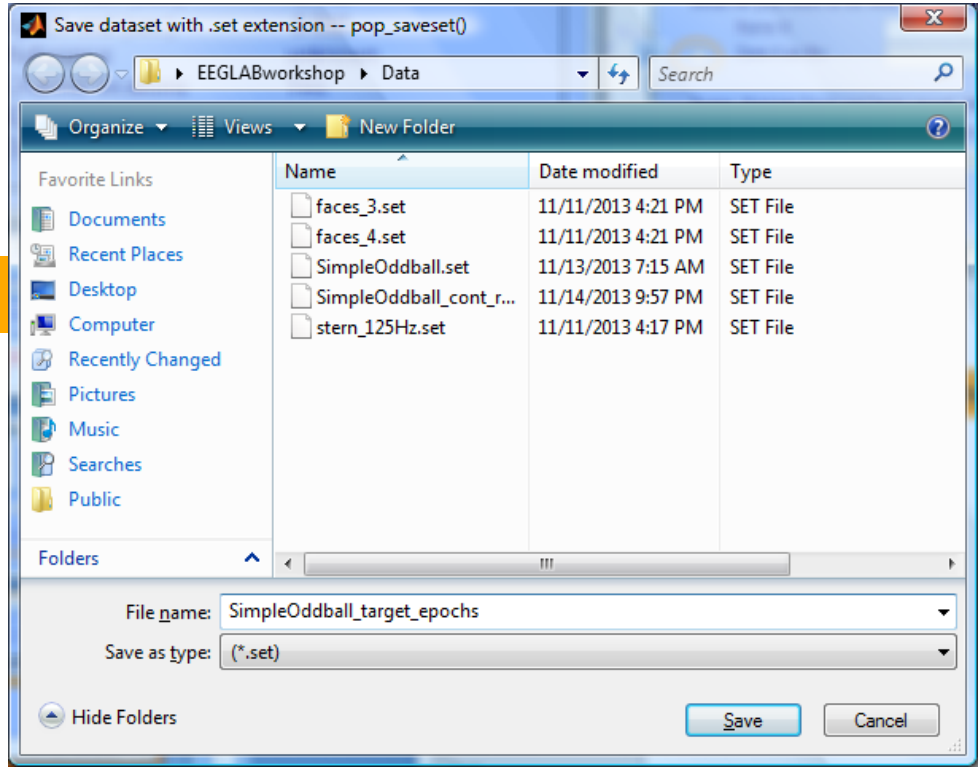


'Do not overwrite current dataset'

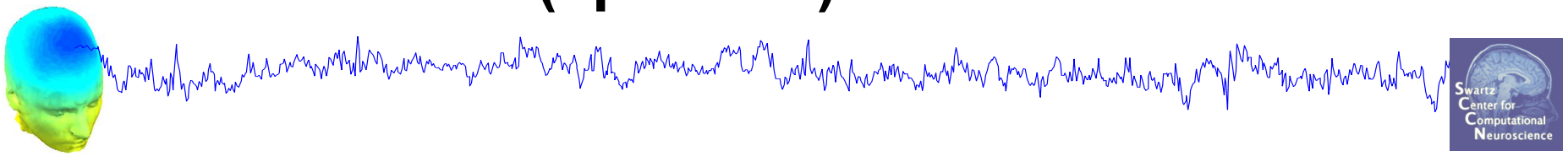
Save dataset (optional)



Or save later from menu



Scroll (epoched) channel data



EEGLAB v11.0.5.4b

File Edit Tools Plot Study Datasets Help

#3: Simp

Filename: n
Channels pe
Frames per
Epochs
Events
Sampling ra
Epoch start
Epoch end (
Reference
Channel loc
ICA weights
Dataset size

Channel loca
Channel data
Channel spe
Channel pro
Channel ERP
Channel ERP
ERP map seri
Sum/Compa
Component
Component
Component
Component
Component
Component
Sum/Compa
Data statistic
Time-freque
Cluster datas

Scroll channel activities -- eegplot()

Figure Display Settings Help

Stack
Norm

2 201 1 2 201 2 2 201 3 2 201 4 2 201 5

Fp1 AF3 F3 F7 FC5 FC1 C3 C7 CP5 CP1 P3 P7 PO7 O1 PO3 Oz PO4 O2 P8 P4 CP6 CP2 C4 C8 T8 FC6 FC2 F4 F8 AF4 Fp2 MEG

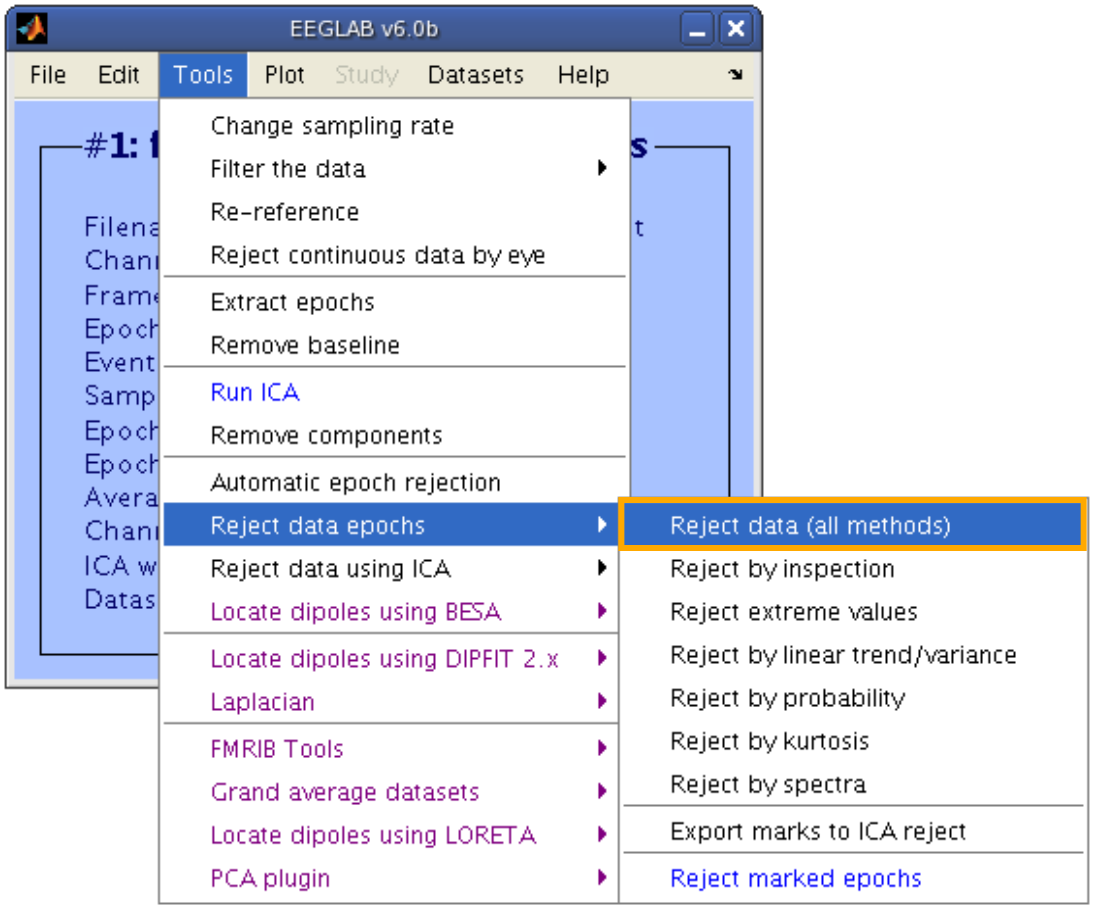
0 250 500 0 250 500 0 250 500 0 250 500 0 250 500

Scale
28

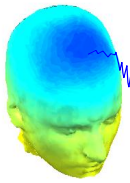
CANCEL Event types << < 1 > >> Chan. Time Value 0.00 0.00 28 + - REJECT

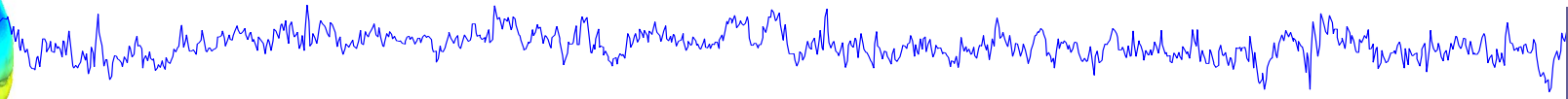
Reject epochs with artifact

Reject data epochs

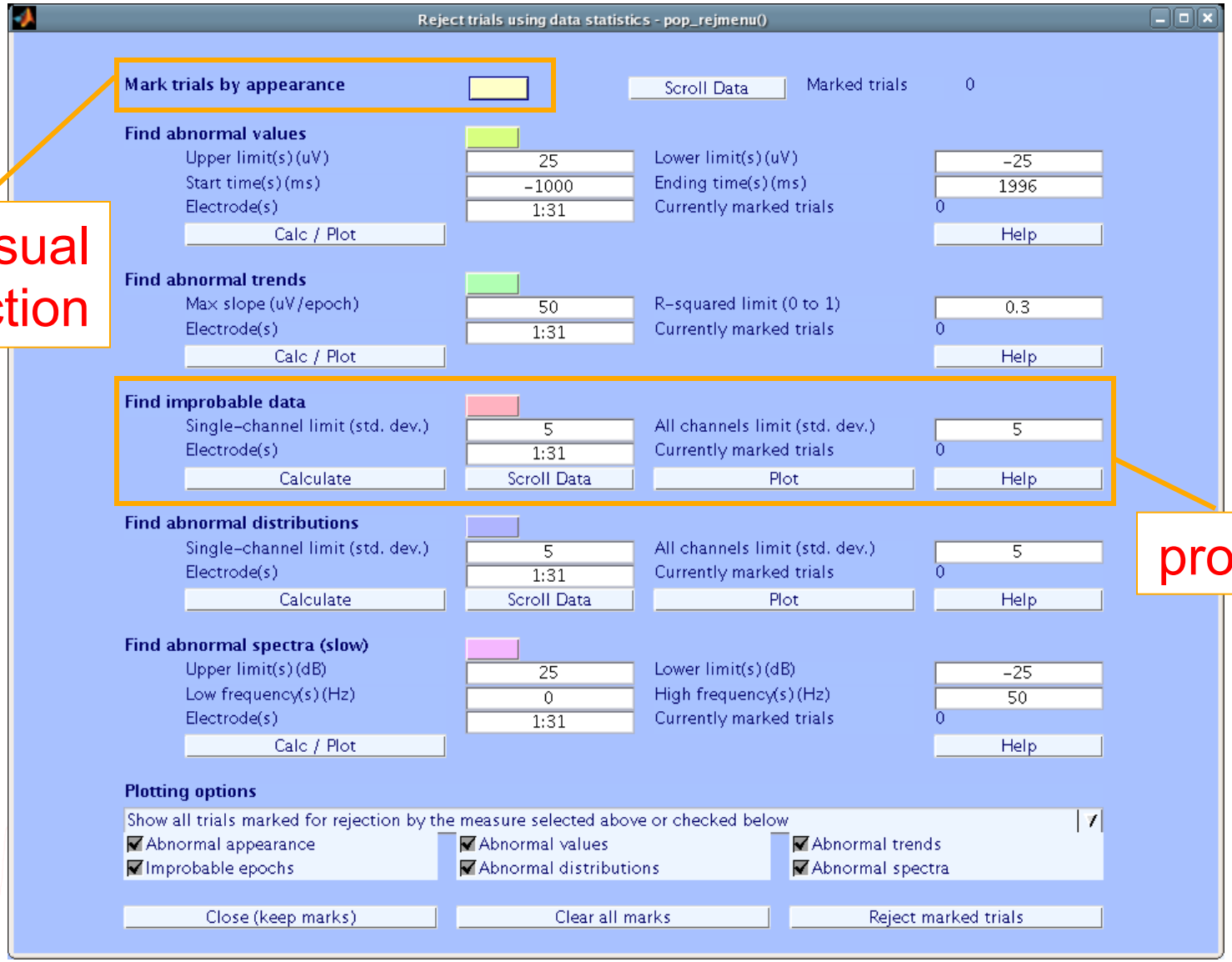


Reject data epochs



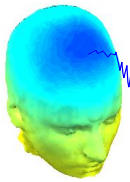


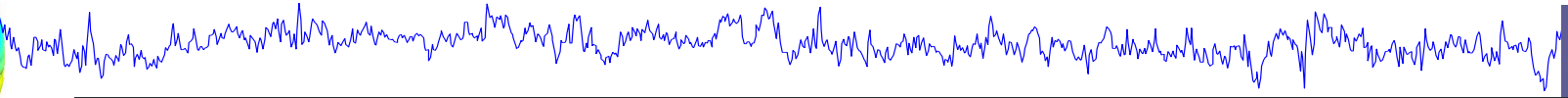
visual inspection

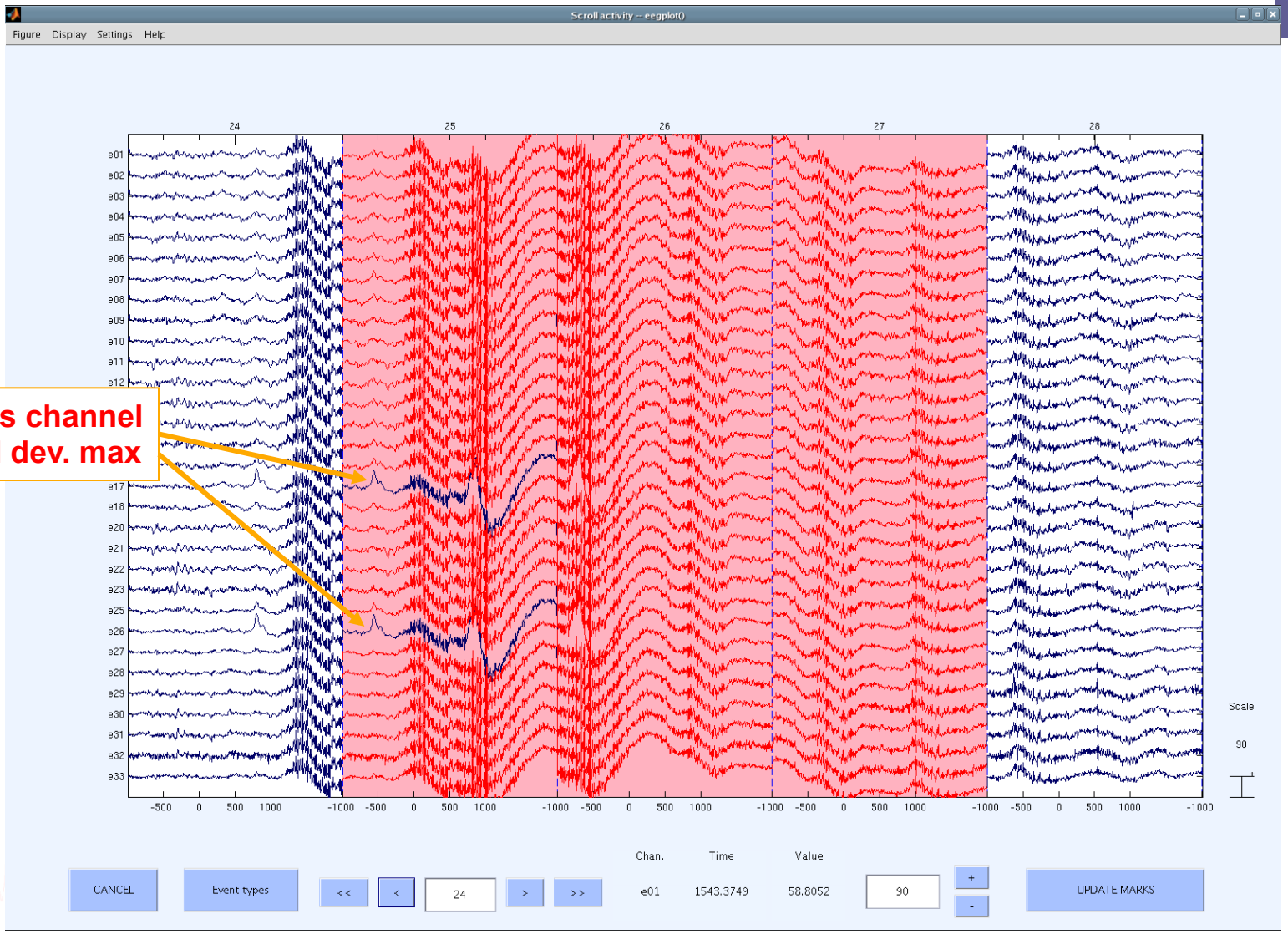


probability

Reject data epochs

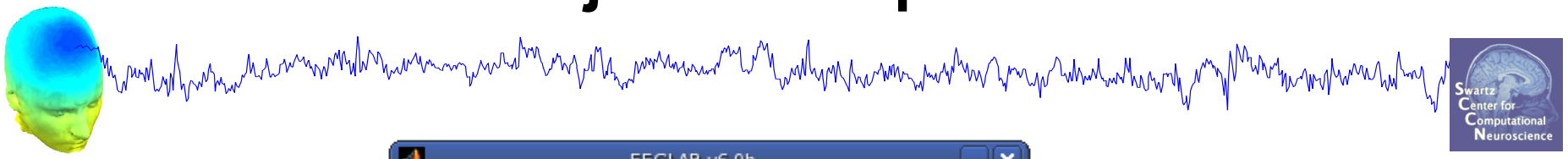




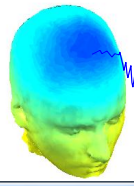


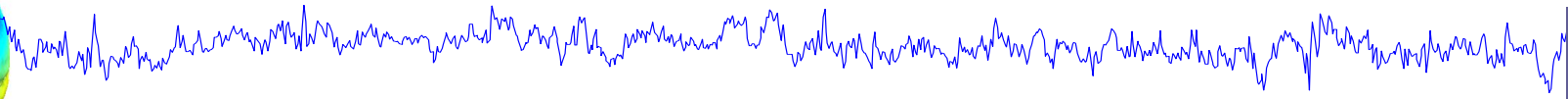
Exceeds channel standard dev. max

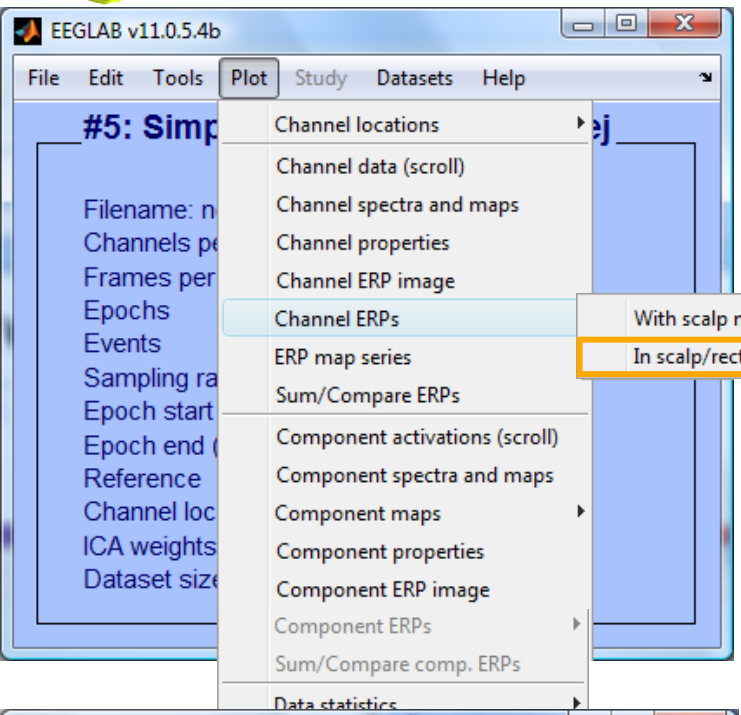
Reject data epochs

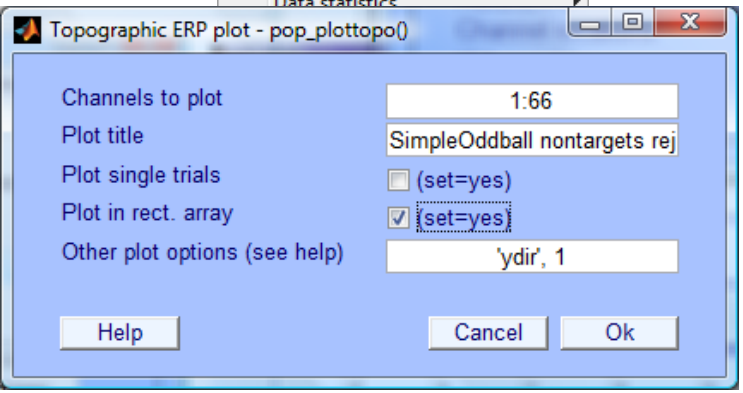


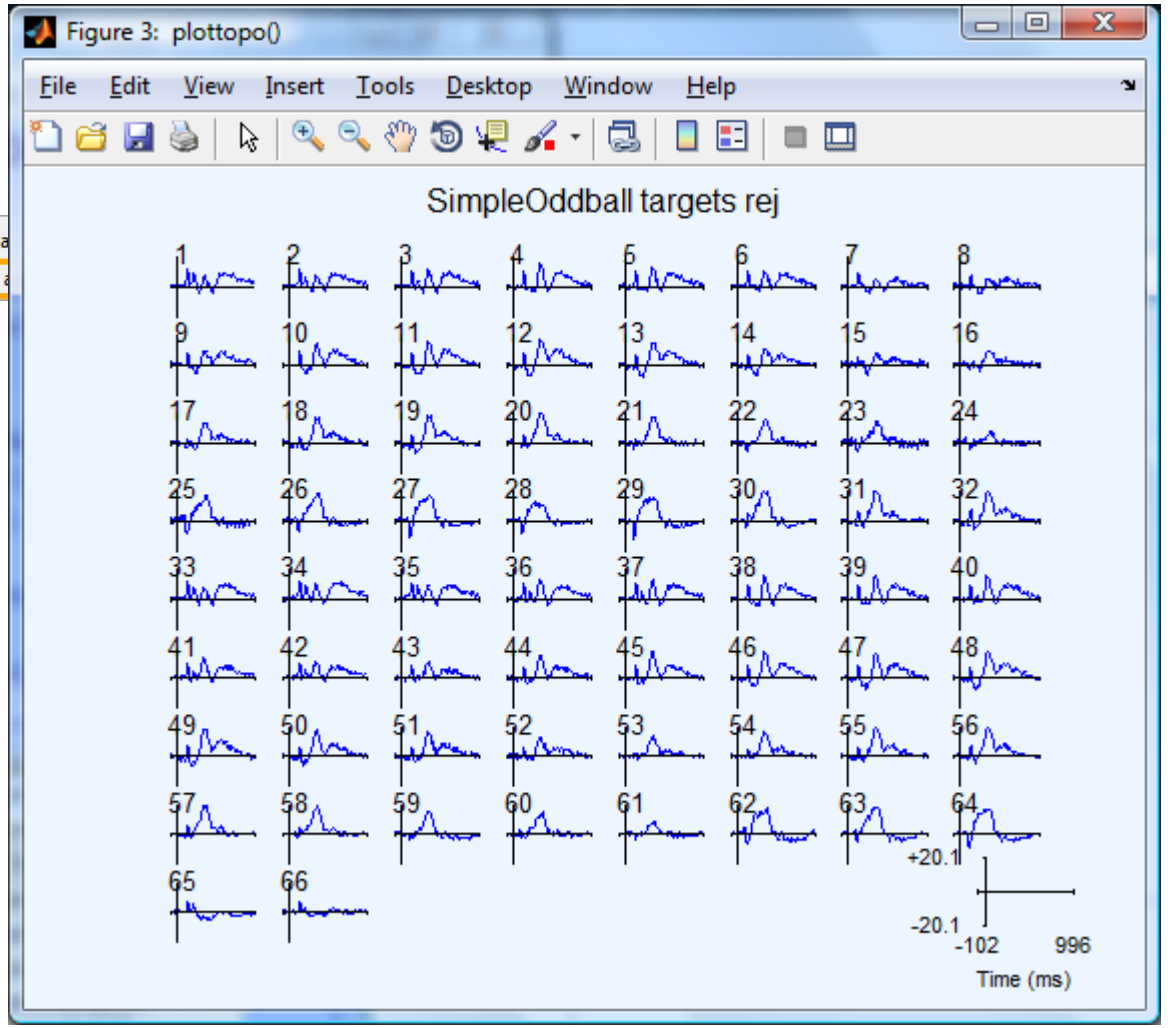
Visualize ERP in rectangular array



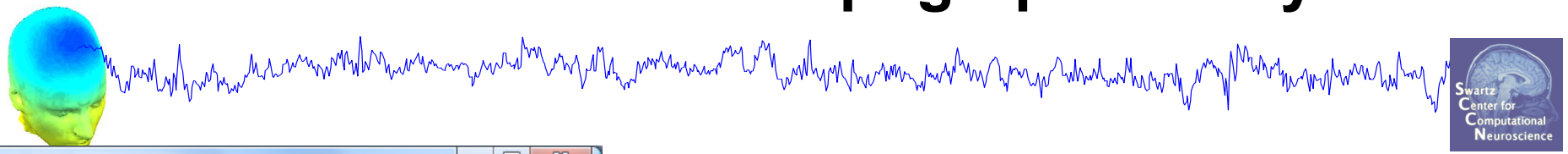


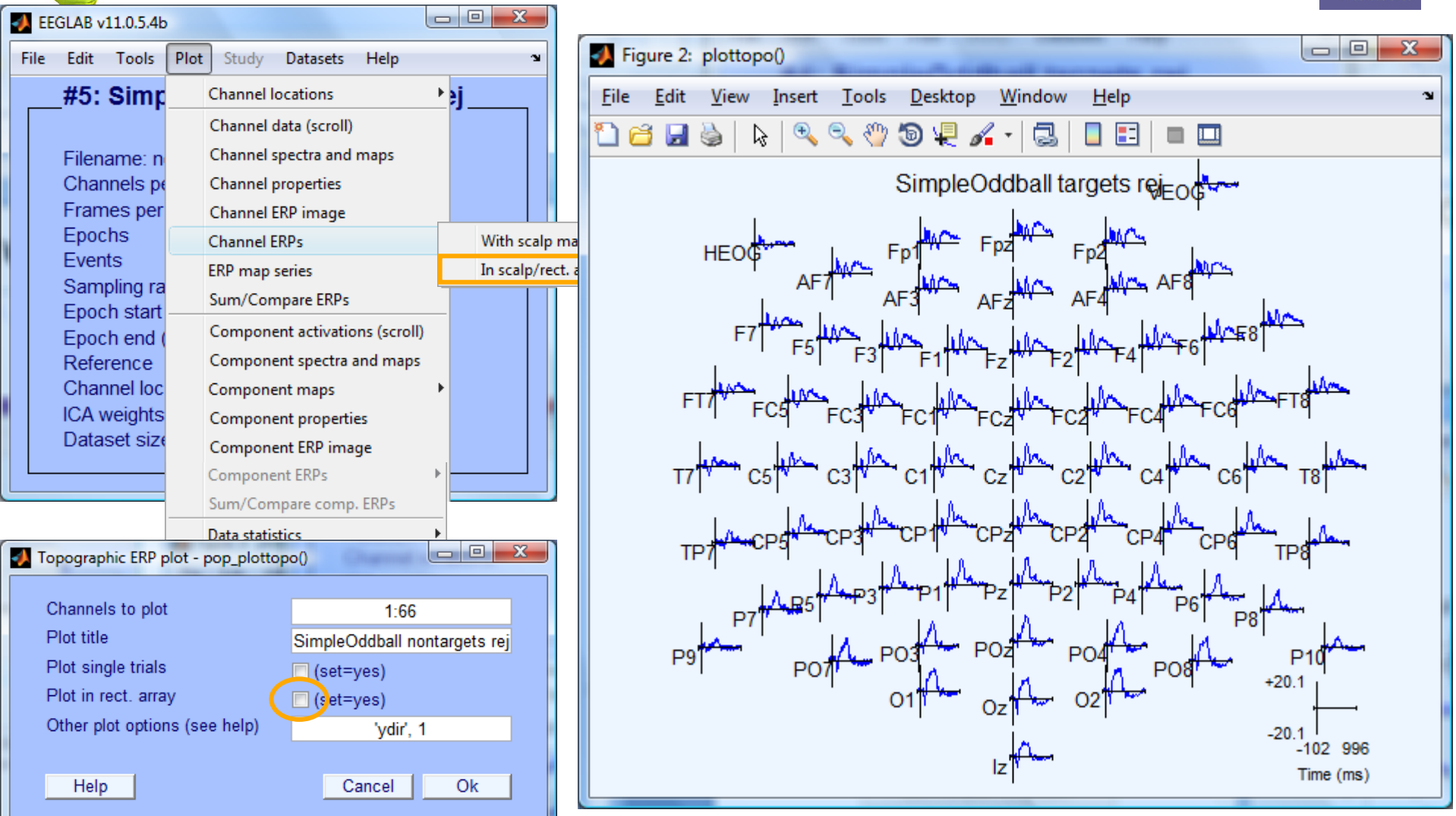




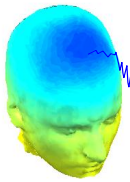


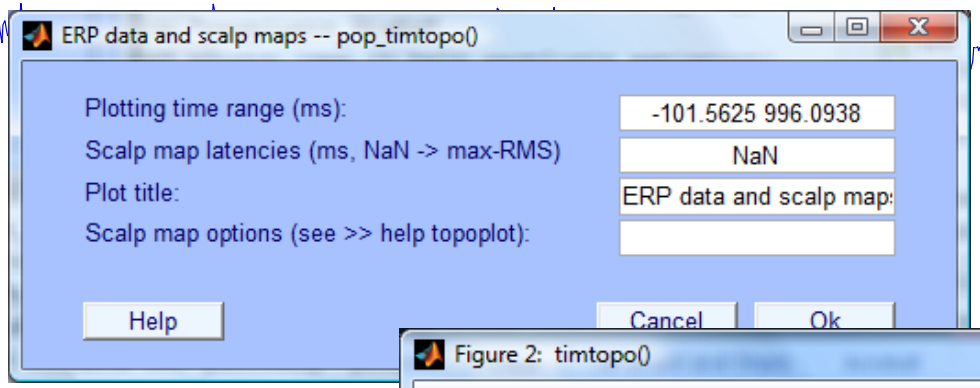
Visualize ERP in topographic array

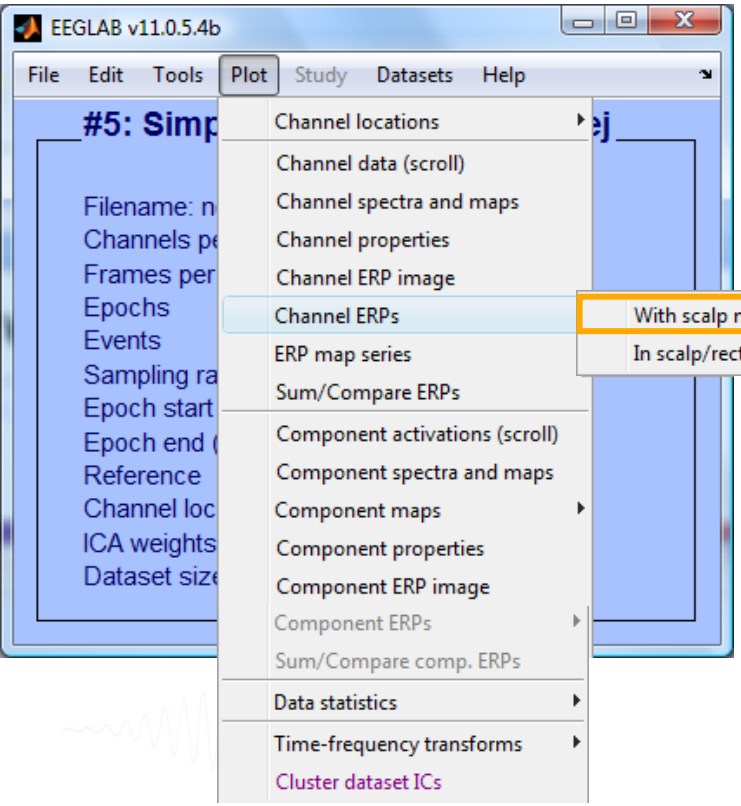


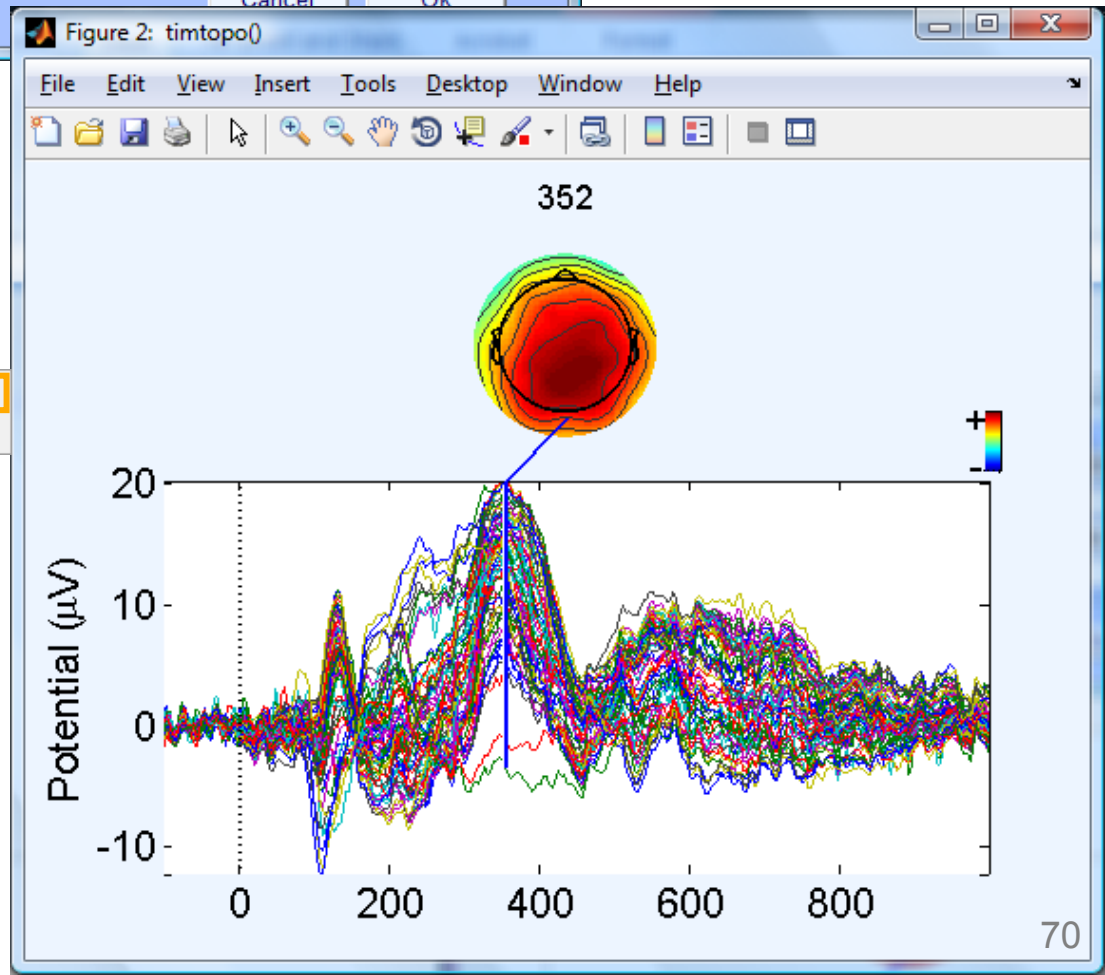


Visualize ERP scalp distribution

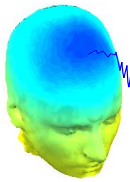


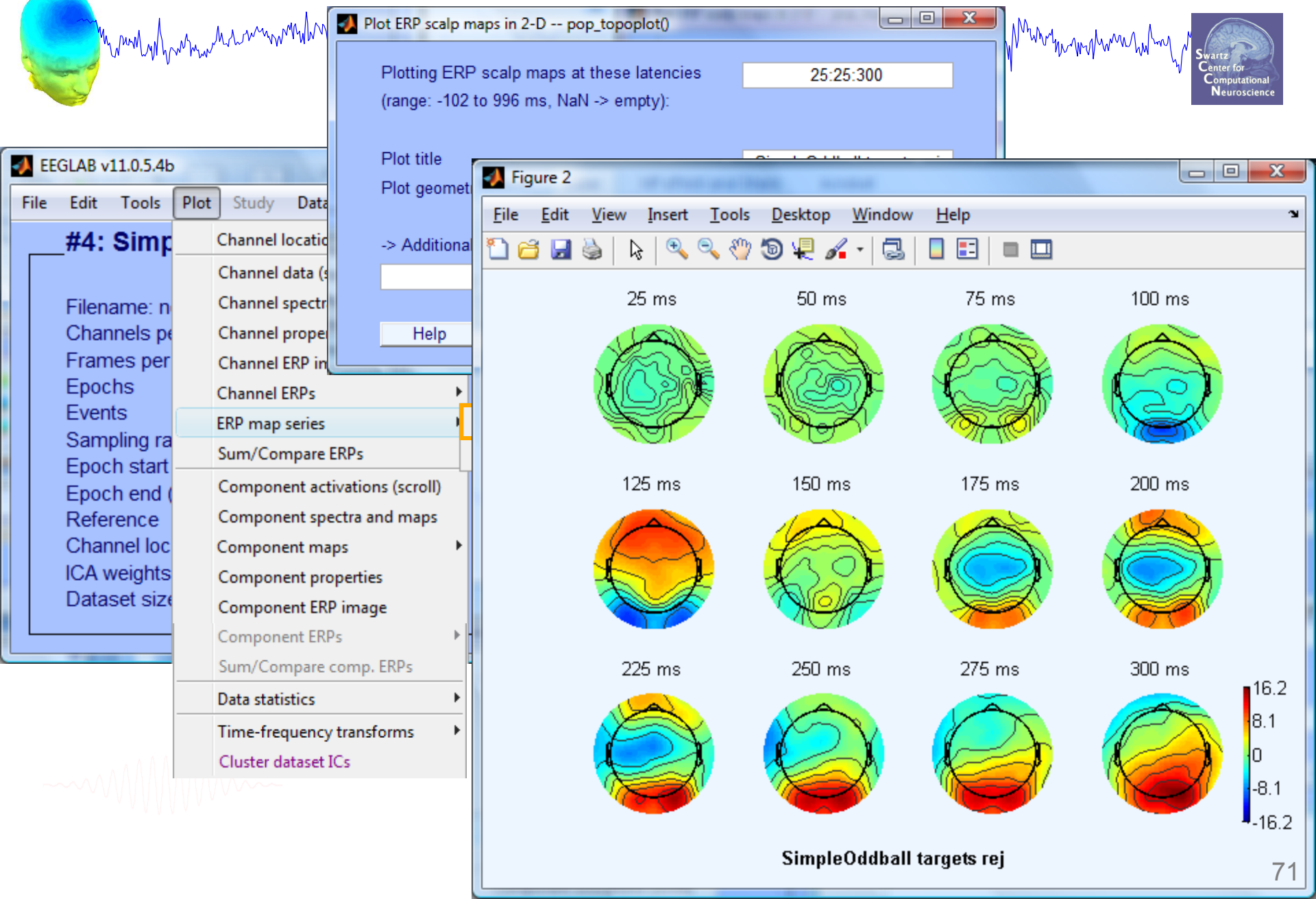




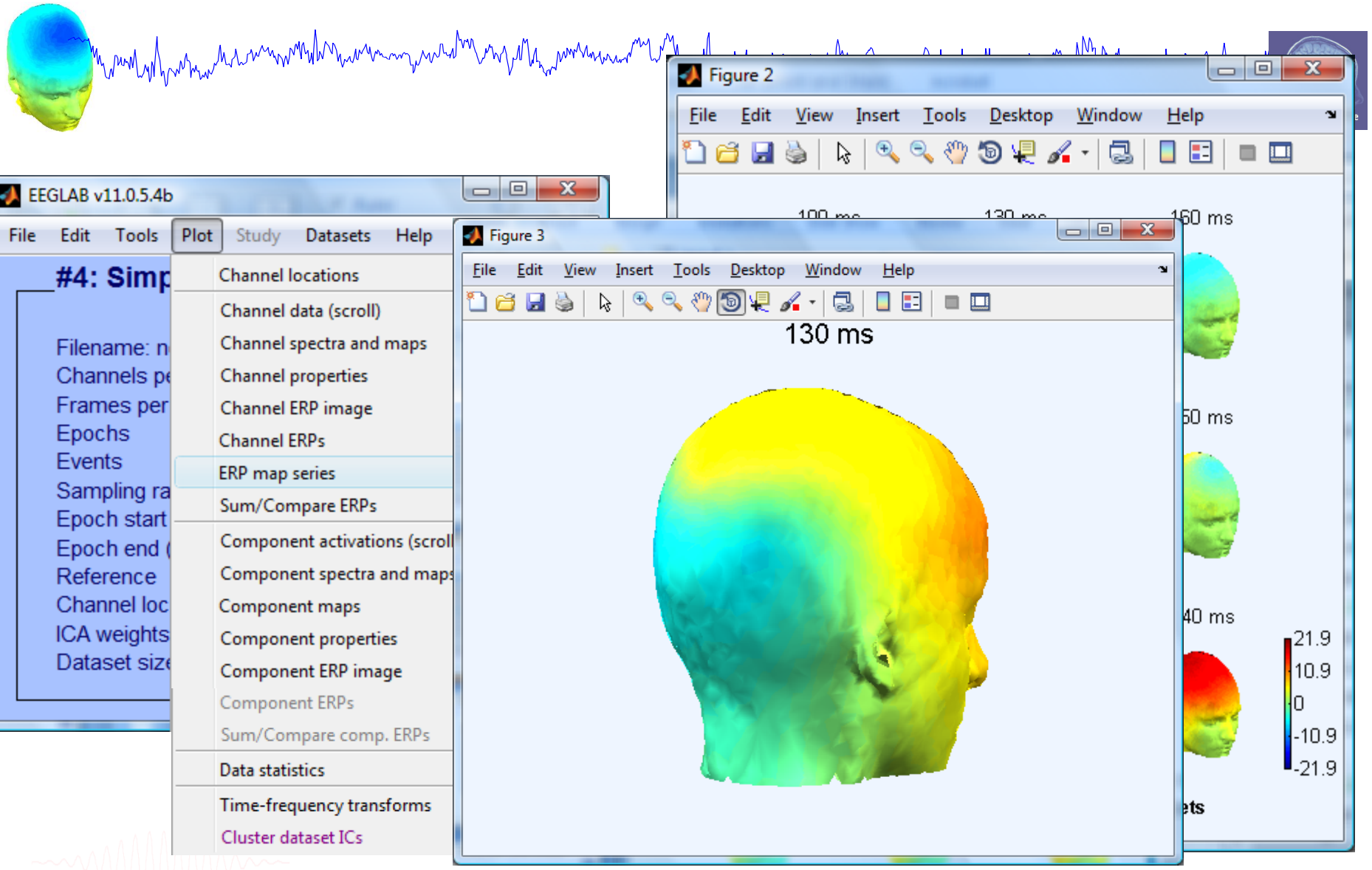


Visualize channel ERPs in 2D

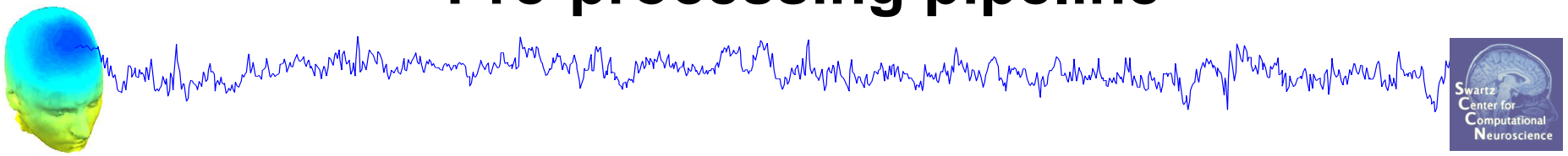




Visualize channel ERPs in 3D



Pre-processing pipeline



Collect high-density EEG data (>30 chan)

Import into EEGLAB

Import event markers and channel locations

Re-reference/ down-sample (if necessary)

High pass filter (~.5 – 1 Hz)

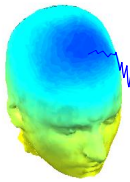
Remove line noise (if necessary)

Identify/reject bad channels

Reject large artifact time points

Run ICA

Exercises



- Preprocess data of your choice or load a previously filtered dataset e.g.SimpleOddball.set
- Identify bad channel(s) using auto-detection tool; plot channel properties of flagged channels
- Identify and remove non-task portions of continuous data; see if the previously flagged channels are still identified as bad
- Epoch on event of interest. Scroll the epoched data and perform visual rejection of epochs
- Explore the automated artifact rejection tools
- Run ICA