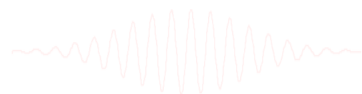


EEG Preprocessing for ICA

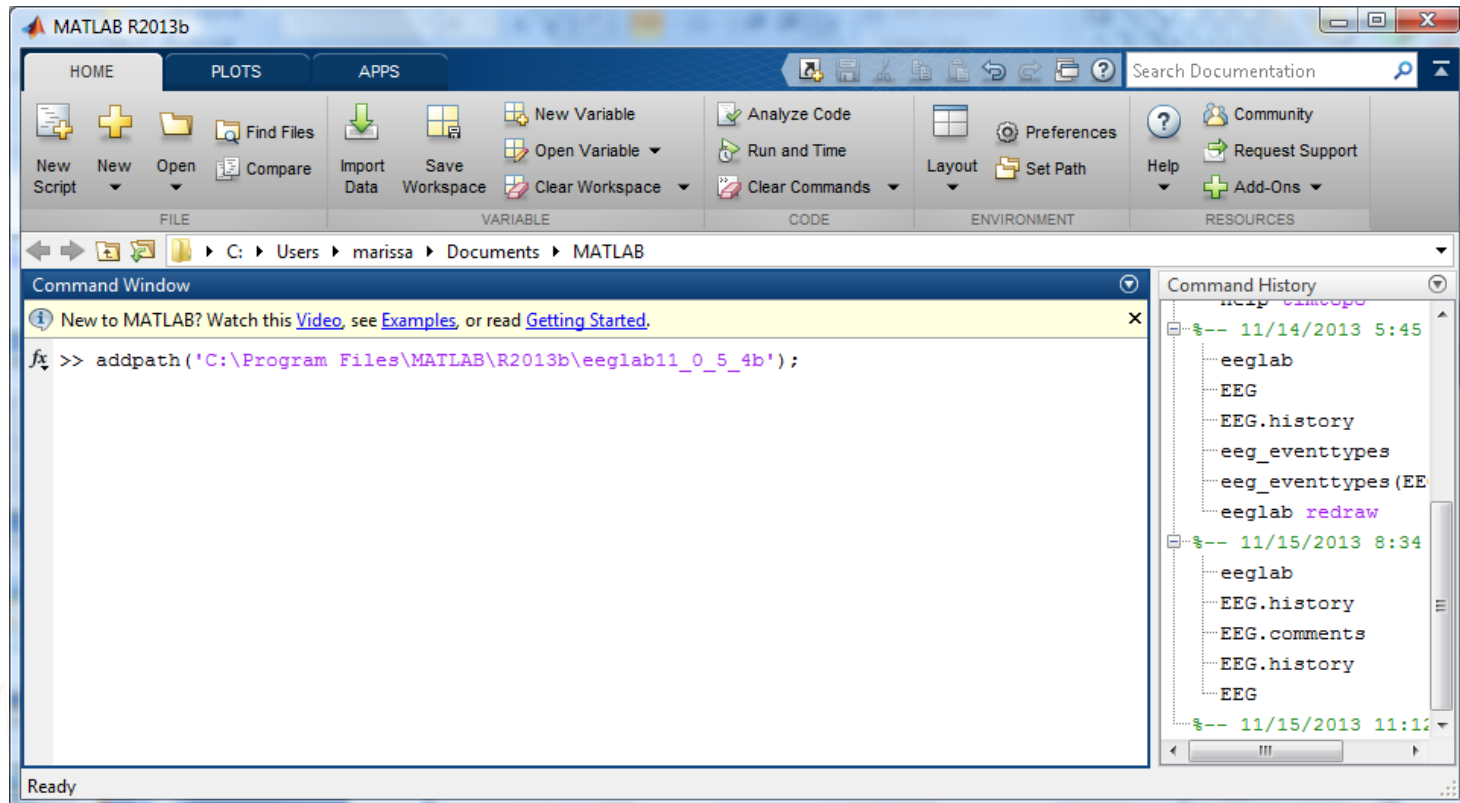
25th EEGLAB Workshop
JAIST Tokyo Satellite, Japan
Day 1
John Iversen



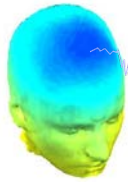
Installing EEGLAB and data folder



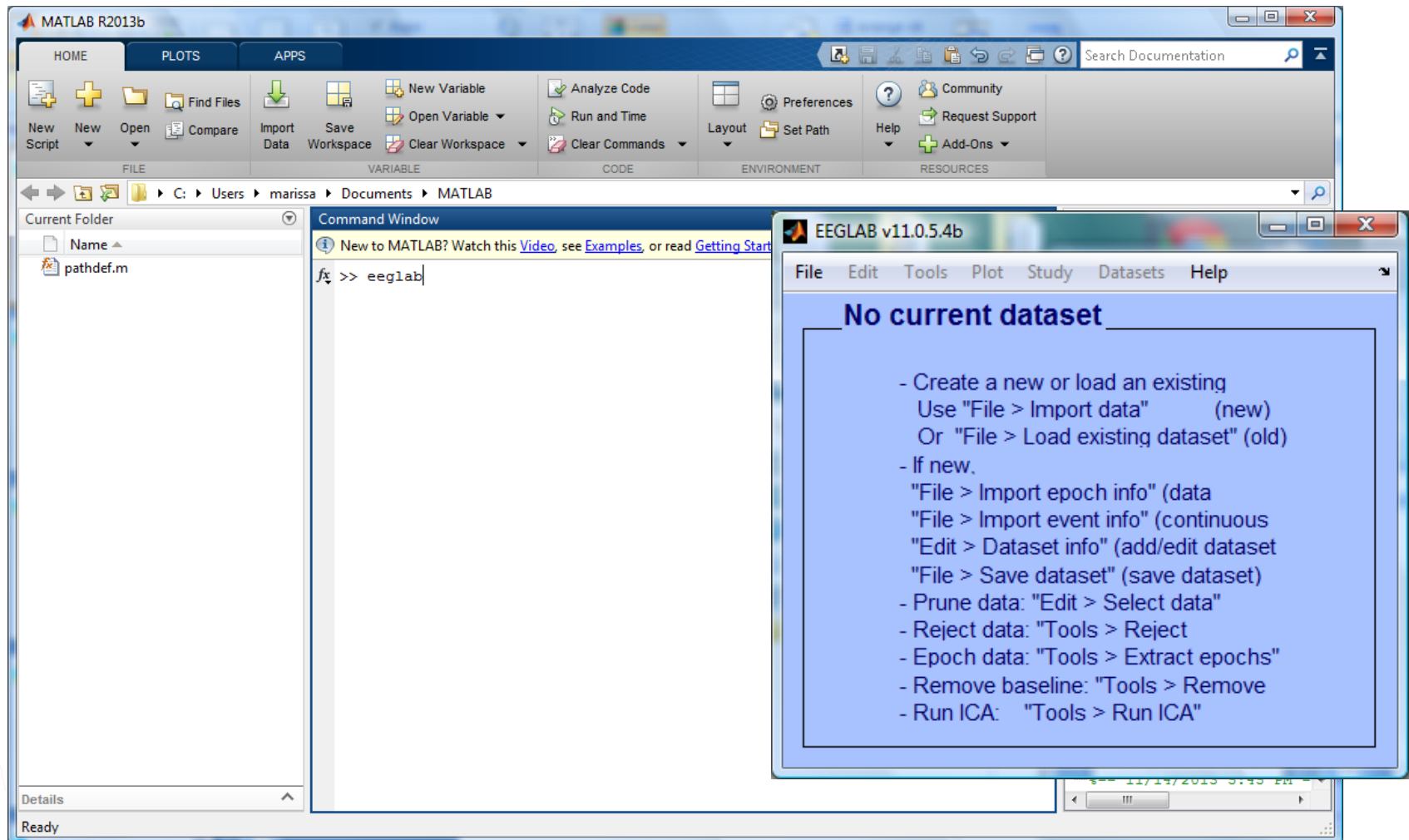
- Start Matlab
- Add the EEGLAB folder to your Matlab path:



The EEGLAB Matlab software



main graphic interface

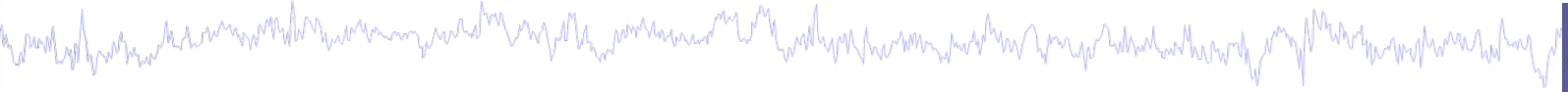
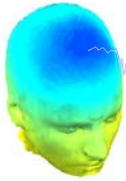


“Secrets” to a good ICA decomposition

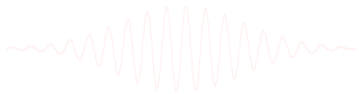


- Garbage in, garbage out (**GIGO**: it's not magic)
- Remove large, non-stereotyped artifacts
- Do you have enough data? (based mostly on time, not frames)
- High-pass filter to remove slow drifts (no low-pass filter needed)
- Remove bad channels
- Data must be in double precision (not single)

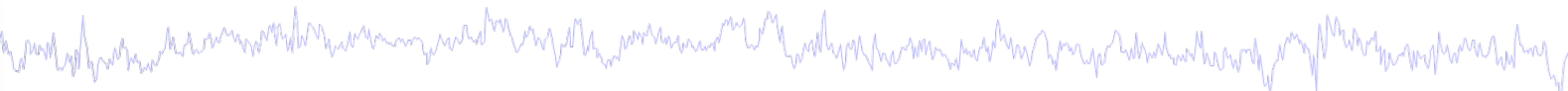
The Goal of Preprocessing



- Create a complete EEGLAB data set with
 - EEG time series signal
 - Channel Locations
 - Event information
- Applying signal processings on EEG time series to help ICA decompositions
 - Data ‘cleaning’—artifact rejection.
 - Removing noisy channels.
 - Removing noisy segments of data.
 - Applying frequency filter.



Pre-processing pipeline



**Import time-series
data into EEGLAB**

**Import event markers
and channel locations**

**High pass filter
(~.5 – 1 Hz)**

**Remove line noise
(if necessary)**

**Identify/reject
bad channels**

Examine raw data

**Reject large artifact
time points**

**Re-reference/
down-sample
(if necessary)**

Run ICA

Importing a dataset



EEGLAB v11.0.5.4b

File Edit Tools Plot Study Datasets Help

Import data

- Import epoch info
- Import event info
- Export

Load existing dataset

- Save current dataset(s)
- Save current dataset as
- Clear dataset(s)

Create study

- Load existing study
- Save current study
- Save current study as
- Clear study

Memory and other options

History scripts

Quit

Using EEGLAB functions and plugins

- Using the FILE-IO interface
- Using the BIOSIG interface
- Troubleshooting data formats...

From ASCII/float file or Matlab array

- From Netstation .mff (FILE-IO toolbox)
- From Netstation binary simple file
- From Multiple seg. Netstation files
- From Netstation Matlab files

From BCI2000 ASCII file

From Snapmaster .SMA file

From Neuroscan .CNT file

From Neuroscan .EEG file

From Biosemi BDF file (BIOSIG toolbox)

From Biosemi BDF and EDF files (BDF plugin)

From EDF/EDF+/GDF files (BIOSIG toolbox)

From ANT EEProbe .CNT file

From ANT EEProbe .AVR file

From BCI2000 .DAT file

From BIOPAC MATLAB files

From Brain Vis. Rec. .vhdr file

From Brain Vis. Anal. Matlab file

From CTF folder (MEG)

From ERPSS .RAW or .RDF file

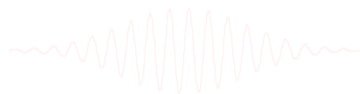
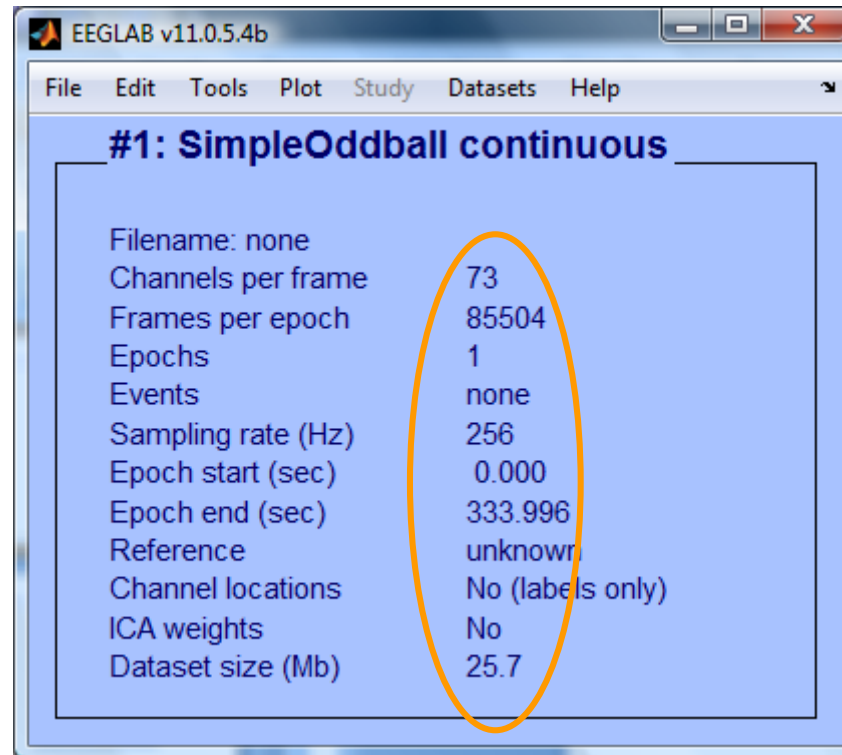
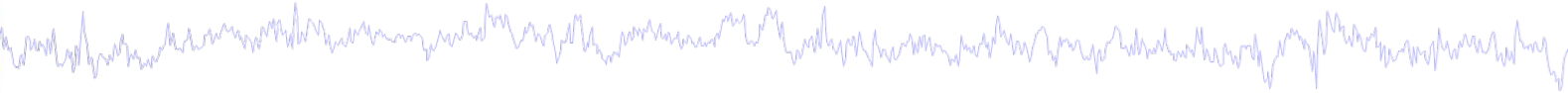
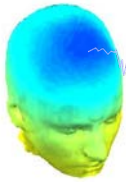
From INStep .ASC file

From 4D .m4d pdf file

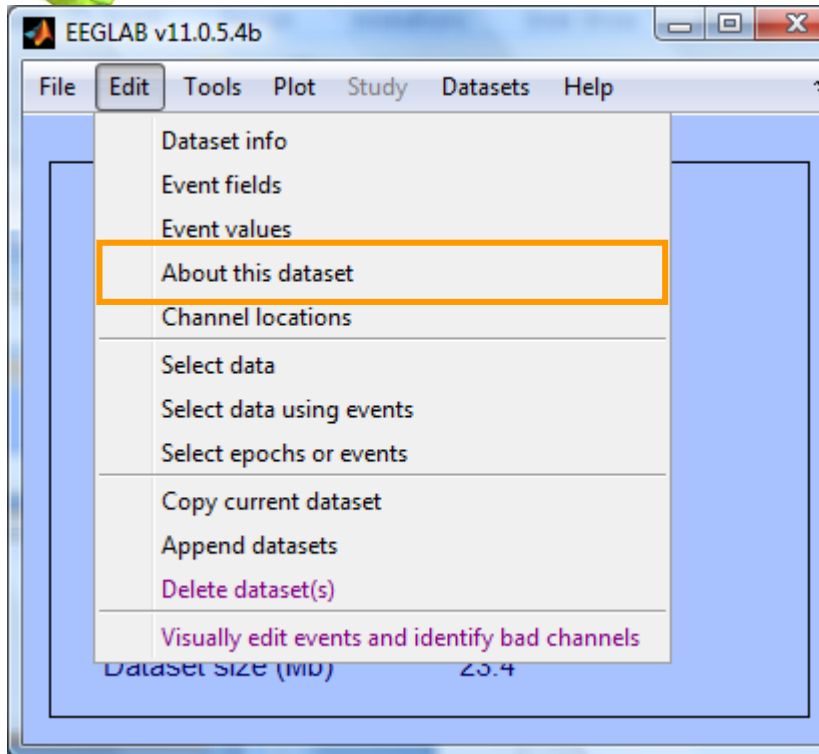
From Procom Infinity Text File

Tip for Biosemi users:
Use the 'BDF plugin' version
of the Biosemi BDF/EDF importer

Imported EEG data



Comments and dataset history

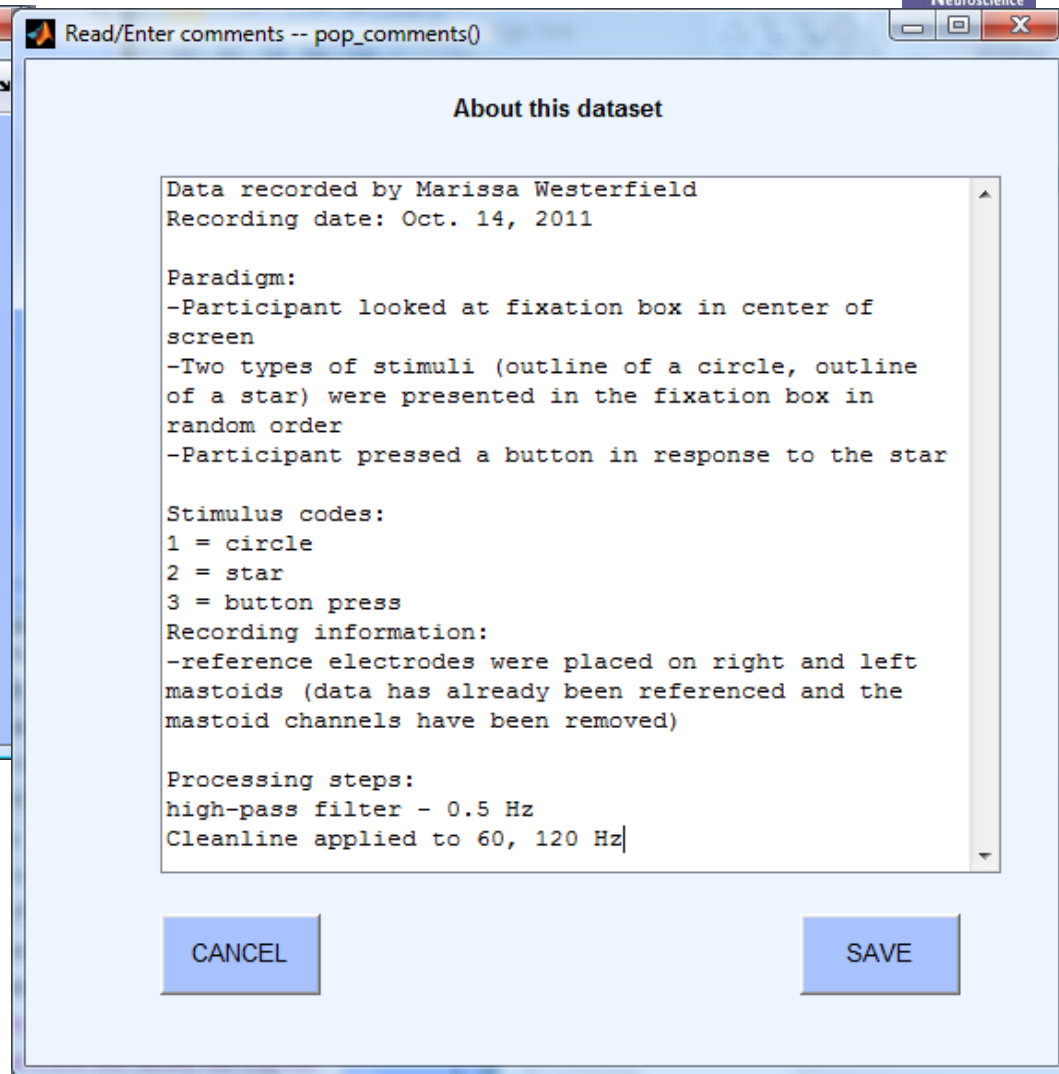


Also:

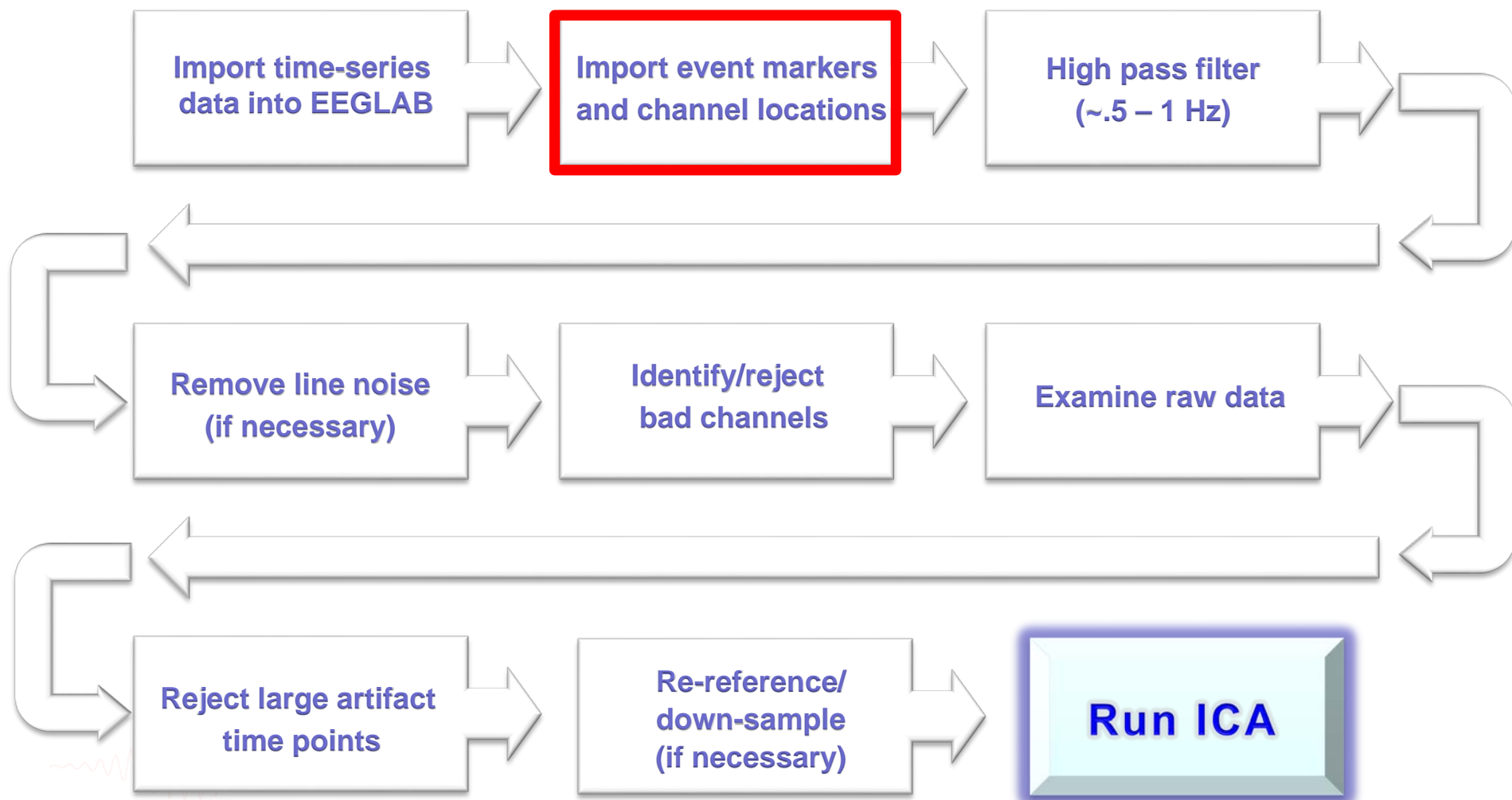
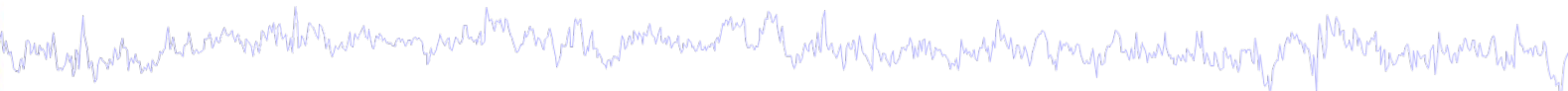
`>> EEG.comments`

and

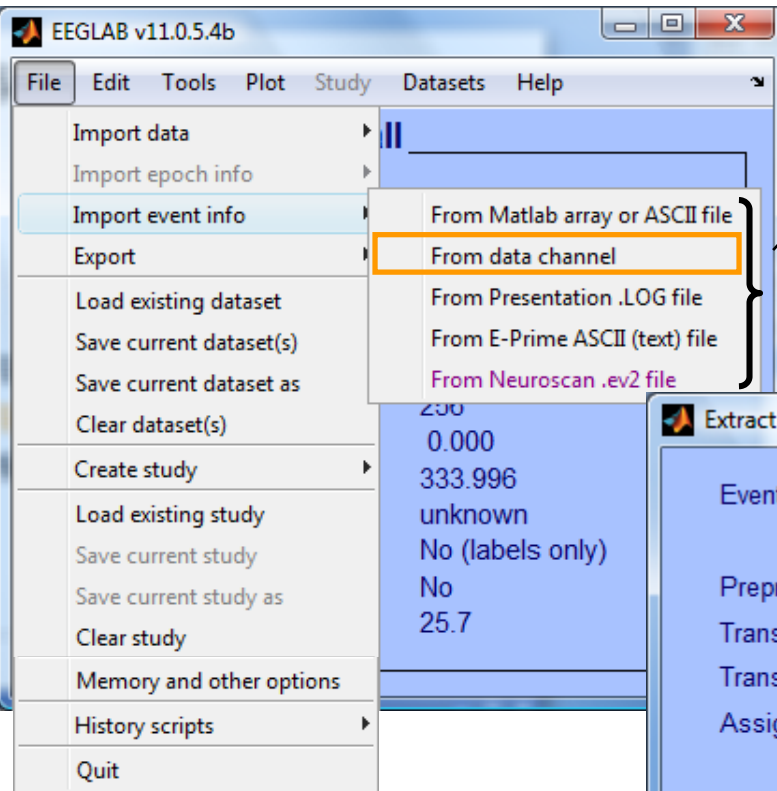
`>> EEG.history`



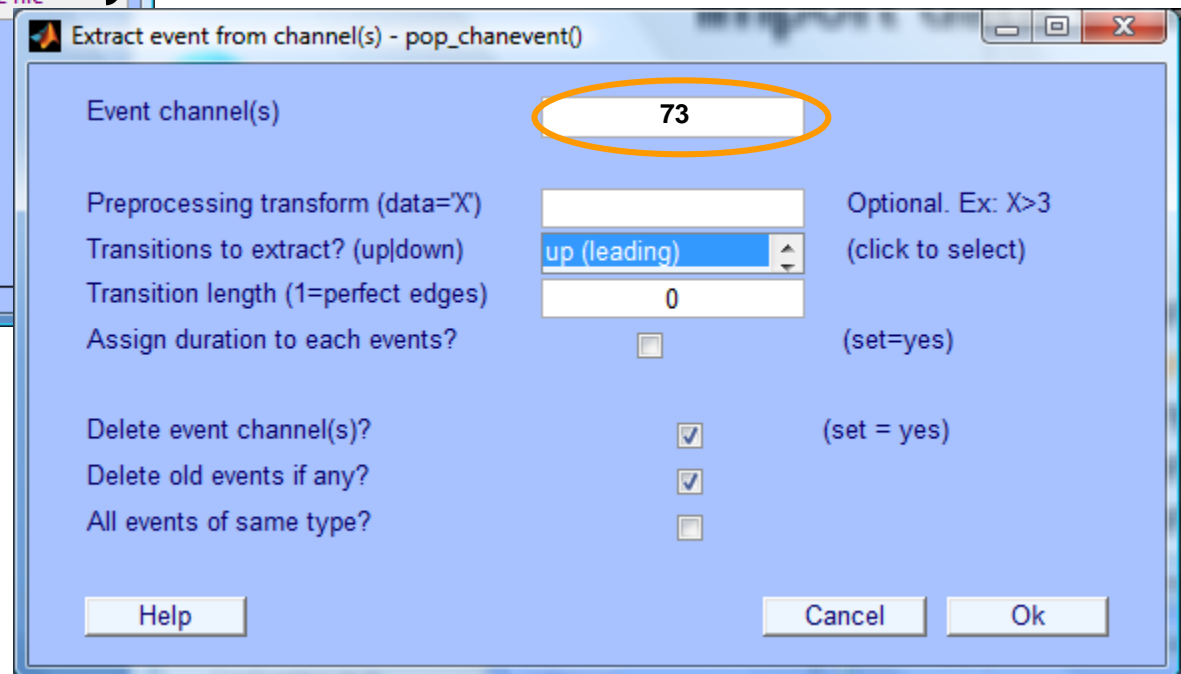
Pre-processing pipeline



Import data events

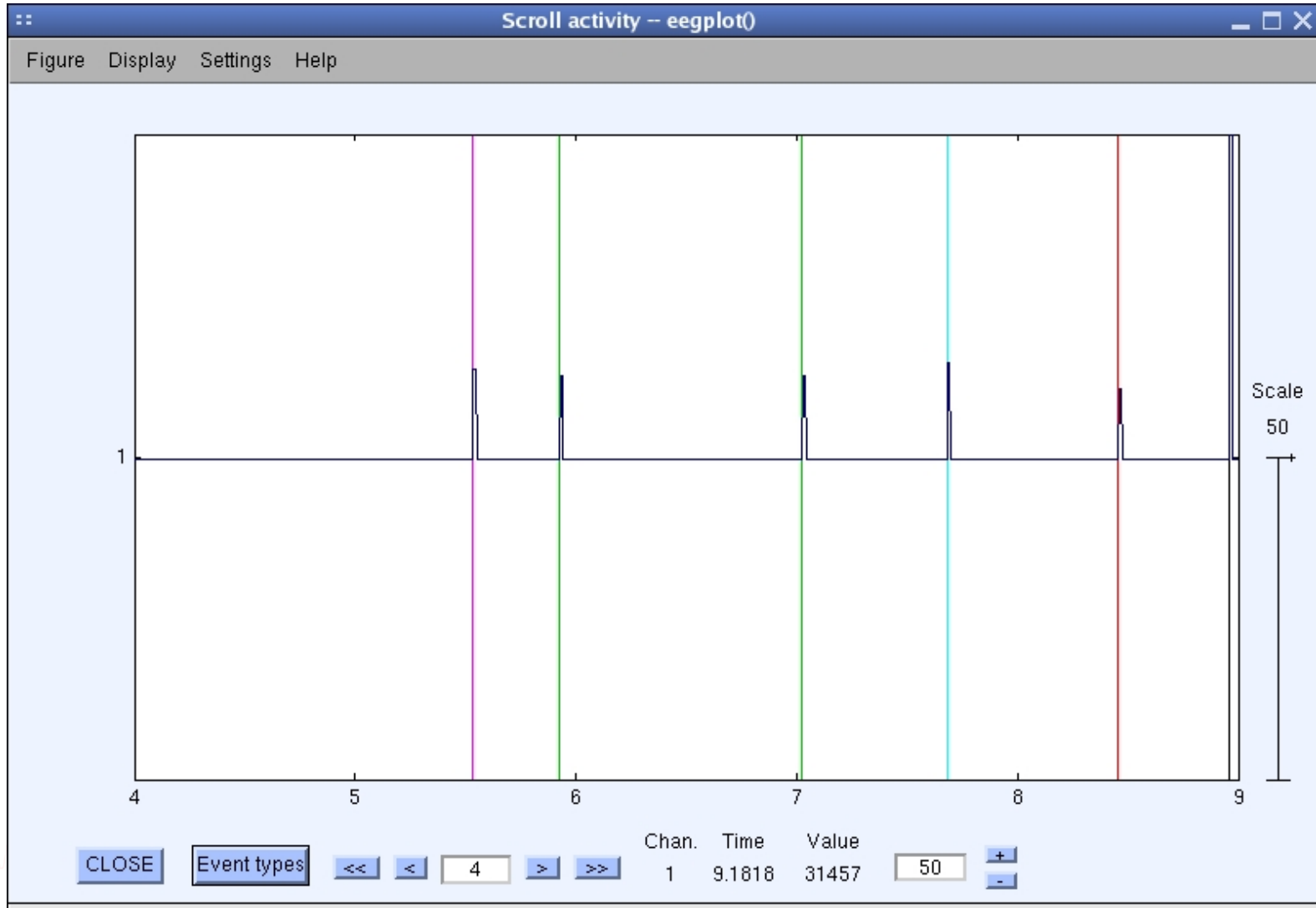
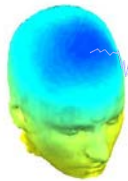


- Import events from Matlab array or ASCII file
- **Import events from data channel**
- Import from Presentation event file
- Import events from E-Prime event file
- Import events from Neuroscan event file

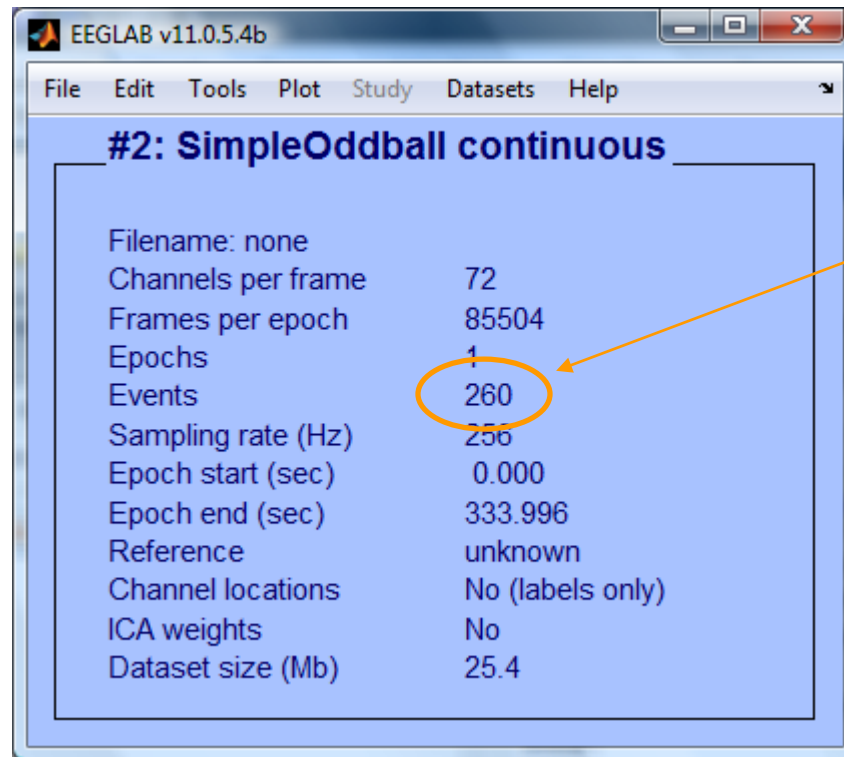
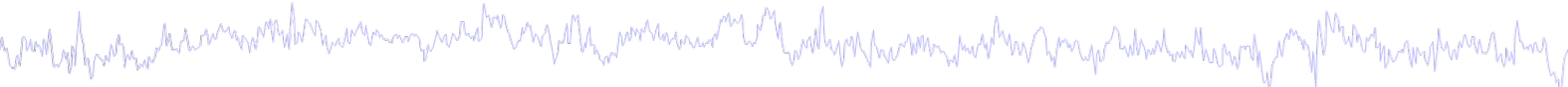


(Often imported automatically
during data import)

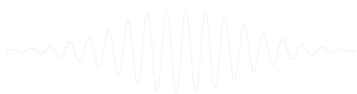
Appearance of an event channel in raw data



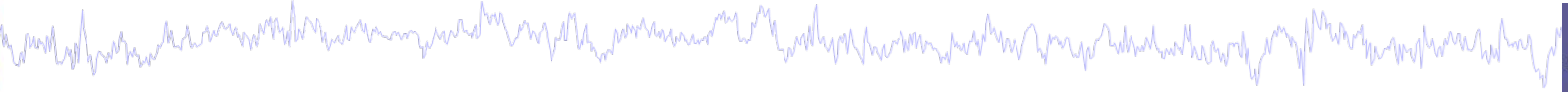
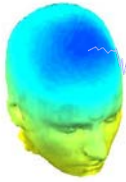
Imported data events



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



Sample data: basic P300 paradigm



File

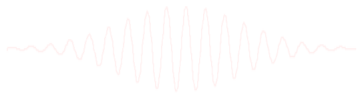
SimpleOddball.set

Data

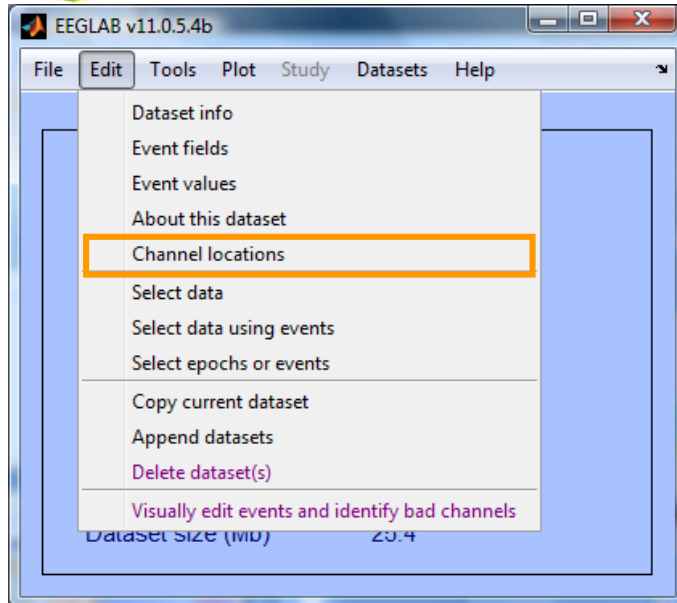
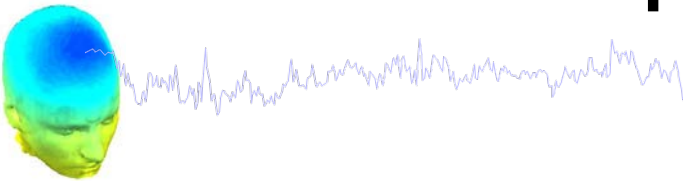
68 channel EEG, 256 Hz sampling rate, Biosemi system, re-referenced during import to averaged left and right mastoid electrodes

Task

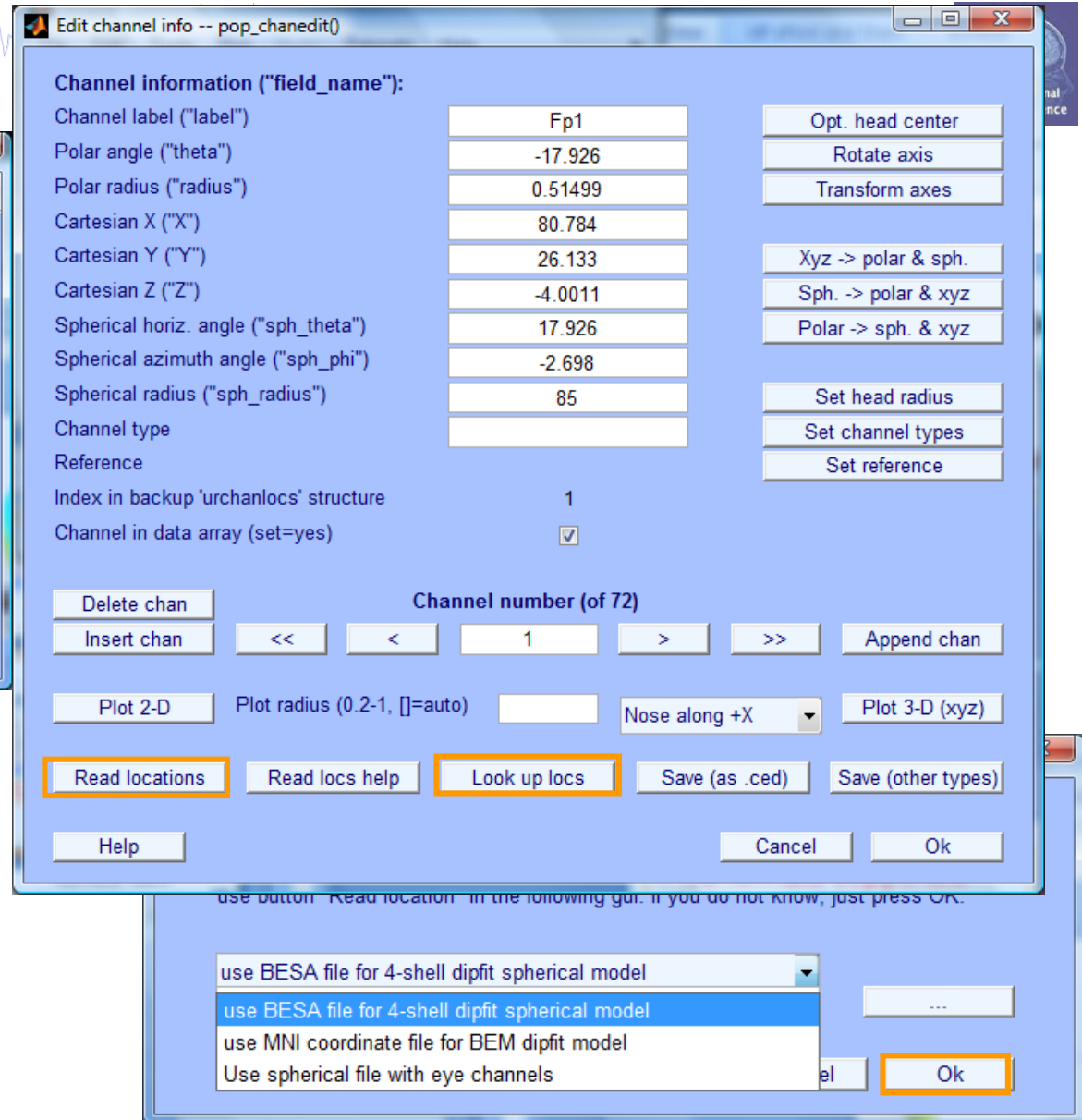
speeded button press response to star shape (no response to circle shape), 100 ms presentation duration, 200 trials



Import channel locations



7 file formats supported
(Polhemus, BESA, ...)



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)	<input checked="" type="checkbox"/>

Channel number (of 71)

Channel indices: 1:71
Type (e.g. EEG): EEG

Buttons: Delete chan, Insert chan, <<, <, >, >>, 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

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

Convert channel locations -- pop_chancenter()

Optimize center location ☒ or specify center 0 0 0

Channel indices to ignore for best-sphere matching

Help Cancel Ok

Force electrode location -- forclocs()

XY value Coordinate Electrode list

0 X (rotate X-Z plane) Cz Pick

Help Cancel Ok

Set channel ...

Channel indices 1:71
Type (e.g. EEG) EEG

Help Cancel Ok

Edit channel info -- pop_chanedit()

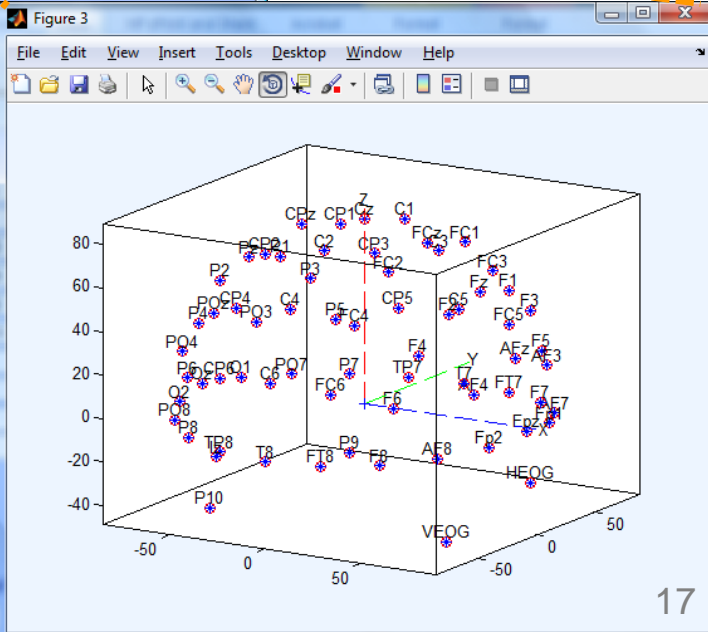
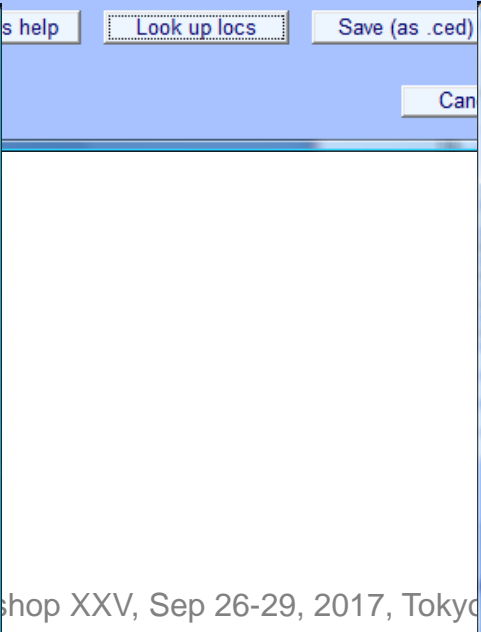
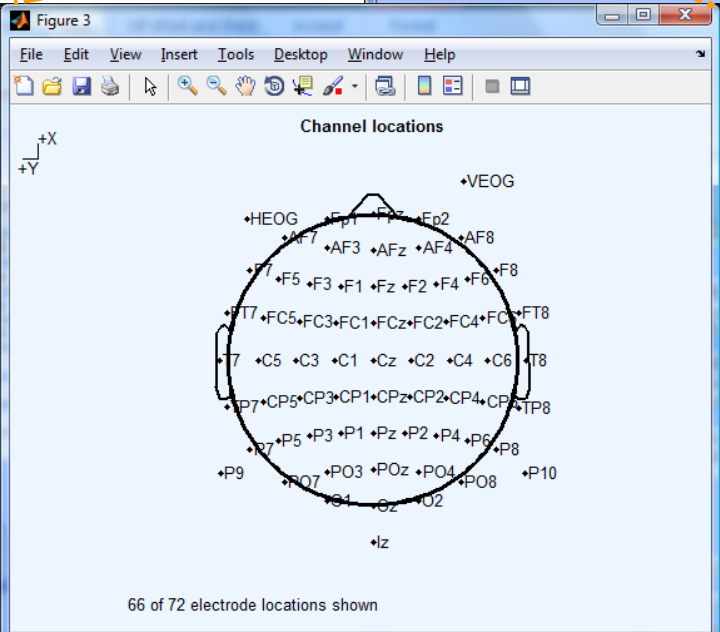
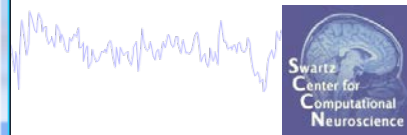
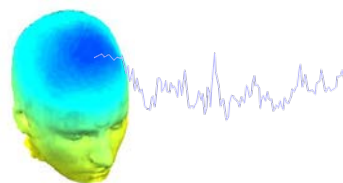
Channel information ("field_name"):

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	XYZ -> polar & sph.
Cartesian Y ("Y")	50.2186	Sph. -> polar & xyz
Cartesian Z ("Z")	-39.9051	Polar -> sph. & xyz
Spherical horiz. angle ("sph_theta")	42	Set head radius
Spherical azimuth angle ("sph_phi")	-28	Set channel types
Spherical radius ("sph_radius")	85	Set reference
Channel type		
Reference		
Index in backup 'urchanlocs' structure	68	
Channel in data array (set=yes)	<input checked="" type="checkbox"/>	

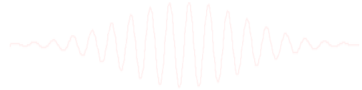
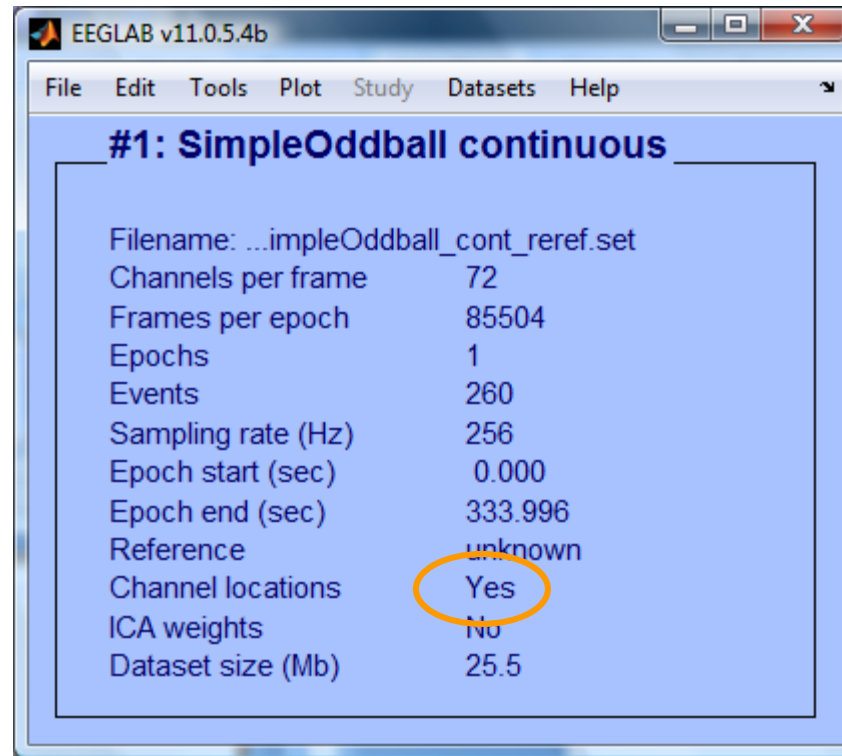
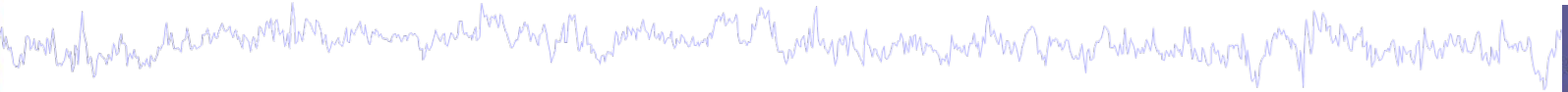
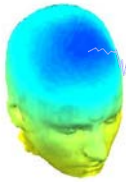
Channel number (of 72)

68

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



Imported channel locations



Remove unwanted channels

EEGLAB v11.0.5.4b

File Edit Tools Plot Study Datasets Help

Dataset info
Event fields
Event values
About this dataset
Channel locations
Select data
Select data using events
Select epochs of
Copy current data
Append dataset
Delete dataset(s)
Visually edit event
Dataset size (mb)

reref

Select data -- pop_select()

Select data in: Input desired range on->remove these

Time range [min max] (s)

Point range (ex: [1 10])

Epoch range (ex: 3:2:10)

Channel range EXG5 EXG6 EXG7 EXG8

Scroll dataset

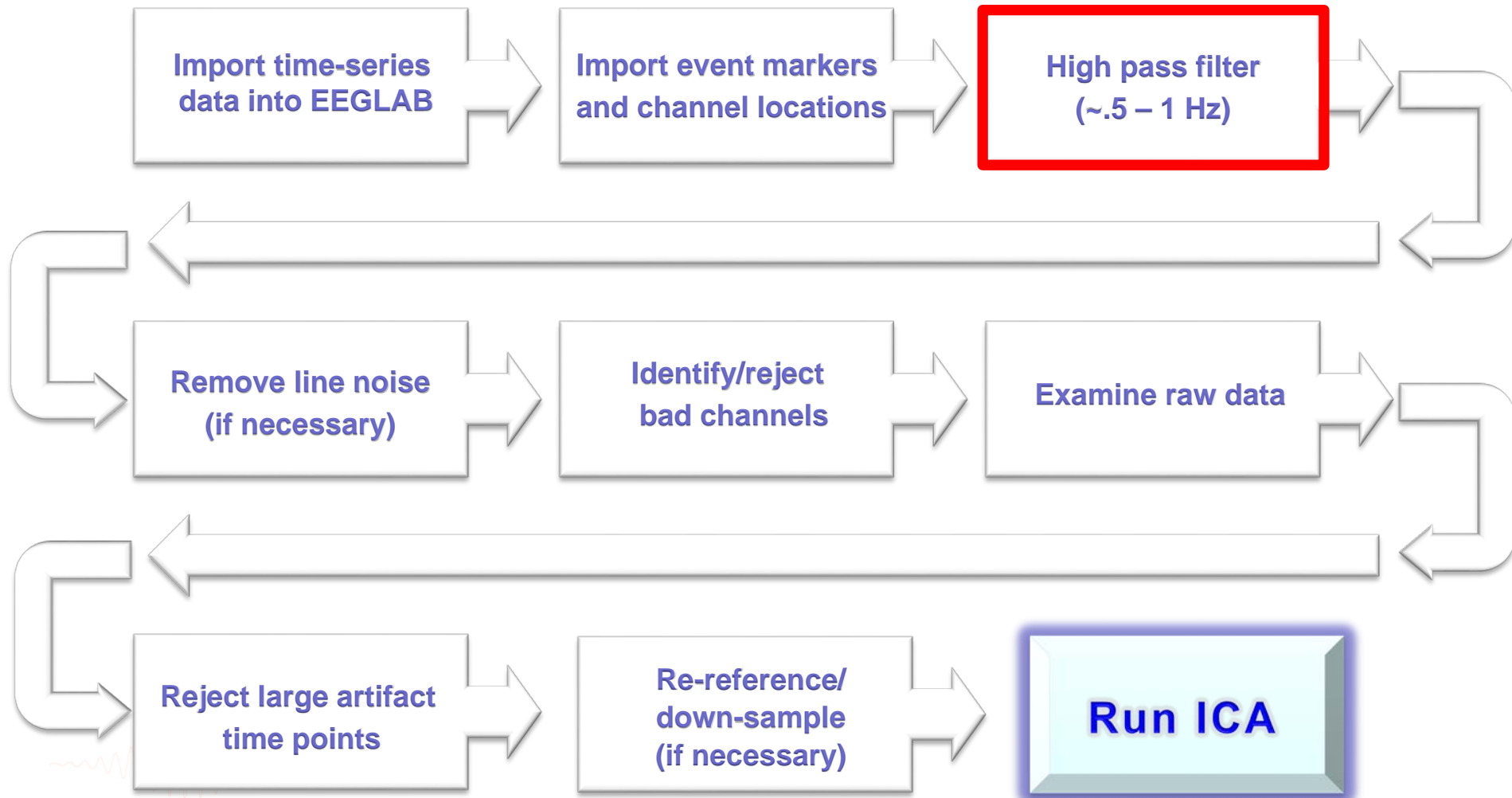
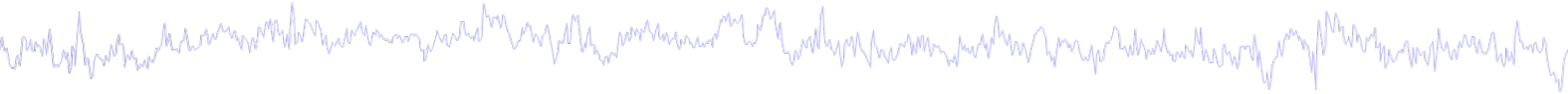
Help Cancel Ok

(use shift|ctrl to select several)

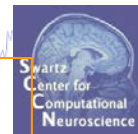
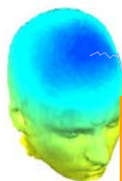
- 45 - FC4
- 46 - FC2
- 47 - FCz
- 48 - Cz
- 49 - C2
- 50 - C4
- 51 - C6
- 52 - T8
- 53 - TP8
- 54 - CP6
- 55 - CP4
- 56 - CP2
- 57 - P2
- 58 - P4
- 59 - P6
- 60 - P8
- 61 - P10
- 62 - PO8
- 63 - PO4
- 64 - O2
- 65 - VEOG
- 66 - HEOG
- 67 - EXG5
- 68 - EXG6
- 69 - EXG7
- 70 - EXG8

Cancel Ok

Pre-processing pipeline

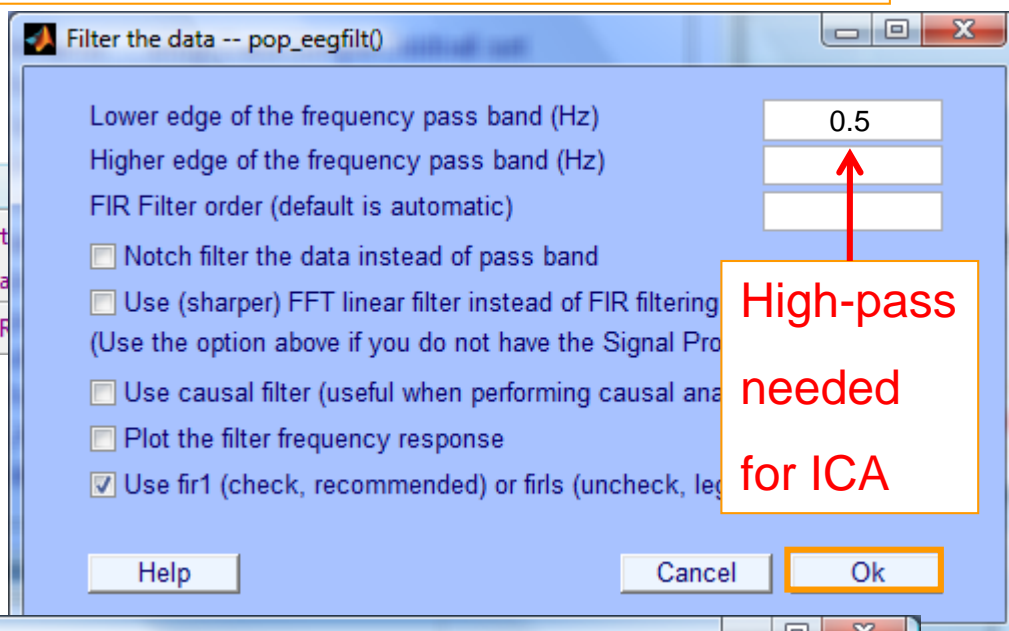
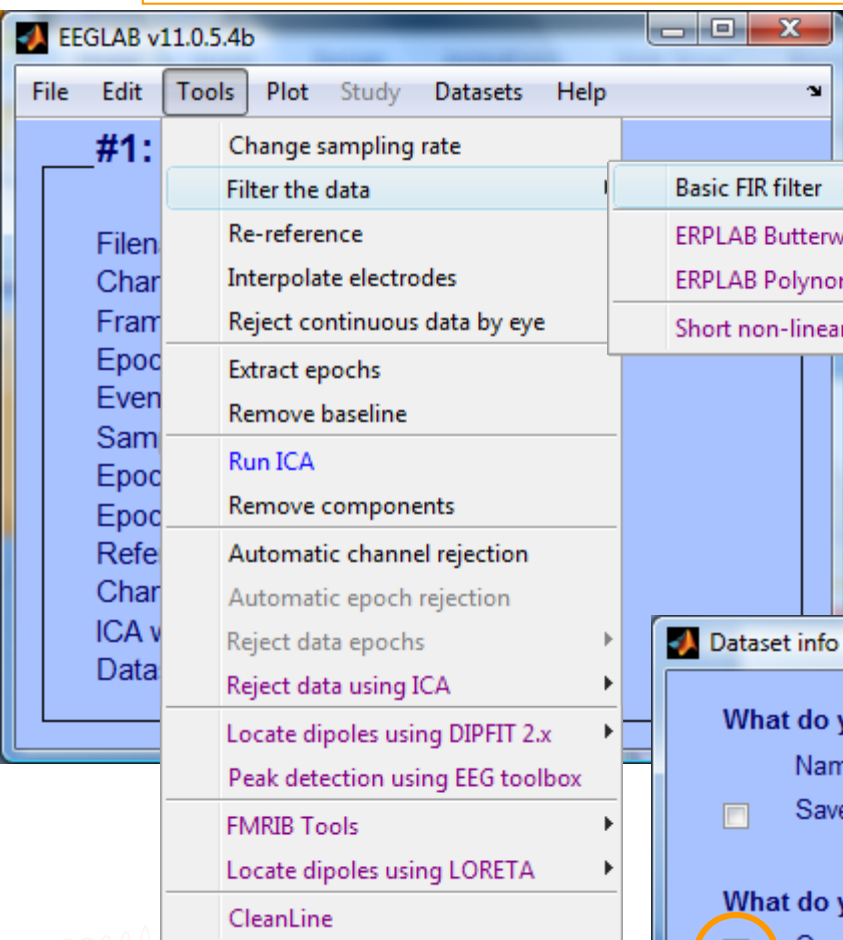


High-Pass Filter the data

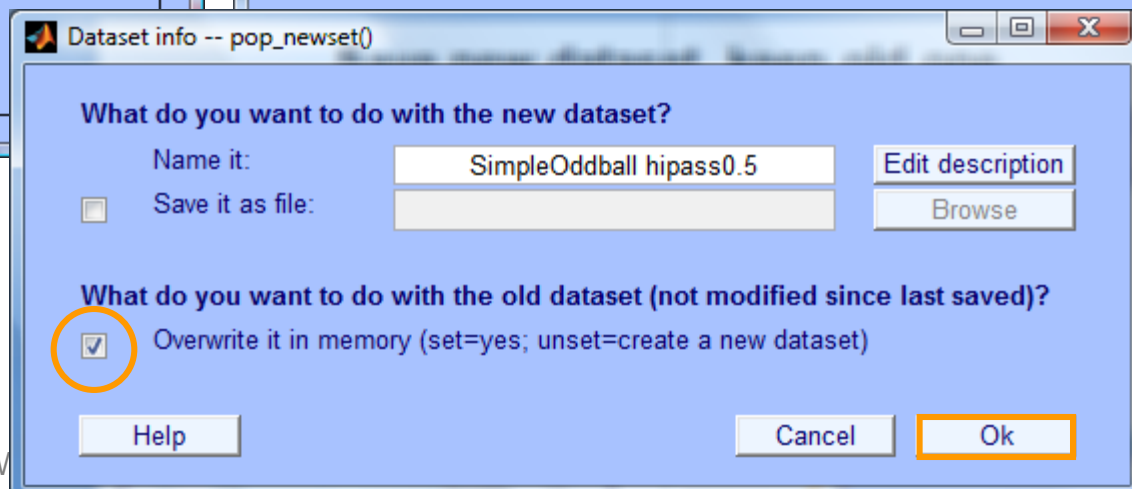


Reason: To improve data stationarity.

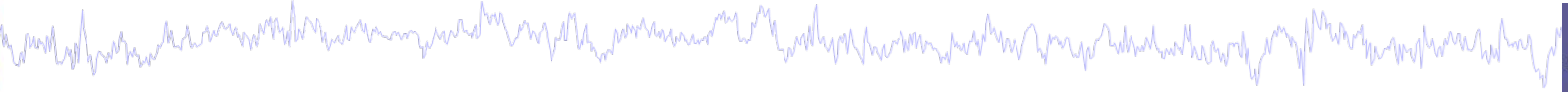
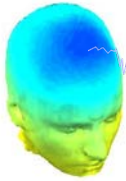
ICA is biased to amplitude, and EEG data has $1/f$ power spectrum density.



High-pass
needed
for ICA



Good resource for learning filters

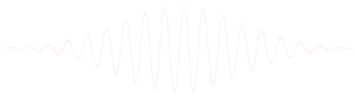


https://sccn.ucsd.edu/wiki/Firfilt_FAQ

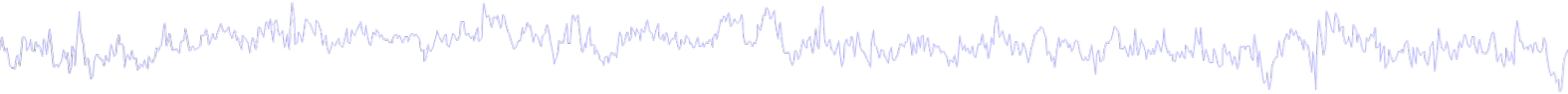
<https://cloud.github.com/downloads/widmann/firfilt/firfilt.pdf>



Andreas Widmann from Leipzig



Pre-processing pipeline



**Import time-series
data into EEGLAB**

**Import event markers
and channel locations**

**High pass filter
(~.5 – 1 Hz)**

**Remove line noise
(if necessary)**

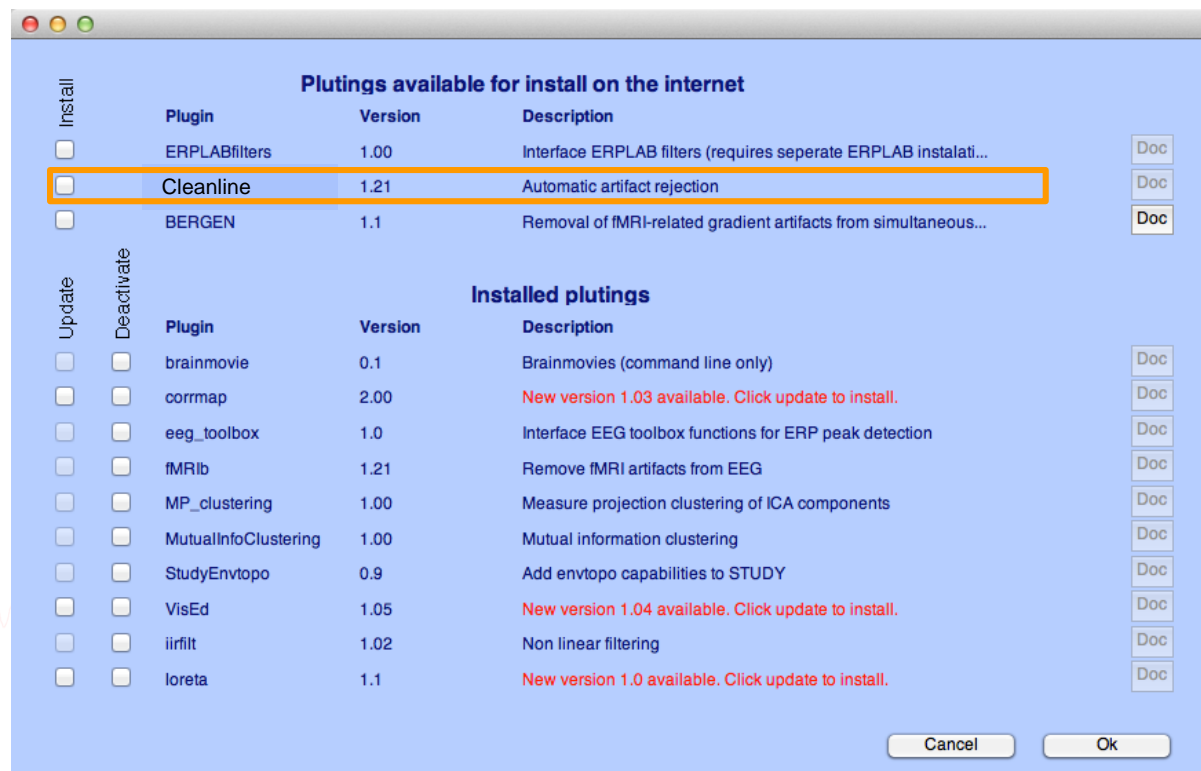
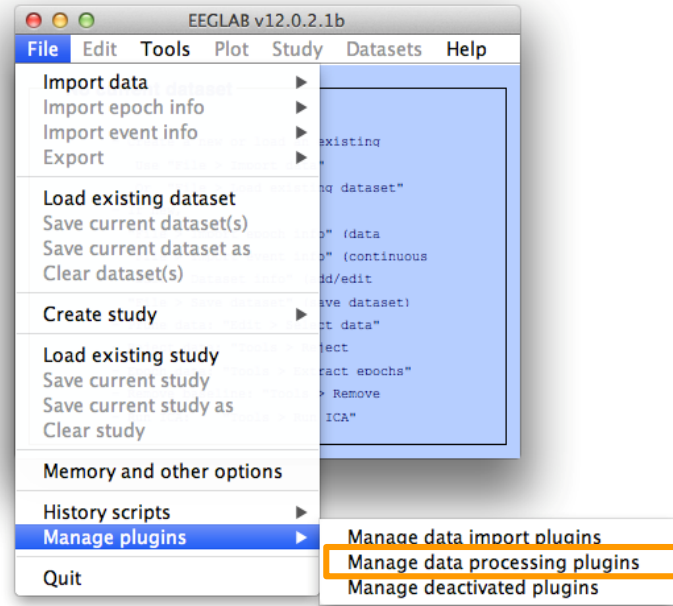
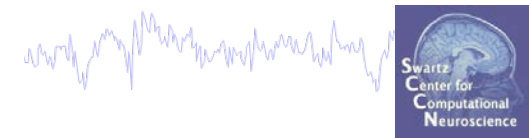
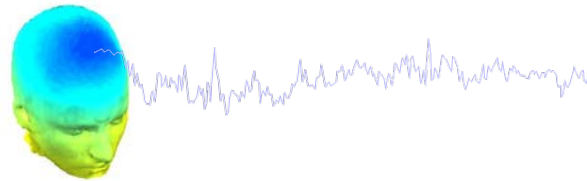
**Identify/reject
bad channels**

Examine raw data

**Reject large artifact
time points**

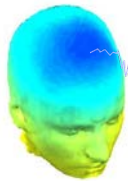
**Re-reference/
down-sample
(if necessary)**

Run ICA



Tim Mullen
Qusp CEO

Remove line noise (Cleanline)



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

CleanLine Options

Line noise frequencies to remove [60 120]
Scan for line noise ☒ (set)
p-value for detection of significant sinusoid 0.01
Bandwidth (Hz) 2
Type of signal to clean Channels
Indices of Channels/Components to clean '1:66'
Sliding window length (sec) 4
Sliding window step size (sec) 2
Window overlap smoothing factor 100
FFT padding factor 2
Visualize Original and Cleaned Spectra ☐ (set)
Normalize log spectrum by detrending ☐ (set)
Produce verbose output ☒ (set)
Plot Individual Figures ☐ (set)

check

Help Cancel Ok

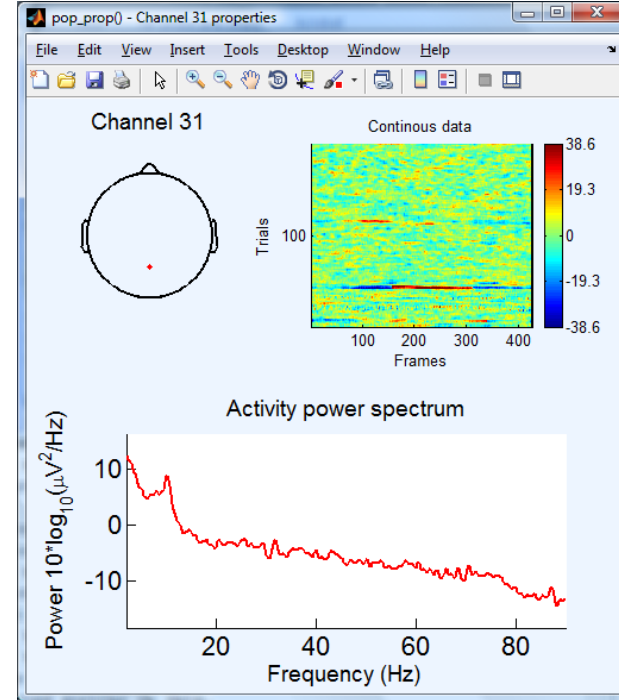
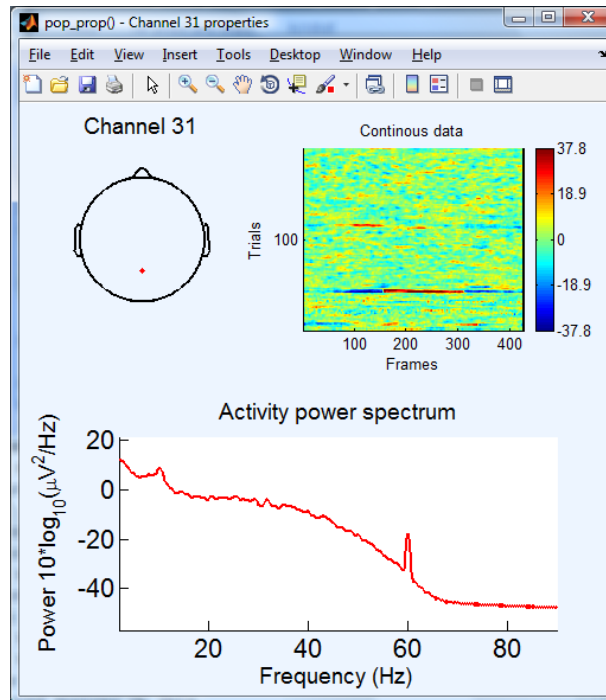
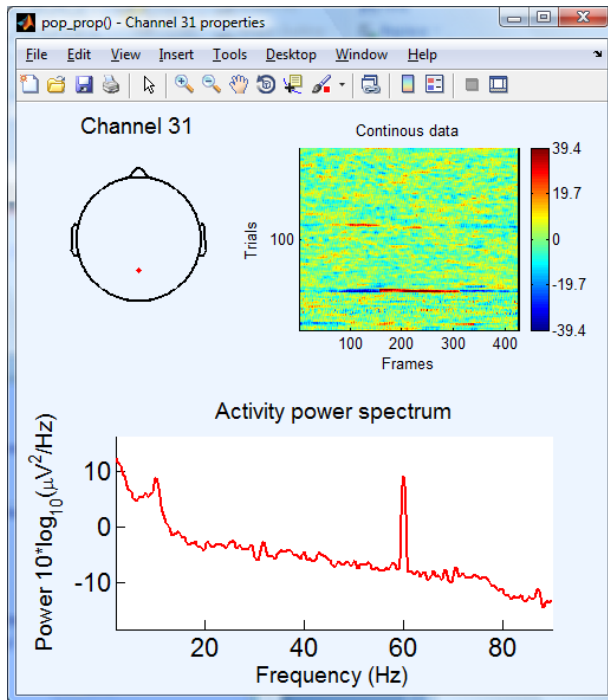
Filter comparisons

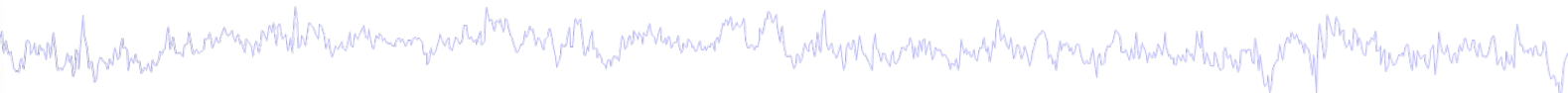


0.5 Hz high-pass filter

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

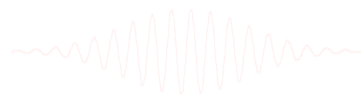
0.5 Hz high-pass filter
Cleanline



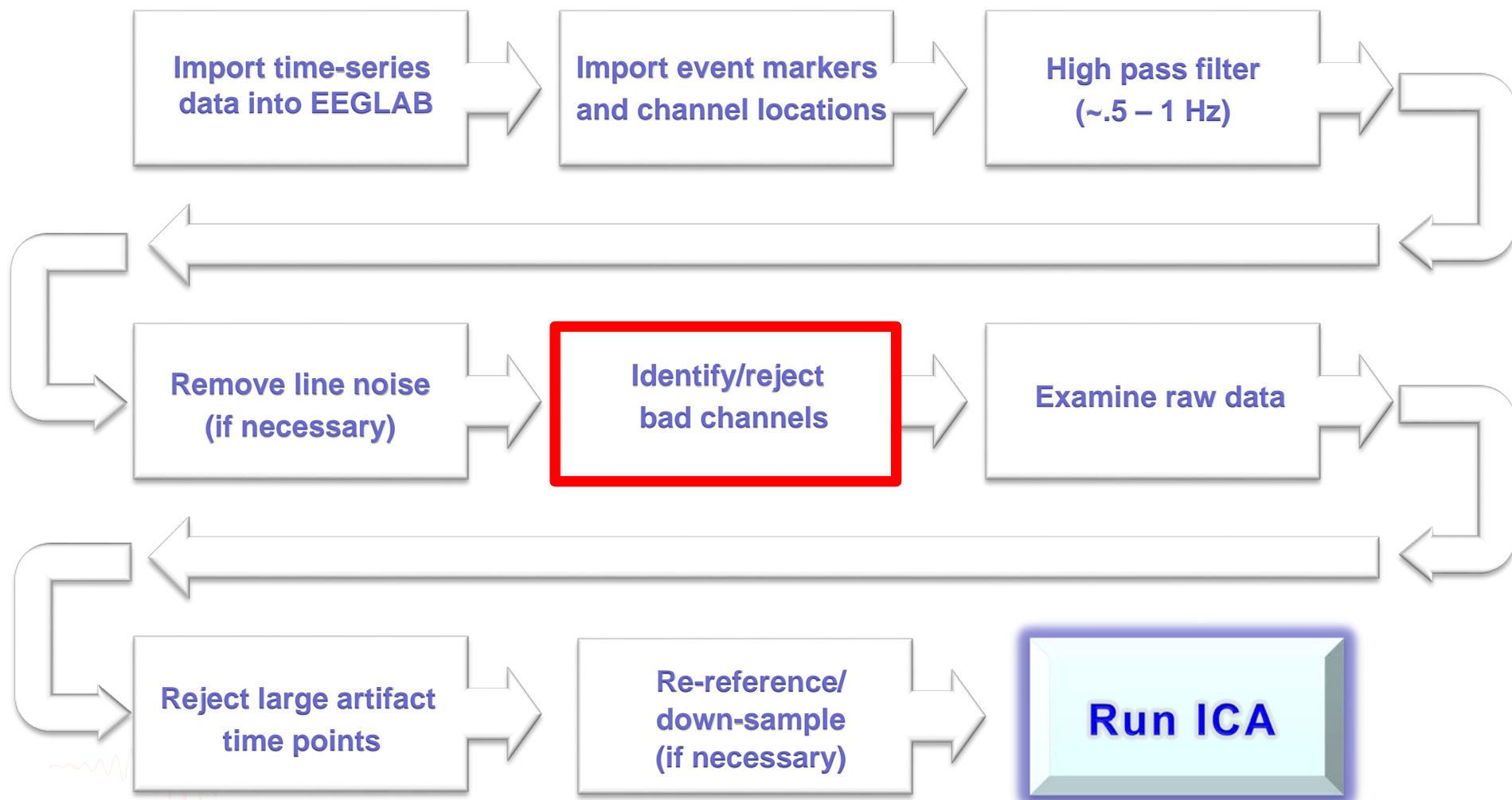
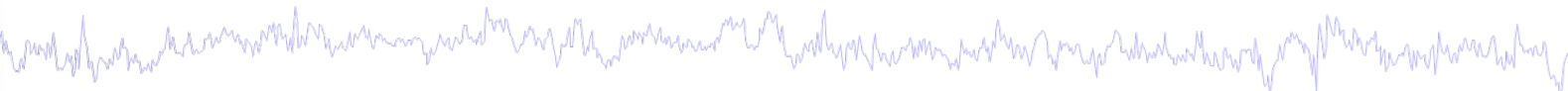


Data Cleaning for ICA

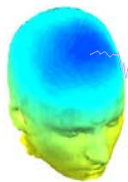
1. Continuous Data



Pre-processing pipeline



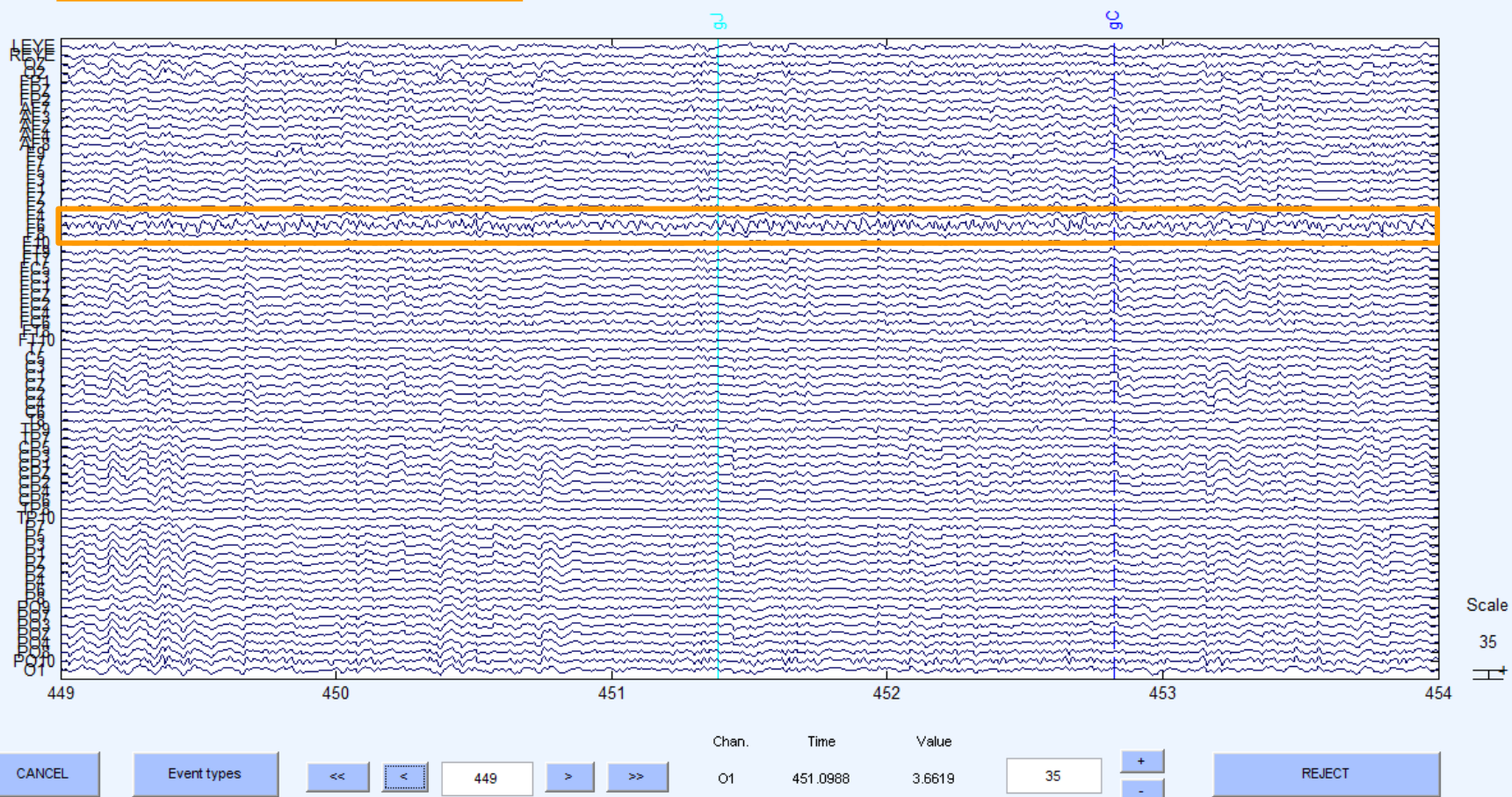
Manually identifying bad channels



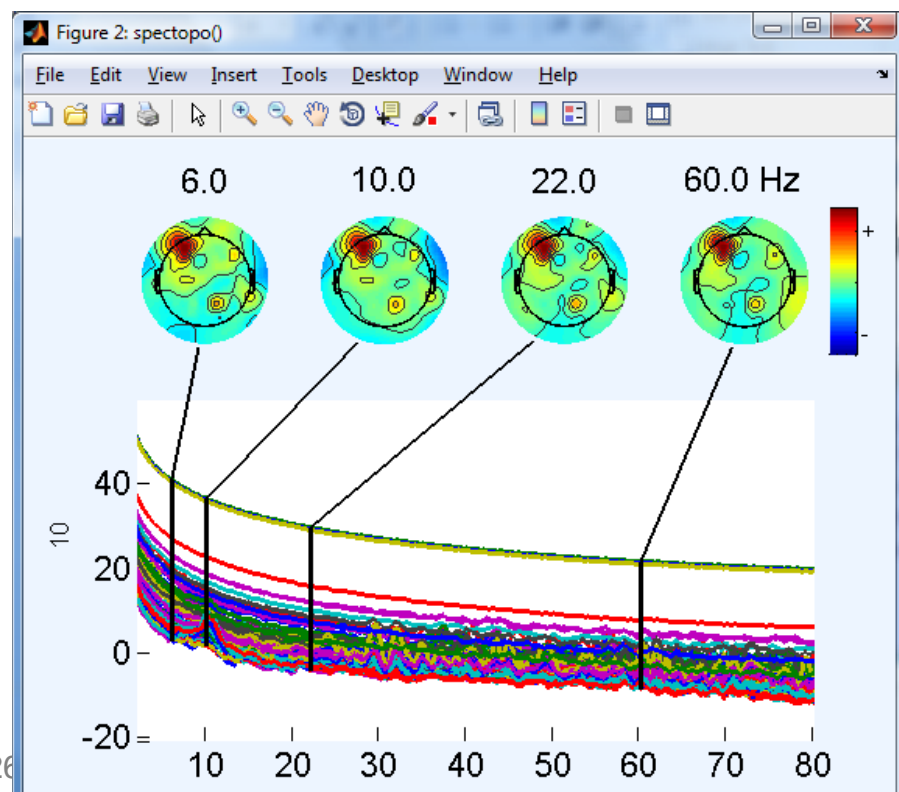
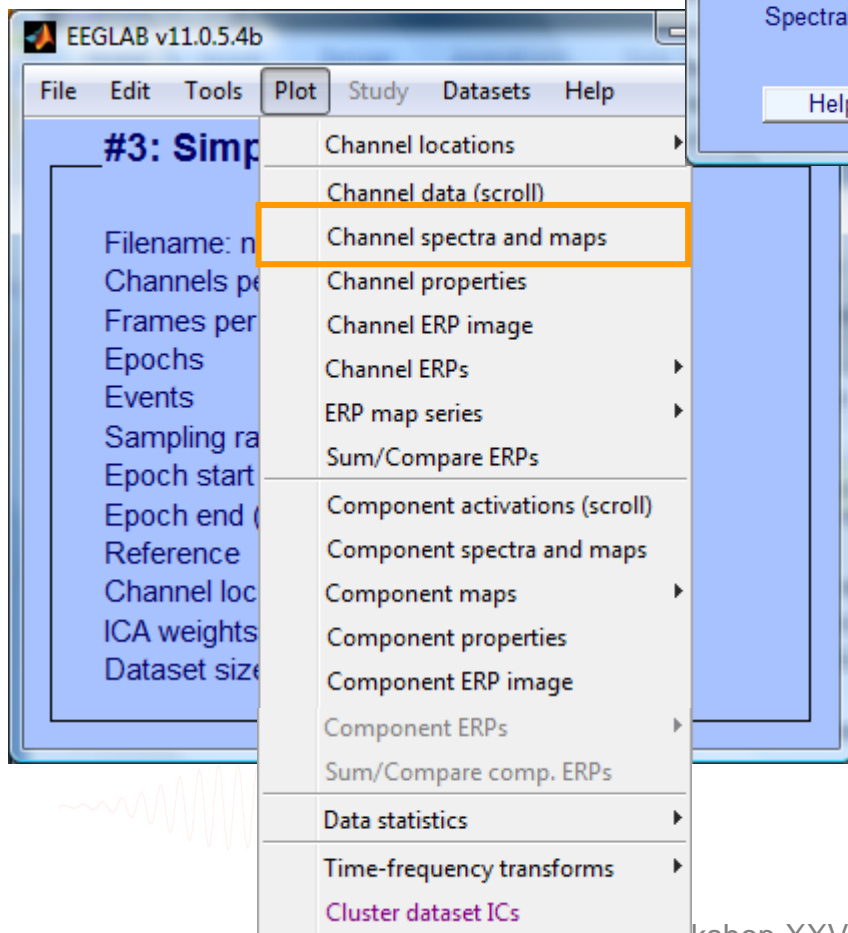
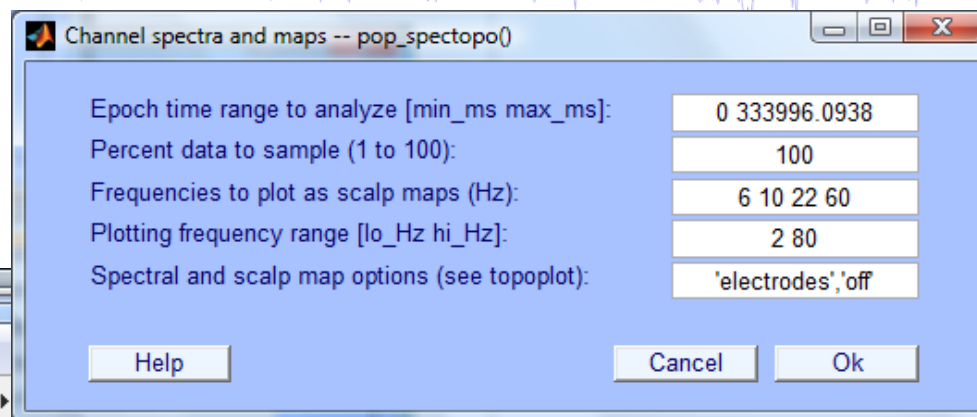
Scroll channel activities -- eegplot()

Figure Display Settings Help

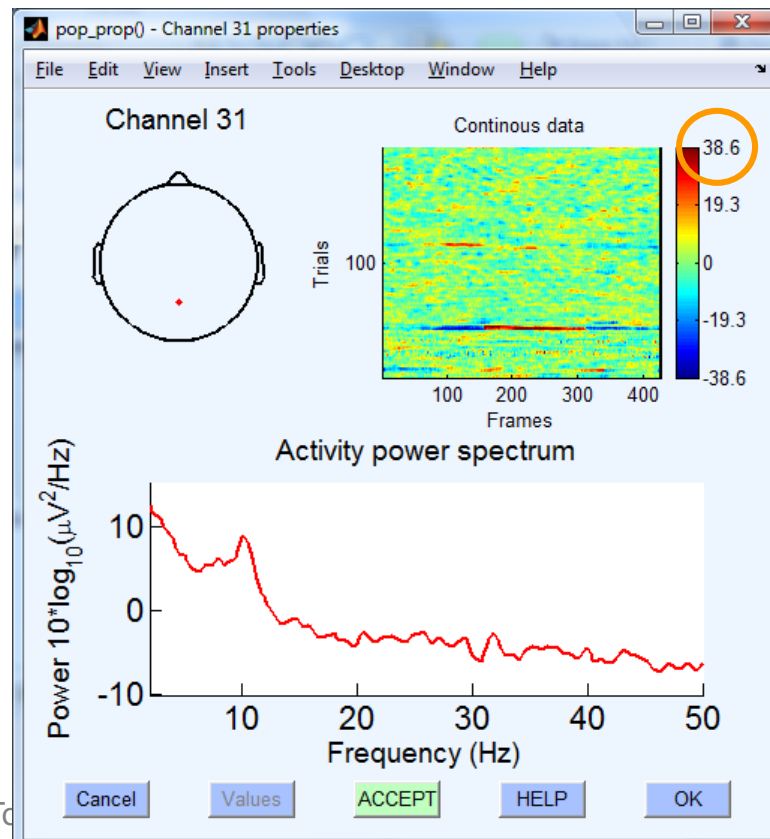
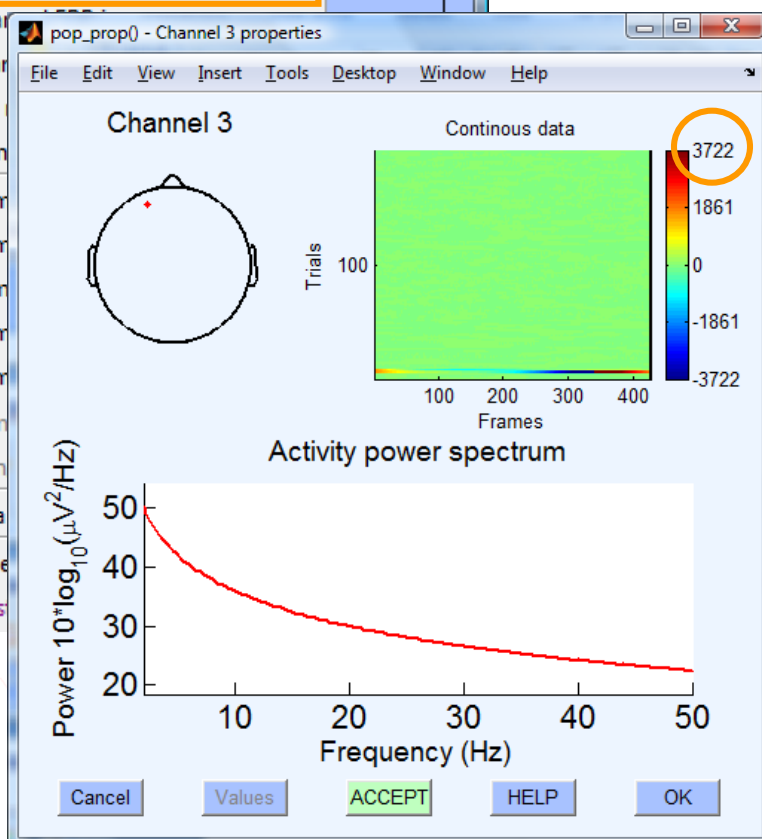
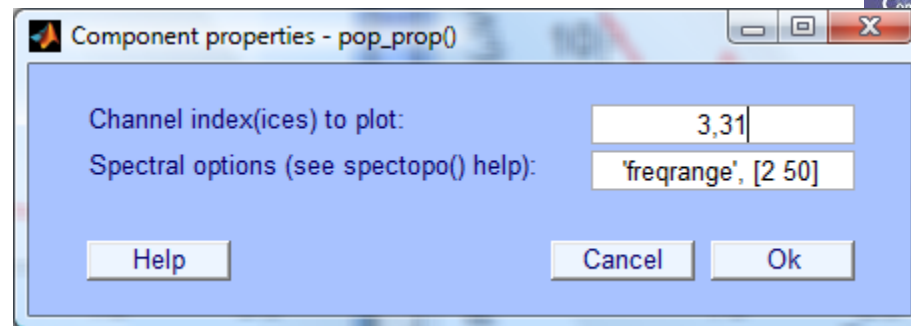
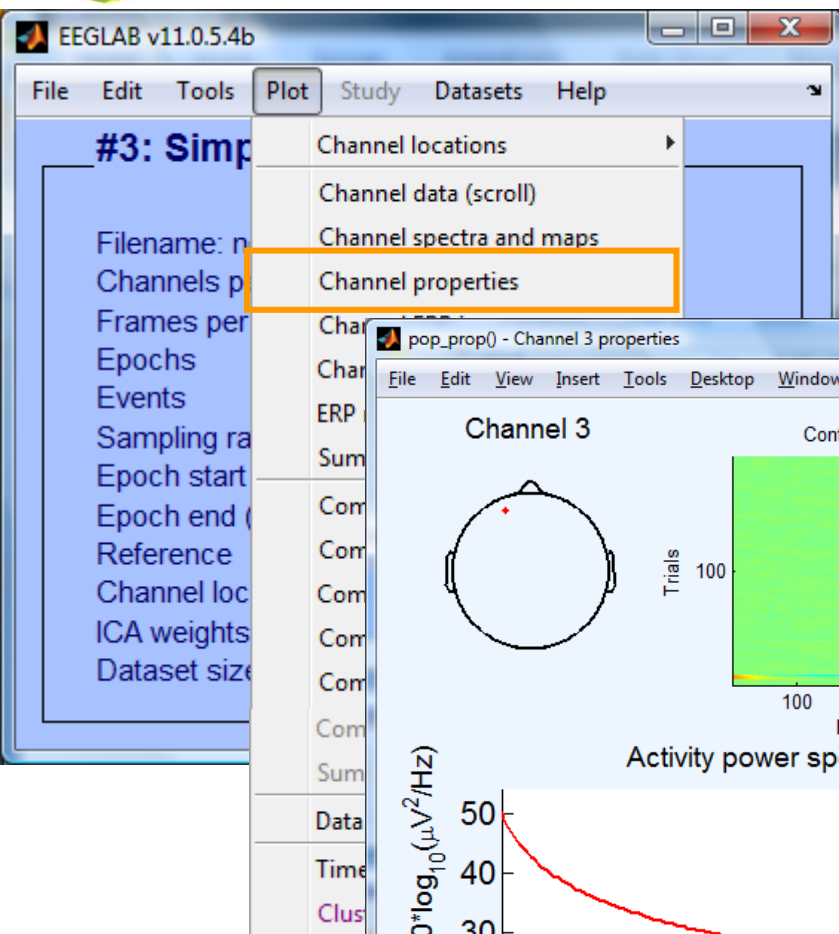
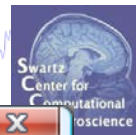
1) Identify bad channel



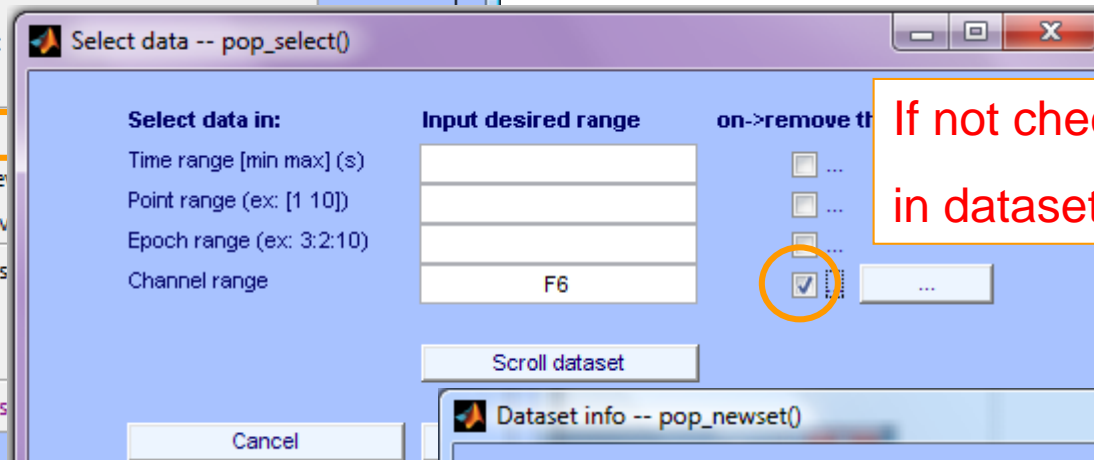
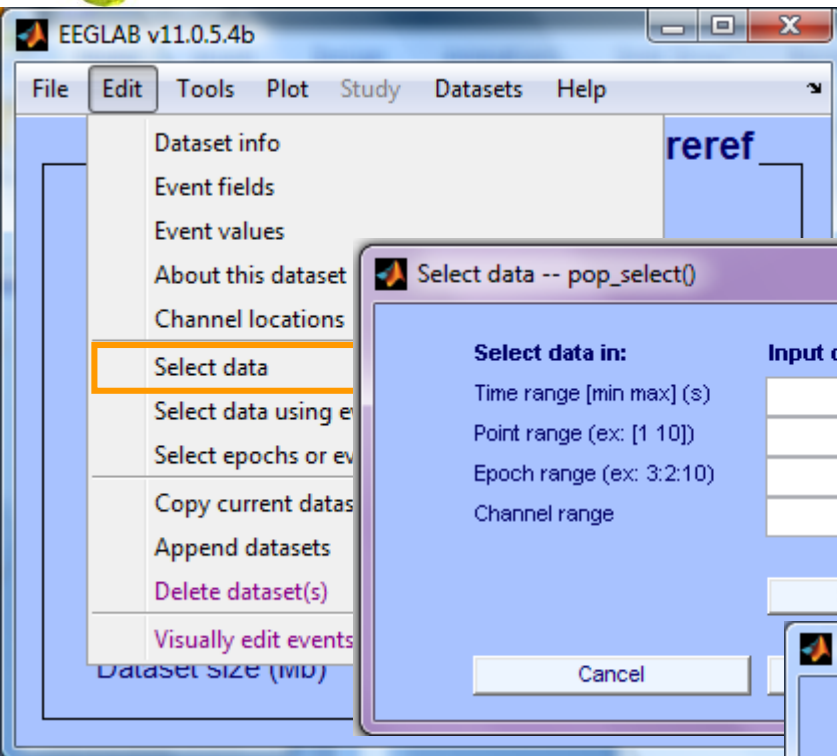
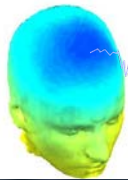
Manually identifying bad channels



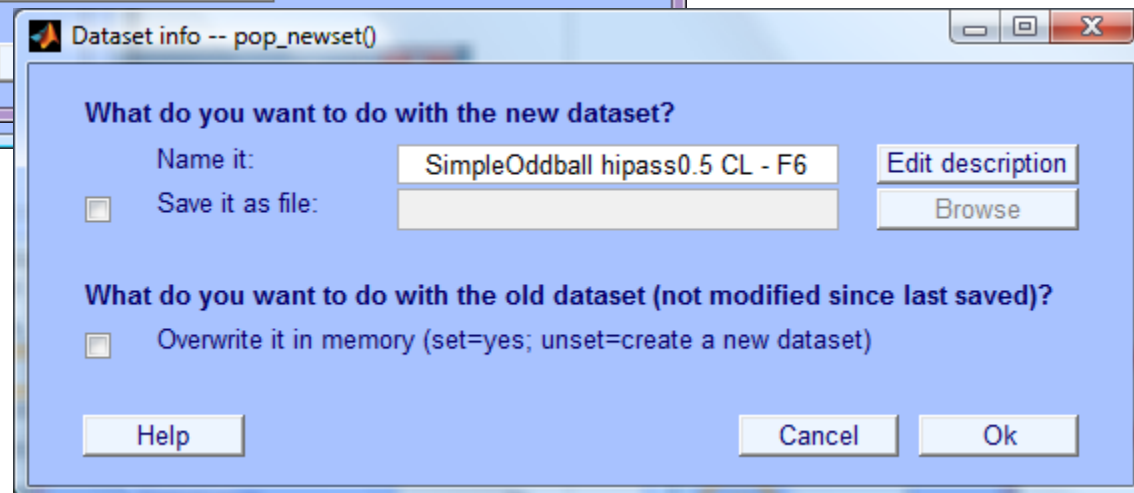
Manually identifying bad channels



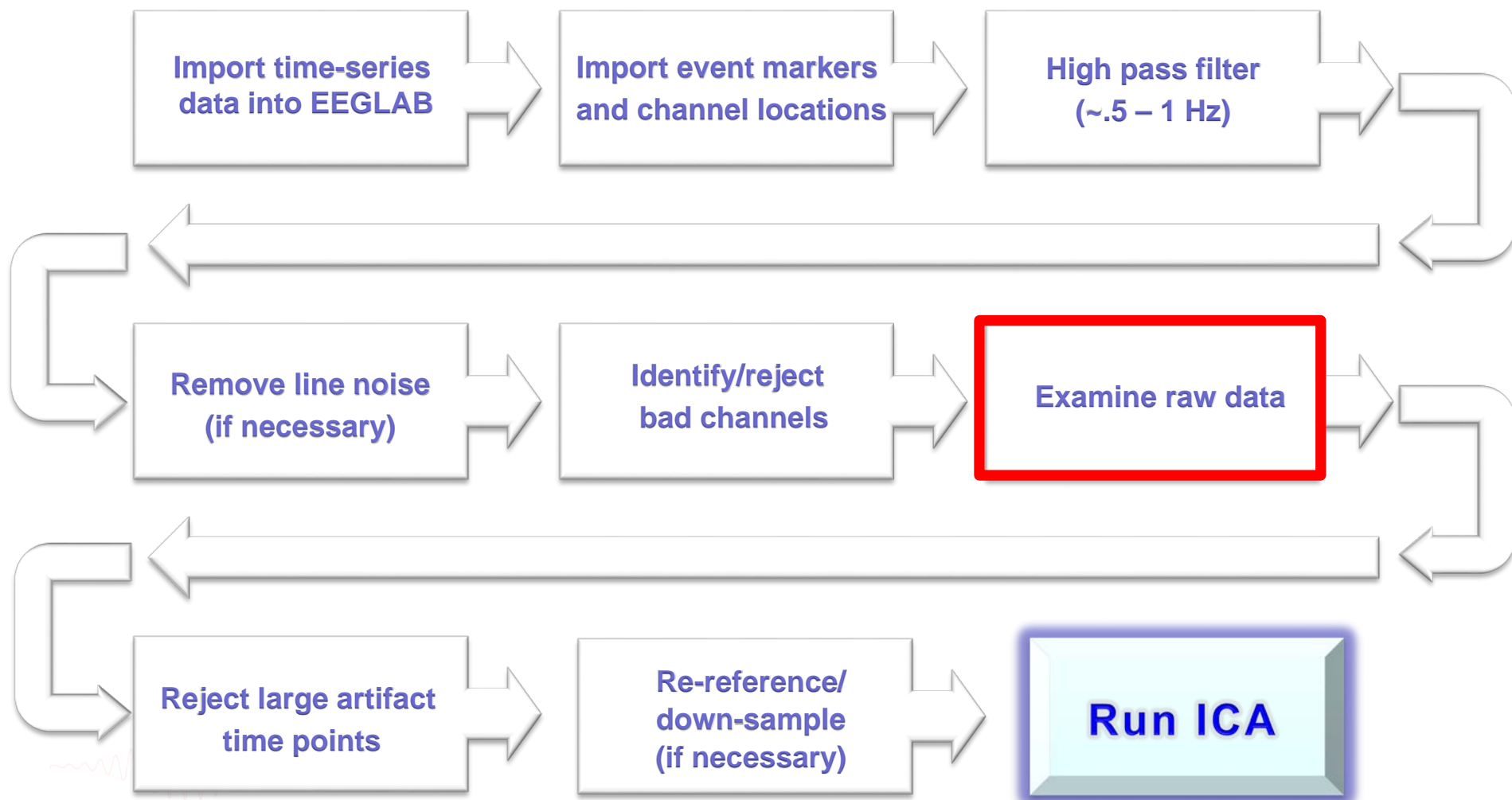
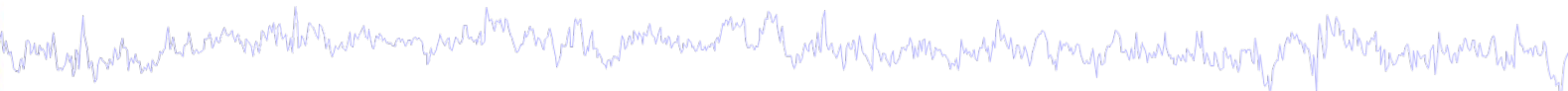
Removing channel(s)



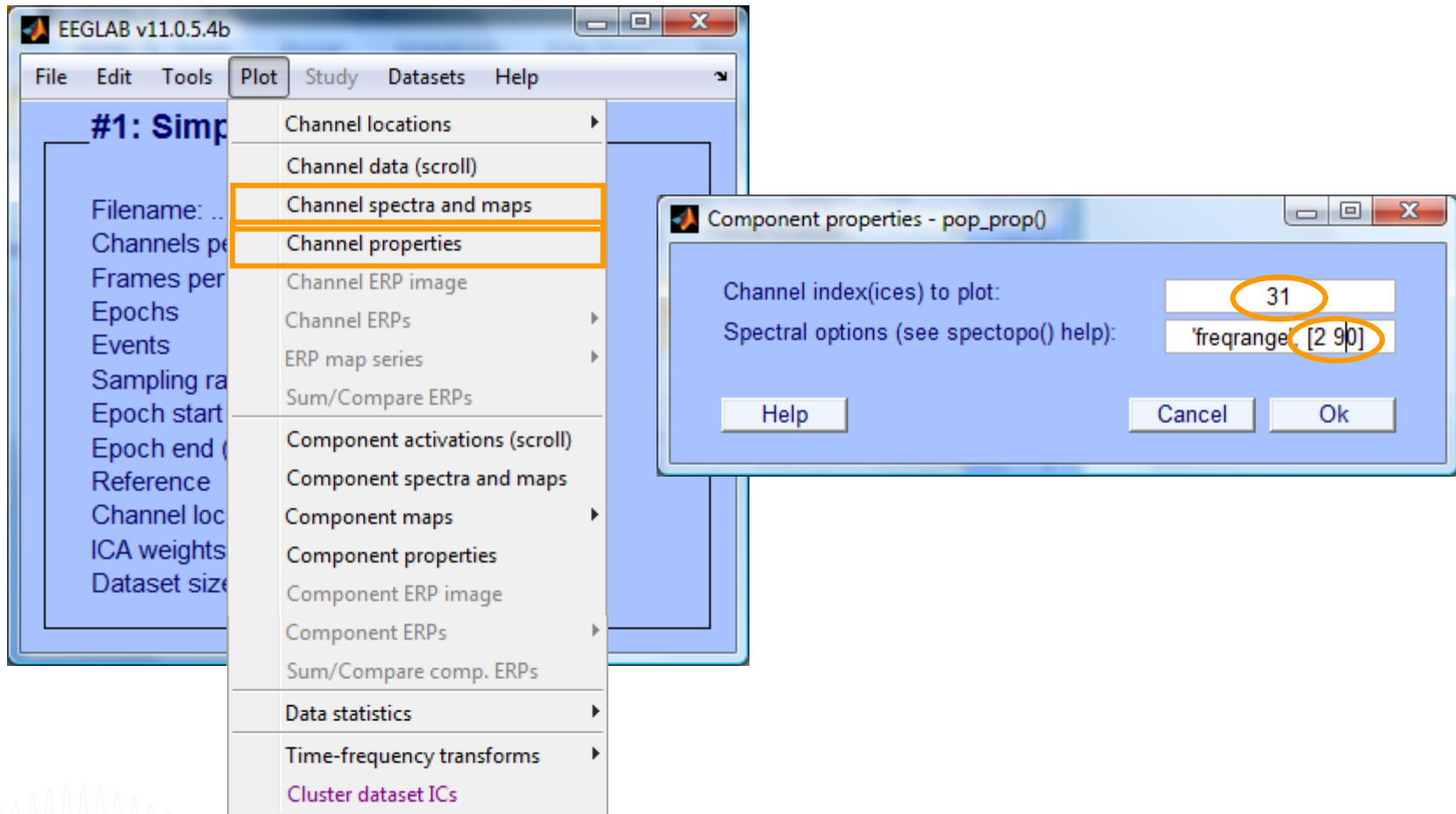
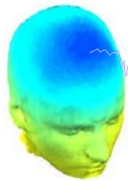
If not checked, will result
in dataset with one channel



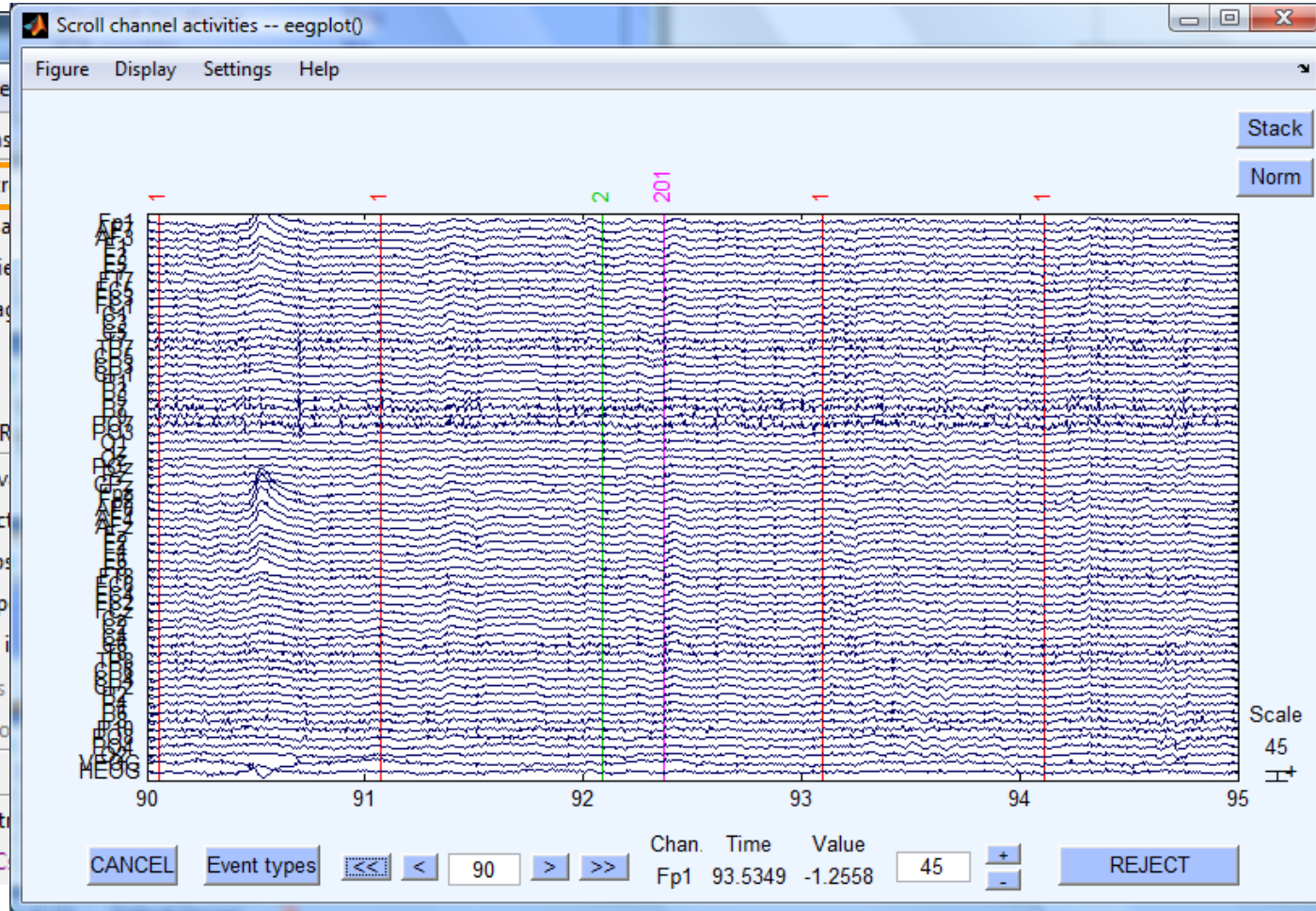
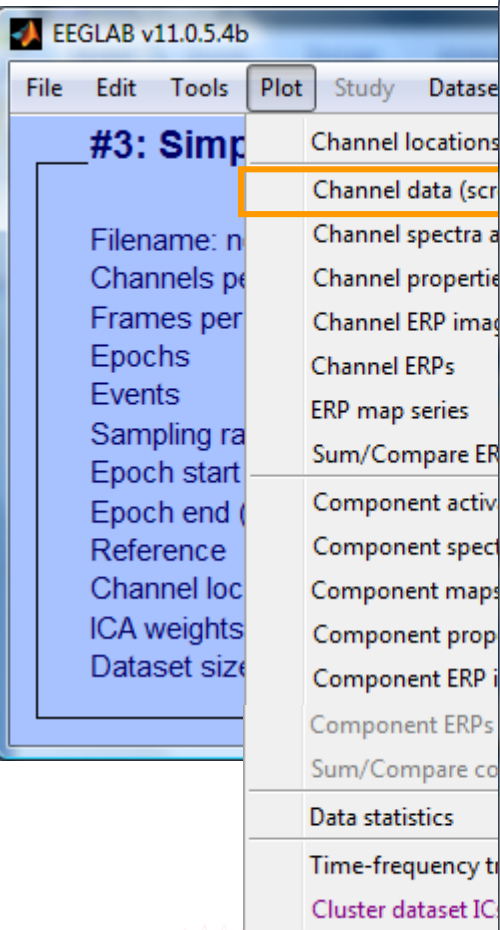
Pre-processing pipeline



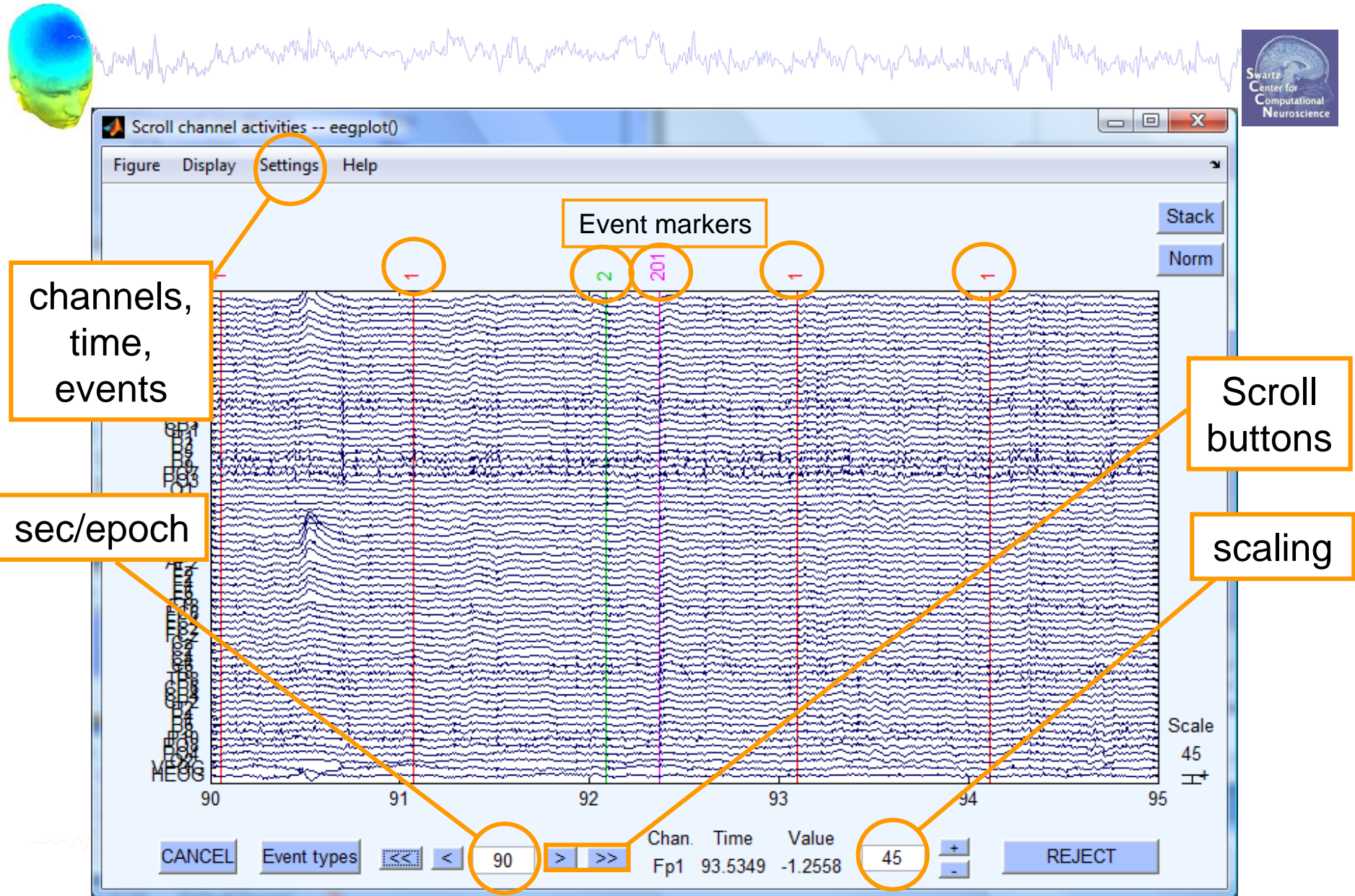
Plot channel spectra



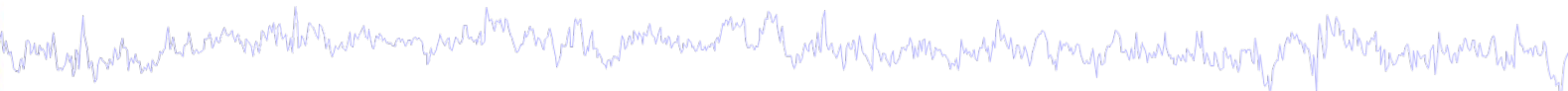
Scroll channel data



Scroll channel data



Pre-processing pipeline



**Import time-series
data into EEGLAB**

**Import event markers
and channel locations**

**High pass filter
(~.5 – 1 Hz)**

**Remove line noise
(if necessary)**

**Identify/reject
bad channels**

Examine raw data

**Reject large artifact
time points**

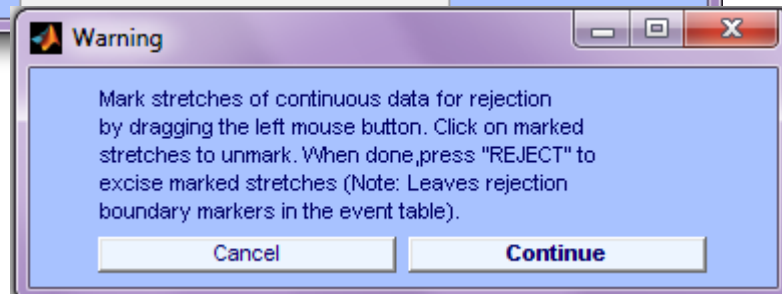
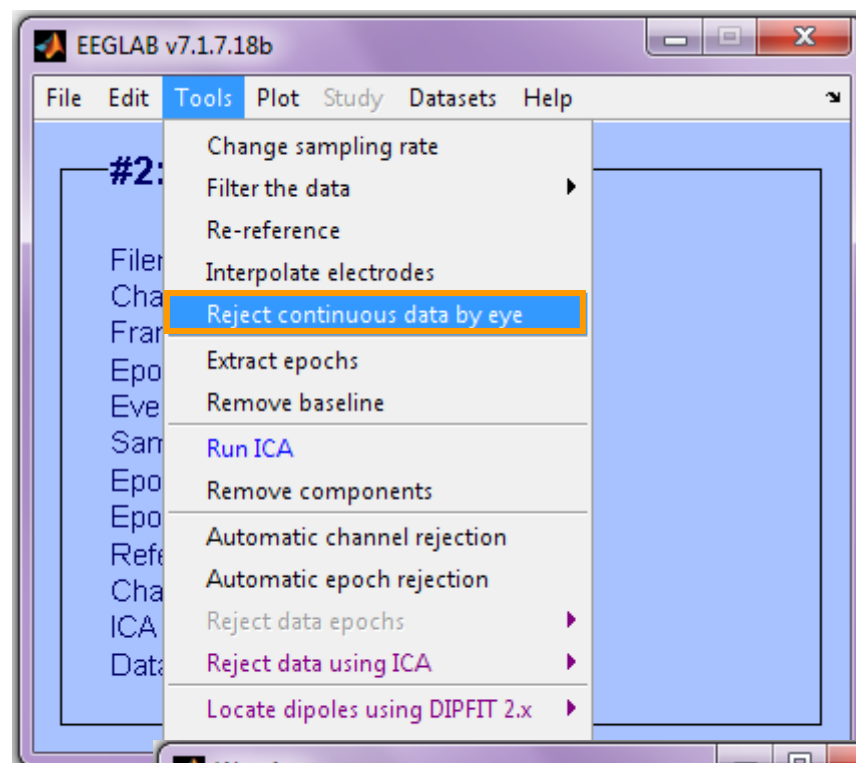
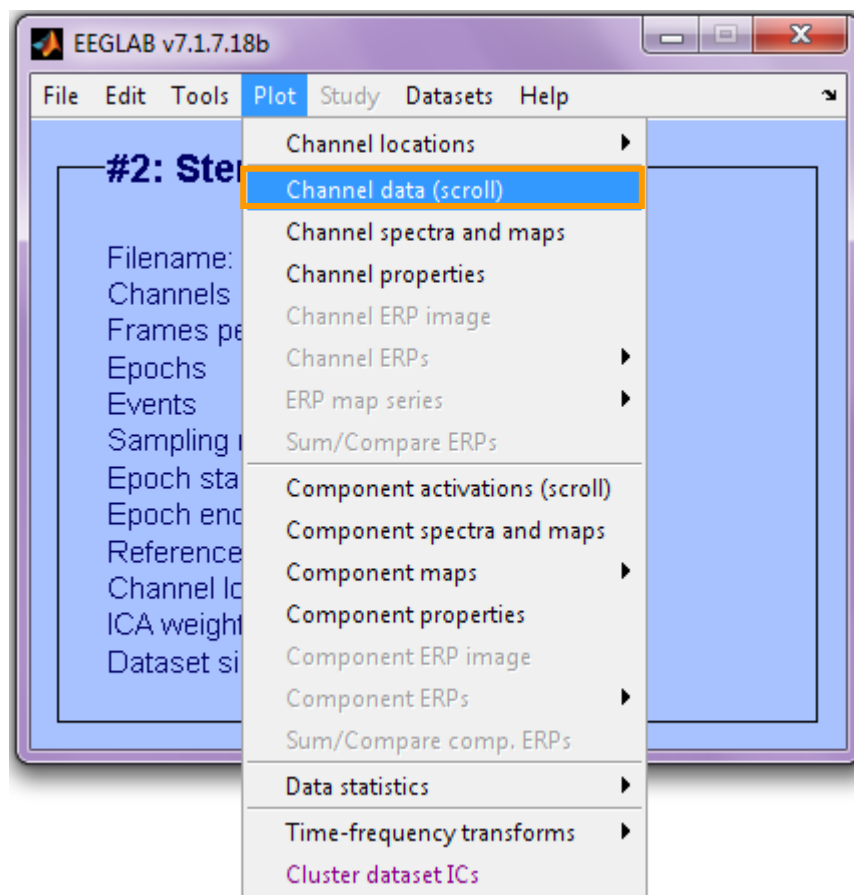
**Re-reference/
down-sample
(if necessary)**

Run ICA

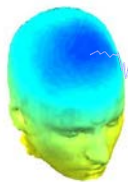
Reject continuous data



Equivalent



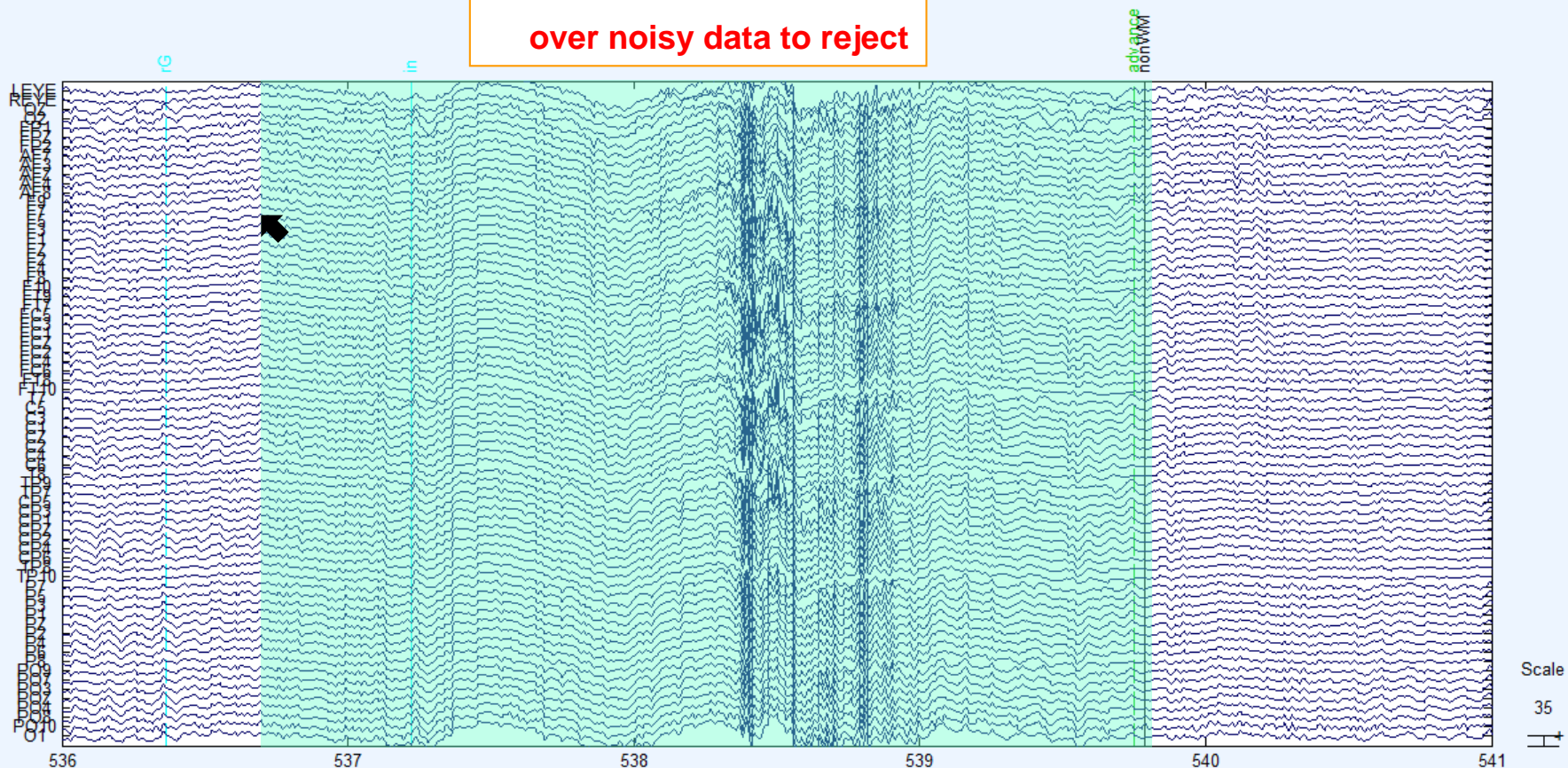
Reject continuous data



Scroll channel activities -- eegplot()

Figure Display Settings Help

Click and drag with mouse
over noisy data to reject



Scale
35
↑
↓

CANCEL

Event types

<<

<

536

>

>>

Chan.

Time

Value

FC6

539.9355

4.8773

35

+

-

REJECT

Rejecting data for ICA

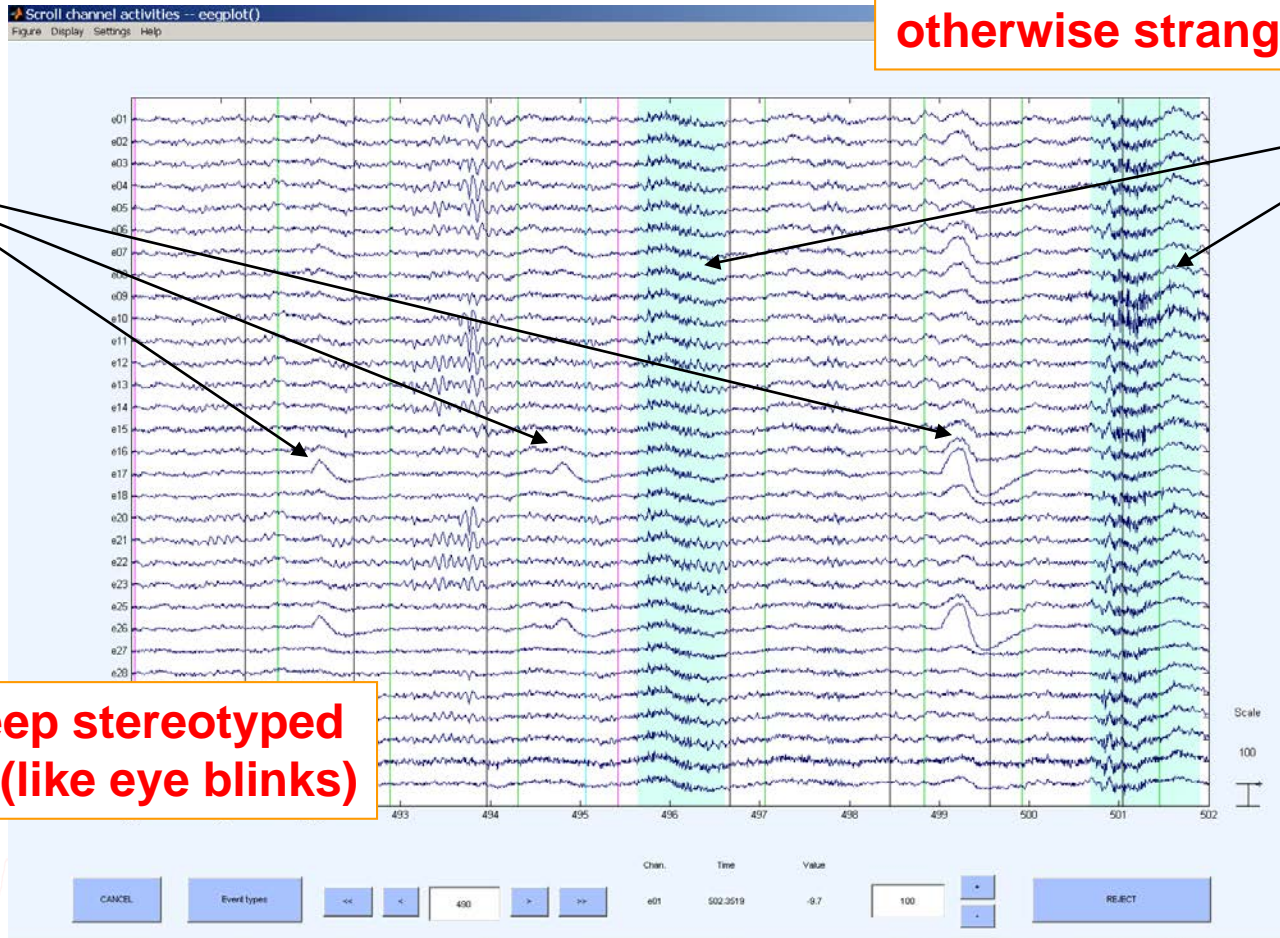
To prepare data for ICA:

Reject large muscle or otherwise strange events...

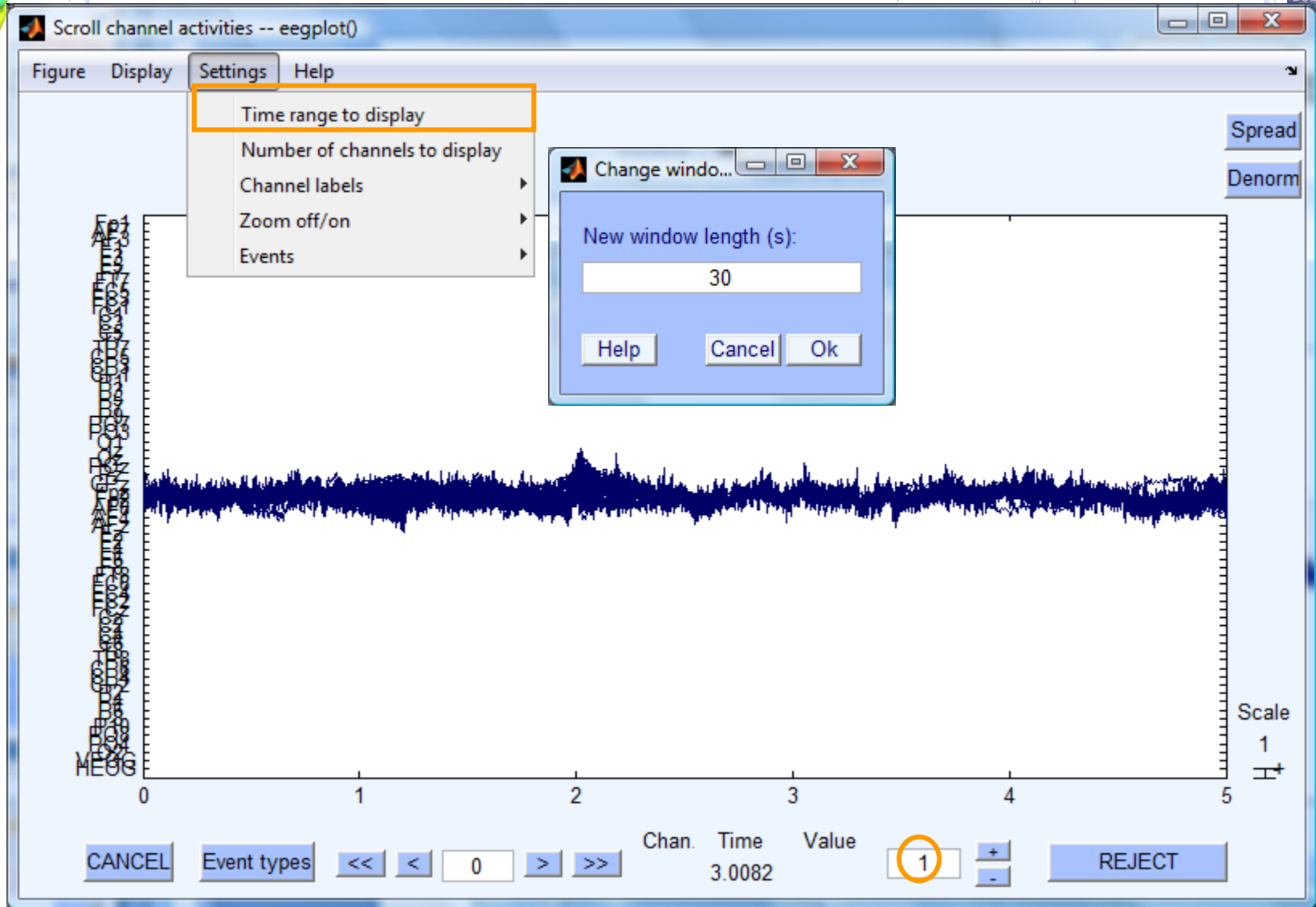
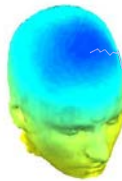
Reject

Keep

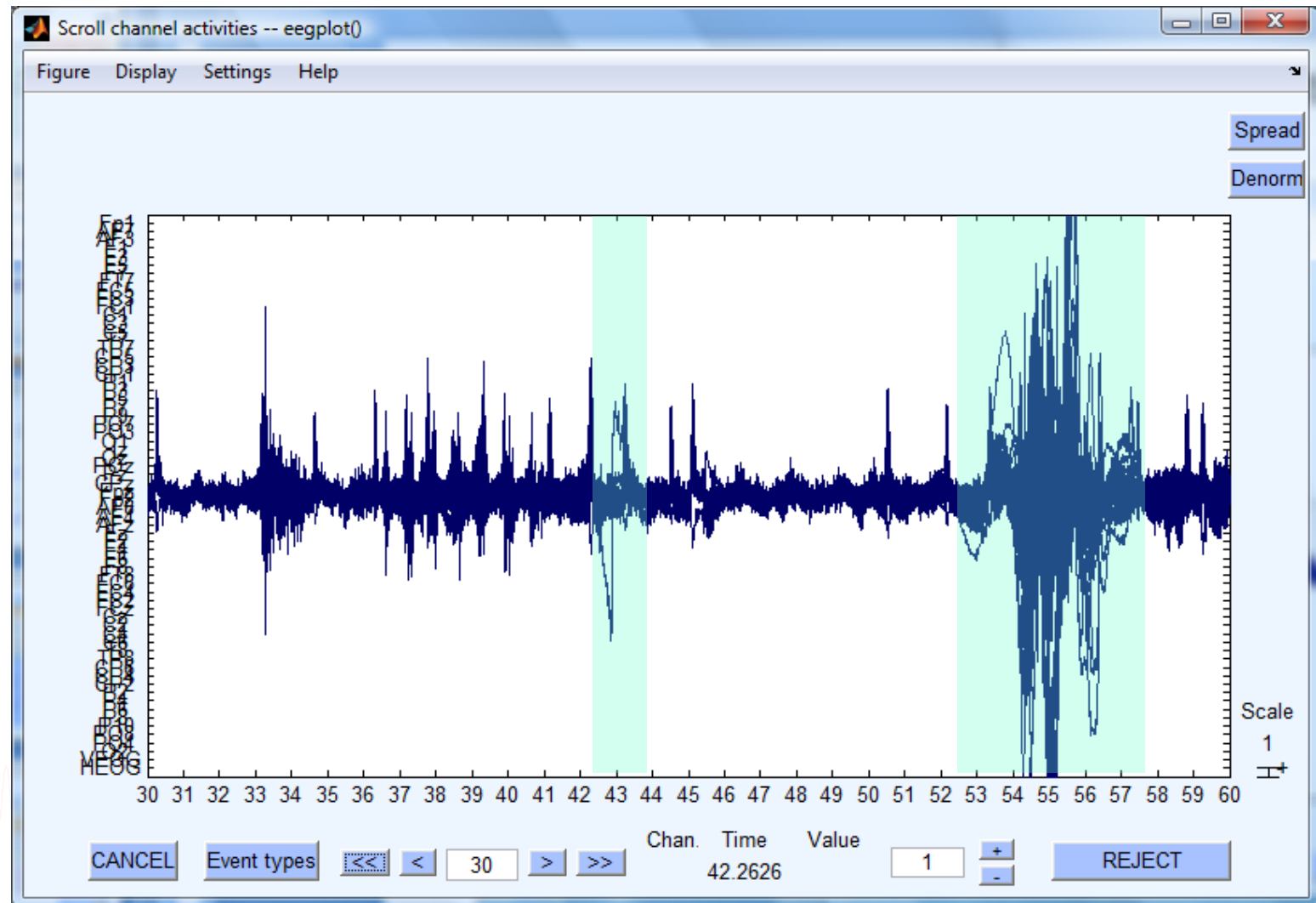
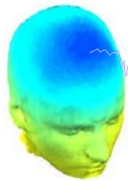
... but keep stereotyped artifacts (like eye blinks)

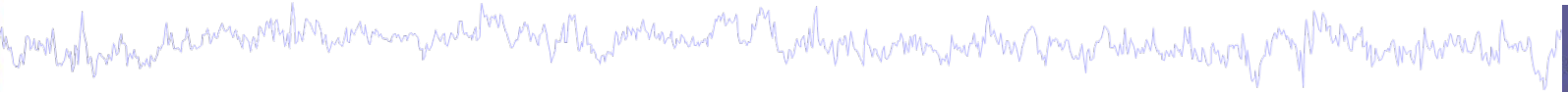
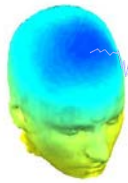


Fast (but sloppy) artifact rejection



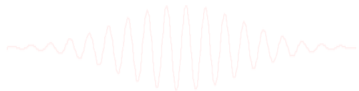
Fast (but sloppy) artifact rejection



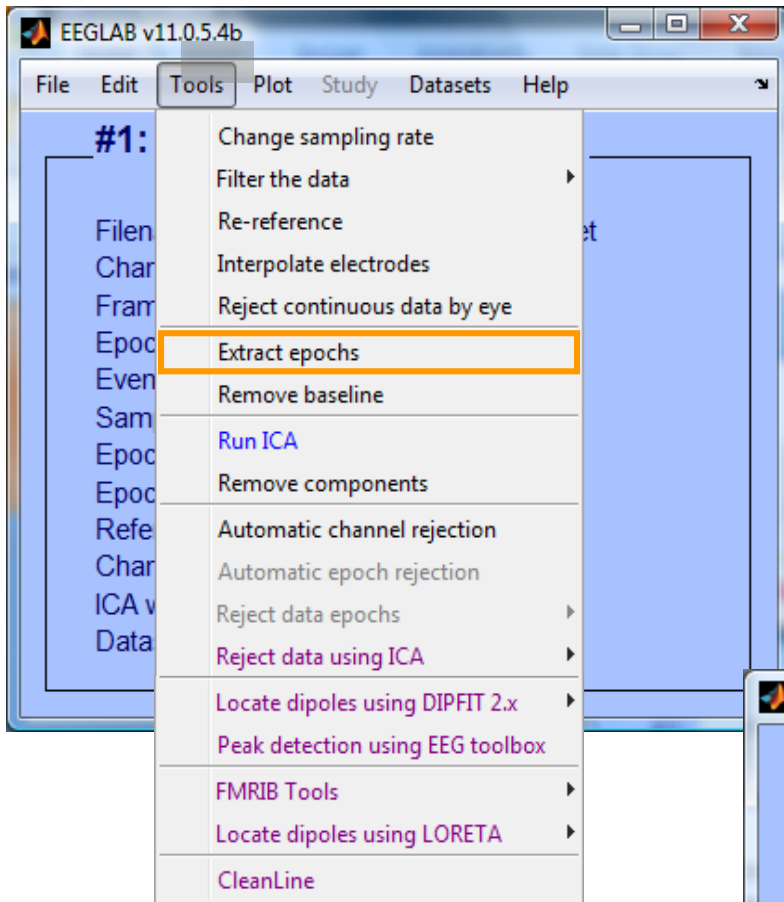


Data Cleaning for ICA

Variant 2: Epoched Data

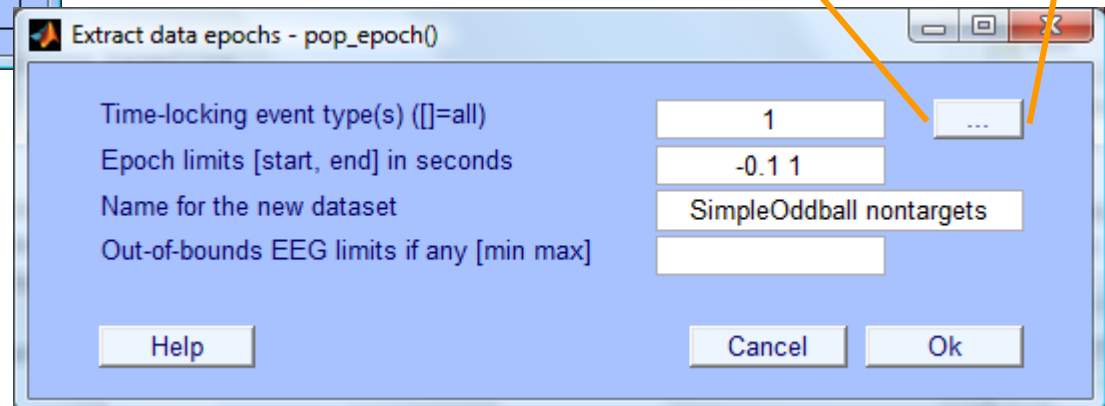
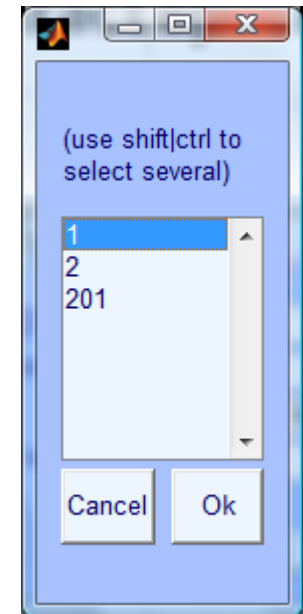


Extract epochs

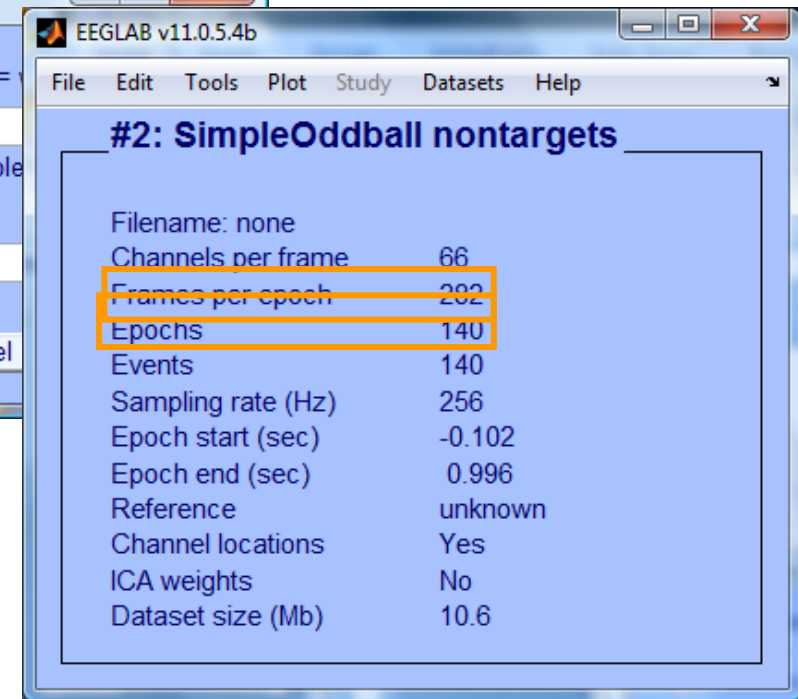
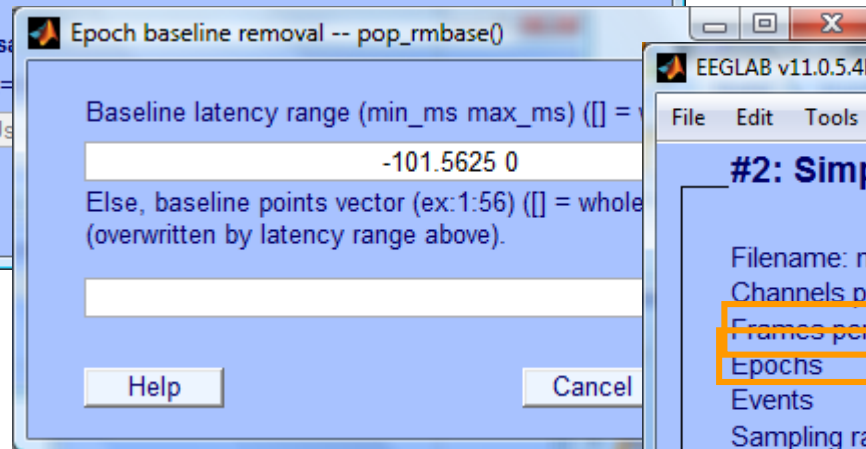
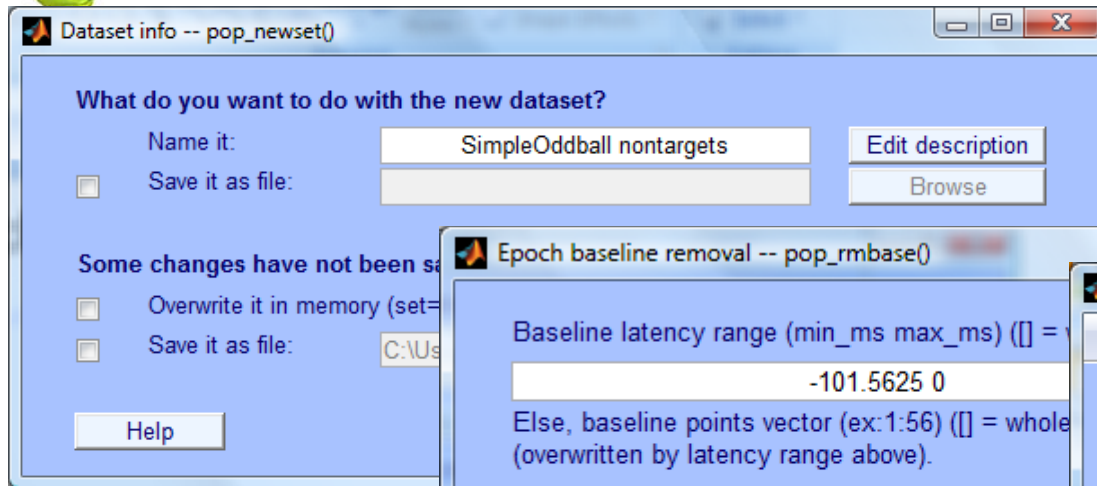


```
>> eeg_eventtypes (EEG)
```

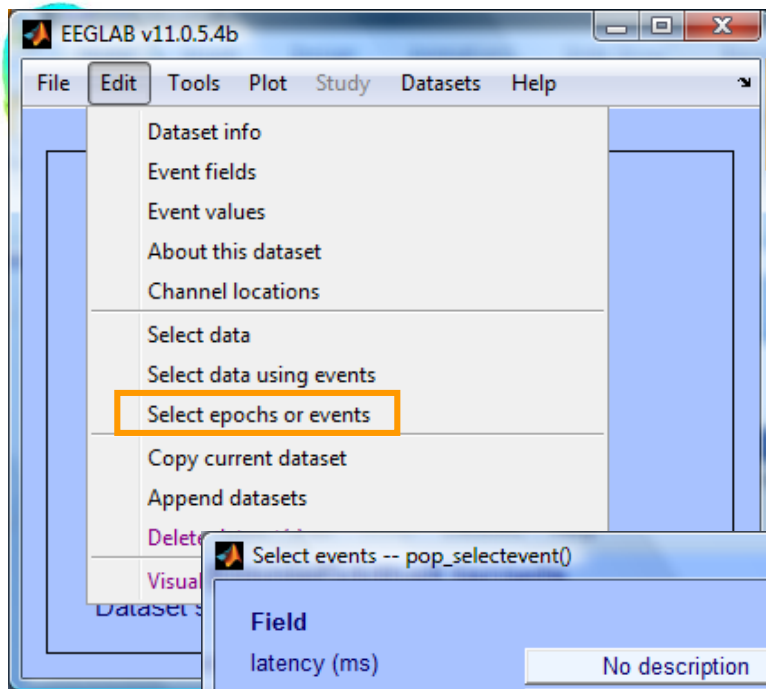
1	140
2	60
201	60



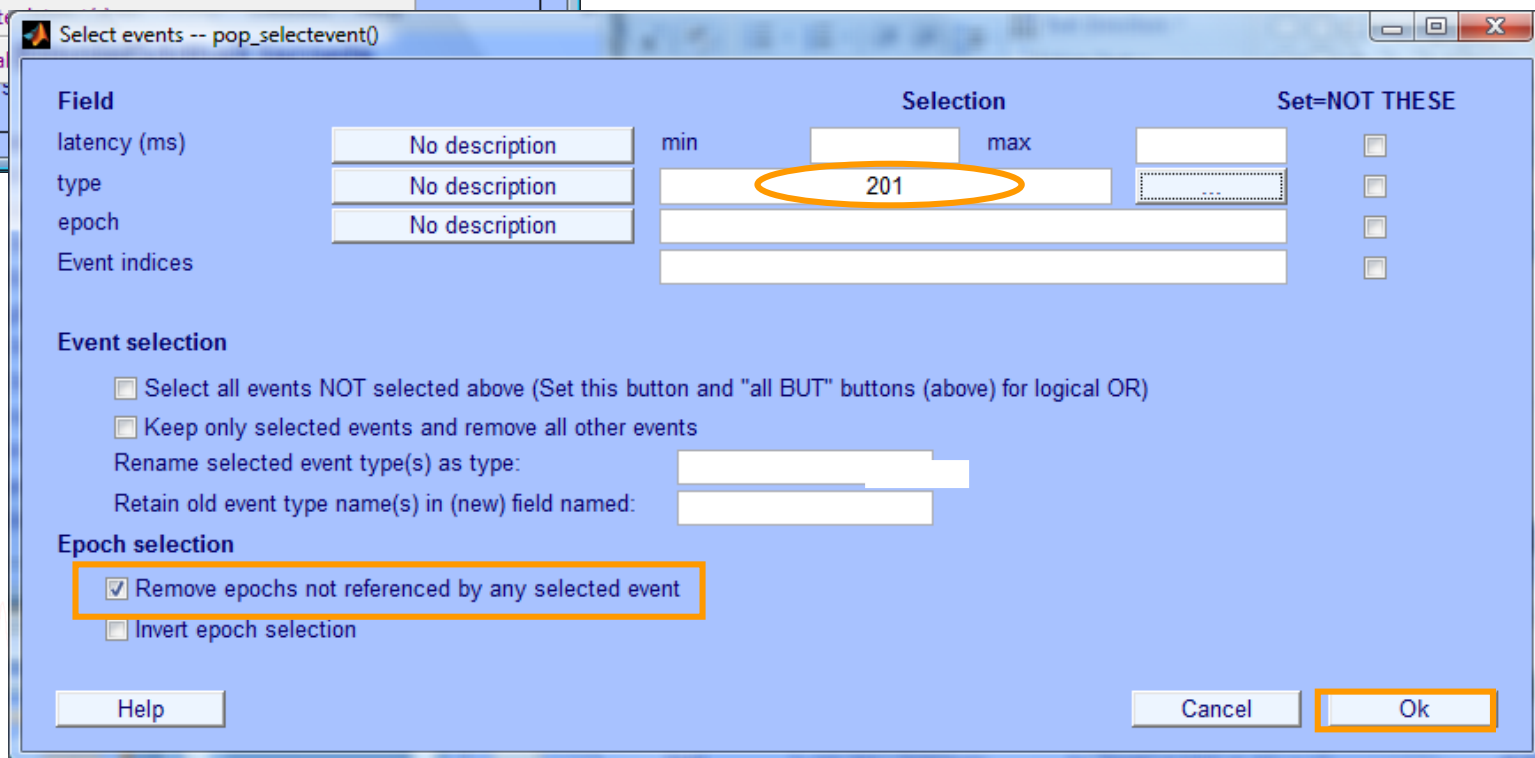
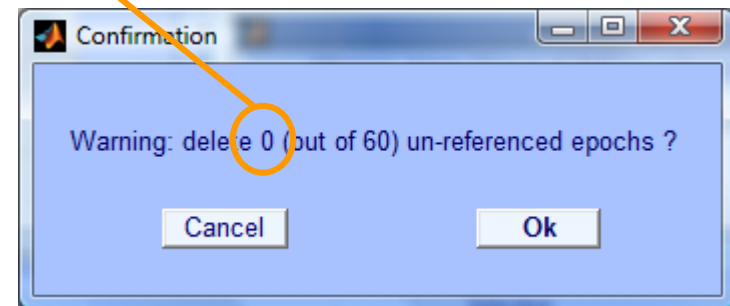
Extract epochs

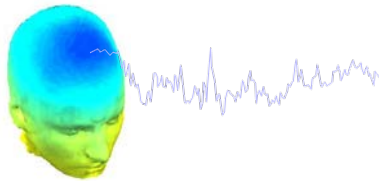


Select a subset of epochs

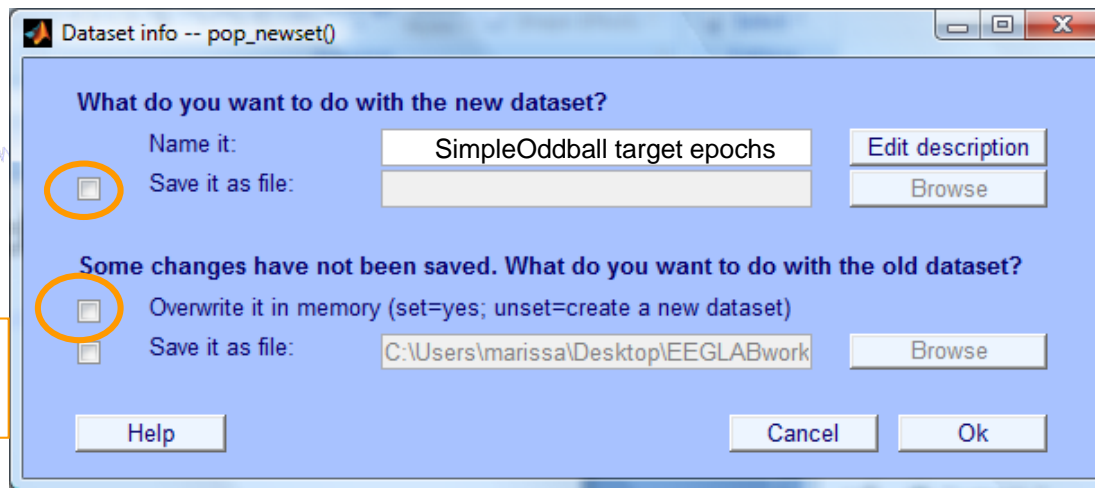


'0' because the subject did not miss any targets

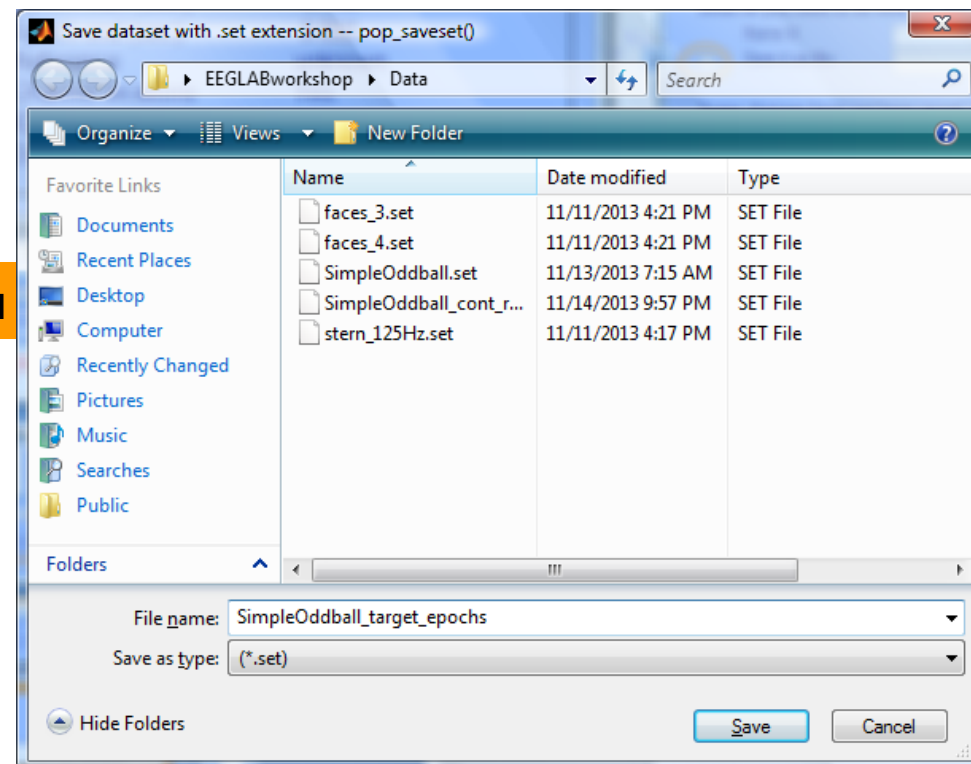
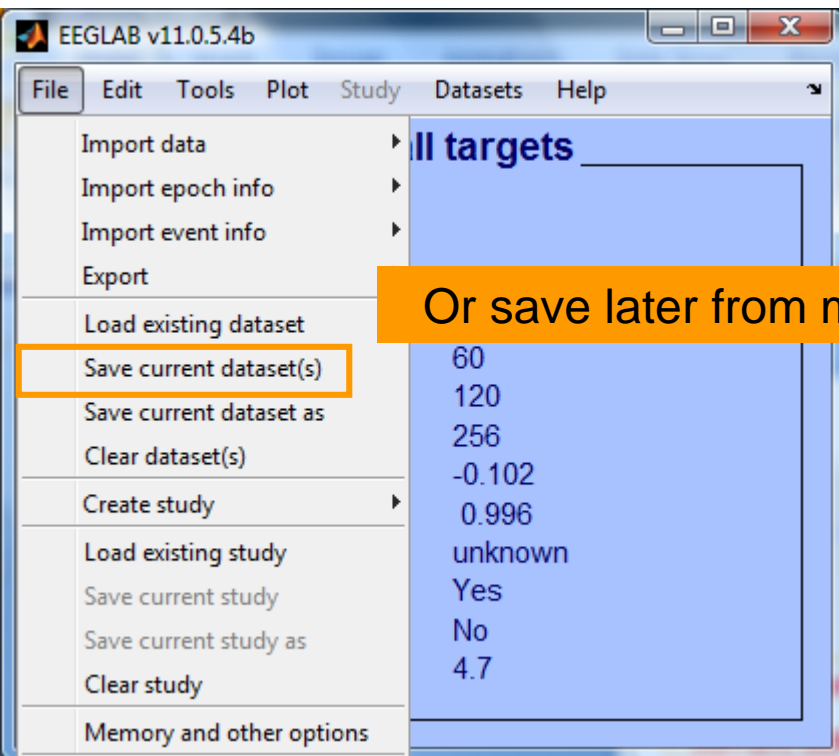




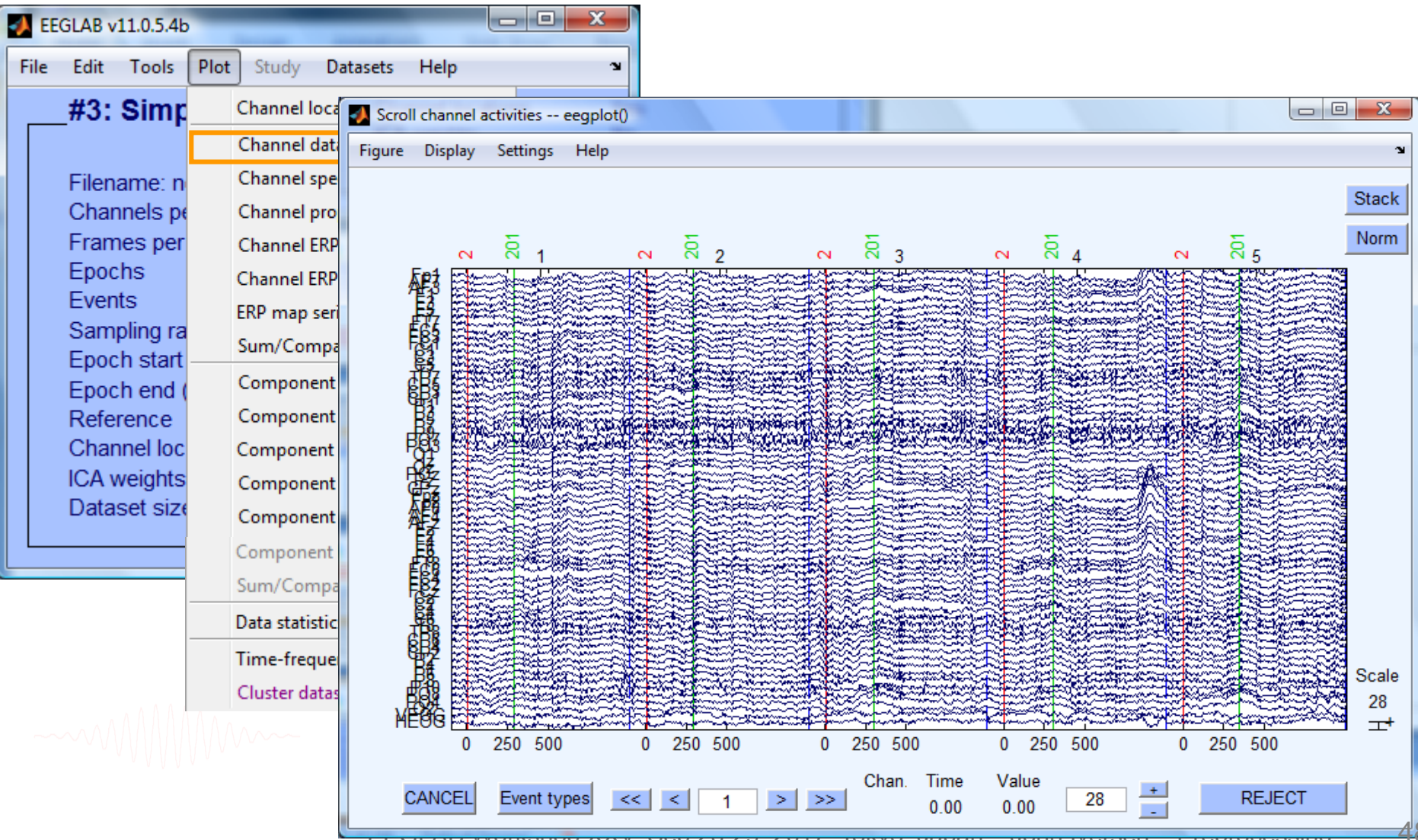
**'Do not overwrite
current dataset'**



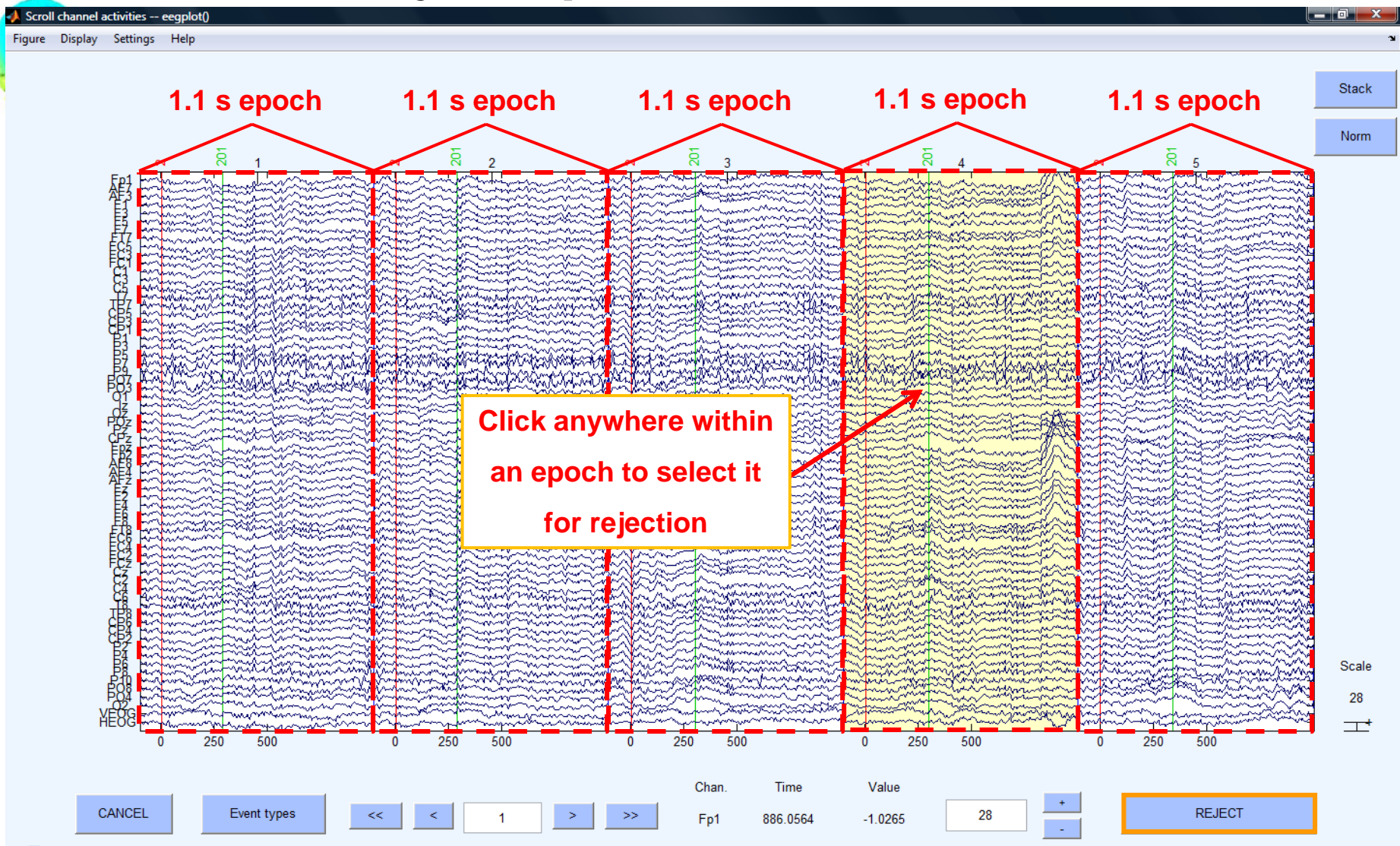
Save dataset (optional)



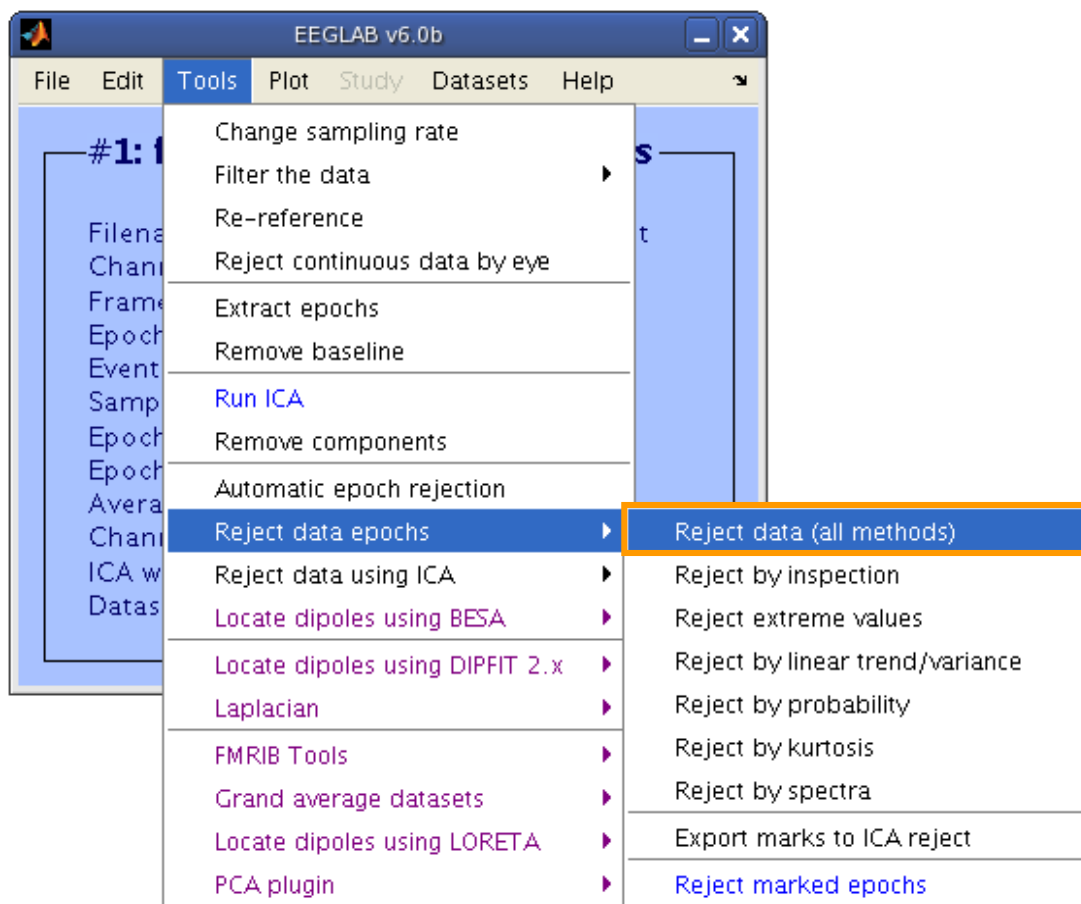
Scroll (epoched) channel data



Reject epochs with artifact



Reject data epochs



Reject data epochs



Reject trials using data statistics - pop_rejmenu()

Mark trials by appearance ☐ Marked trials 0

Find abnormal values ☐

Upper limit(s) (uV) Lower limit(s) (uV)
Start time(s) (ms) Ending time(s) (ms)
Electrode(s) Currently marked trials

Find abnormal trends ☐

Max slope (uV/epoch) R-squared limit (0 to 1)
Electrode(s) Currently marked trials

Find improbable data ☐

Single-channel limit (std. dev.) All channels limit (std. dev.)
Electrode(s) Currently marked trials

Find abnormal distributions ☐

Single-channel limit (std. dev.) All channels limit (std. dev.)
Electrode(s) Currently marked trials

Find abnormal spectra (slow) ☐

Upper limit(s) (dB) Lower limit(s) (dB)
Low frequency(s) (Hz) High frequency(s) (Hz)
Electrode(s) Currently marked trials

Plotting options

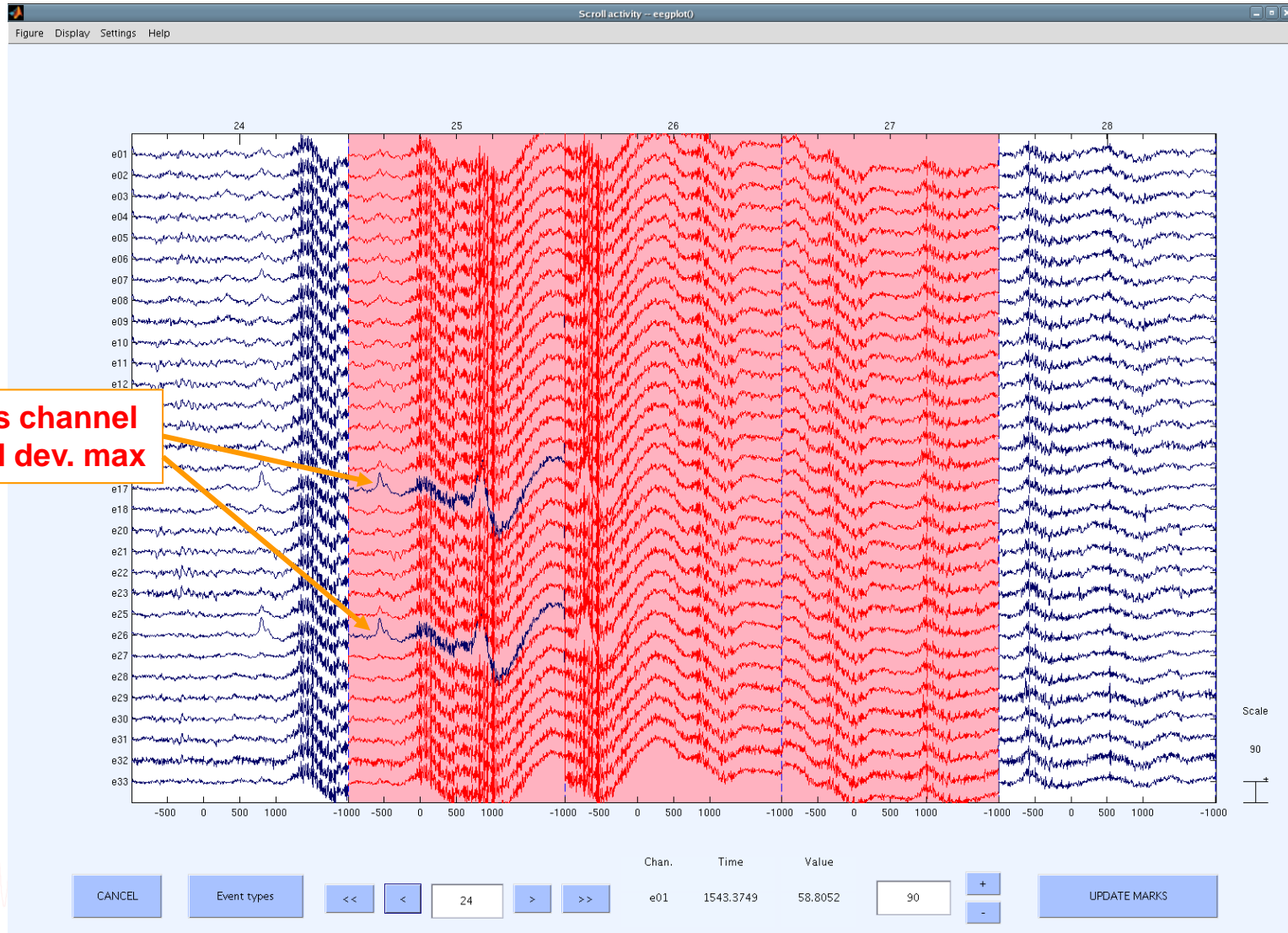
Show all trials marked for rejection by the measure selected above or checked below | /

☒ Abnormal appearance ☒ Abnormal values ☒ Abnormal trends
☒ Improbable epochs ☒ Abnormal distributions ☒ Abnormal spectra

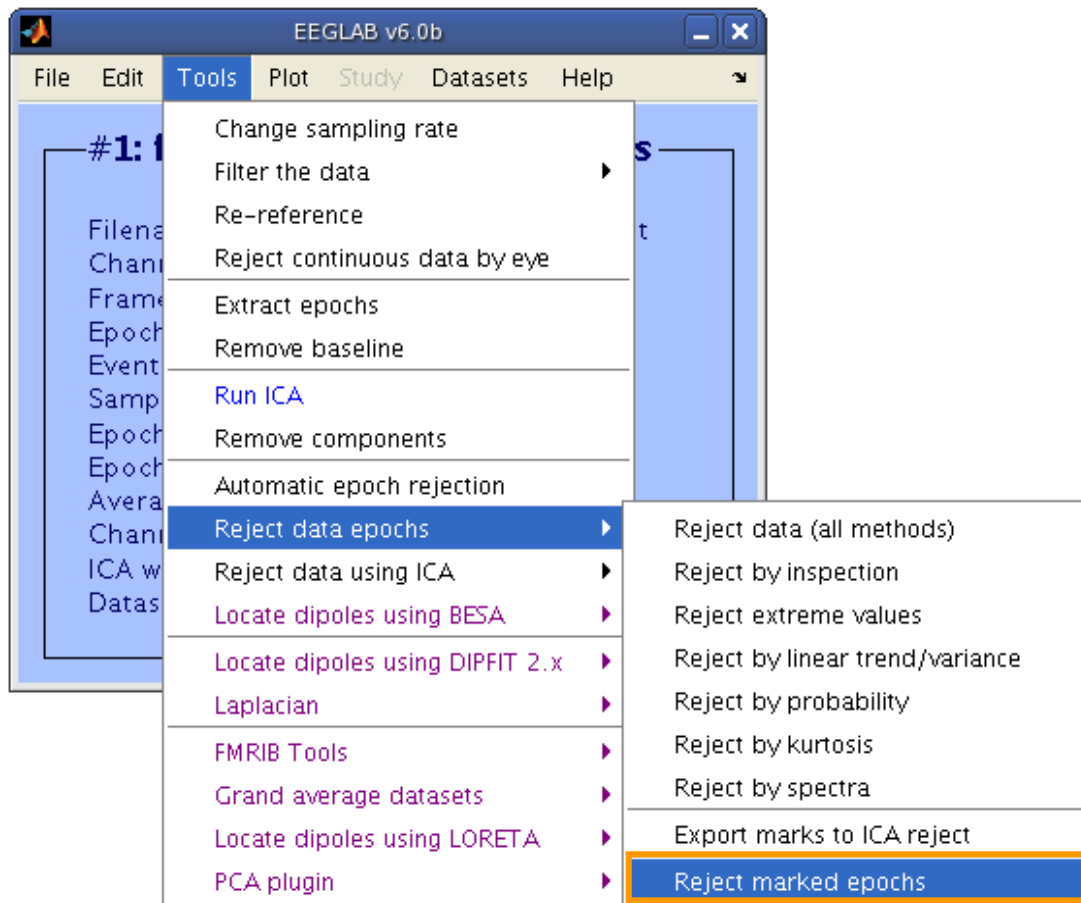
visual
inspection

probability

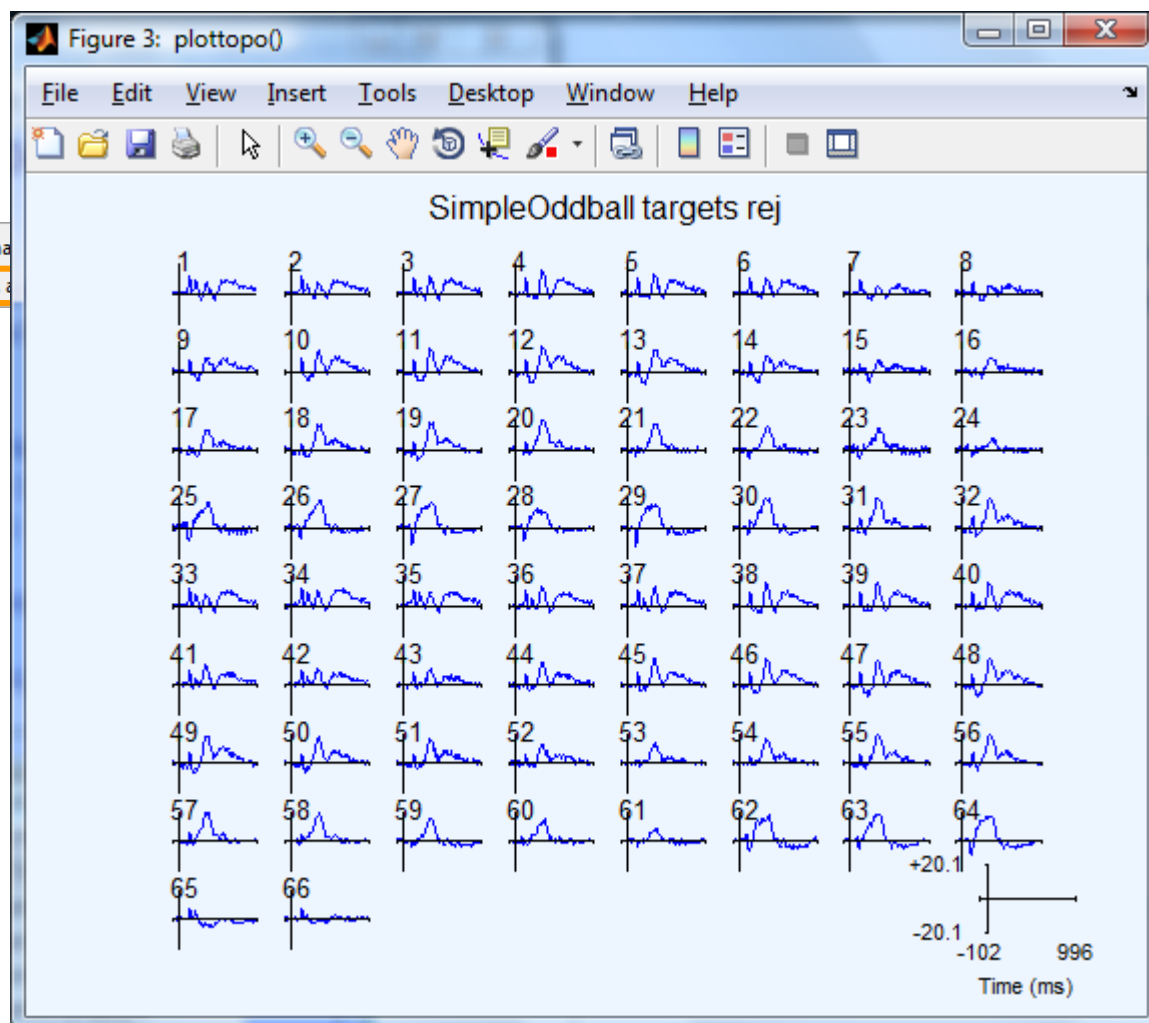
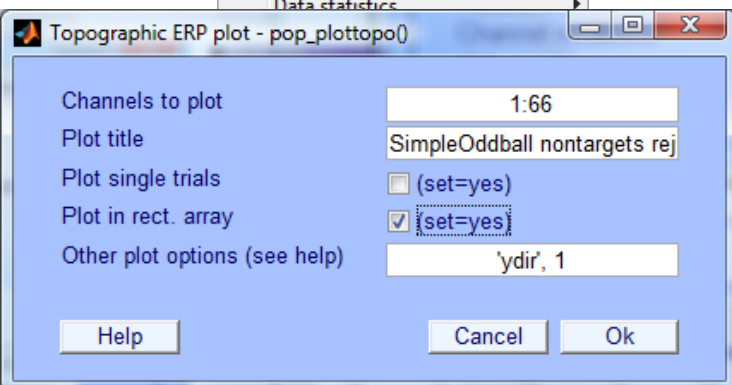
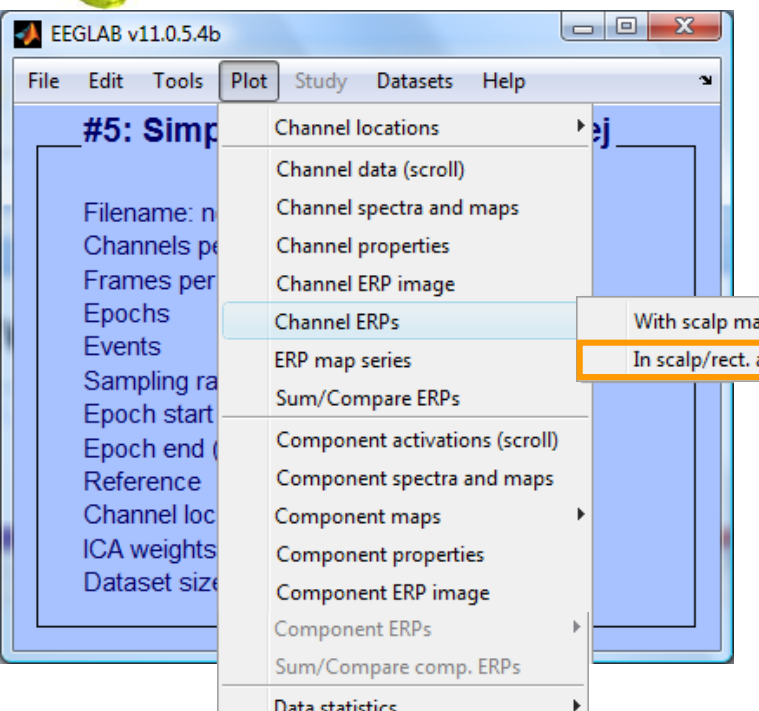
Reject data epochs



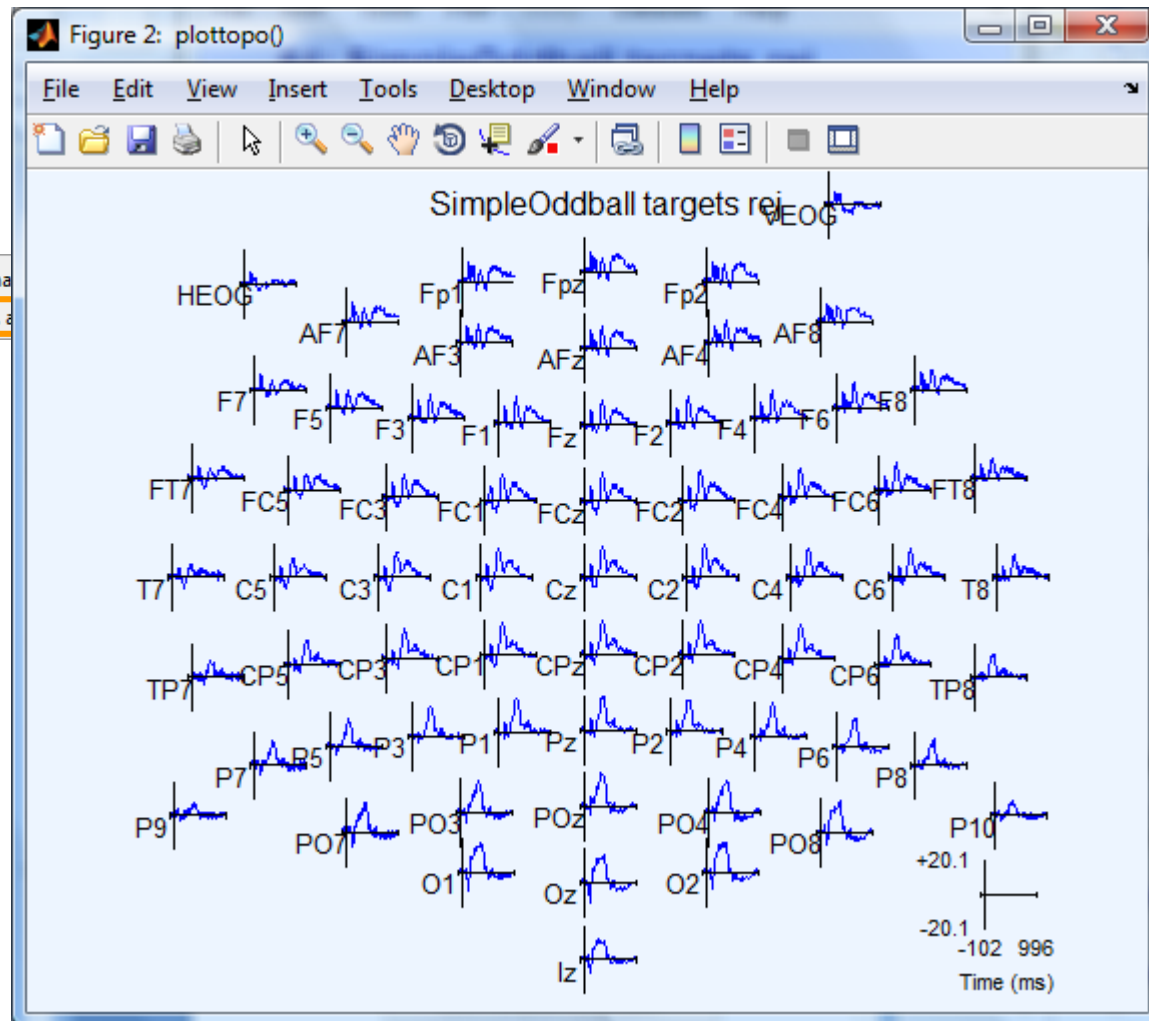
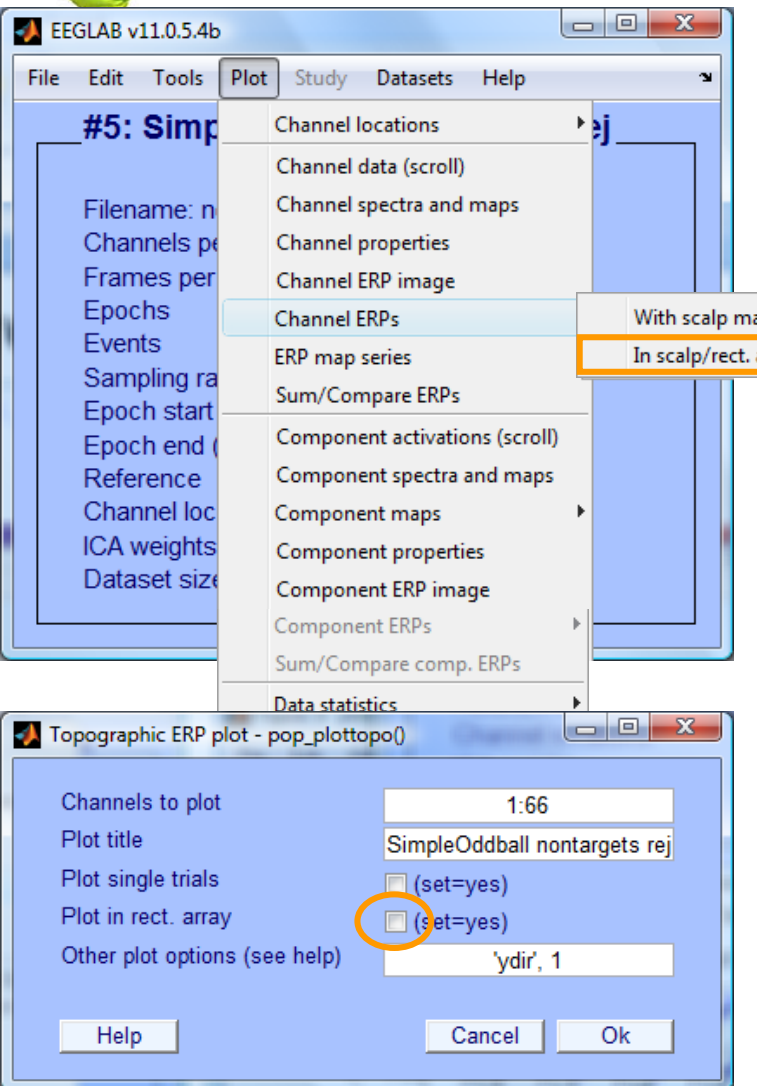
Reject data epochs



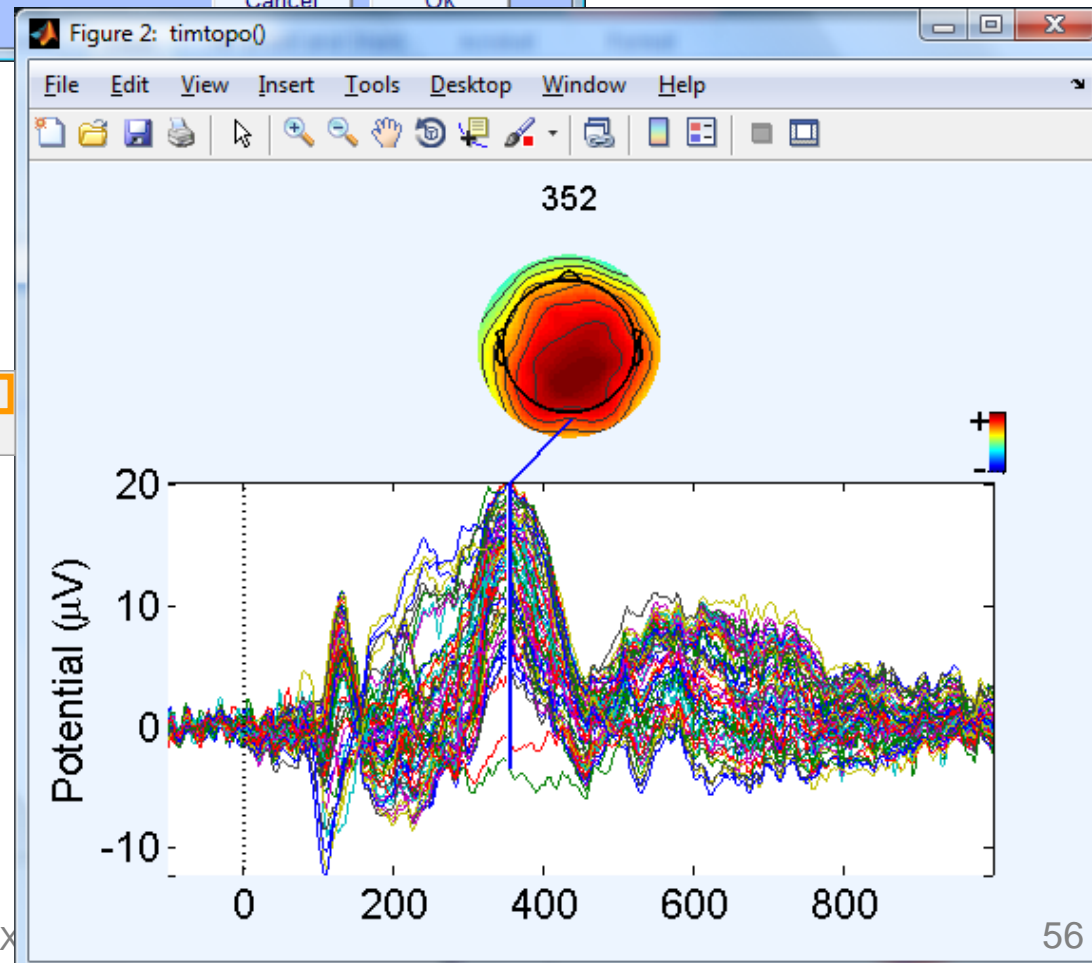
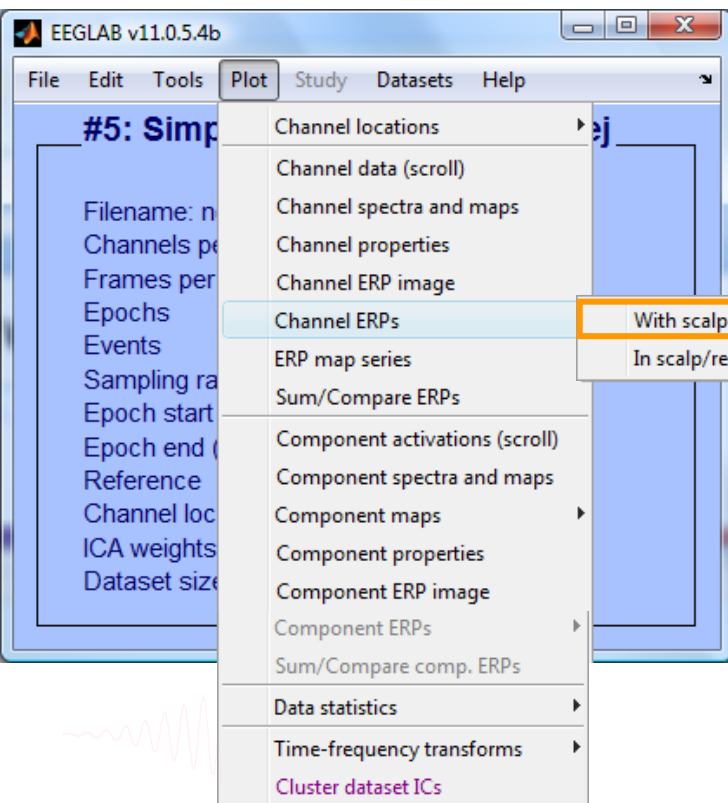
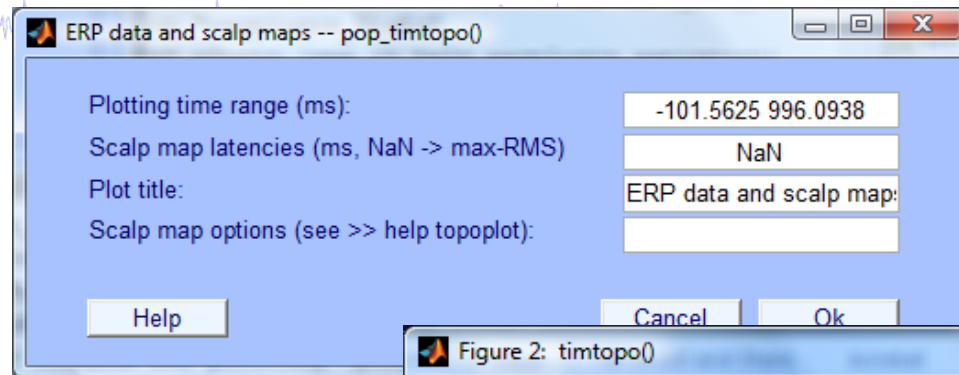
Visualize ERP in rectangular array



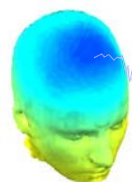
Visualize ERP in topographic array



Visualize ERP scalp distribution



Visualize channel ERPs in 2D



EEGLAB v11.0.5.4b

File Edit Tools **Plot** Study Data

#4: SimpleOddball

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

- Channel location
- Channel data (s...
- Channel spectr...
- Channel proper...
- Channel ERP in...
- Channel ERPs
 - ERP map series
 - Sum/Compare ERPs
- Component activations (scroll)
- Component spectra and maps
- Component maps
- Component properties
- Component ERP image
- Component ERPs
- Sum/Compare comp. ERPs
- Data statistics
- Time-frequency transforms
- Cluster dataset ICs

Plot ERP scalp maps in 2-D -- pop_topoplot()

Plotting ERP scalp maps at these latencies
(range: -102 to 996 ms, NaN -> empty):

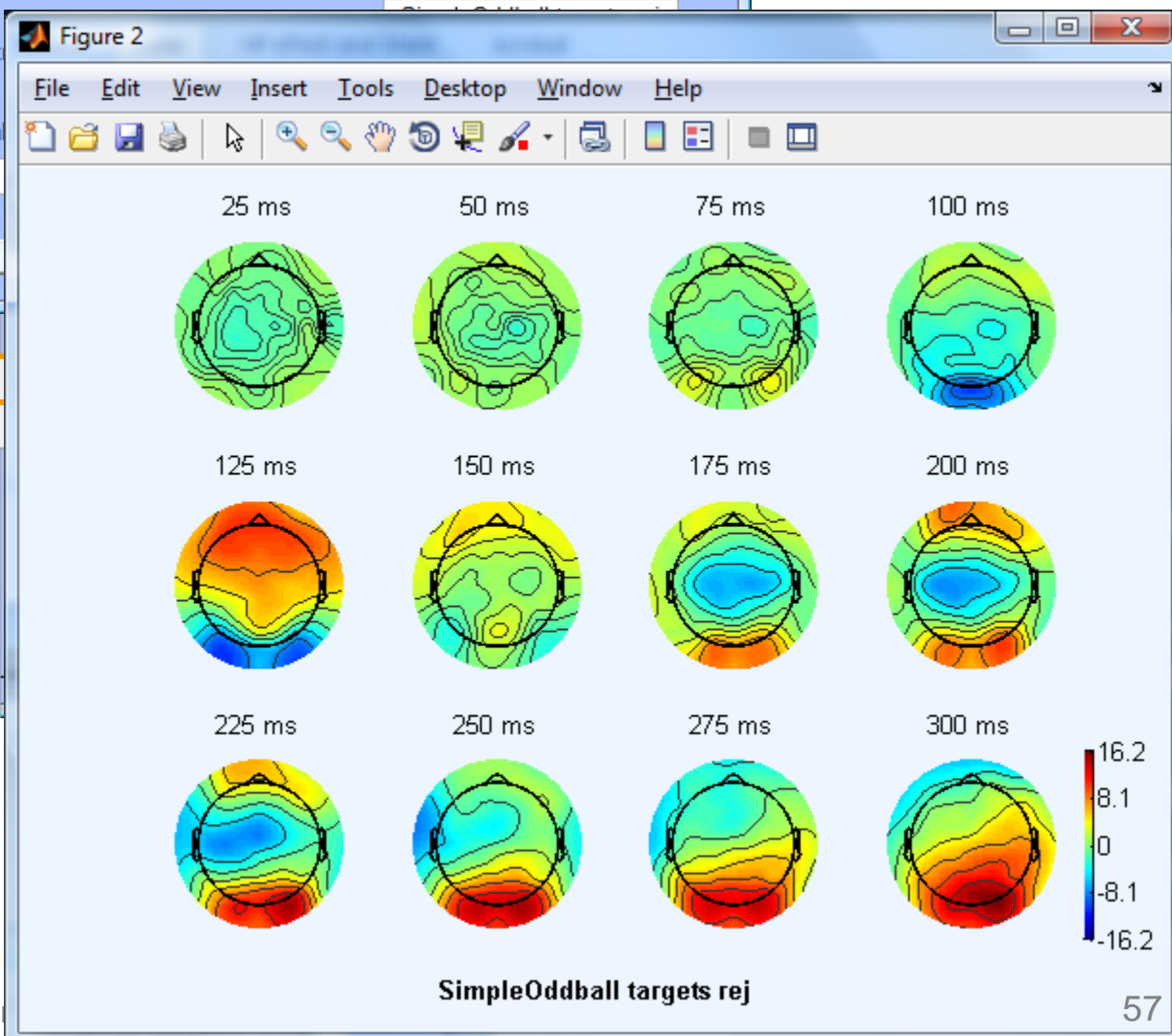
25:25:300

Plot title

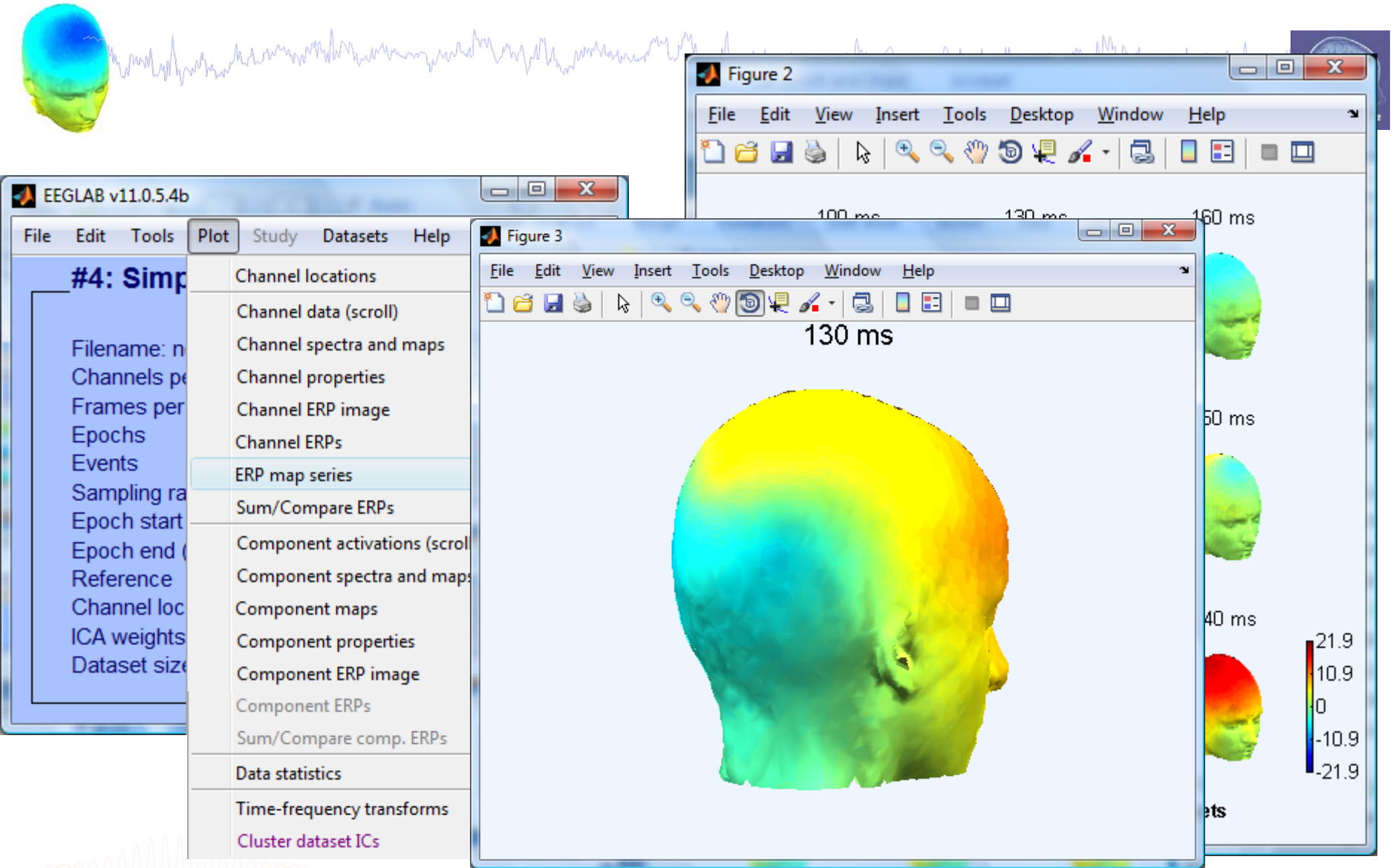
Plot geomet...

-> Additional

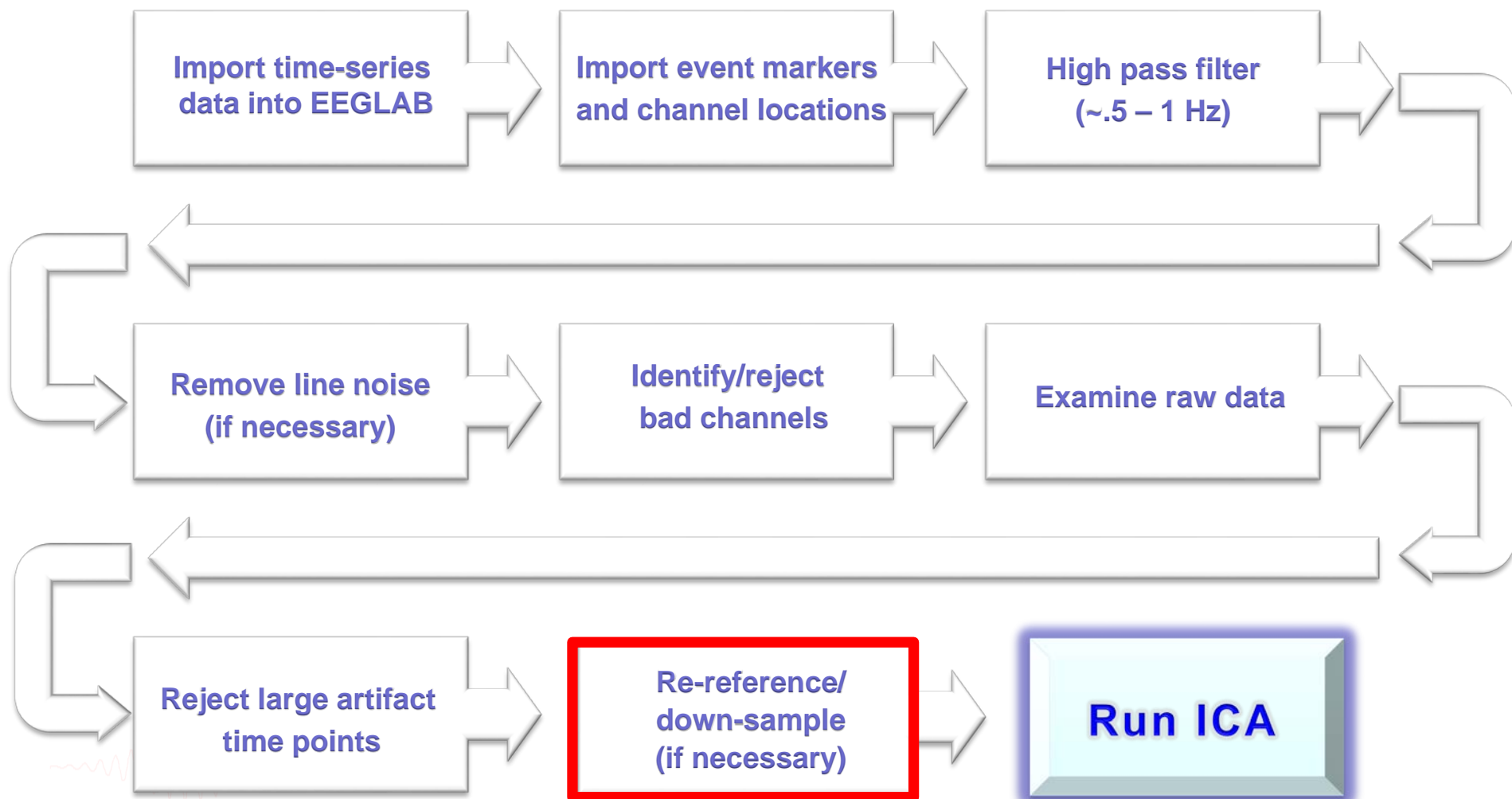
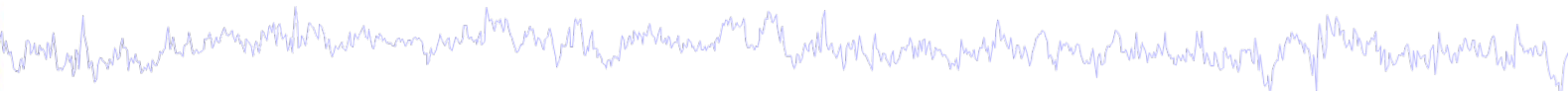
Help



Visualize channel ERPs in 3D



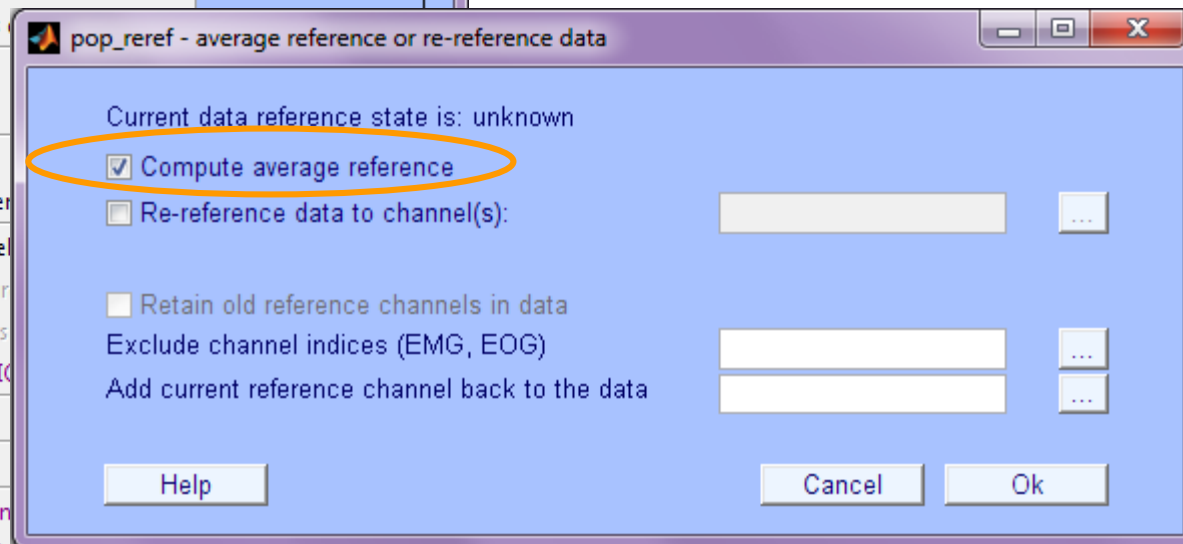
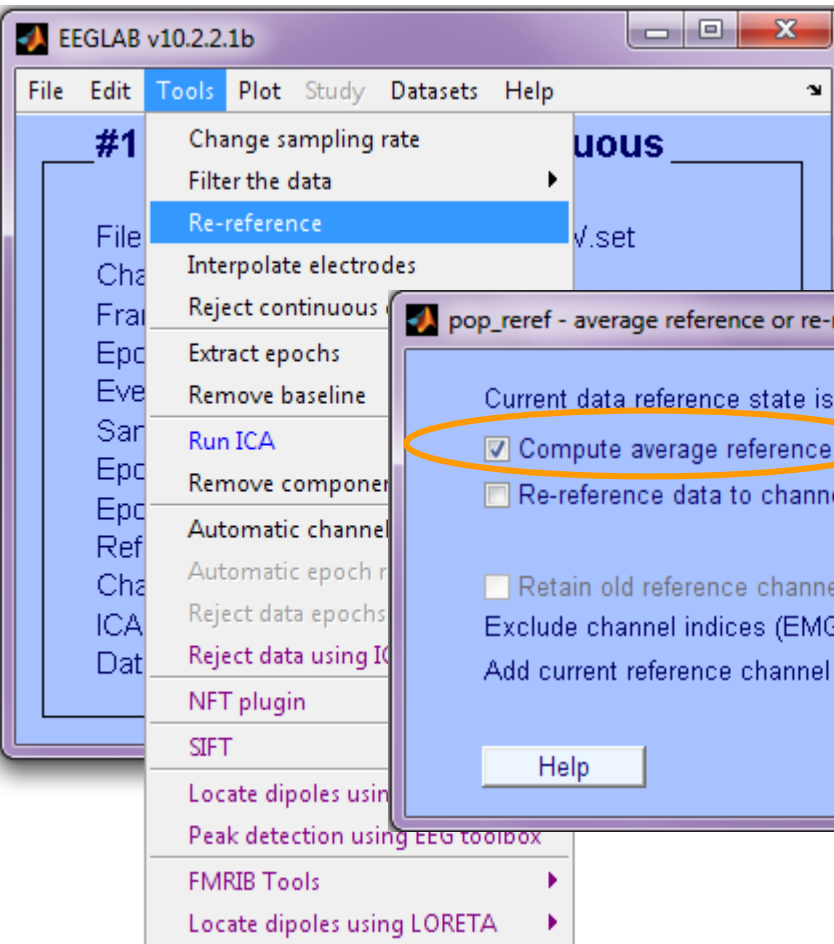
Pre-processing pipeline



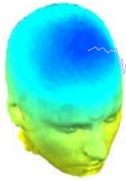
Re-reference data (if necessary/desired)



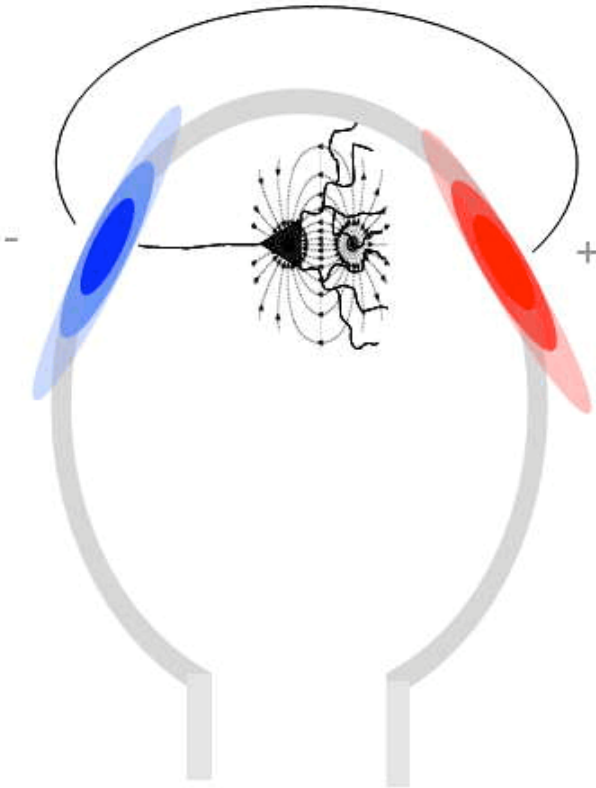
average reference



On Average Referencing



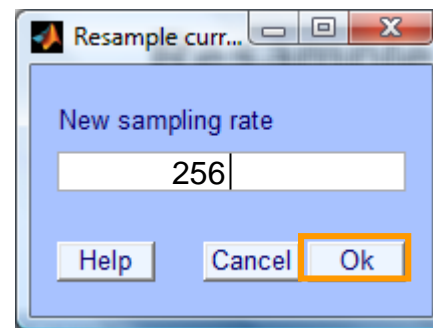
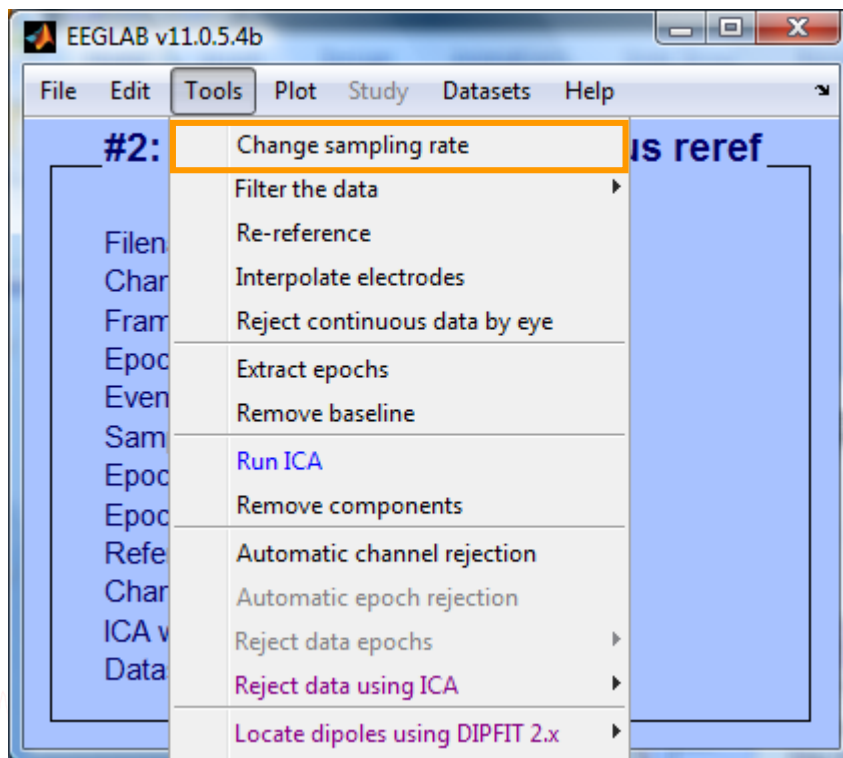
- In theory, positive and negative current across entire head should balance—no net current source or sink: Average referencing enforces this.
- ICA is invariant to re-referencing, except for
 - Effect of rank deficiency
 - DC difference.
- Average referencing reduces data rank by 1, ~~so you must remove one channel (Cz often)~~ See update below.

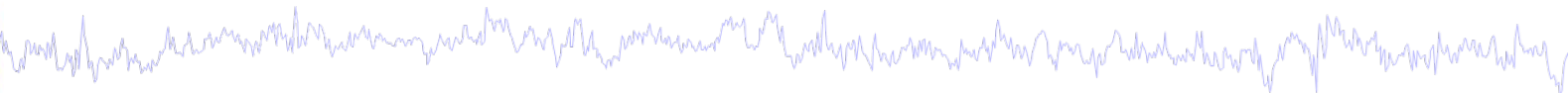


https://sccn.ucsd.edu/wiki/Makoto's_preprocessing_pipeline#Re-reference_the_data_to_average

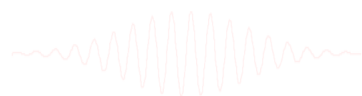
Resample data (if desired)

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





END



Exercises (optional homework)



- Preprocess data of your choice or load a previously filtered dataset e.g. faces_4.set
- 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

