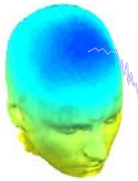


EEG Preprocessing for ICA

25th EEGLAB Workshop

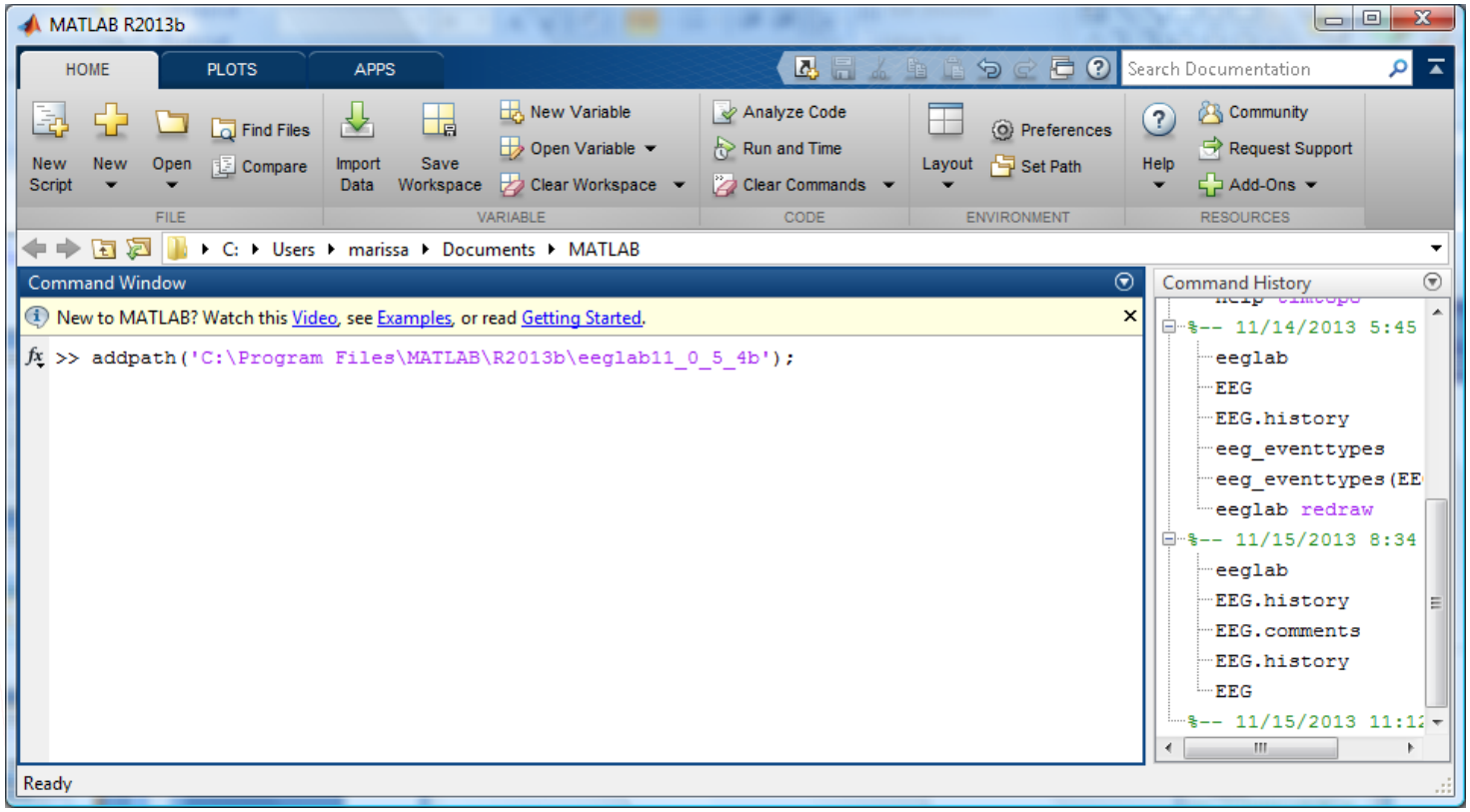
JAIST Tokyo Satellite, Japan

Day 1

John Iversen

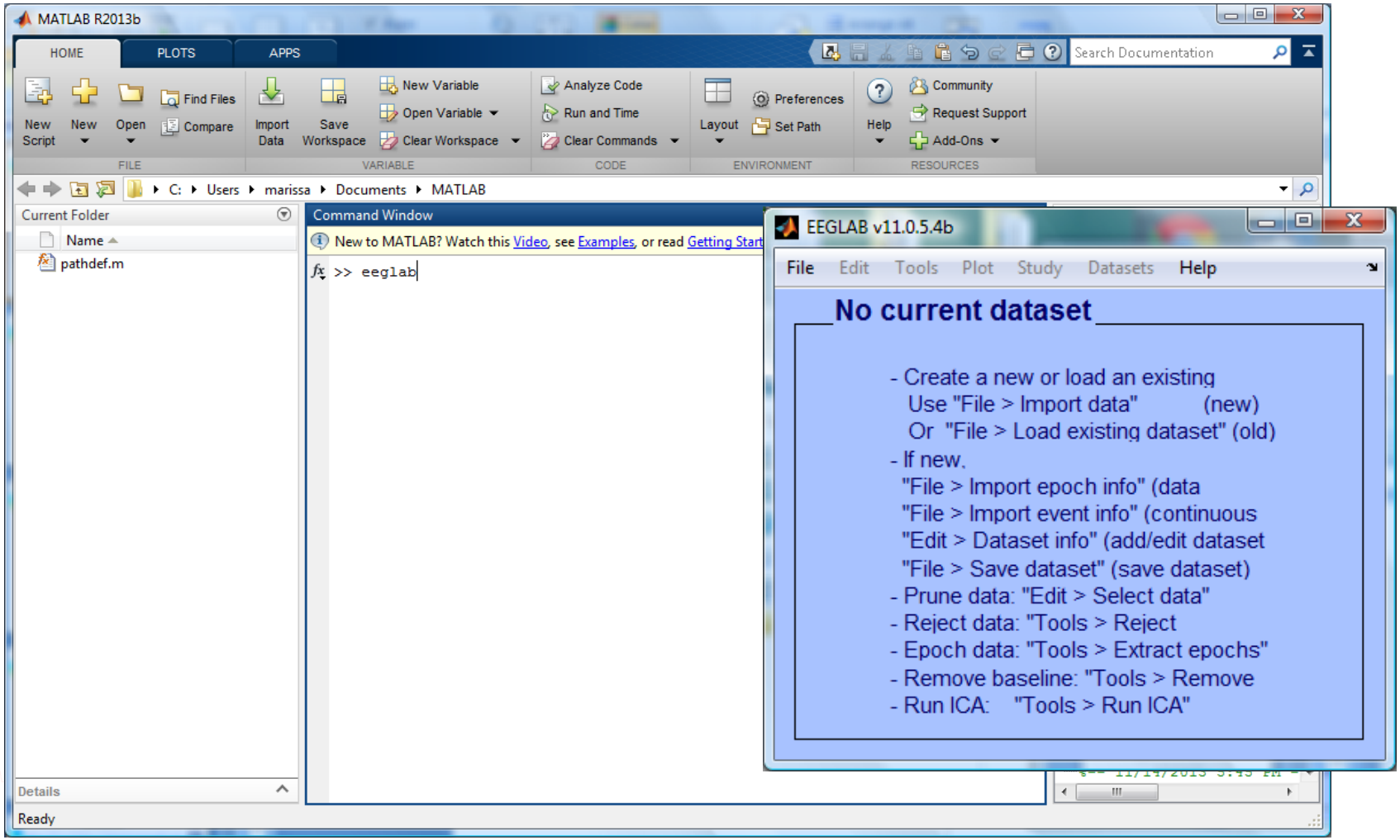
Installing EEGLAB and data folder

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

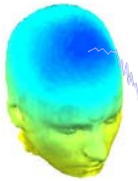


The EEGLAB Matlab software

main graphic interface

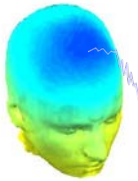


“Secrets” to a good ICA decomposition



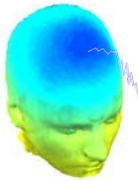
- Garbage in, garbage out (**GIGO**: it’s not magic)
- Remove large, non-stereotyped artifacts
- Do you have enough data? (based mostly on time, not frames)
- High-pass filter to remove slow drifts (no low-pass filter needed)
- Remove bad channels
- Data must be in double precision (not single)

The Goal of Preprocessing



- Create a complete EEGLAB data set with
  - EEG time series signal
  - Channel Locations
  - Event information
- Applying singnal 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

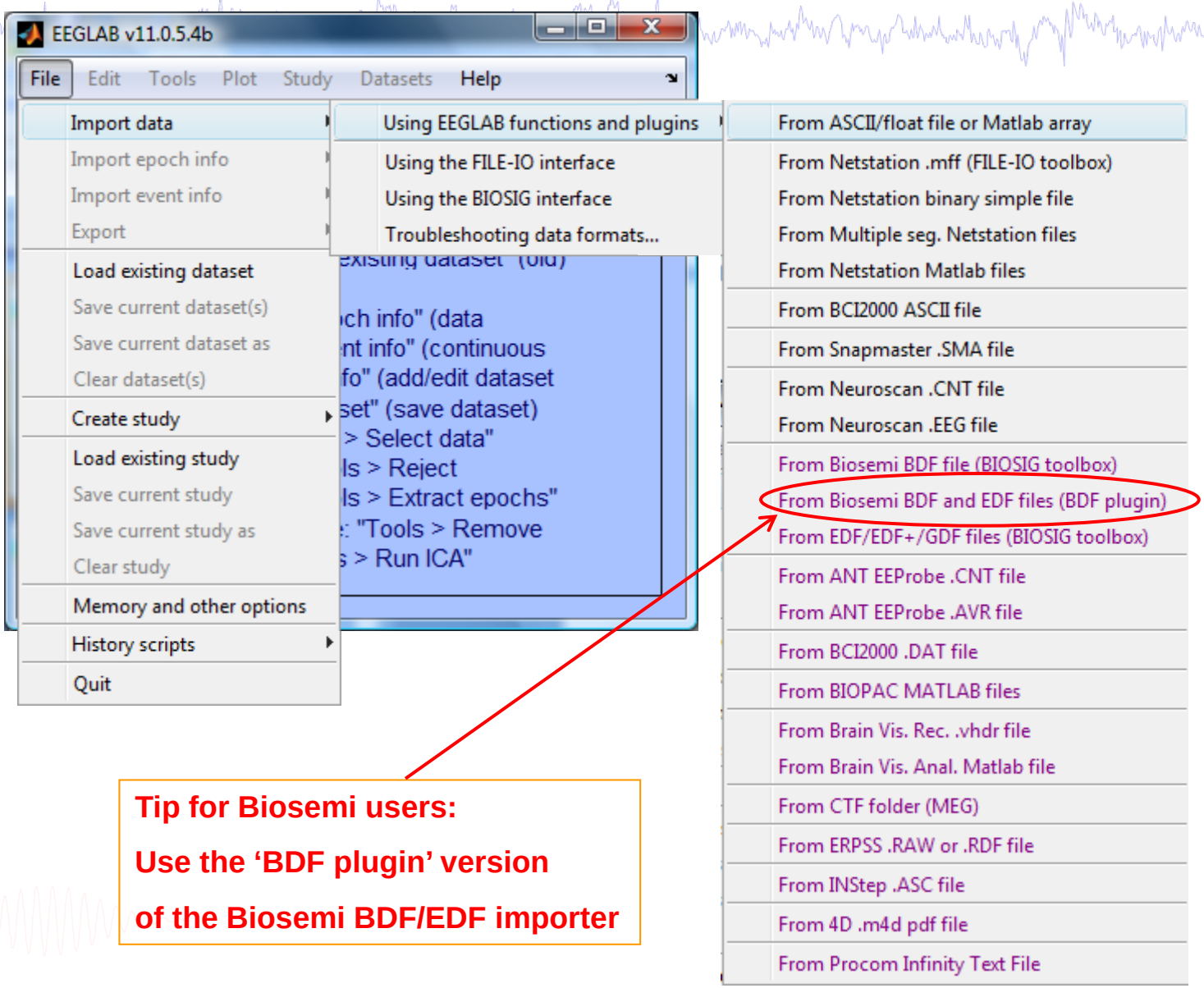
Examine raw data

Reject large artifact time points

Re-reference/ down-sample (if necessary)

Run ICA

Importing a dataset

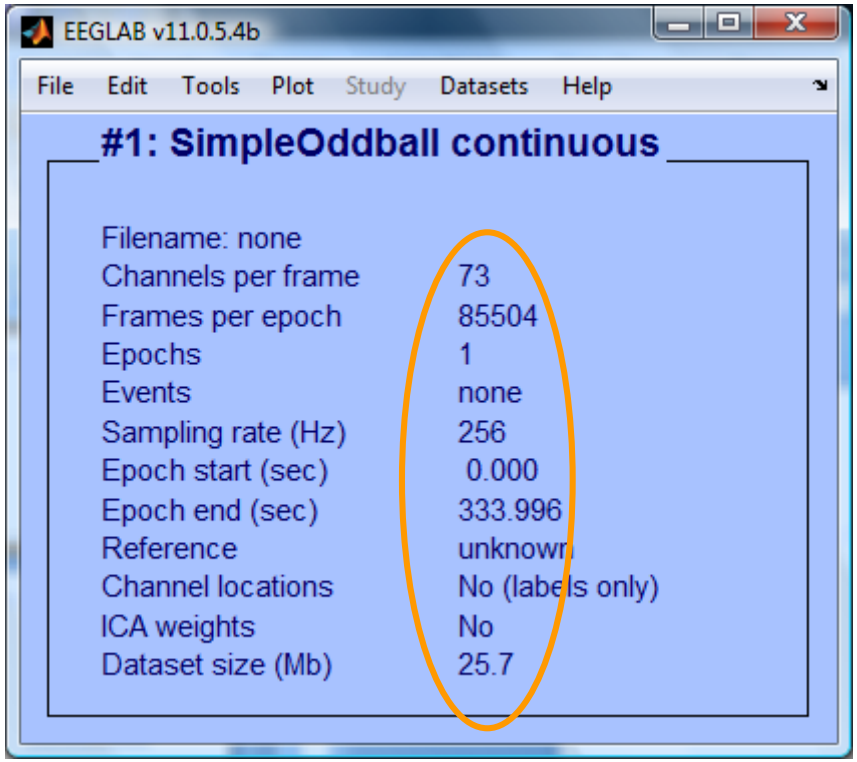


Tip for Biosemi users:

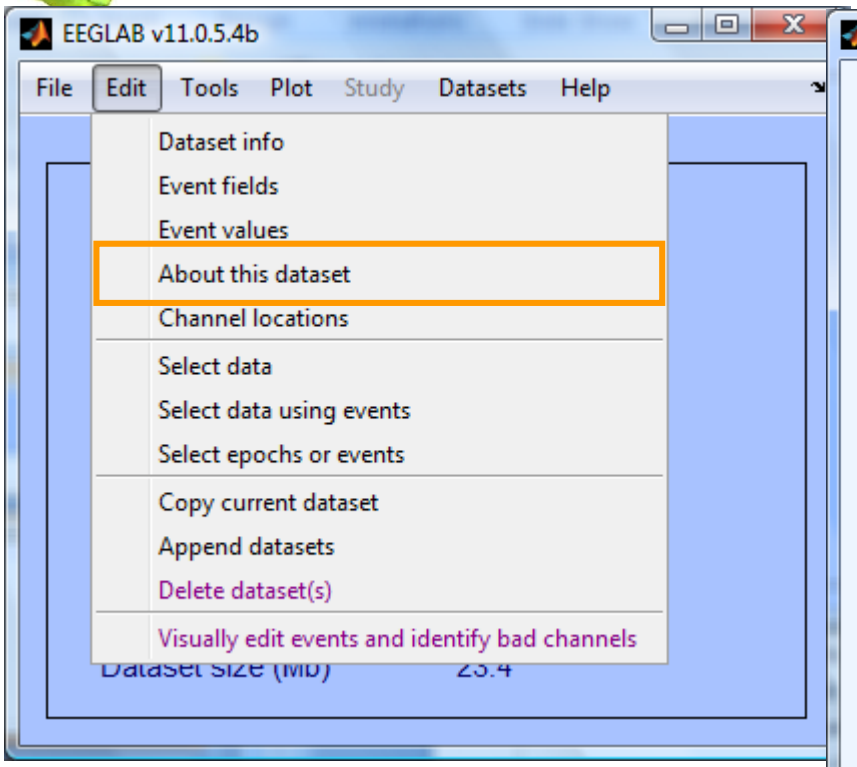
Use the 'BDF plugin' version

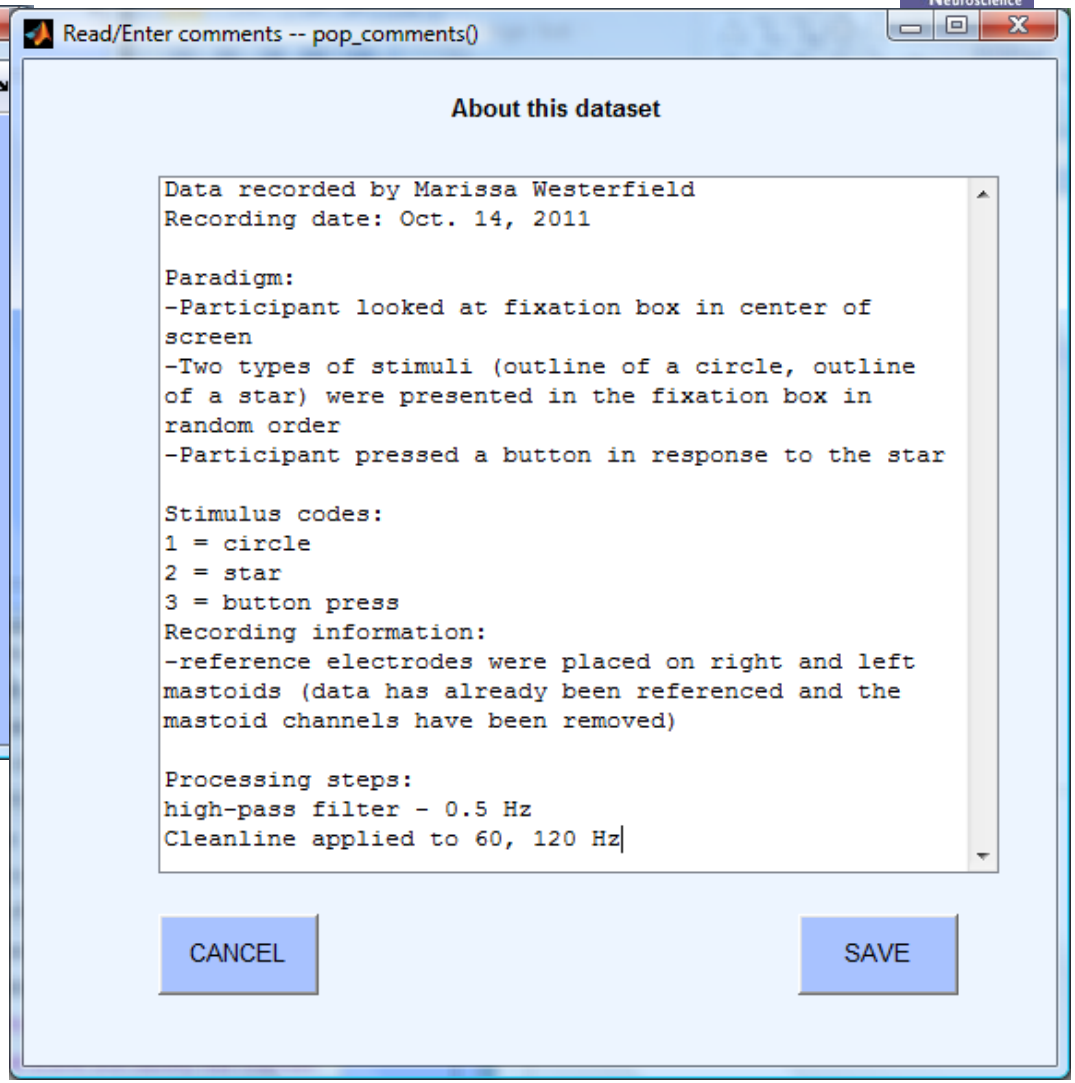
of the Biosemi BDF/EDF importer

Imported EEG data



Comments and dataset history

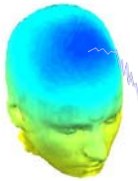




Also:
>> EEG.comments

and
>> EEG.history

Pre-processing pipeline



Import 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

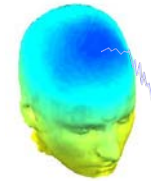
Examine raw data

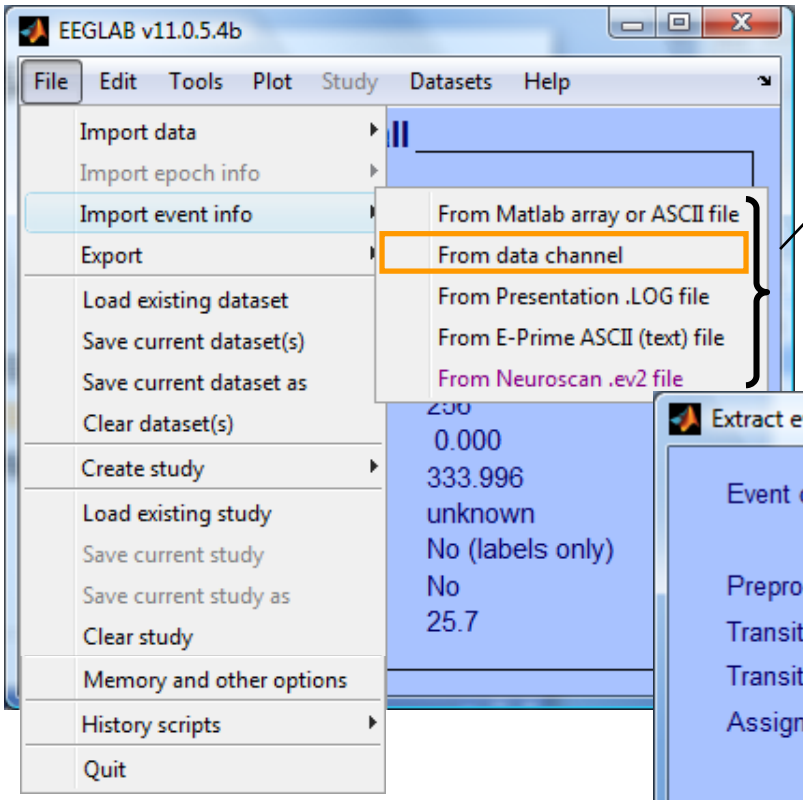
Reject large artifact time points

Re-reference/ down-sample (if necessary)

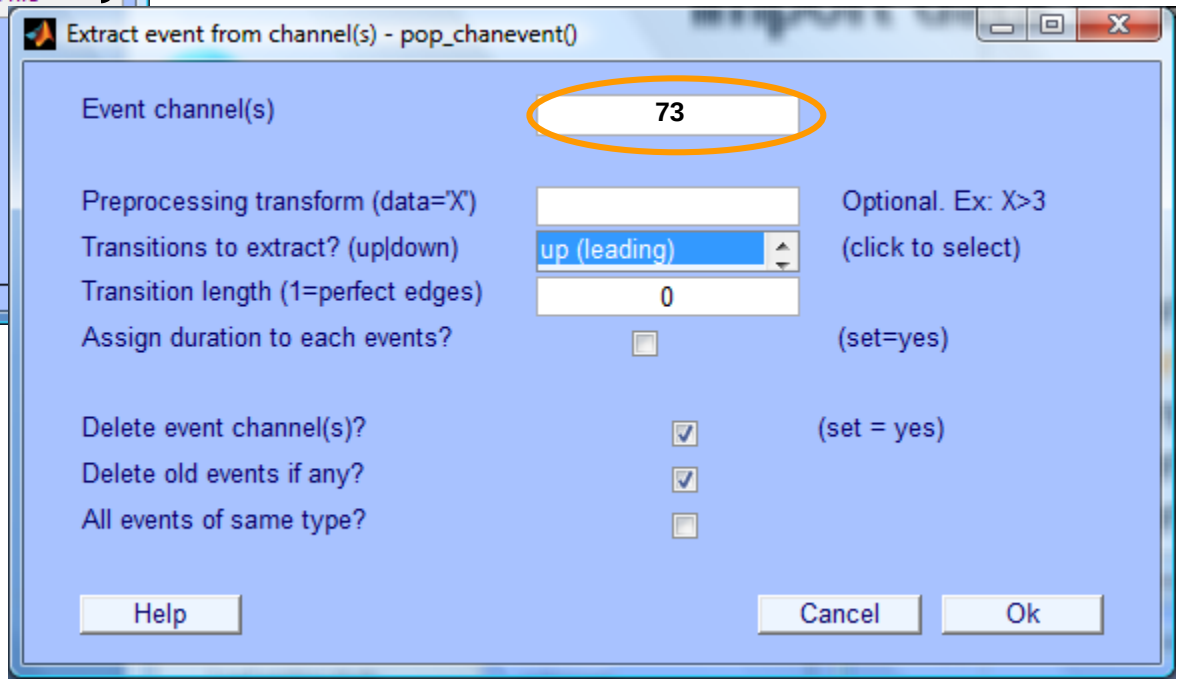
Run ICA

Import data events



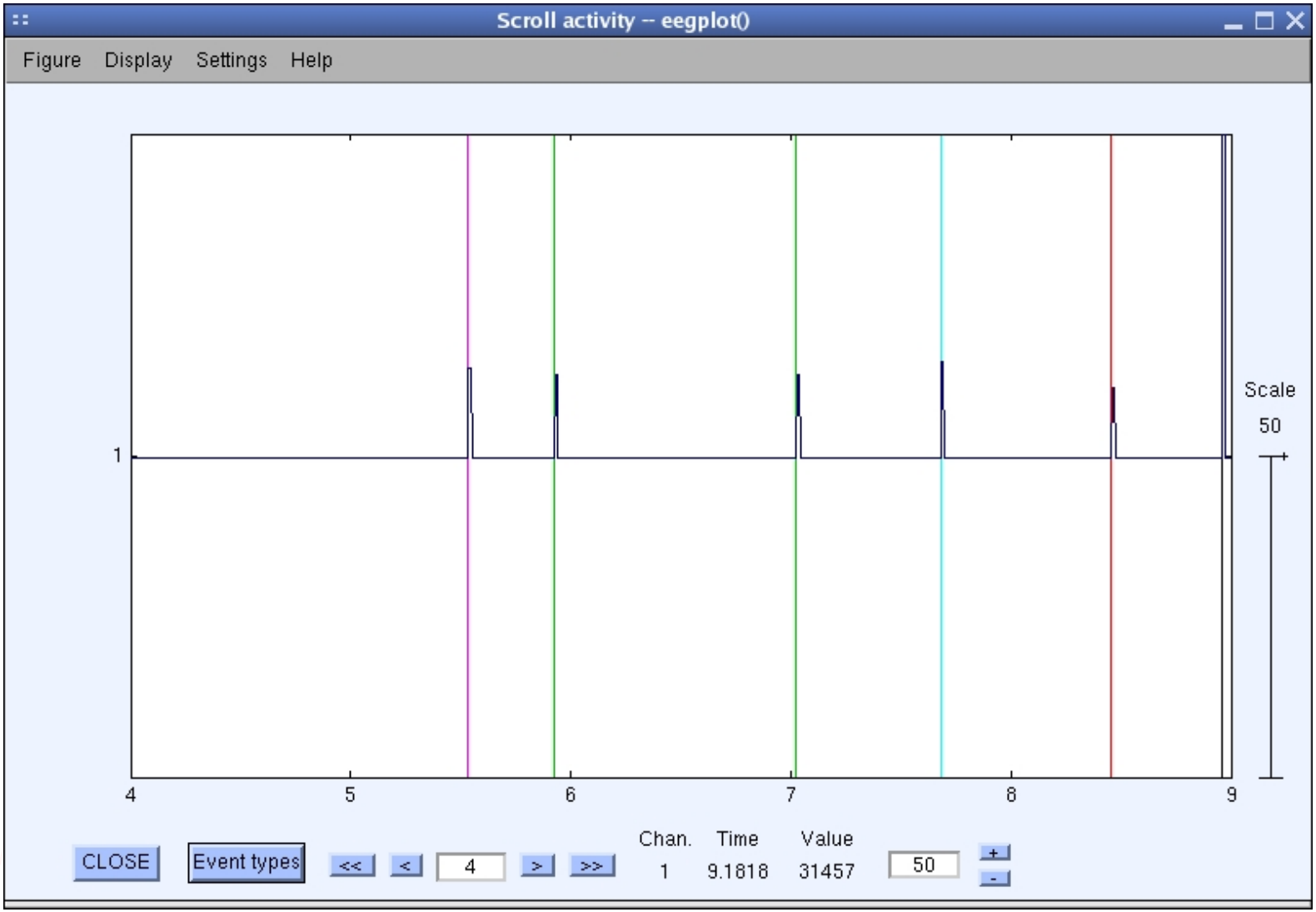


- Import events from Matlab array or ASCII file
- Import events from data channel
- Import from Presentation event file
- Import events from E-Prime event file
- Import events from Neuroscan event file

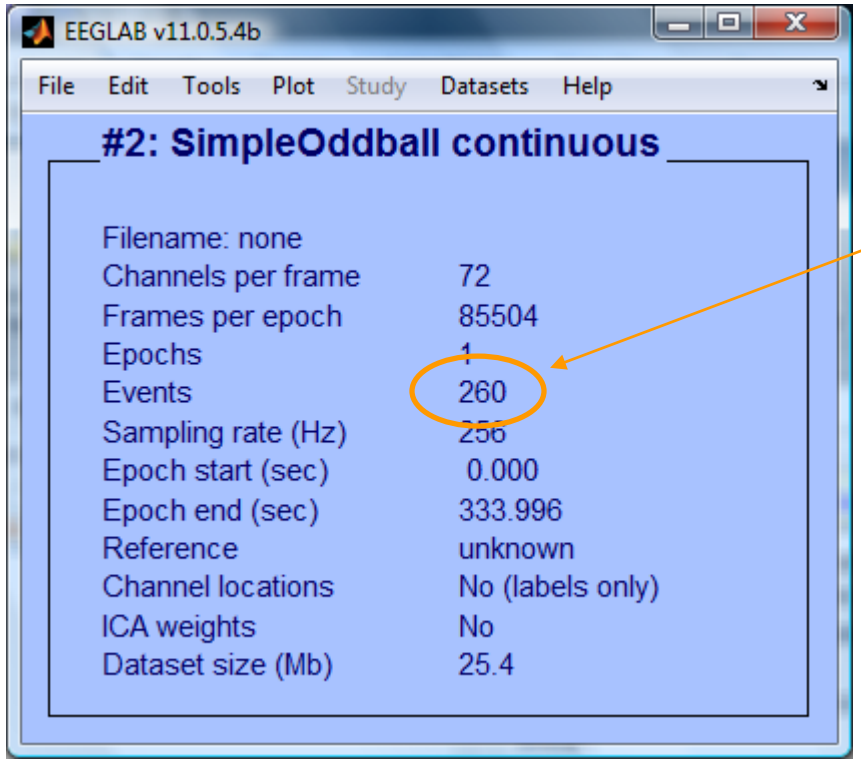


(Often imported automatically during data import)

Appearance of an event channel in raw data

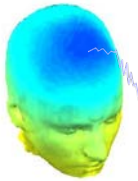


Imported data events



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

Sample data: basic P300 paradigm



File

SimpleOddball.set

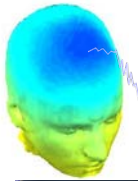
Data

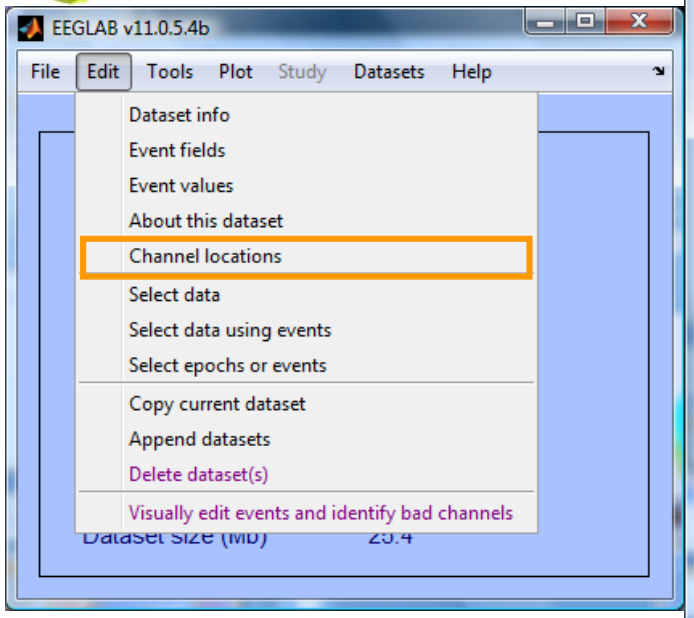
68 channel EEG, 256 Hz sampling rate, Biosemi system, re-referenced during import to averaged left and right mastoid electrodes

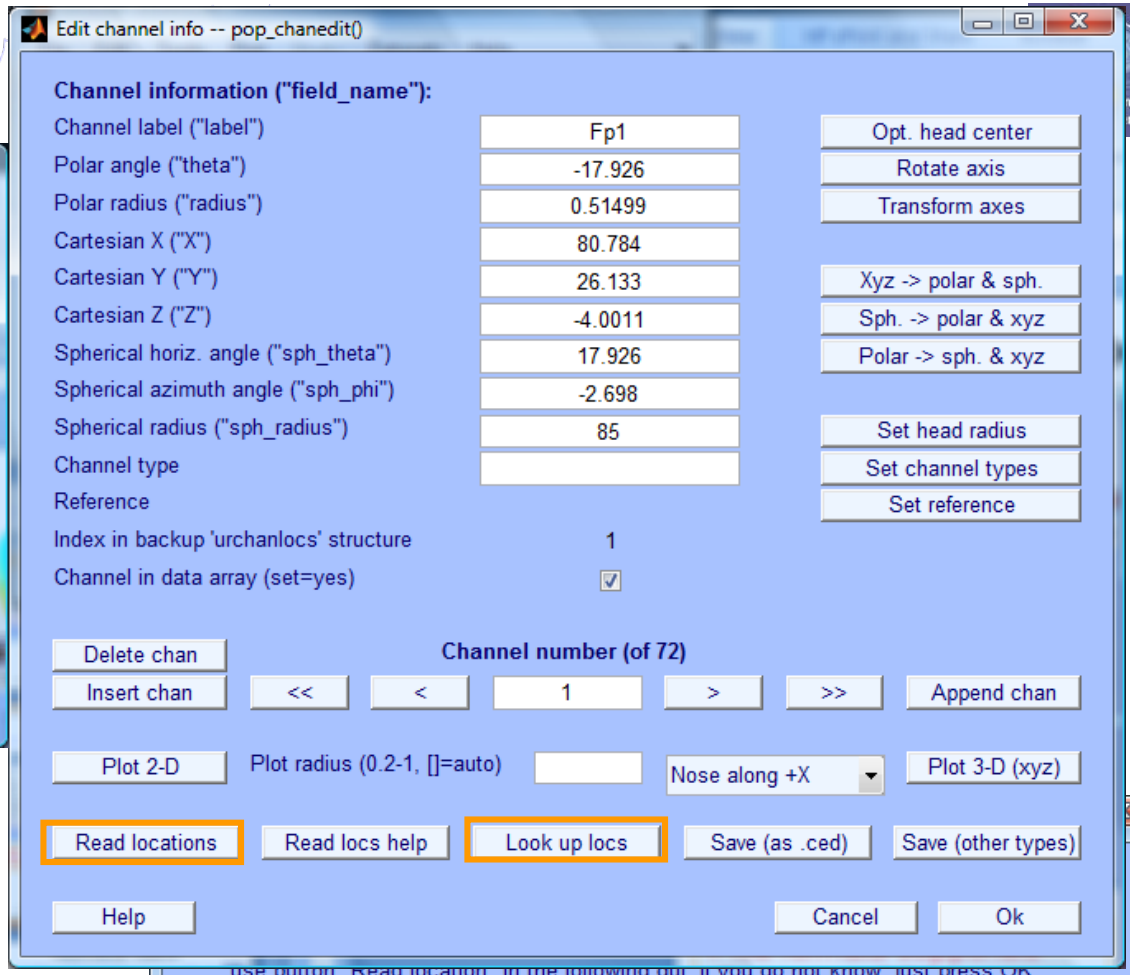
Task

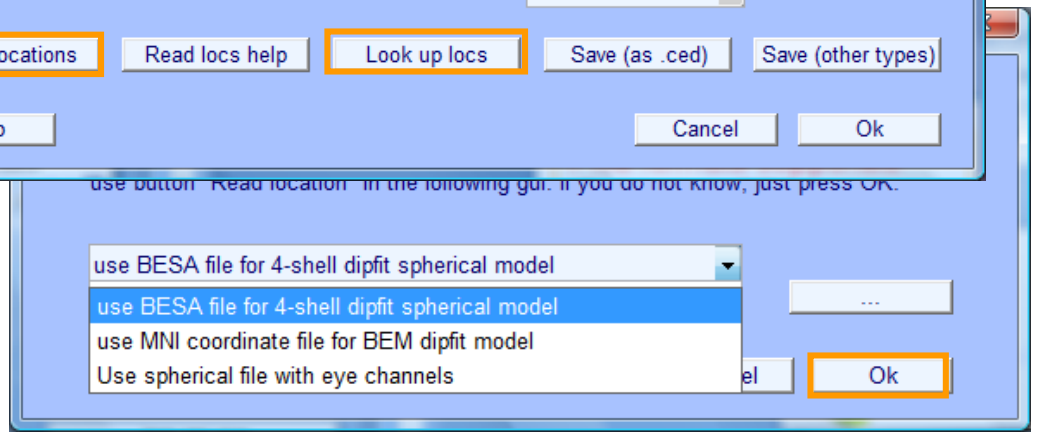
speeded button press response to star shape (no response to circle shape), 100 ms presentation duration, 200 trials

Import channel locations



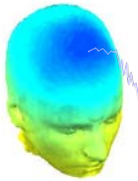


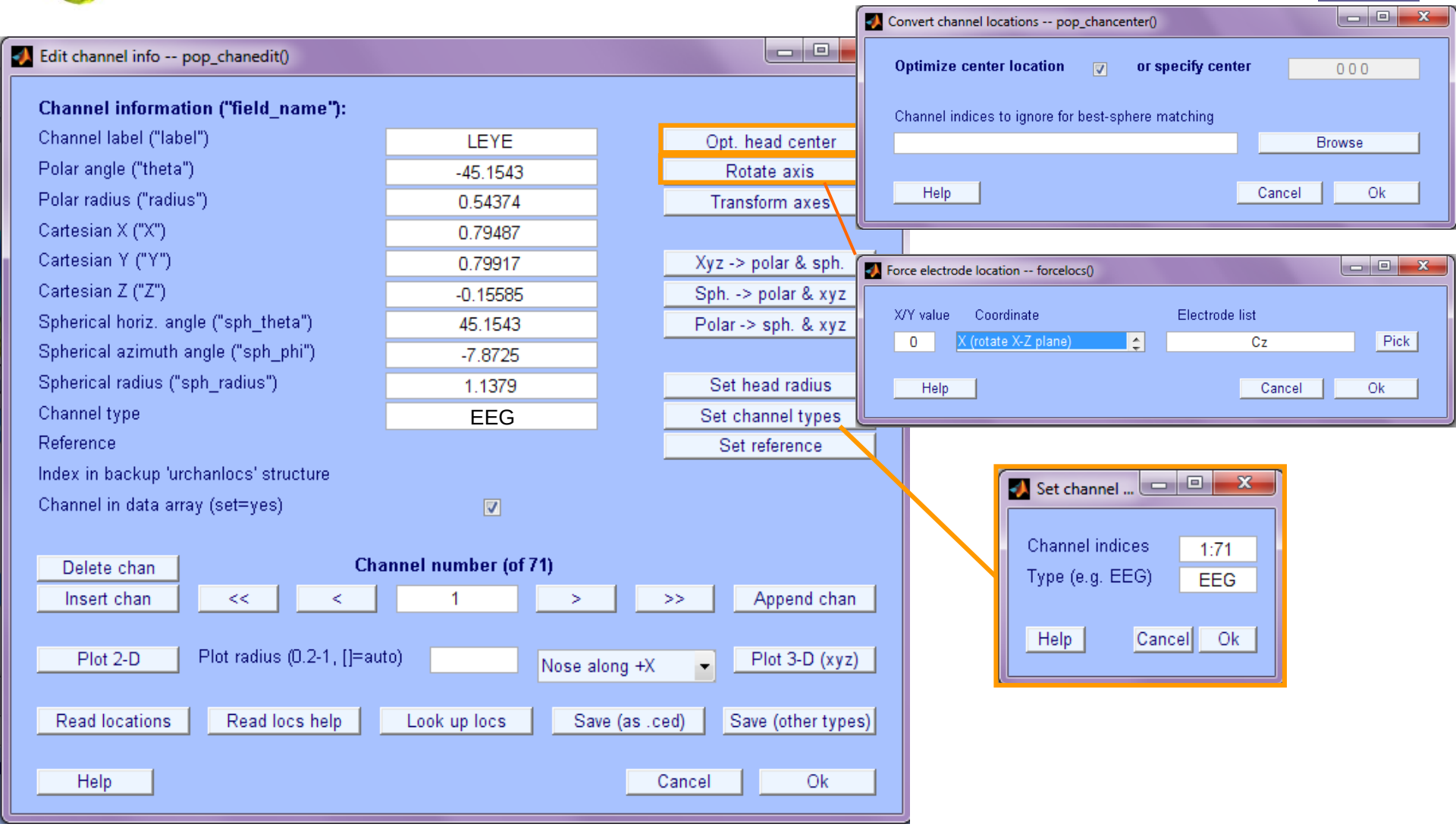


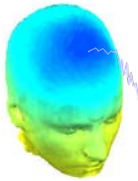


7 file formats supported (Polhemus, BESA, …)

Import channel locations







Edit channel info -- pop_chanedit()

Channel information ("field_name"):

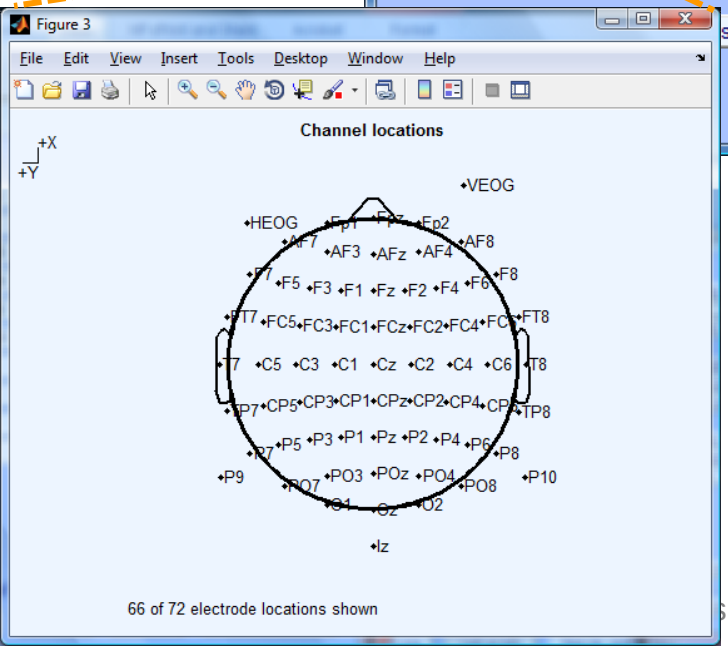
Channel label ("label")	HEOG	Opt. head center
Polar angle ("theta")	-42	Rotate axis
Polar radius ("radius")	0.65556	Transform axes
Cartesian X ("X")	55.7734	
Cartesian Y ("Y")	50.2186	Xyz -> polar & sph.
Cartesian Z ("Z")	-39.9051	Sph. -> polar & xyz
Spherical horiz. angle ("sph_theta")	42	Polar -> sph. & 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)	☑	

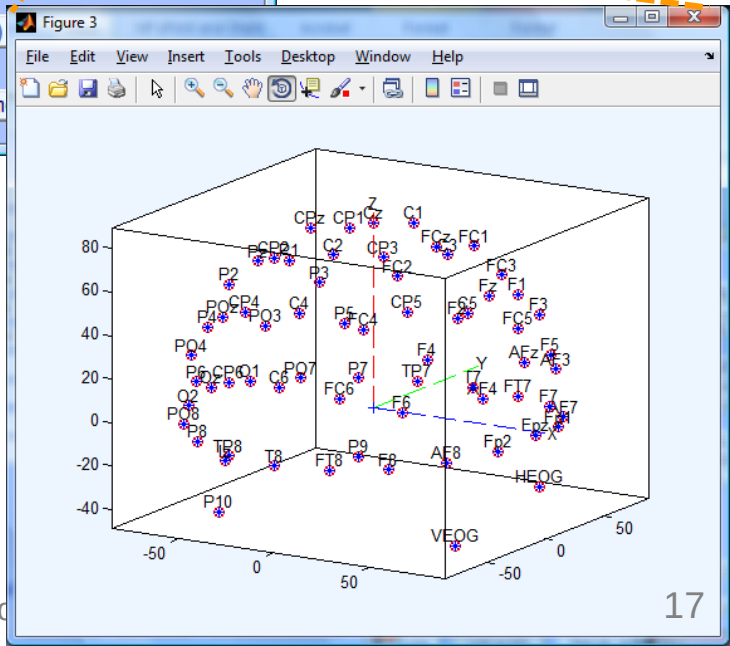
Delete chan | Insert chan | Channel number (of 72): << < 68 > >> | Append chan

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

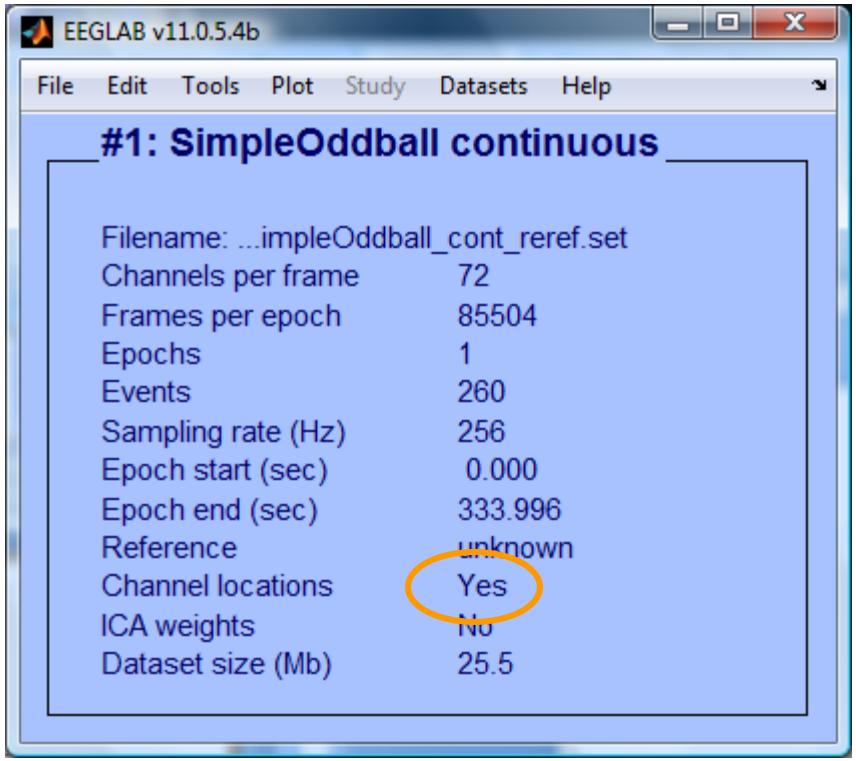
s help | Look up locs | Save (as .ced)

Can

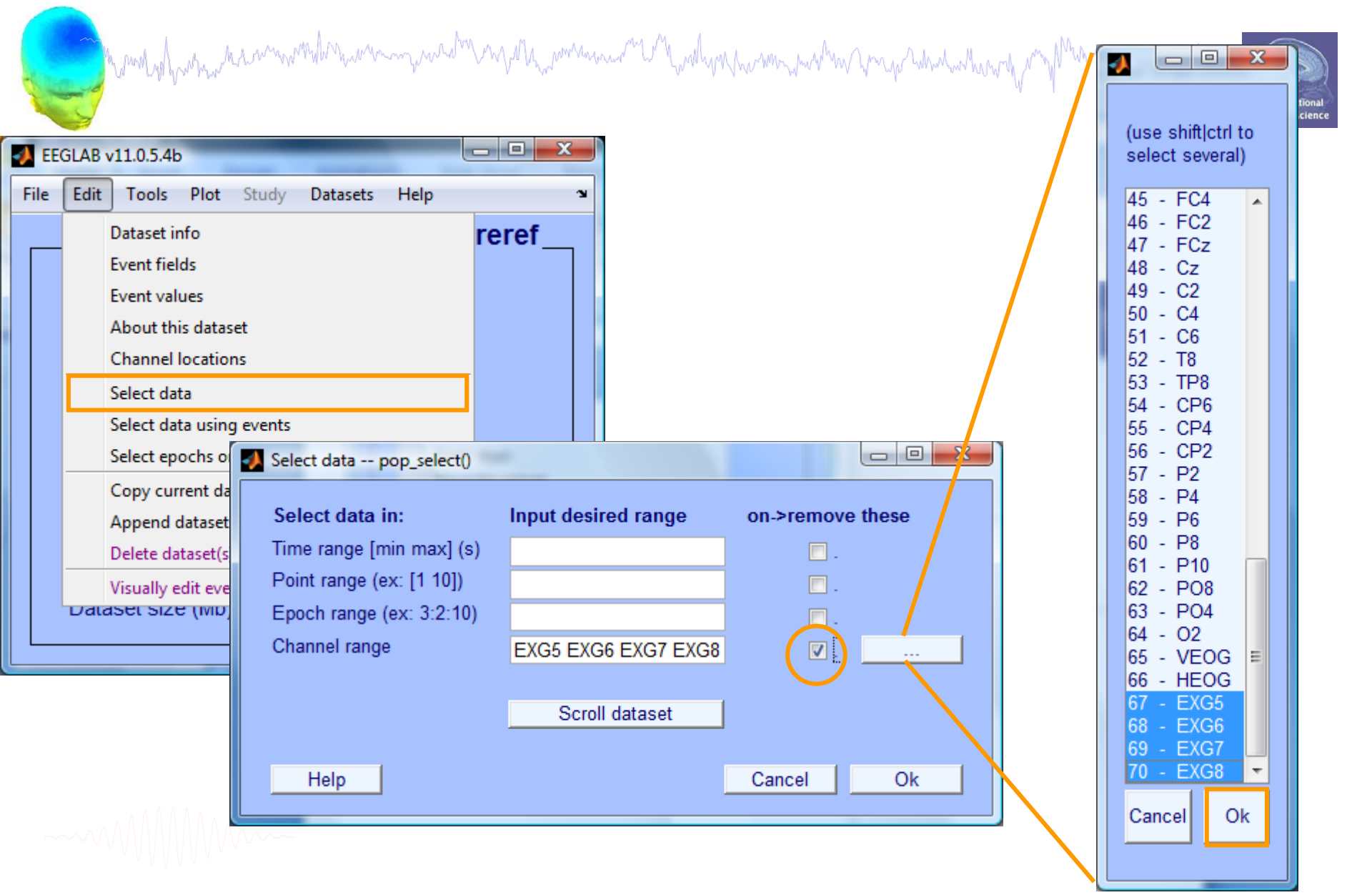




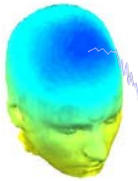
Imported channel locations



Remove unwanted channels



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

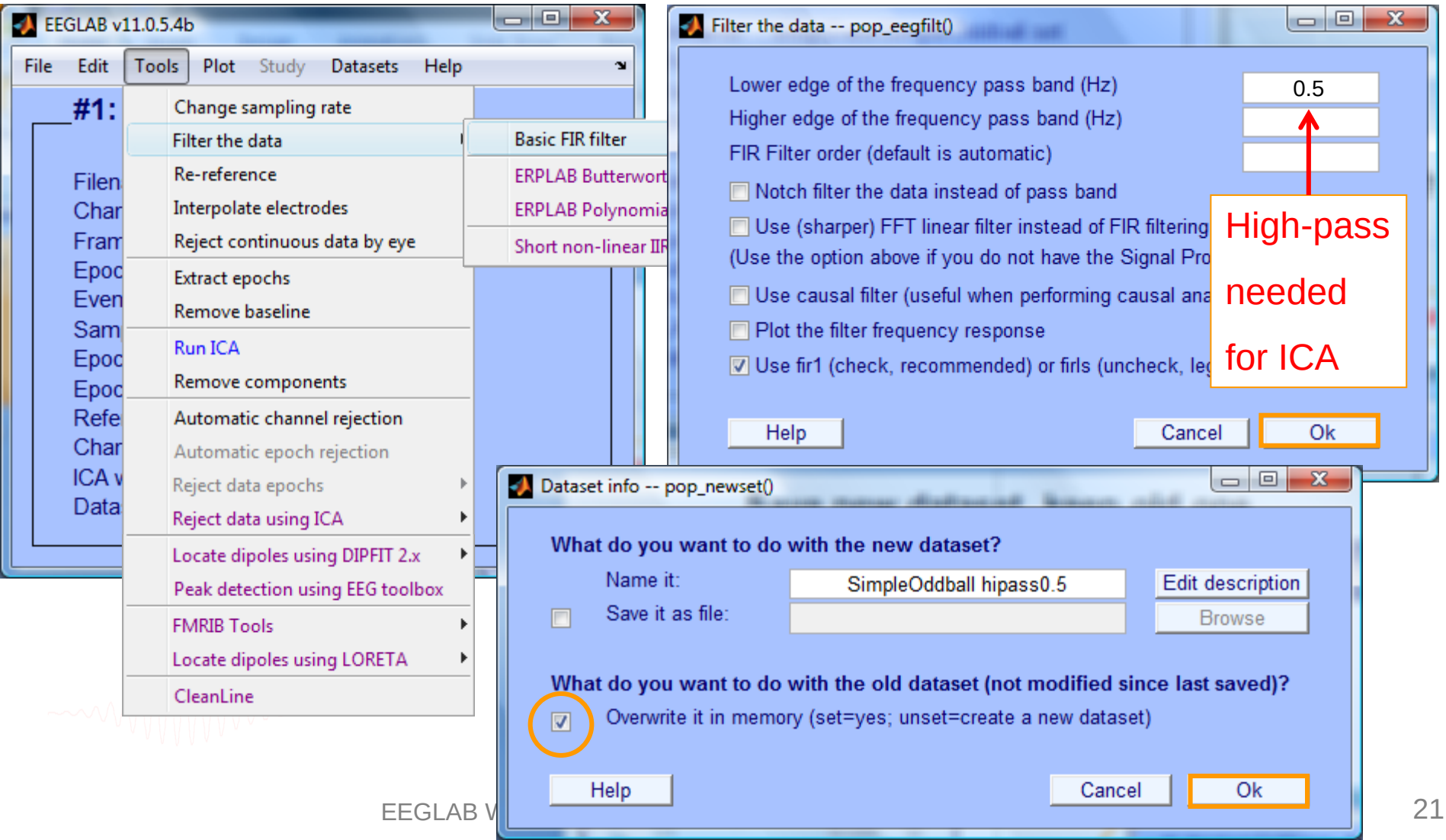
Reject large artifact time points

Re-reference/ down-sample (if necessary)

Run ICA

High-Pass Filter the data

Reason: To improve data stationarity.
ICA is biased to amplitude, and EEG data has 1/f power specrum density.



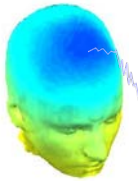
Good resource for learning filters

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)

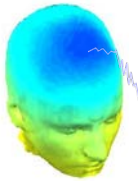
Identify/reject bad channels

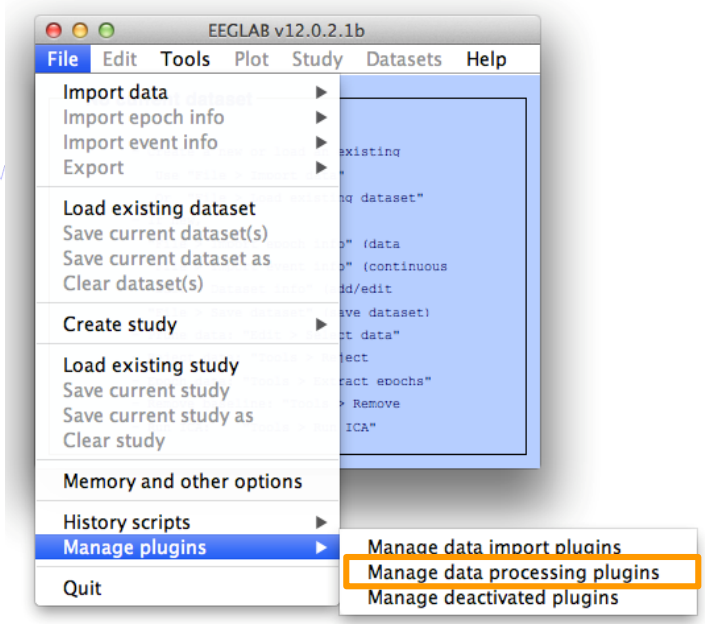
Examine raw data

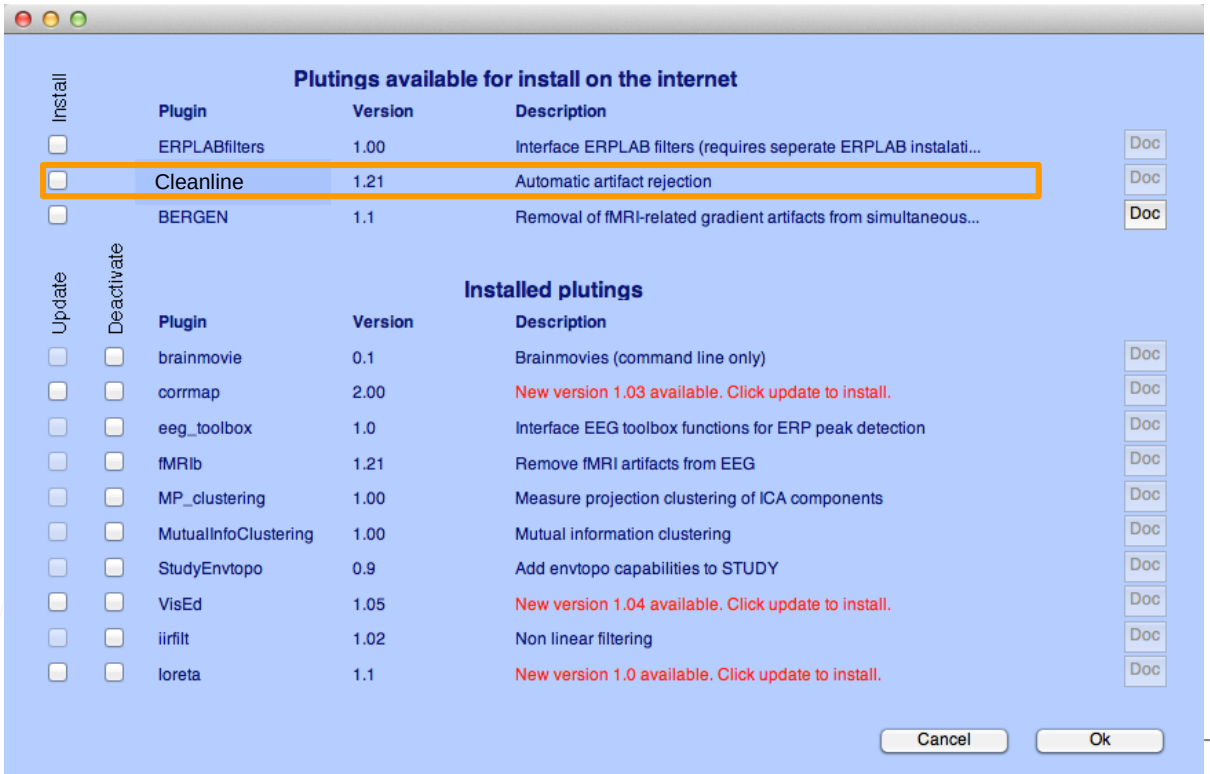
Reject large artifact time points

Re-reference/ down-sample (if necessary)

Run ICA



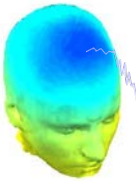


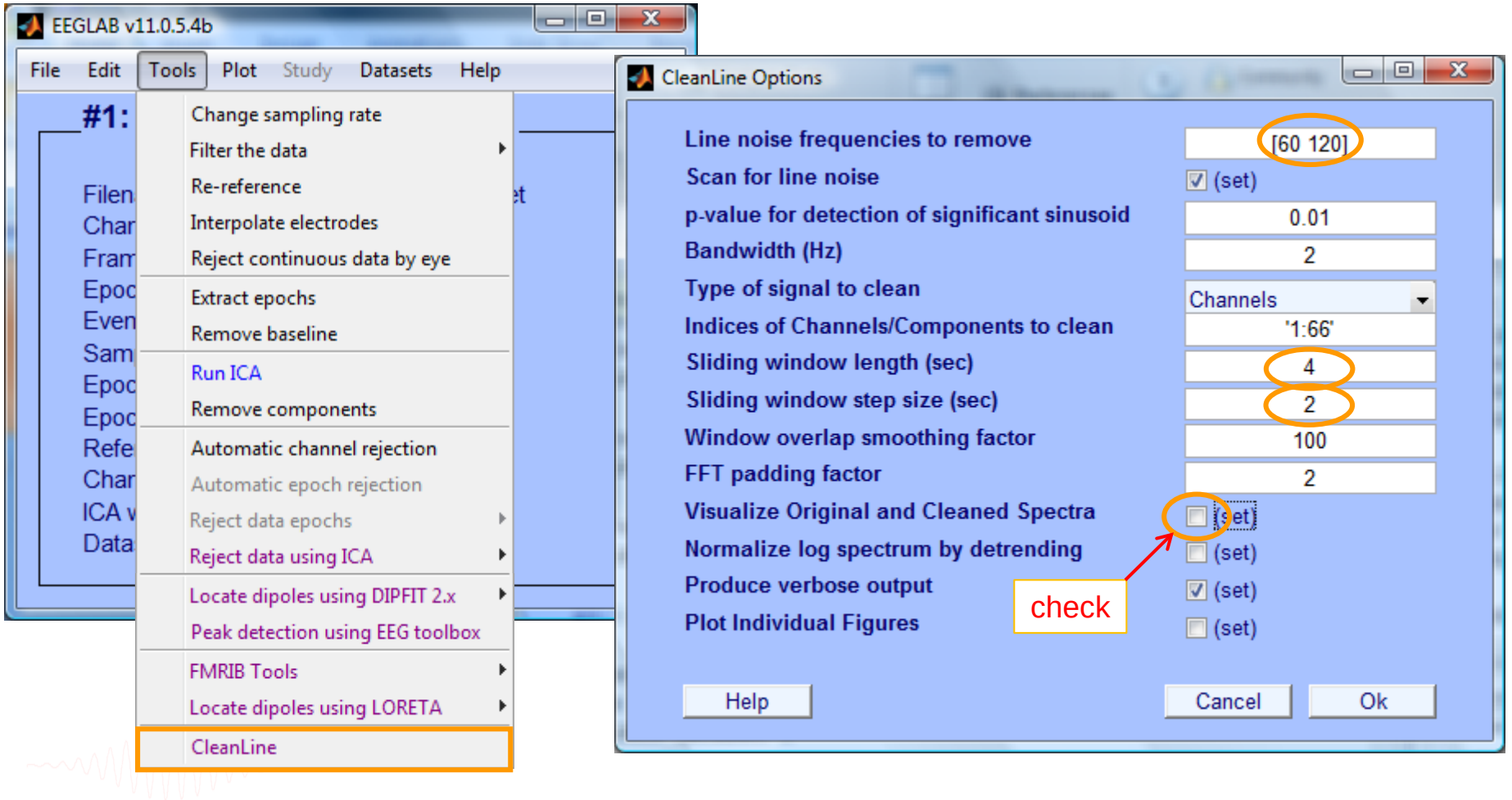


Tim Mullen

Qusp CEO

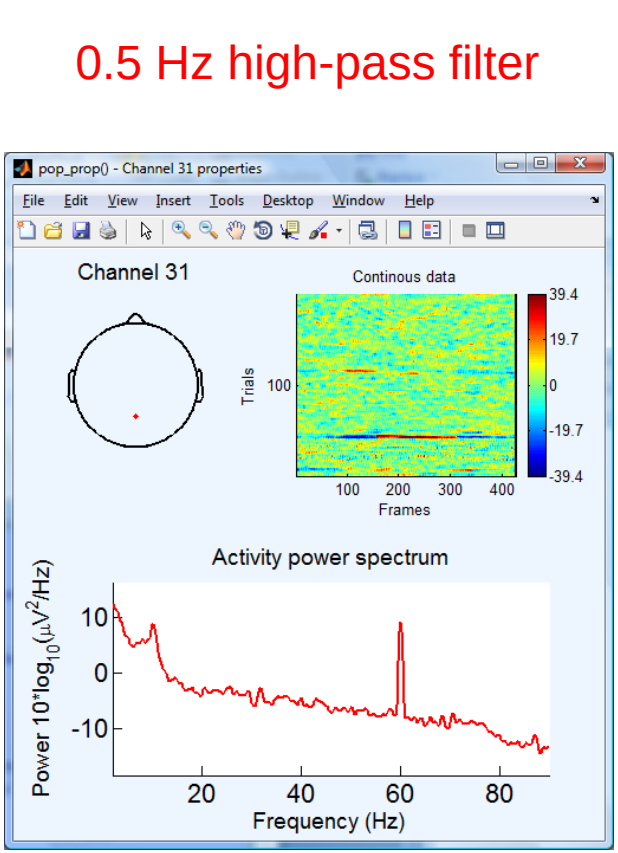
Remove line noise (Cleanline)



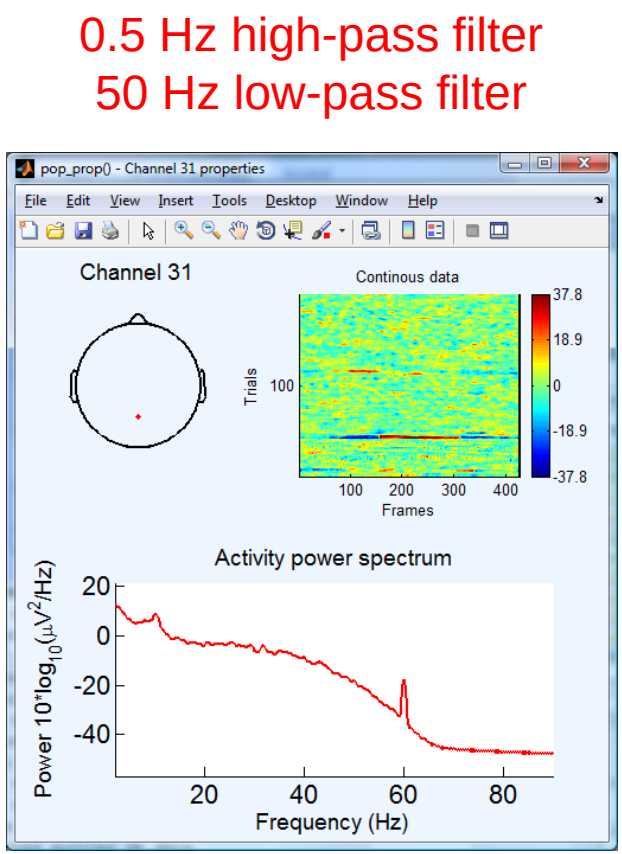


Filter comparisons

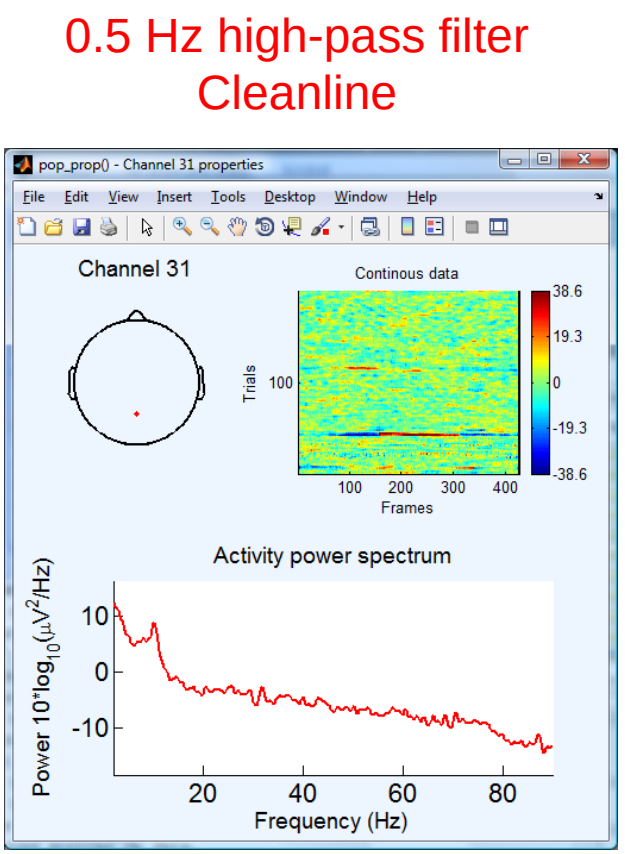
0.5 Hz high-pass filter

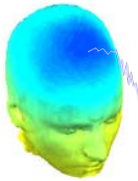


0.5 Hz high-pass filter
50 Hz low-pass filter



0.5 Hz high-pass filter
Cleanline

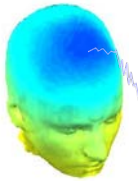




Data Cleaning for ICA

1. Continuous Data

Pre-processing pipeline



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

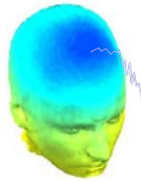
Examine raw data

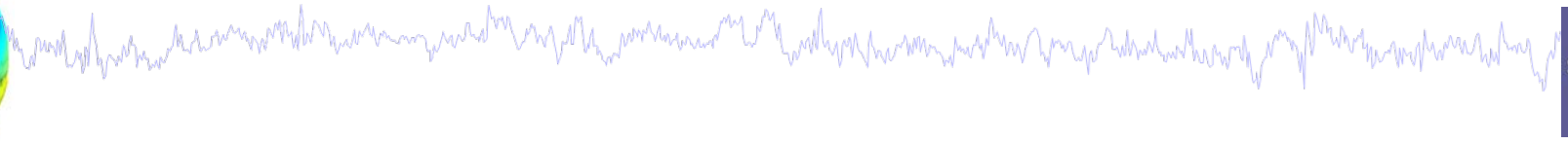
Reject large artifact time points

Re-reference/ down-sample (if necessary)

Run ICA

Manually identifying bad channels

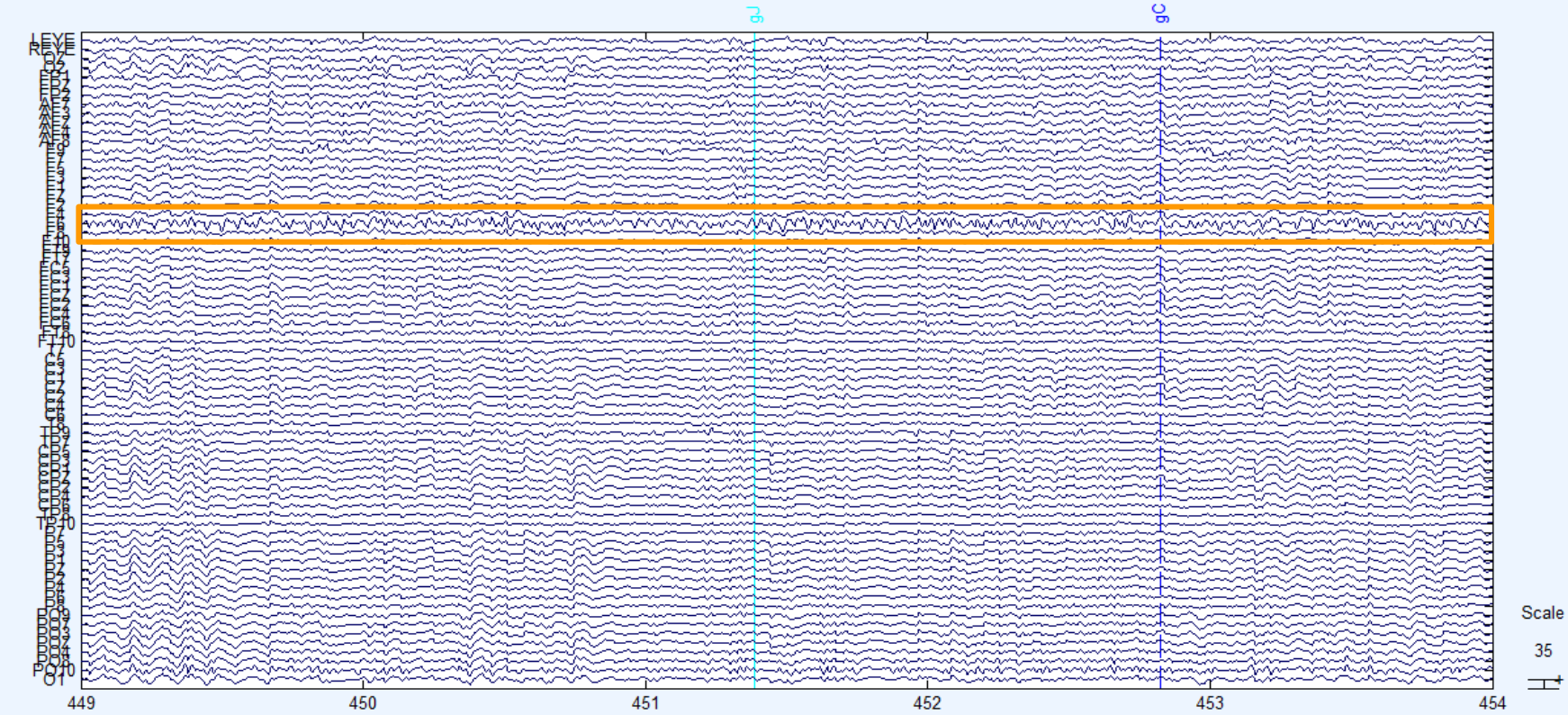




Scroll channel activities -- eegplot()

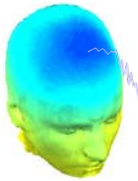
Figure Display Settings Help

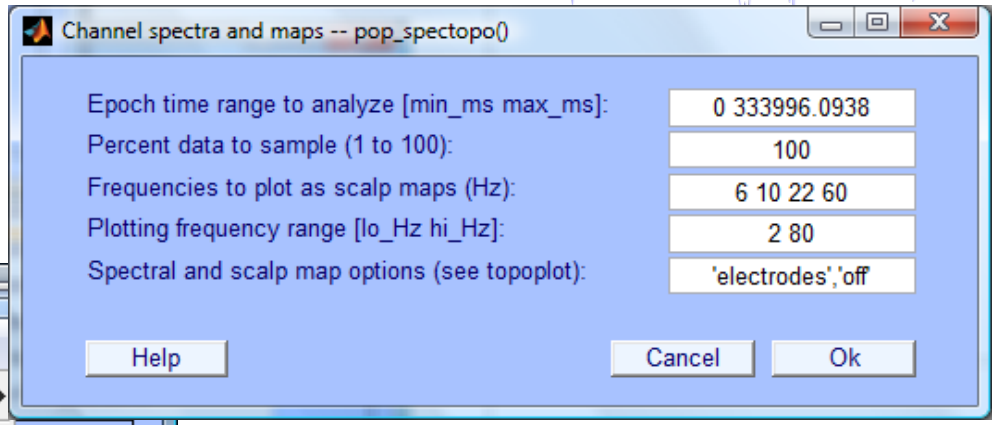
1) Identify bad channel

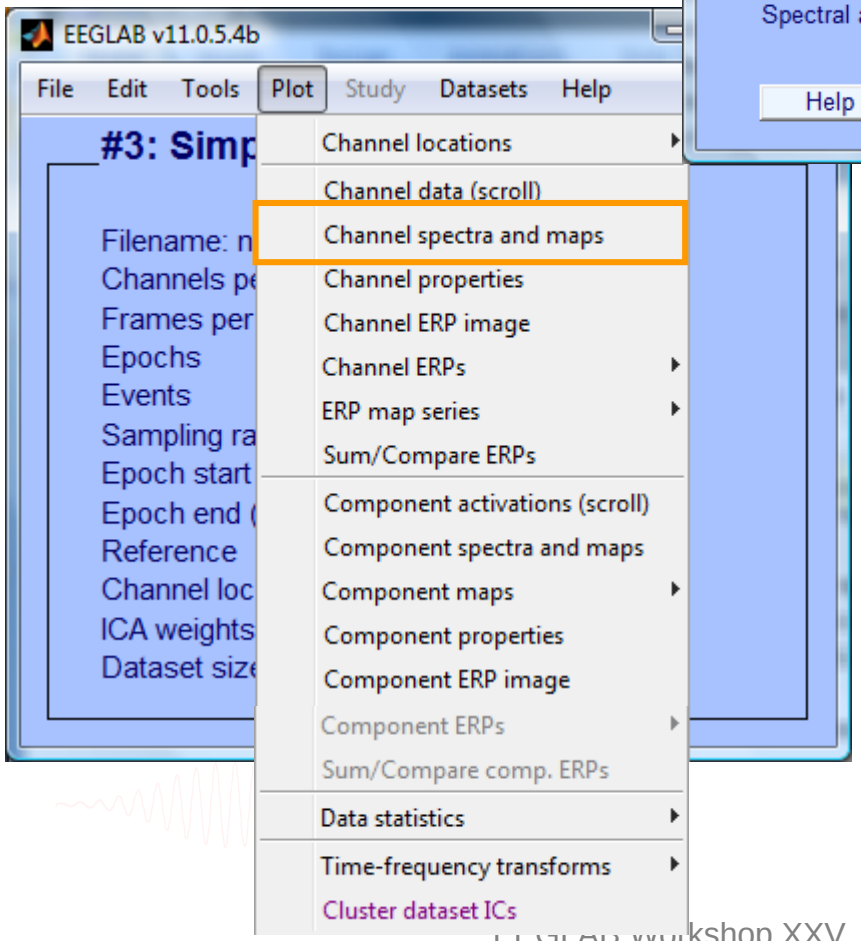


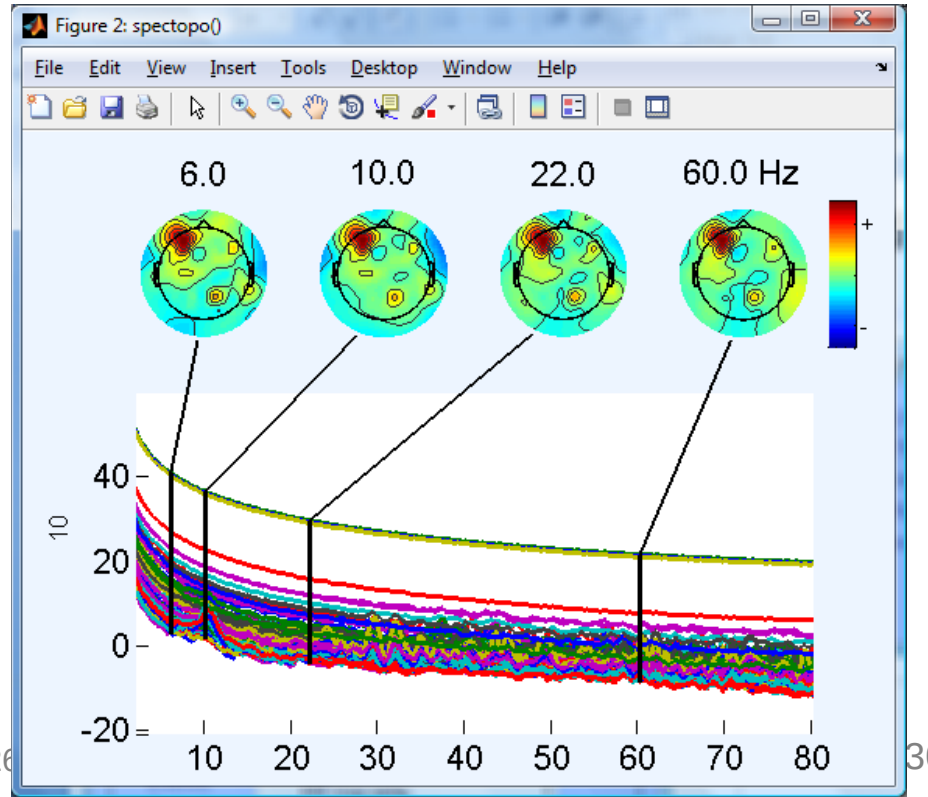
CANCEL | Event types | << | < | 449 | > | >> | Chan. O1 | Time 451.0988 | Value 3.6619 | 35 | + | - | REJECT

Manually identifying bad channels

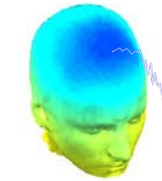


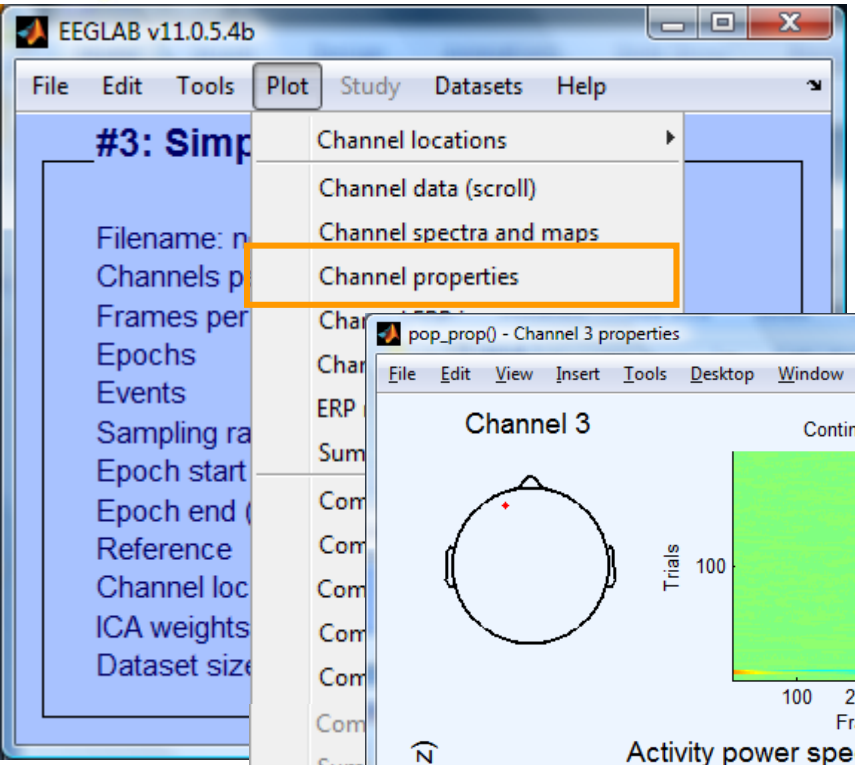


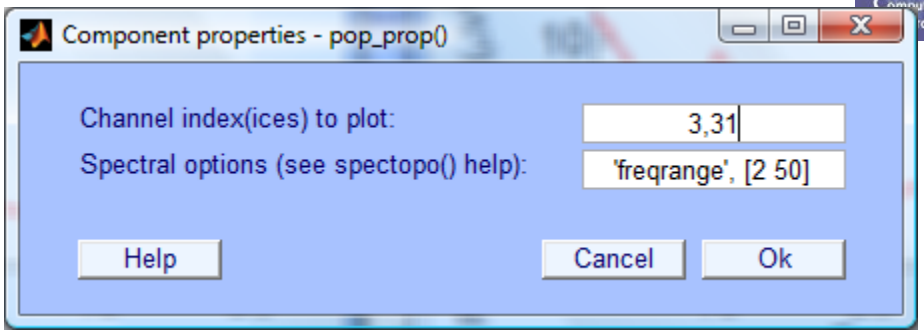


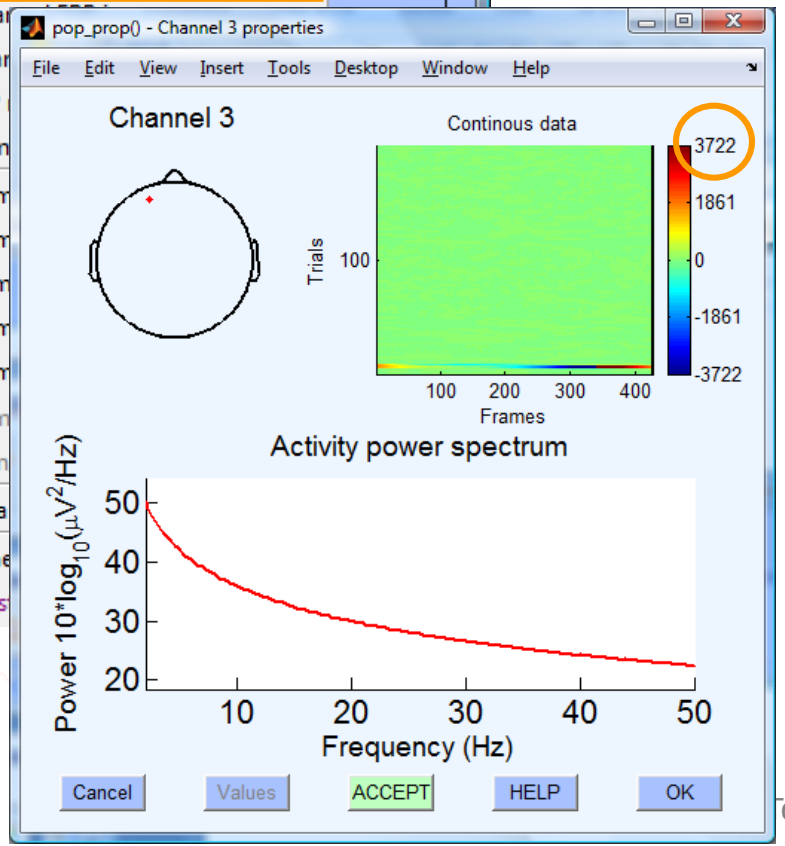


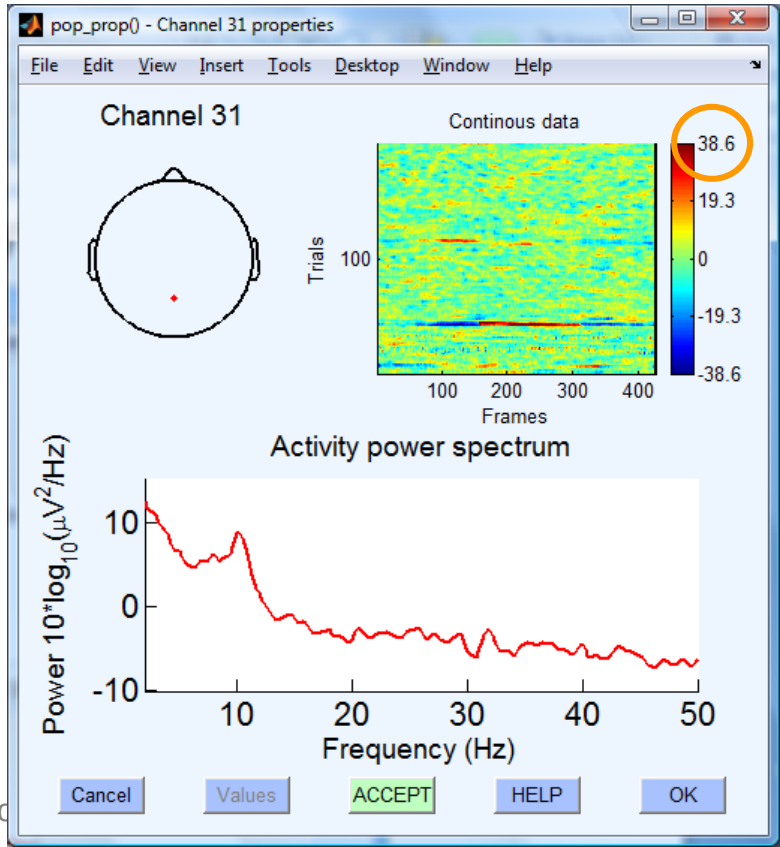
Manually identifying bad channels



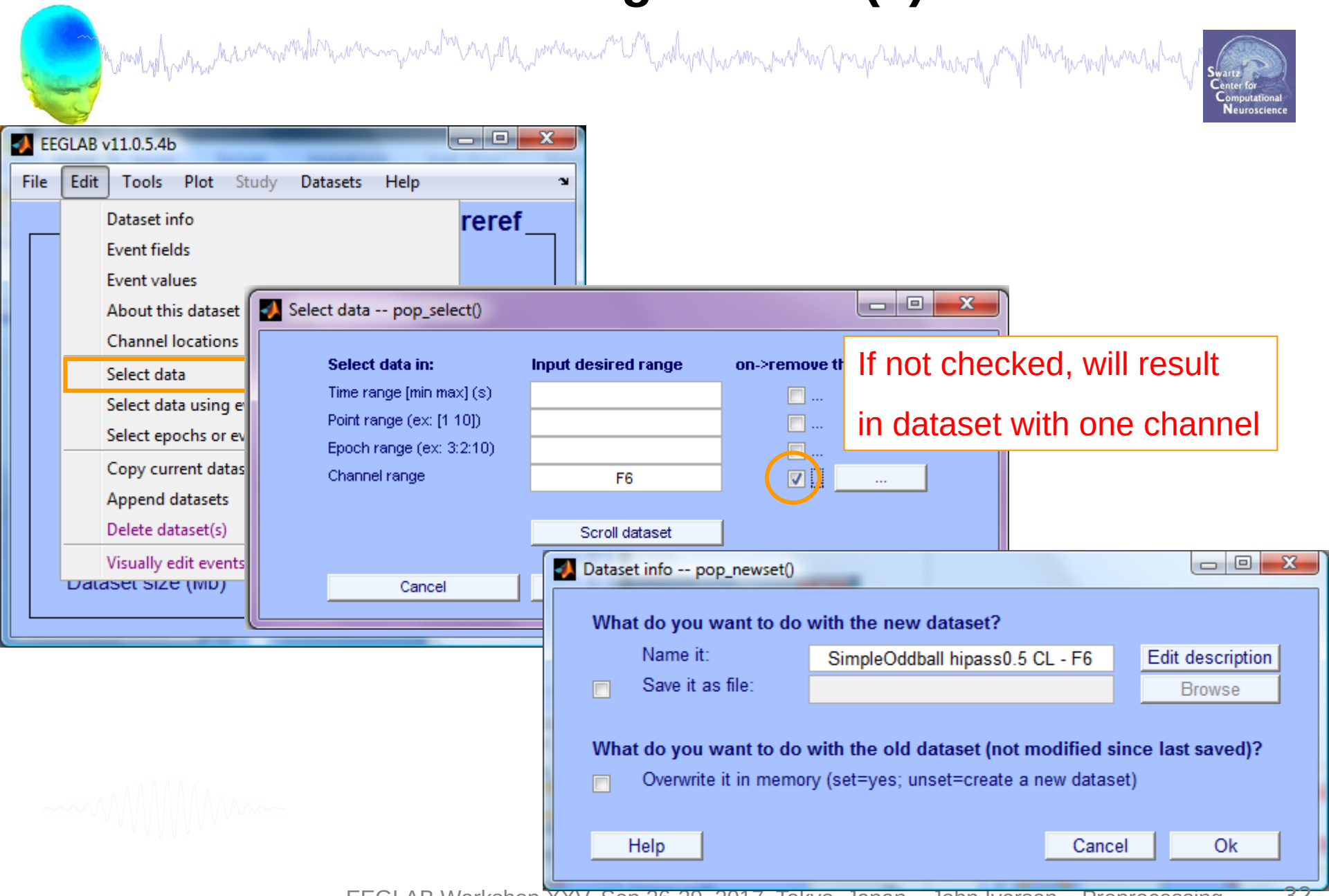




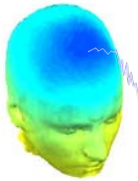




Removing channel(s)



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

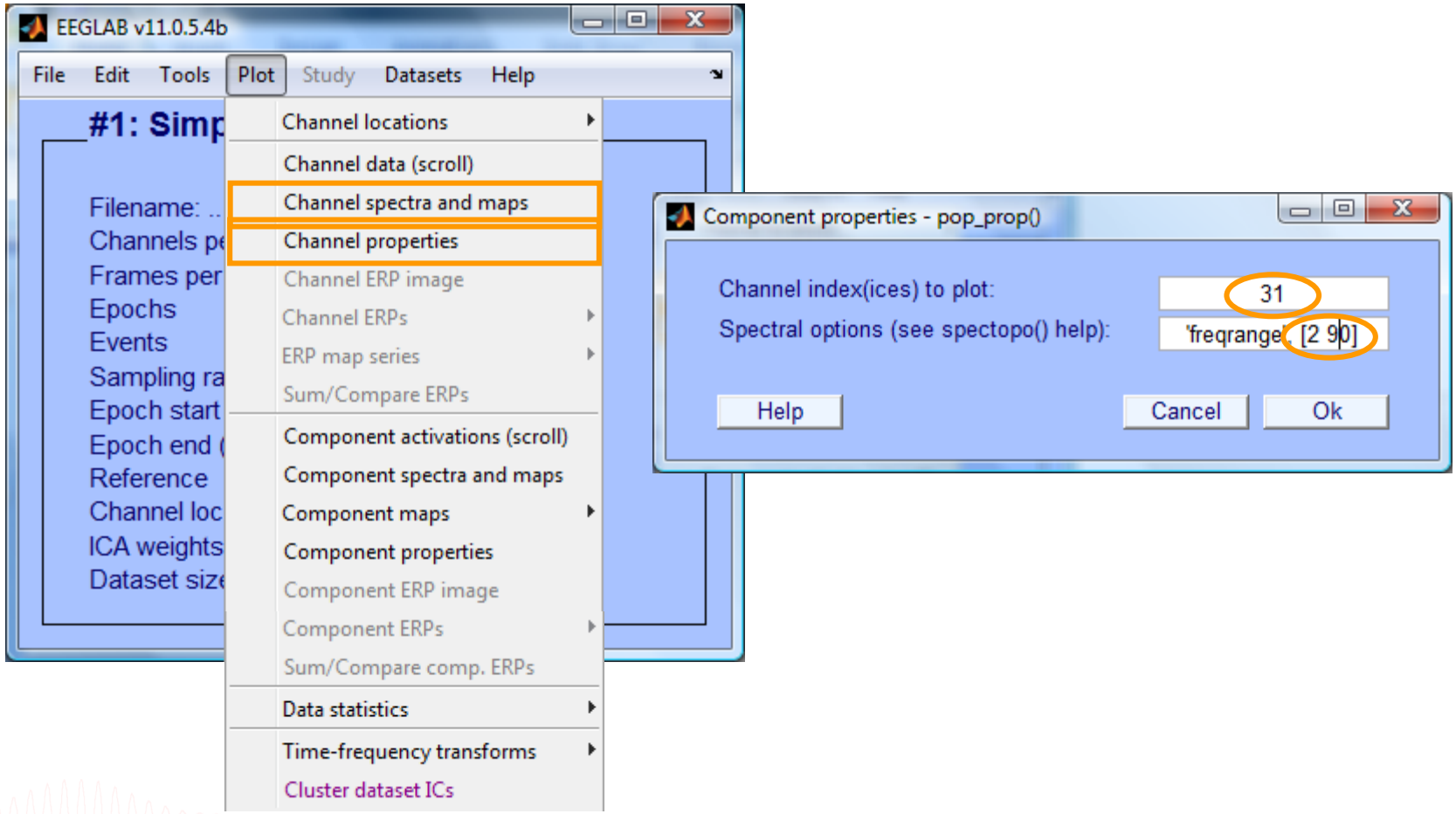
Examine raw data

Reject large artifact time points

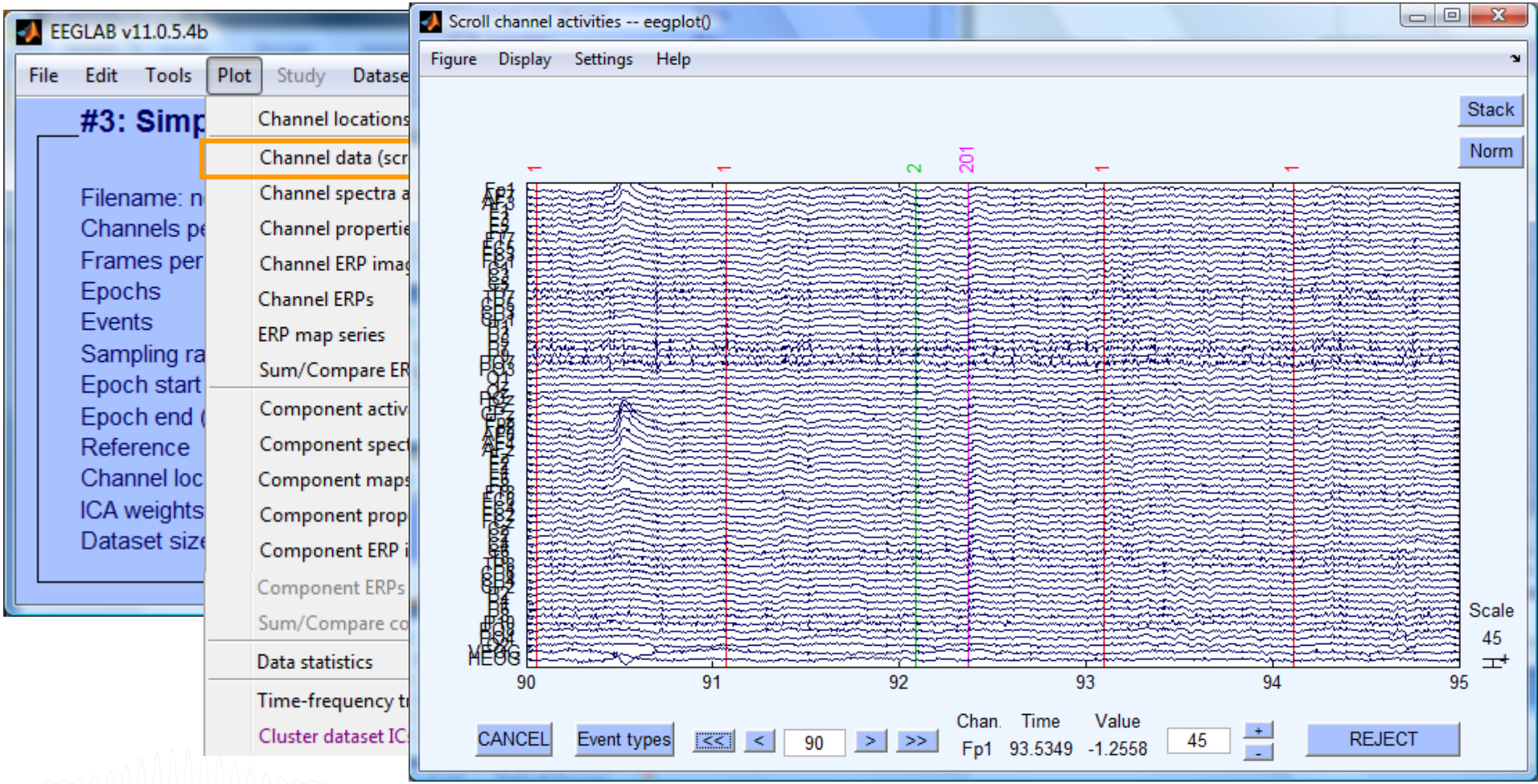
Re-reference/ down-sample (if necessary)

Run ICA

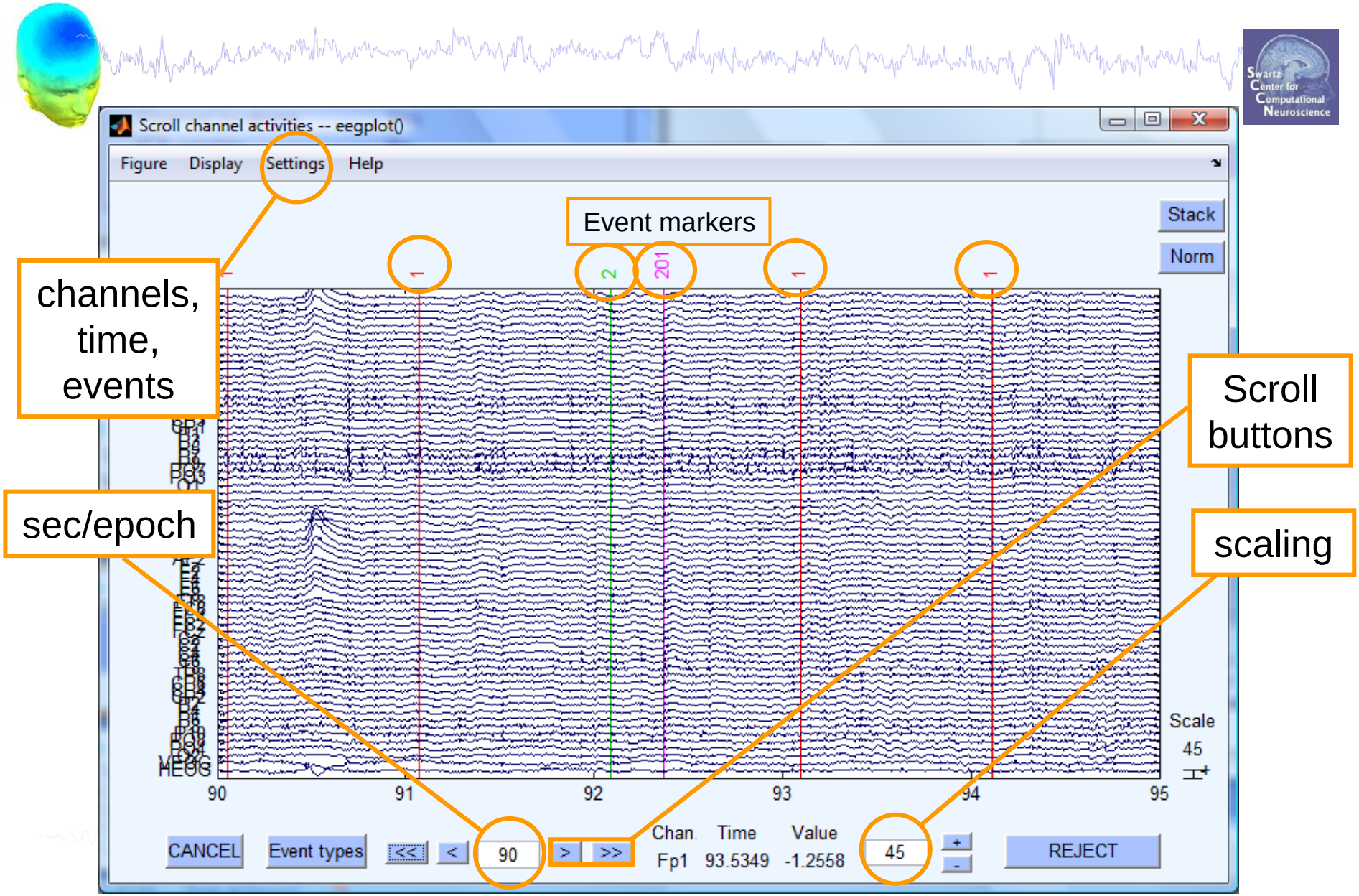
Plot channel spectra



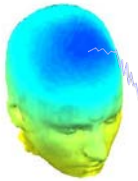
Scroll channel data



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

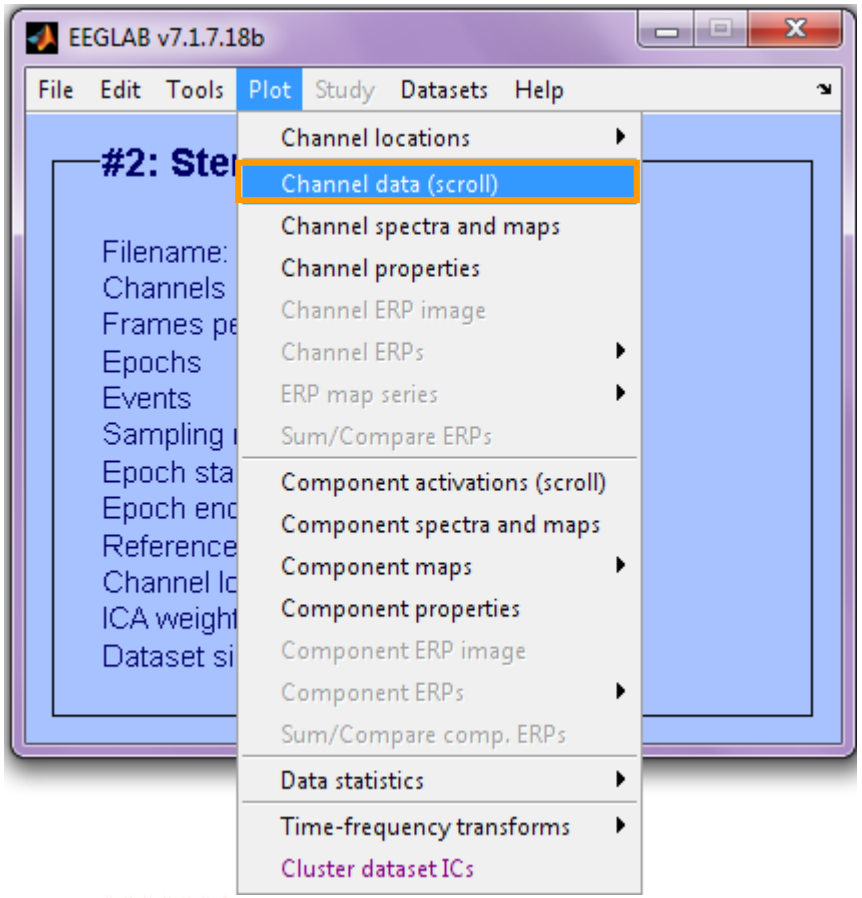
Reject large artifact time points

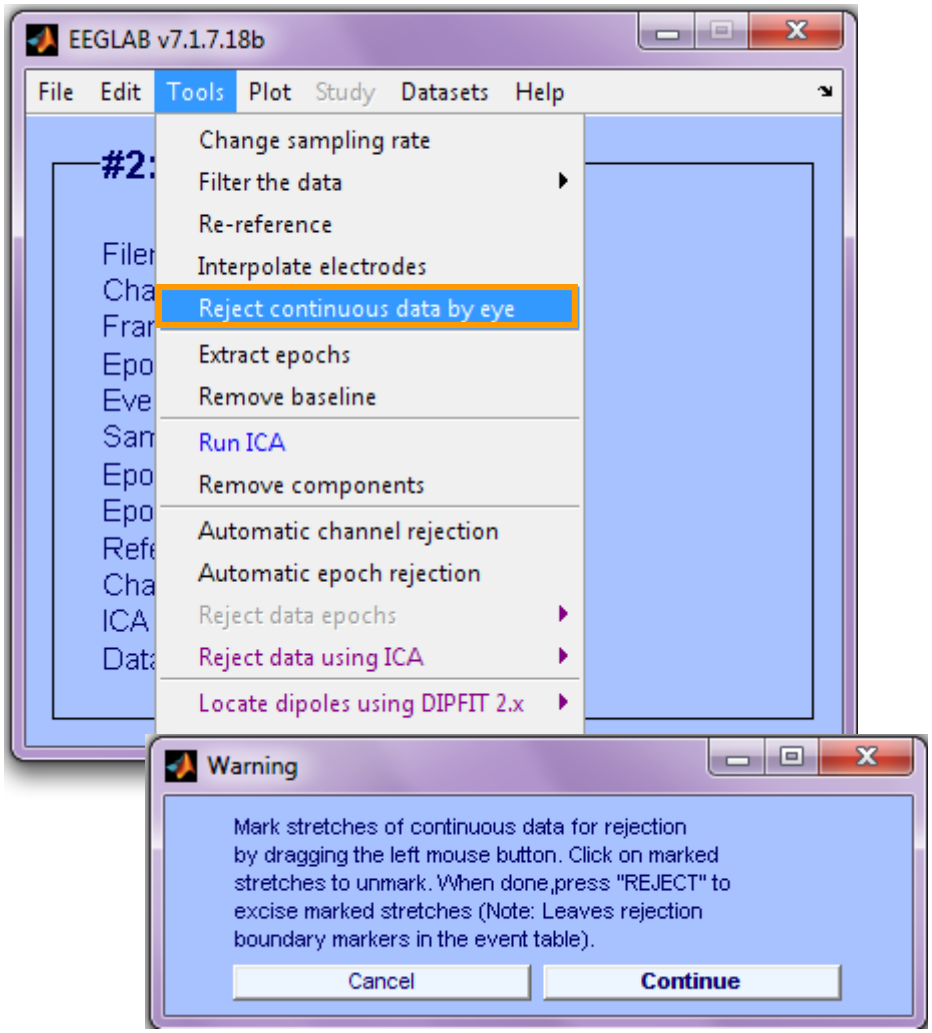
Re-reference/ down-sample (if necessary)

Run ICA

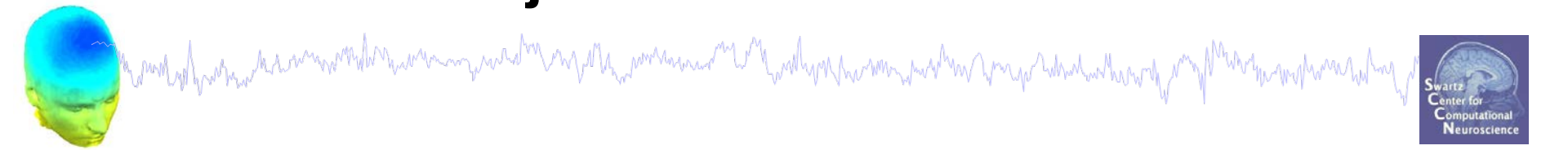
Reject continuous data

Equivalent





Reject continuous data



Scroll channel activities -- eegplot()

Figure Display Settings Help

Click and drag with mouse over noisy data to reject

536 537 538 539 540 541

Scale 35

CANCEL | Event types | << | < | 536 | > | >>

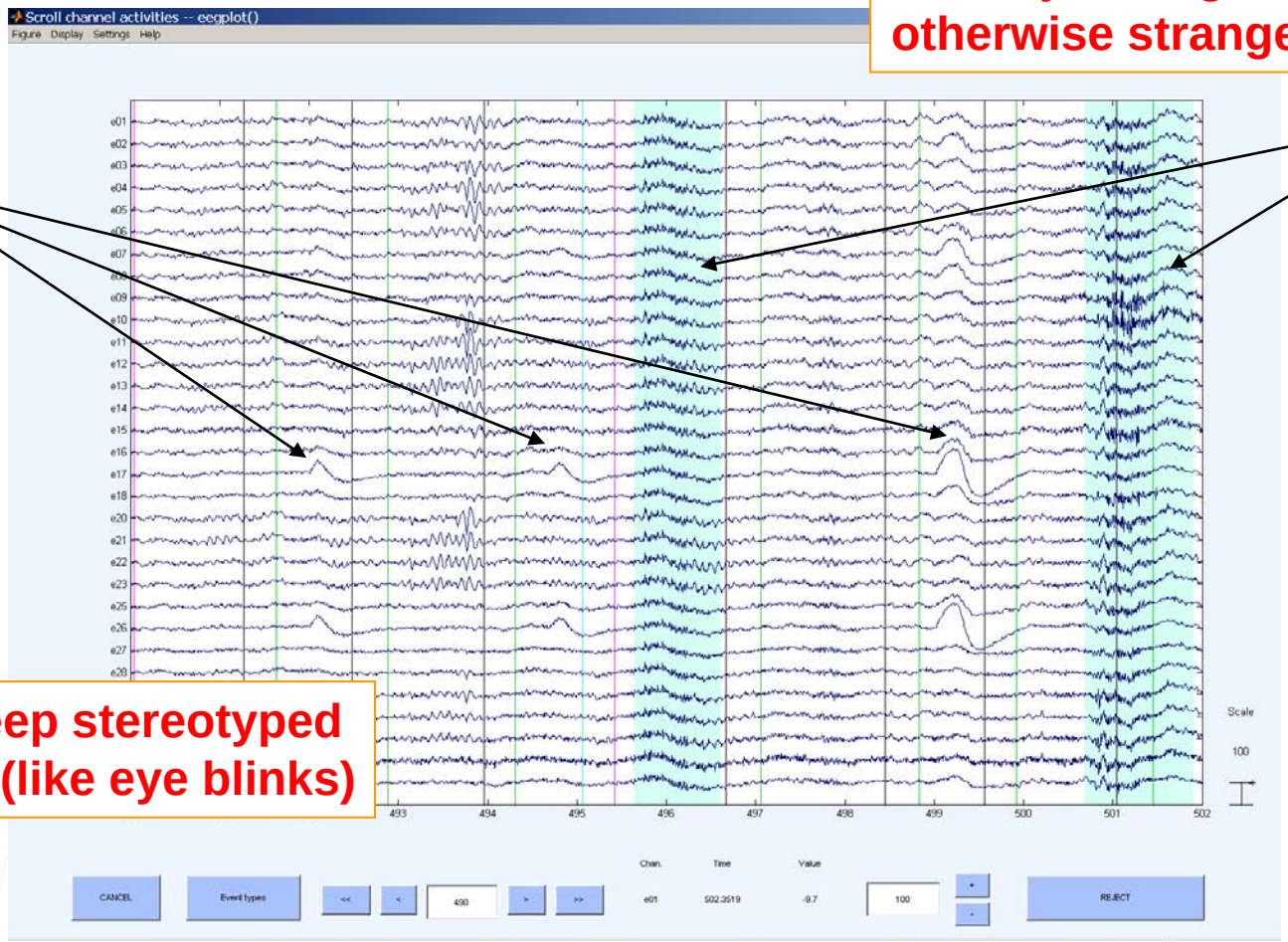
Chan. FC6 | Time 539.9355 | Value 4.8773

35 | + | - | REJECT

Rejecting data for ICA

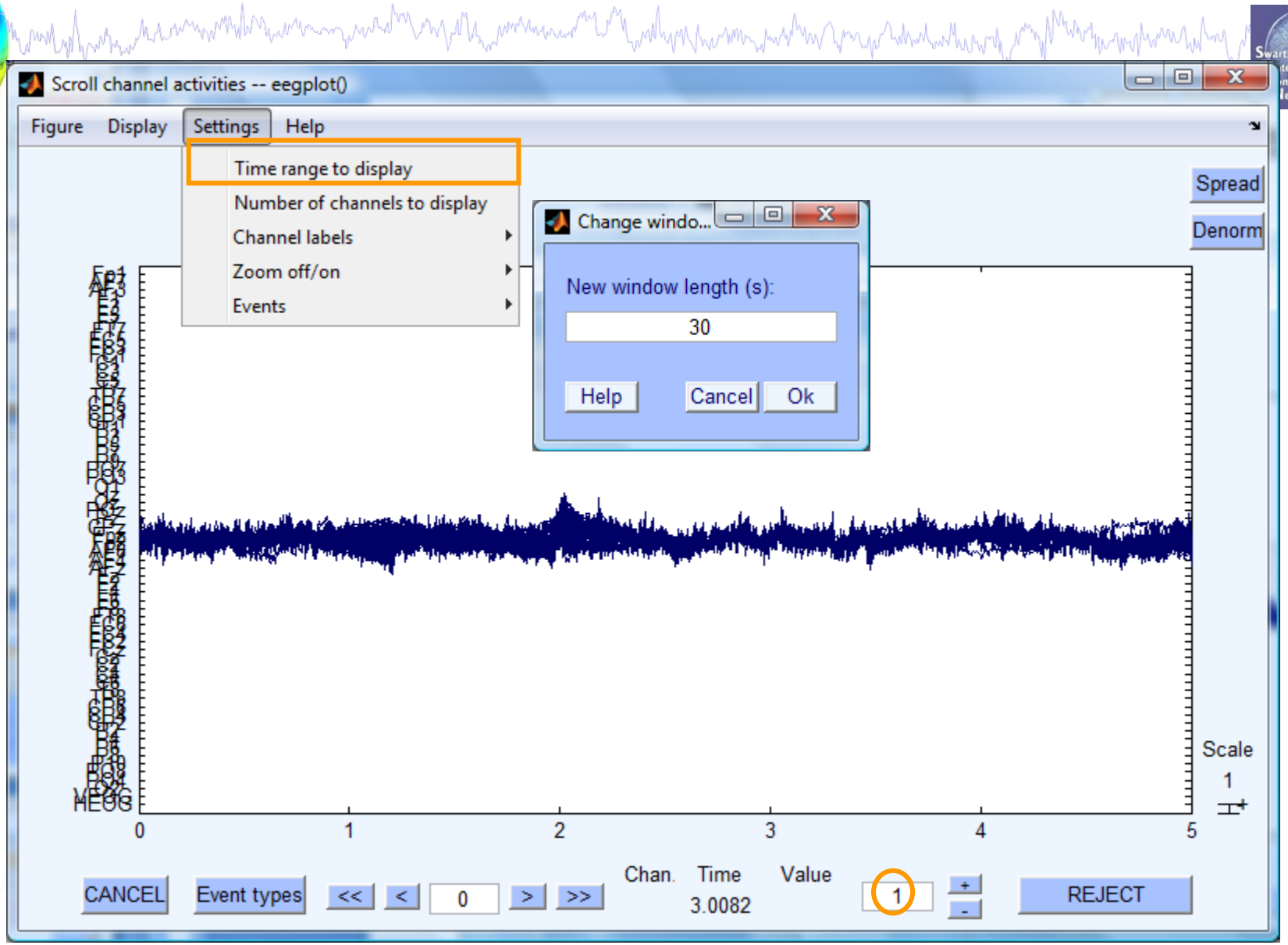
To prepare data for ICA:

Reject large muscle or otherwise strange events...

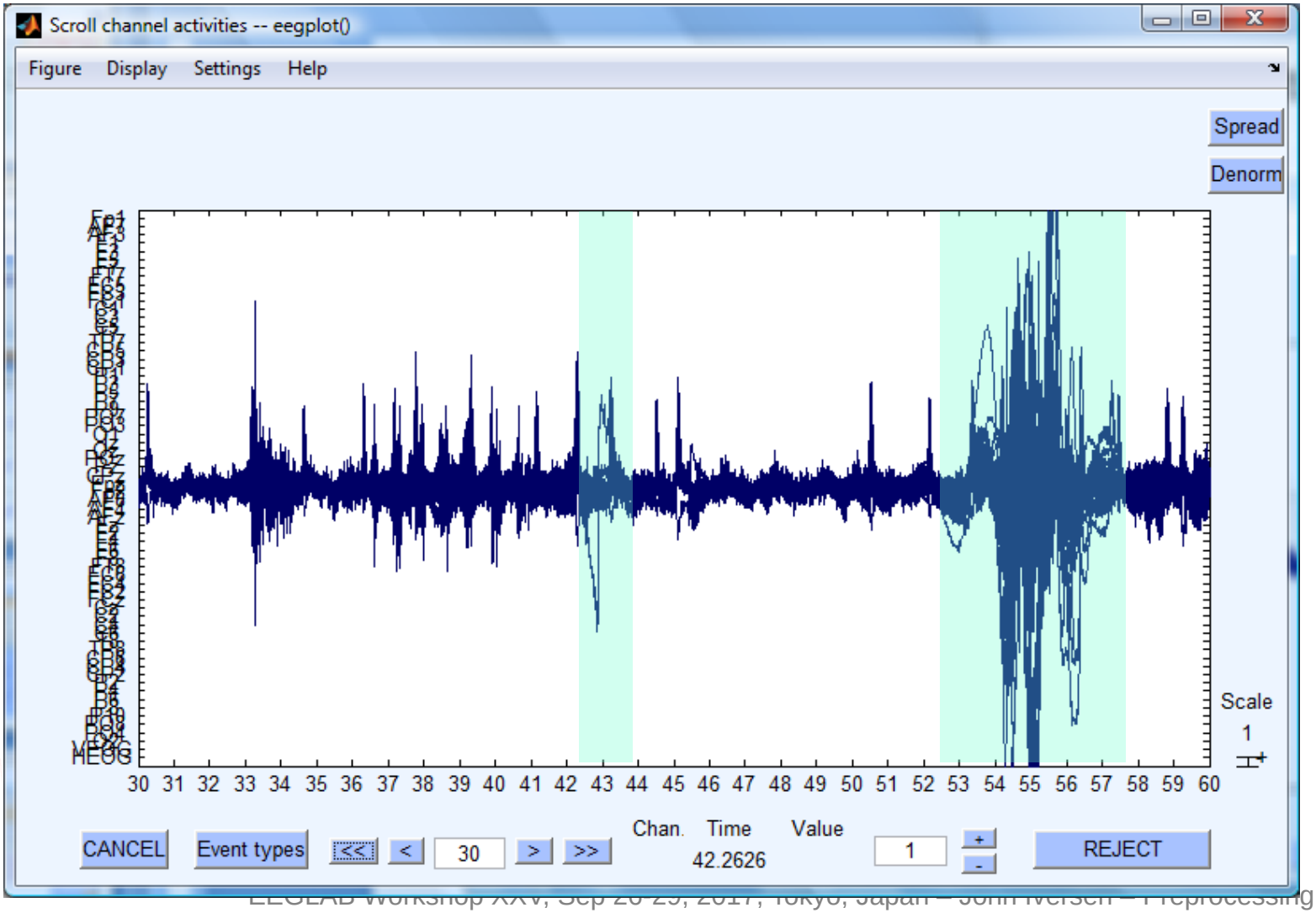


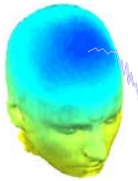
... but keep stereotyped artifacts (like eye blinks)

Fast (but sloppy) artifact rejection



Fast (but sloppy) artifact rejection

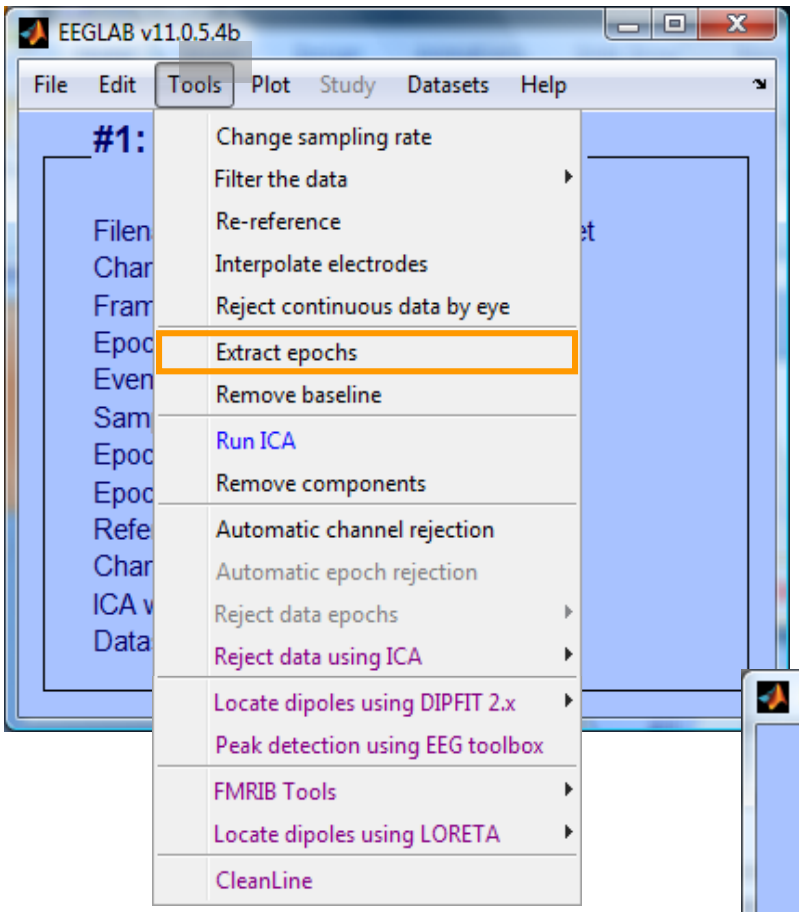




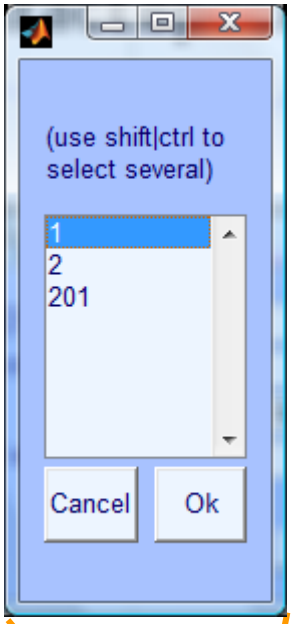
Data Cleaning for ICA

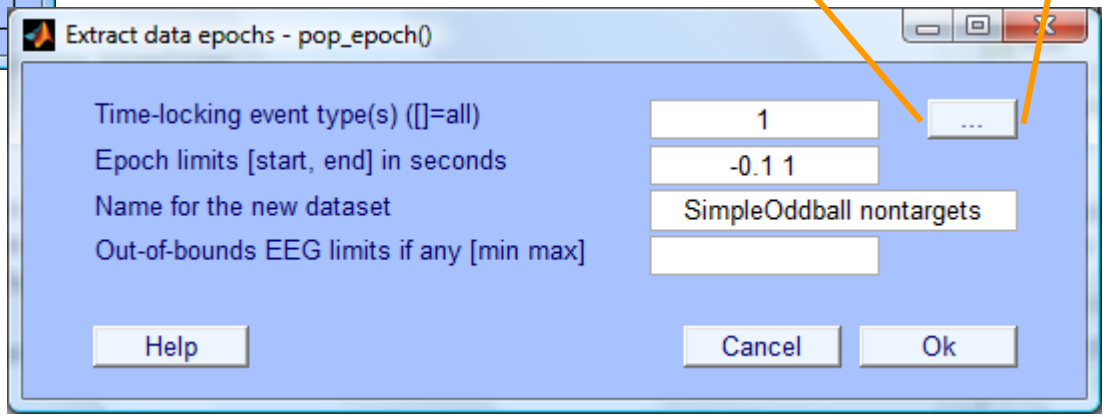
Variant 2: Epoched Data

Extract epochs

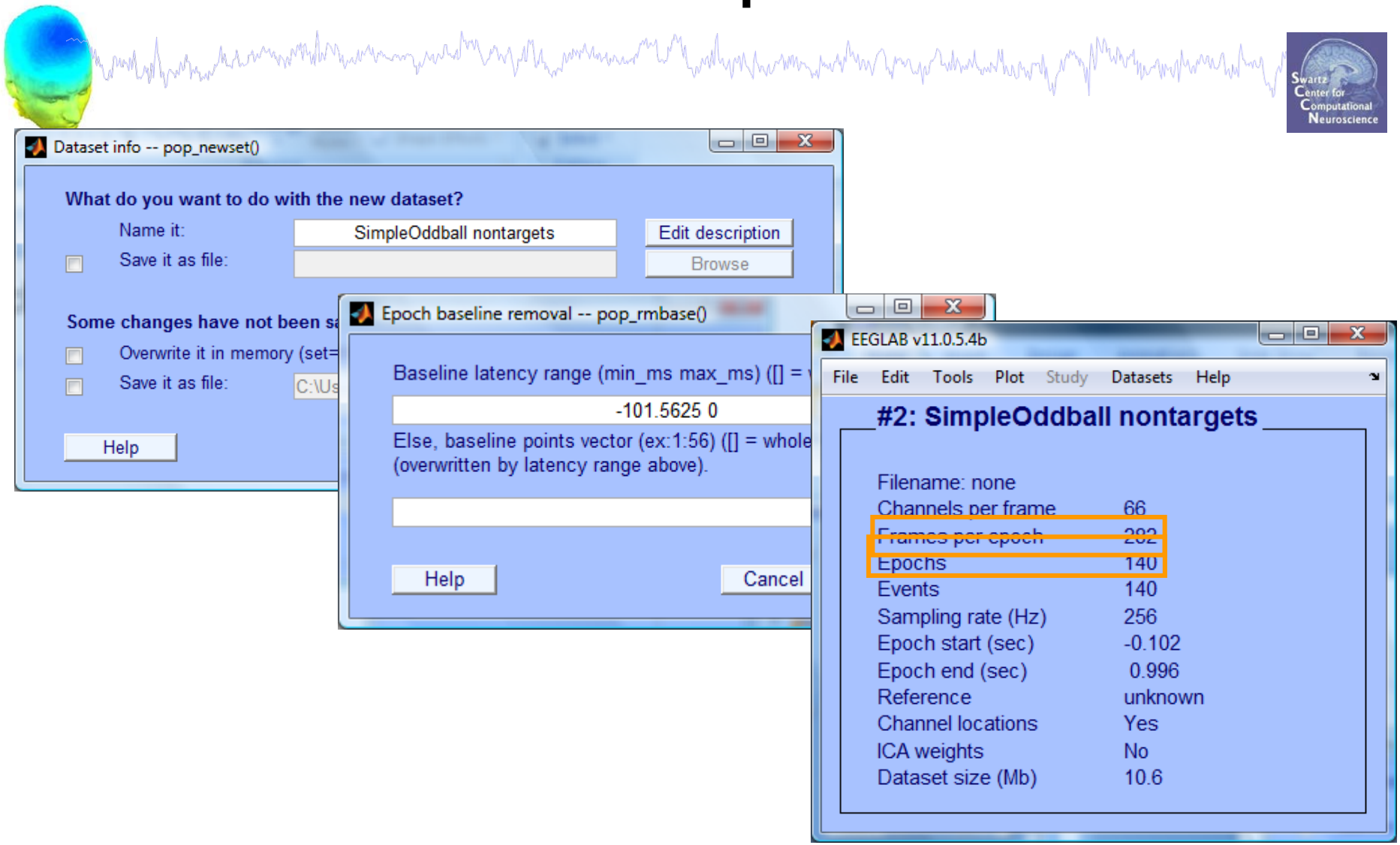



```
>> eeg_eventtypes(EEG)

   1       140

   2       60

   201     60
```

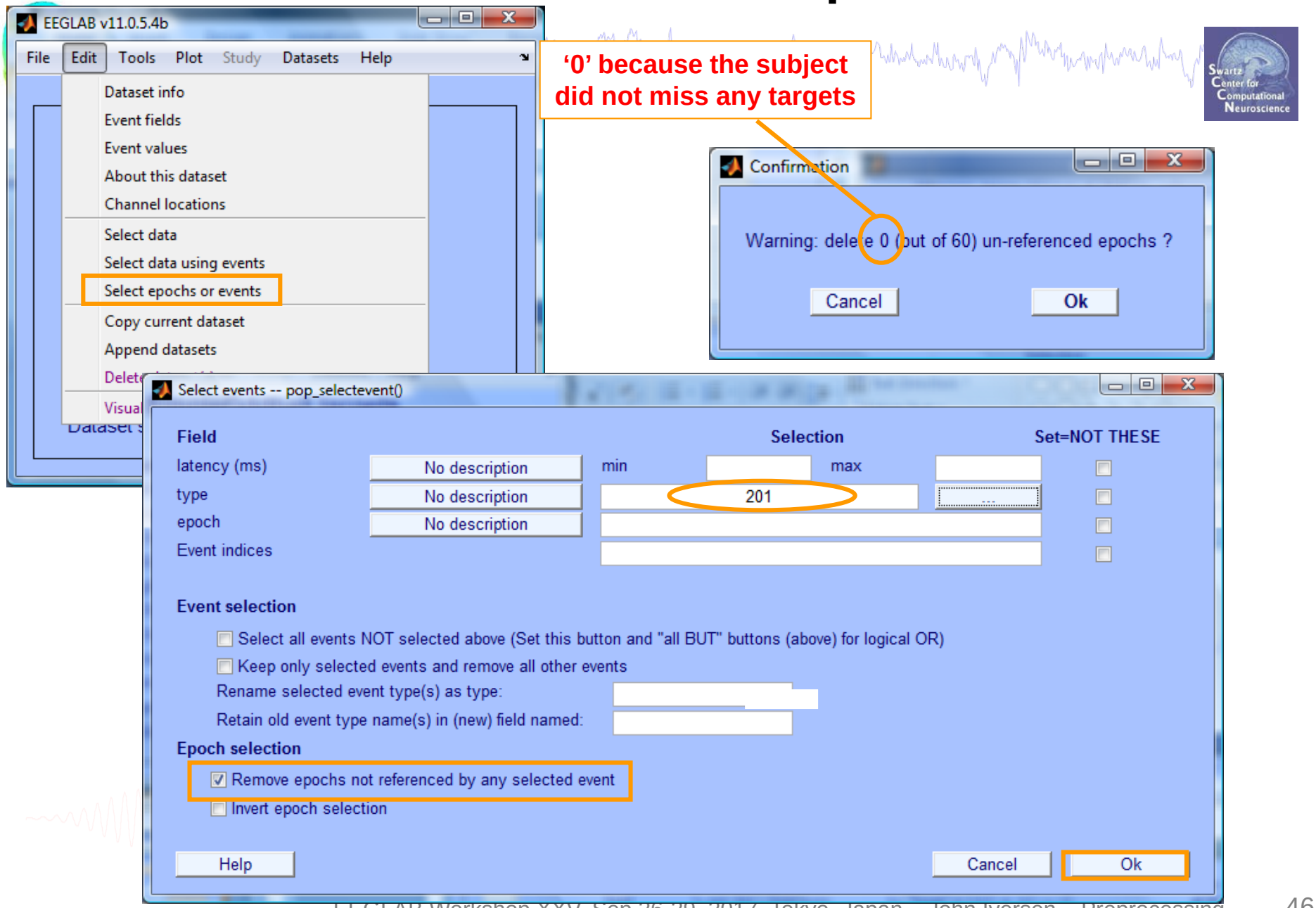


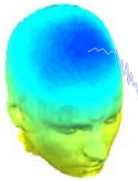


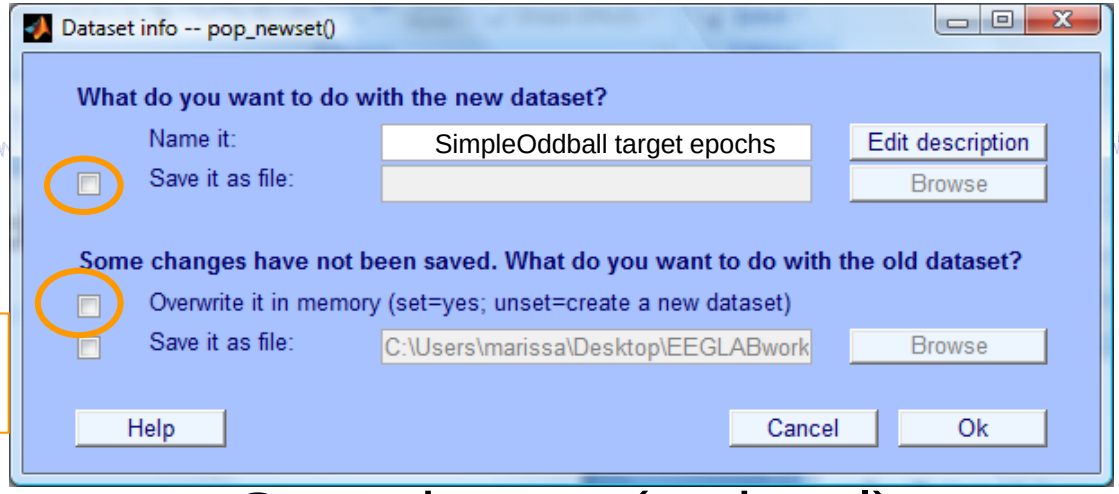
Extract epochs



Select a subset of epochs

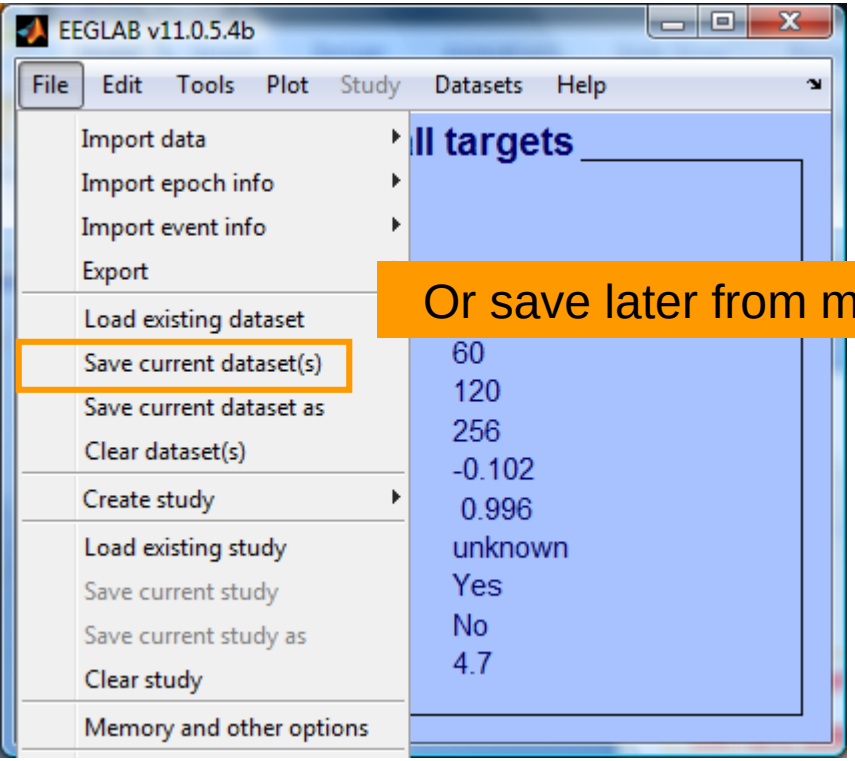




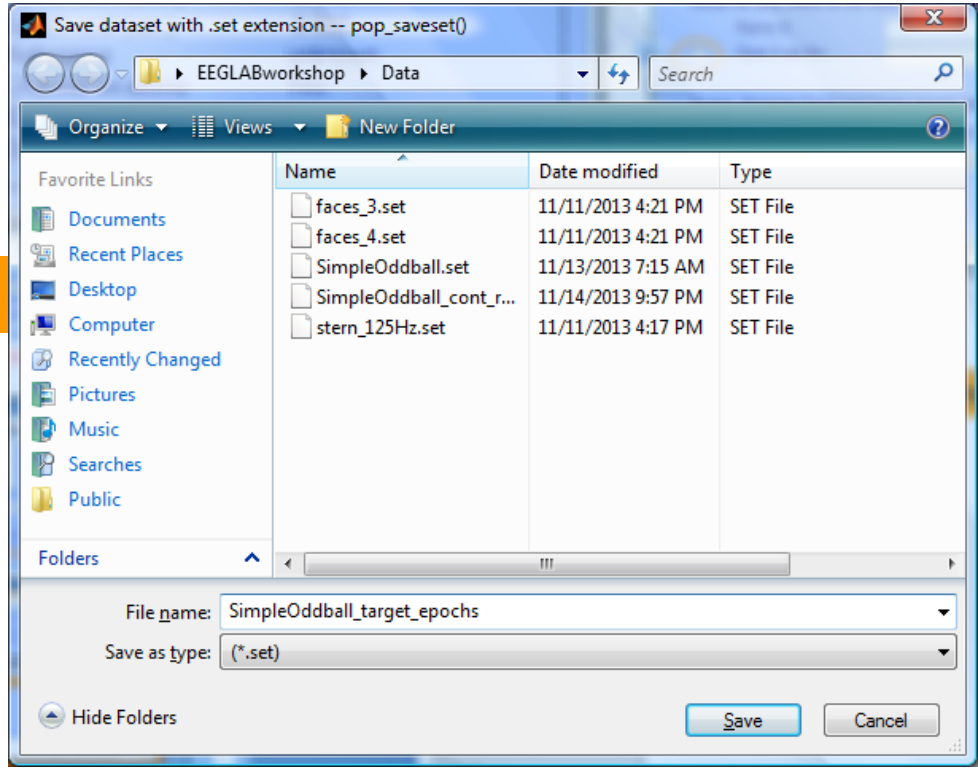


'Do not overwrite current dataset'

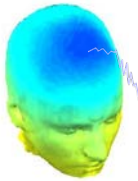
Save dataset (optional)

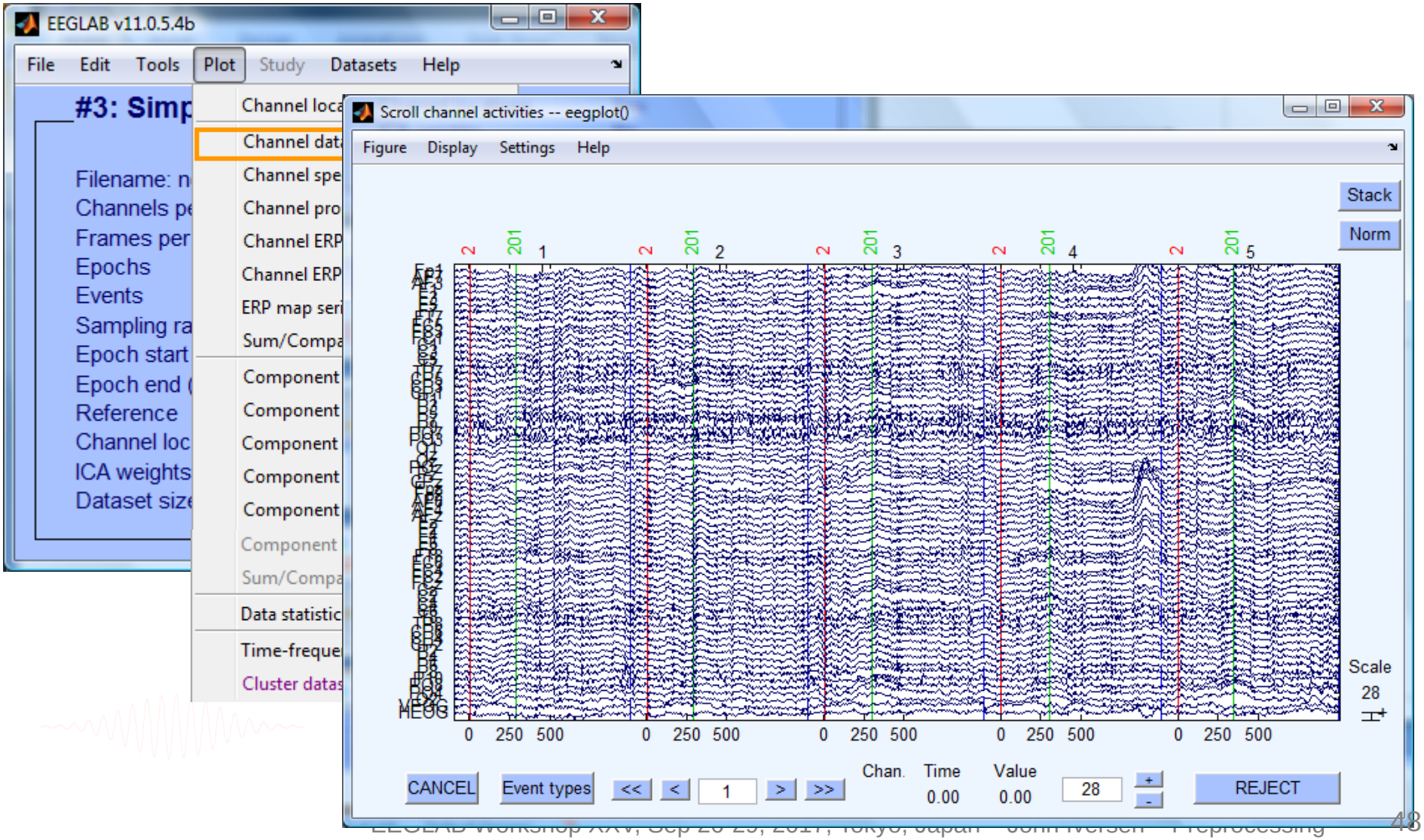


Or save later from menu



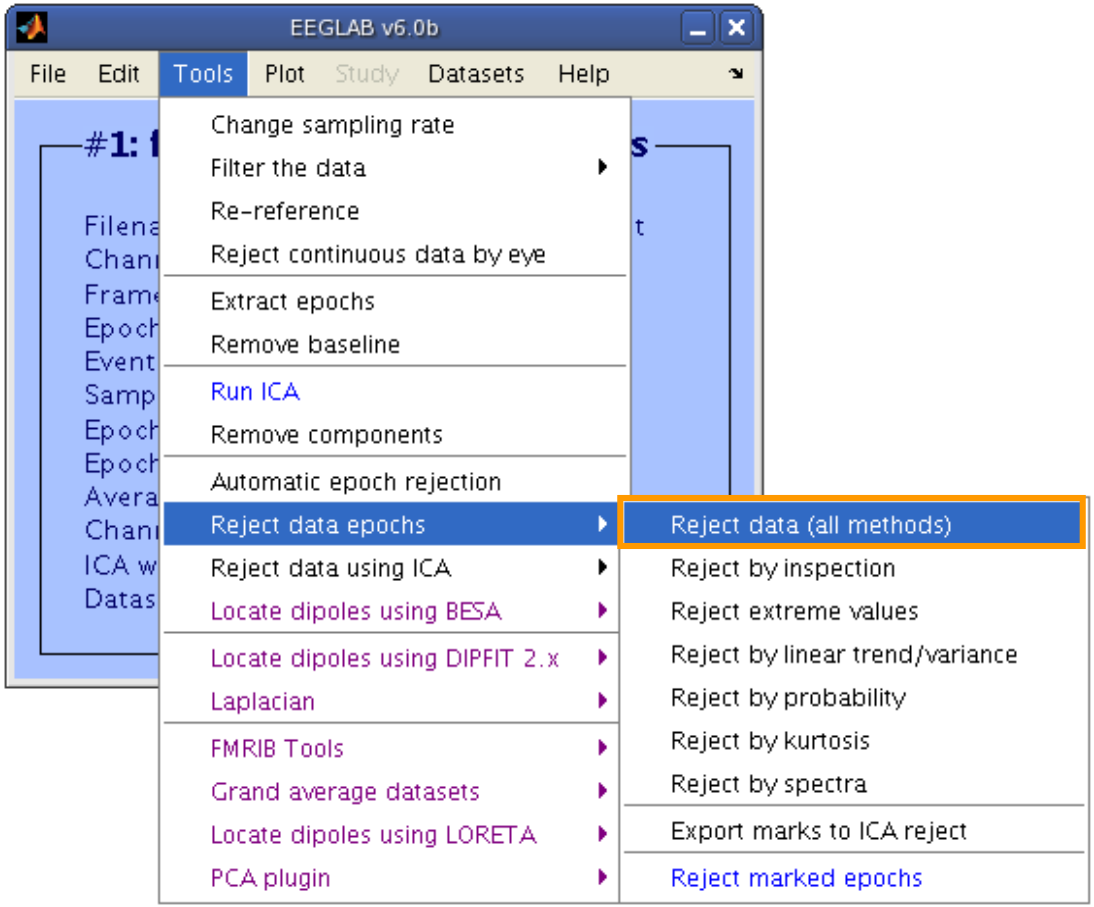
Scroll (epoched) channel data



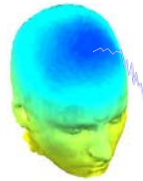


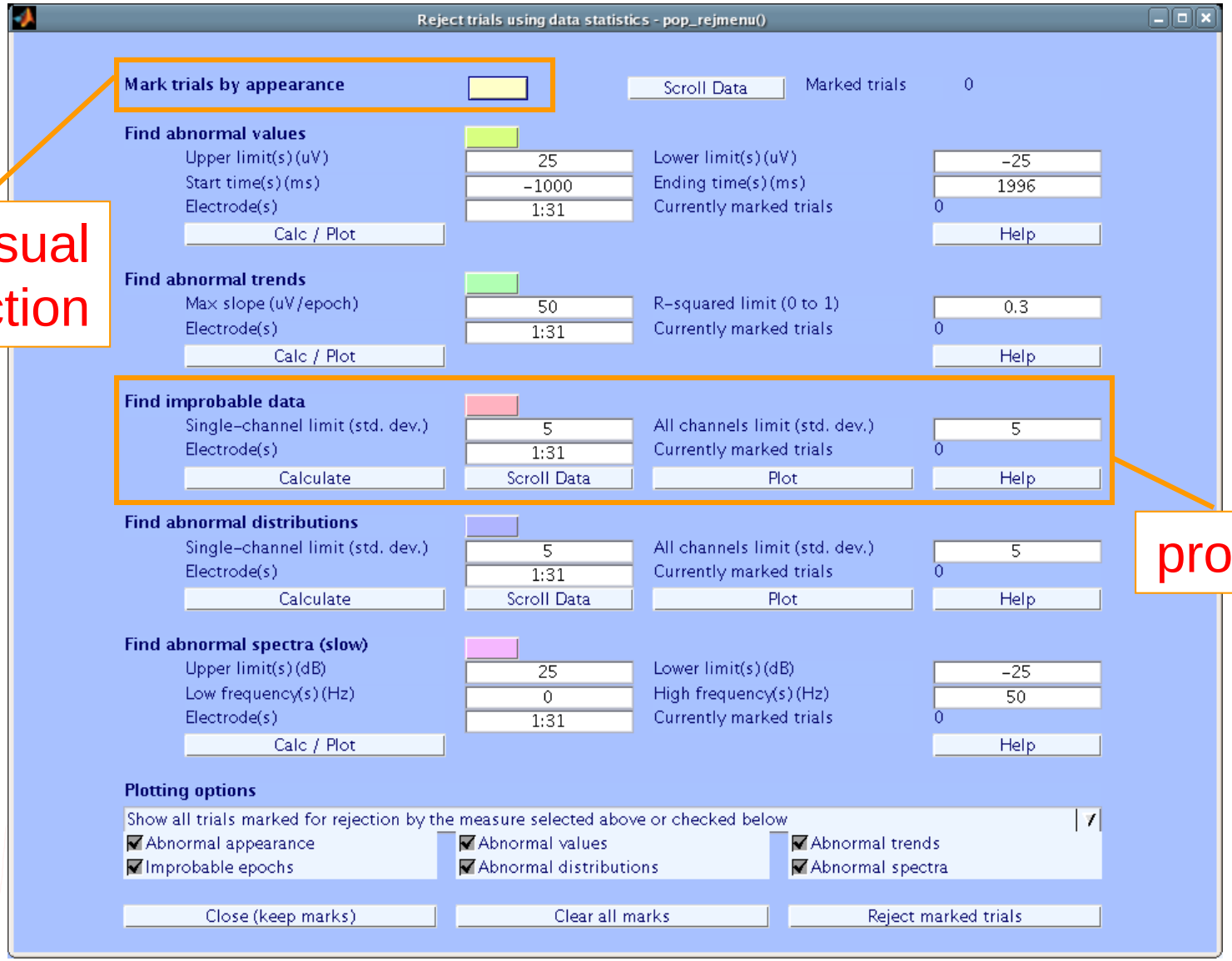
Reject epochs with artifact

Reject data epochs



Reject data epochs

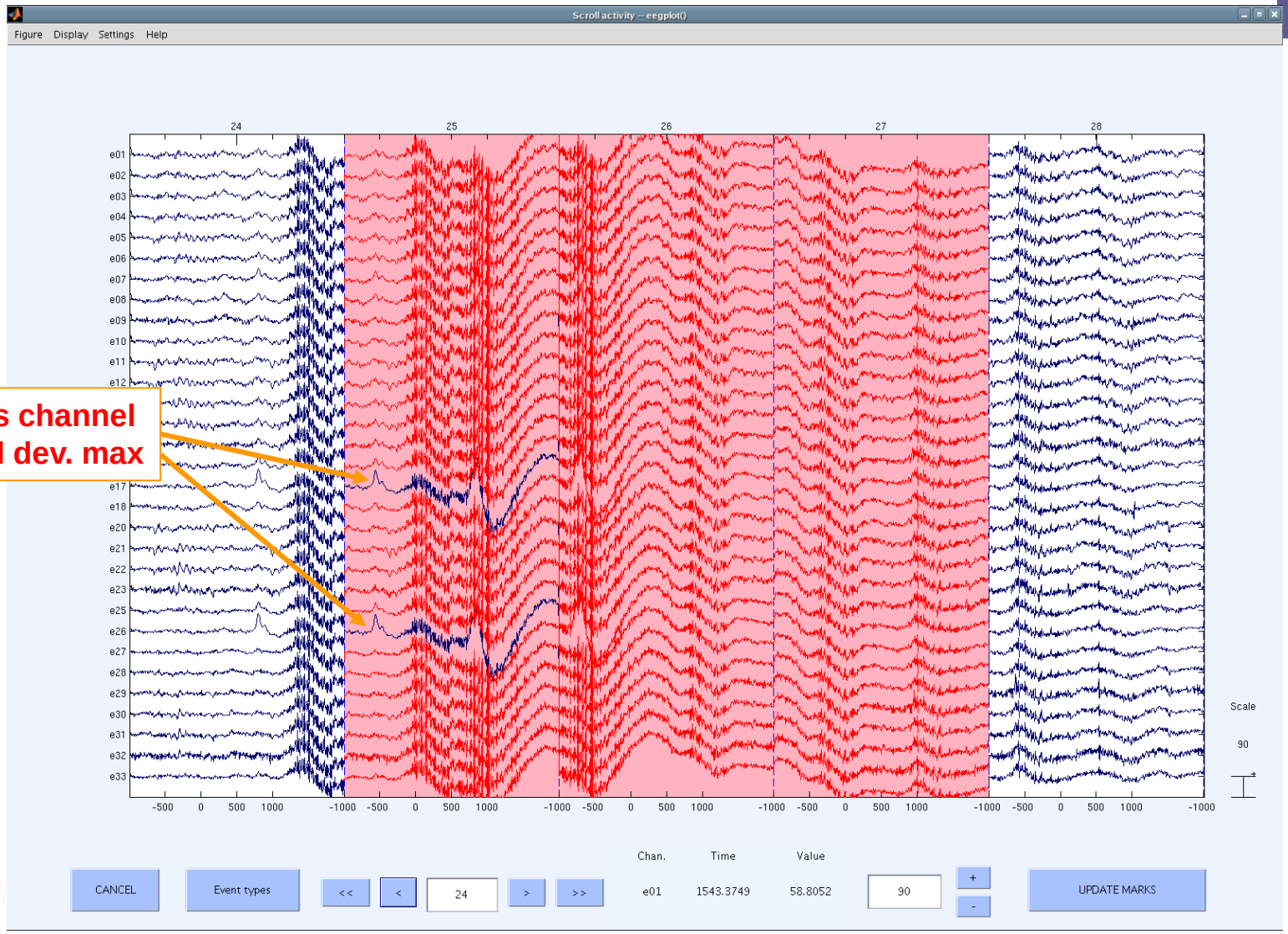




visual inspection

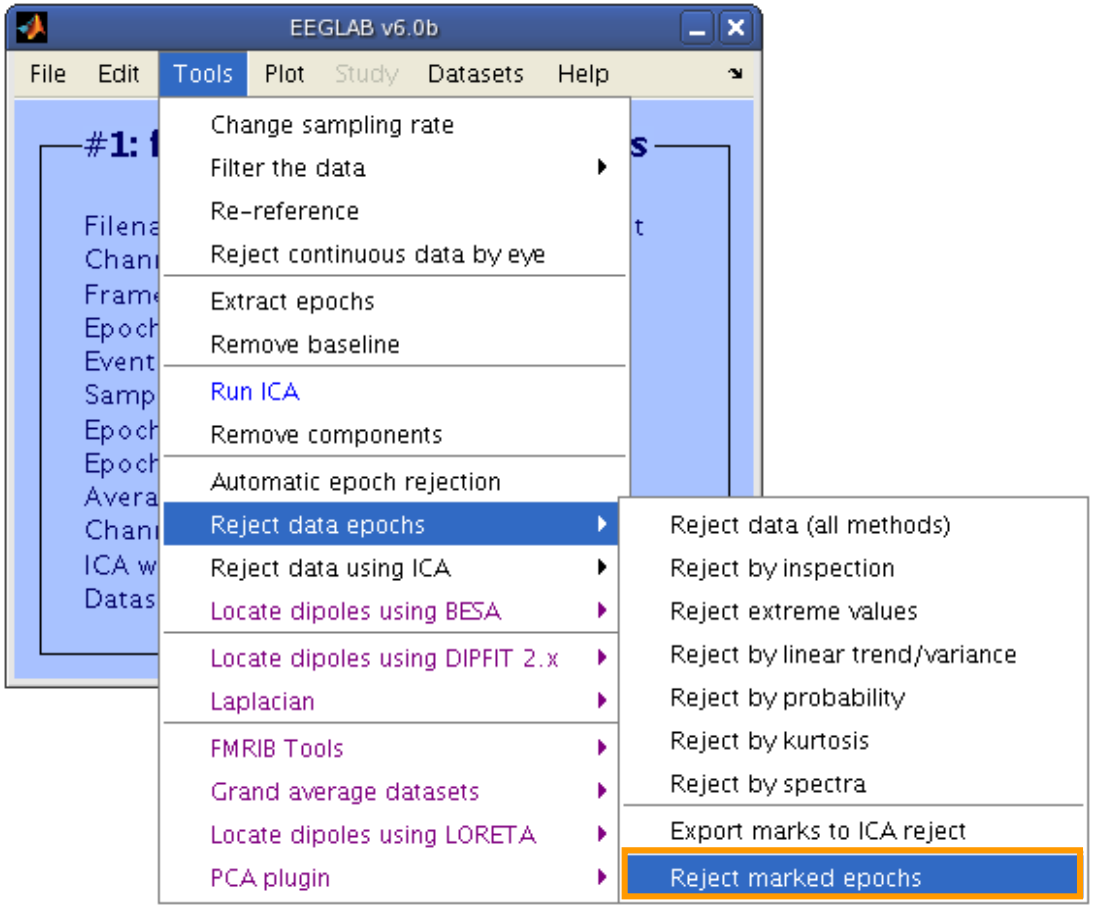
probability

Reject data epochs

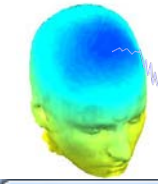


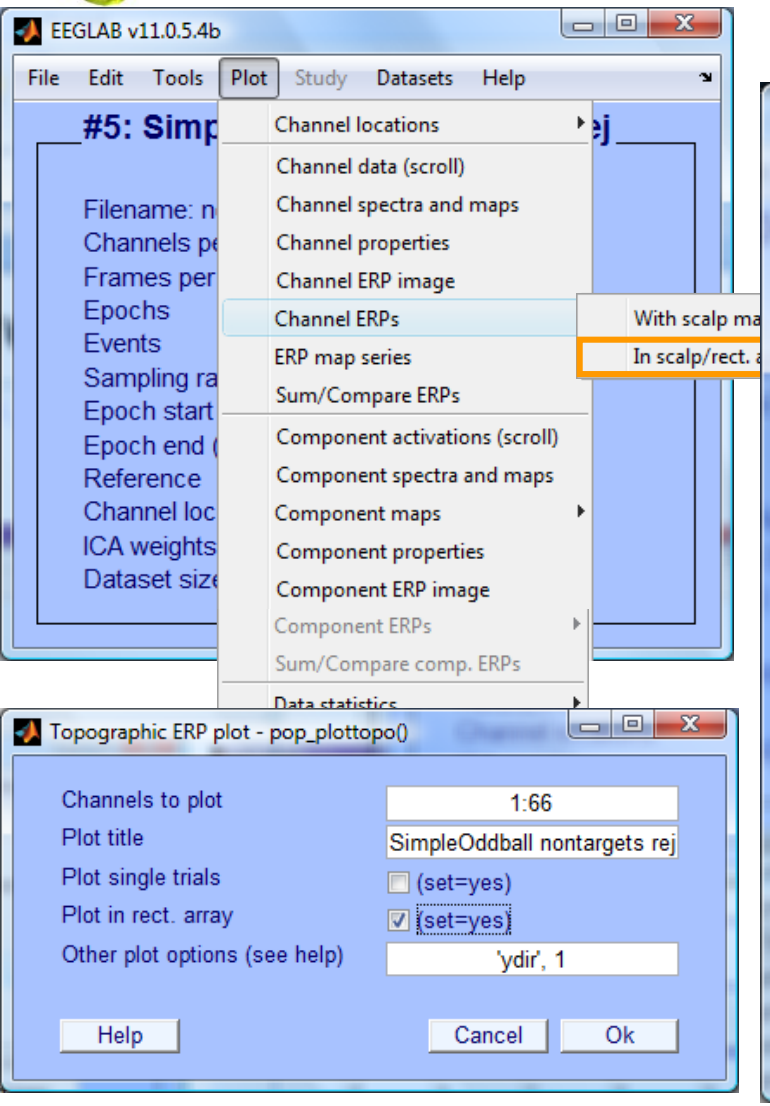
Exceeds channel standard dev. max

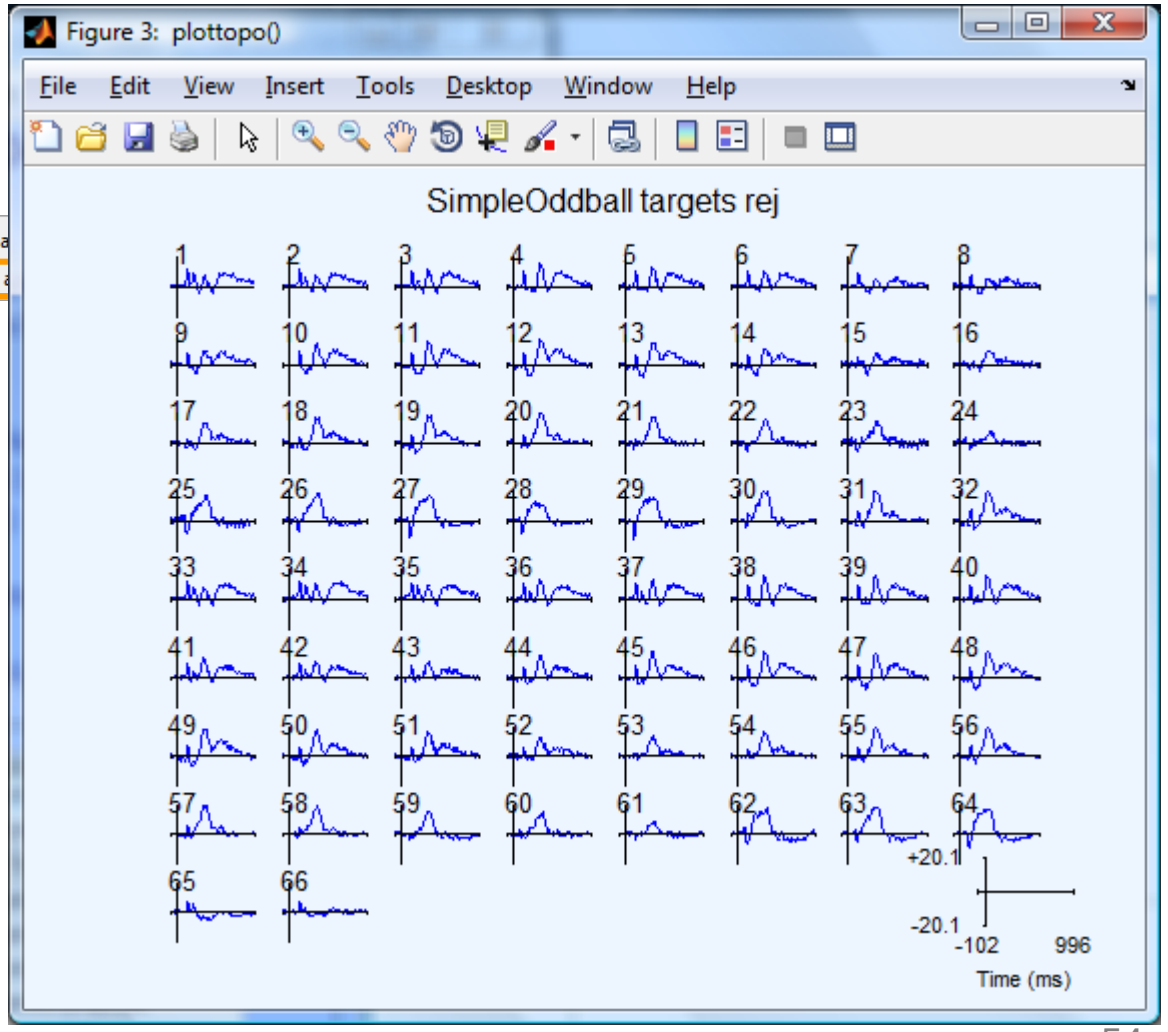
Reject data epochs



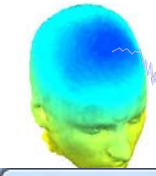
Visualize ERP in rectangular array

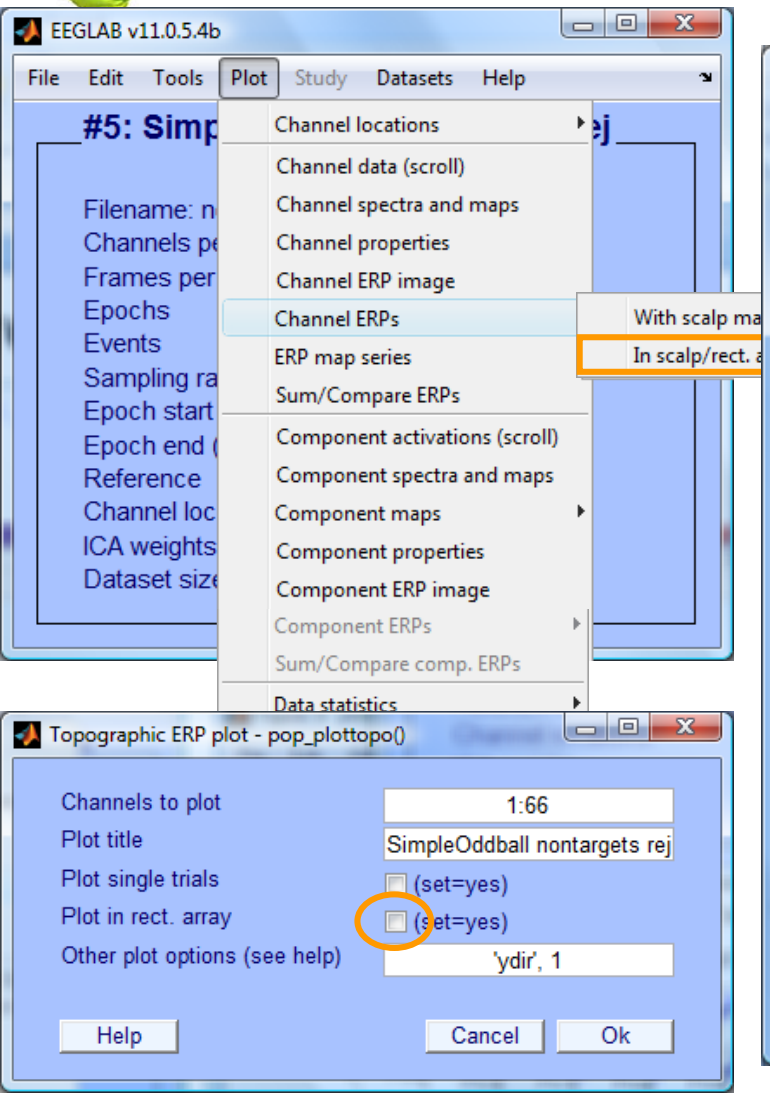


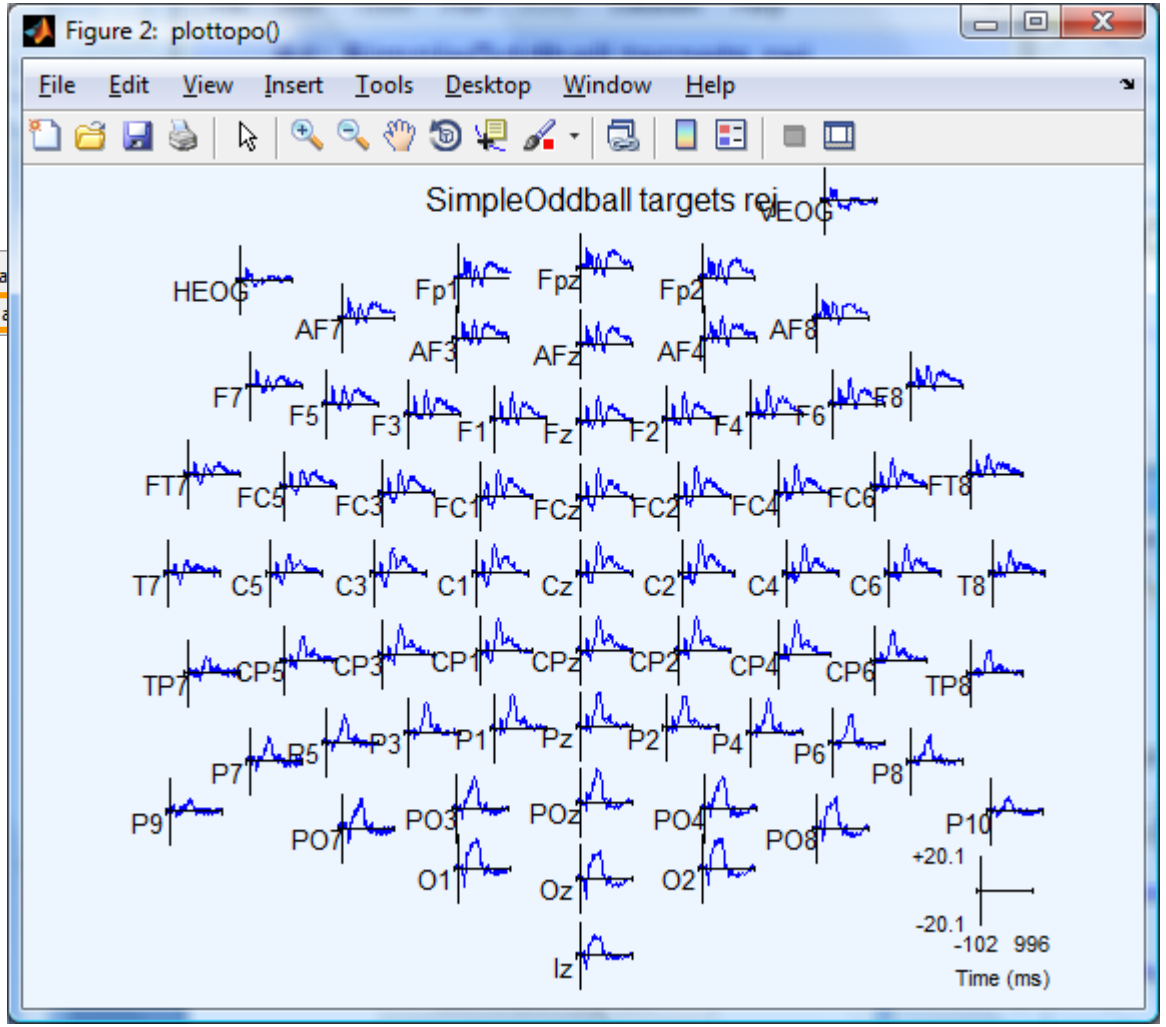




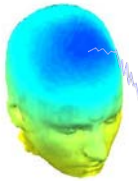
Visualize ERP in topographic array

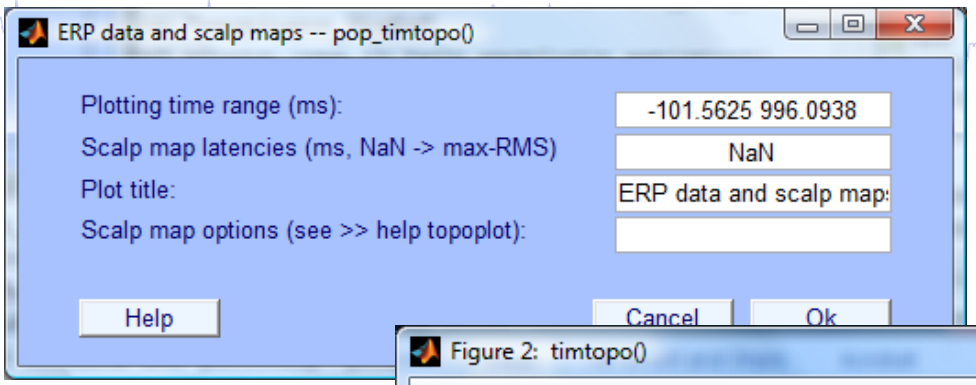


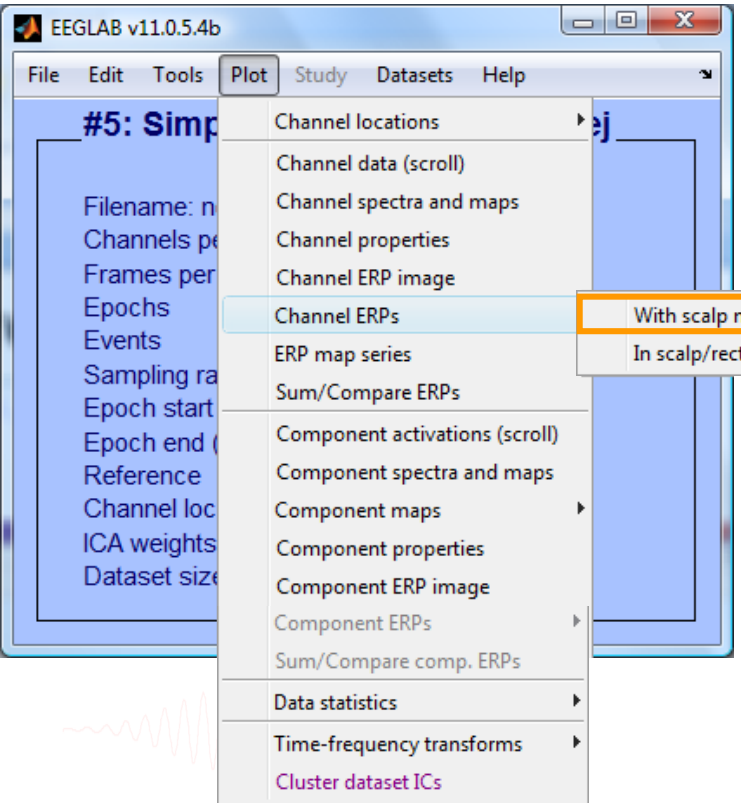


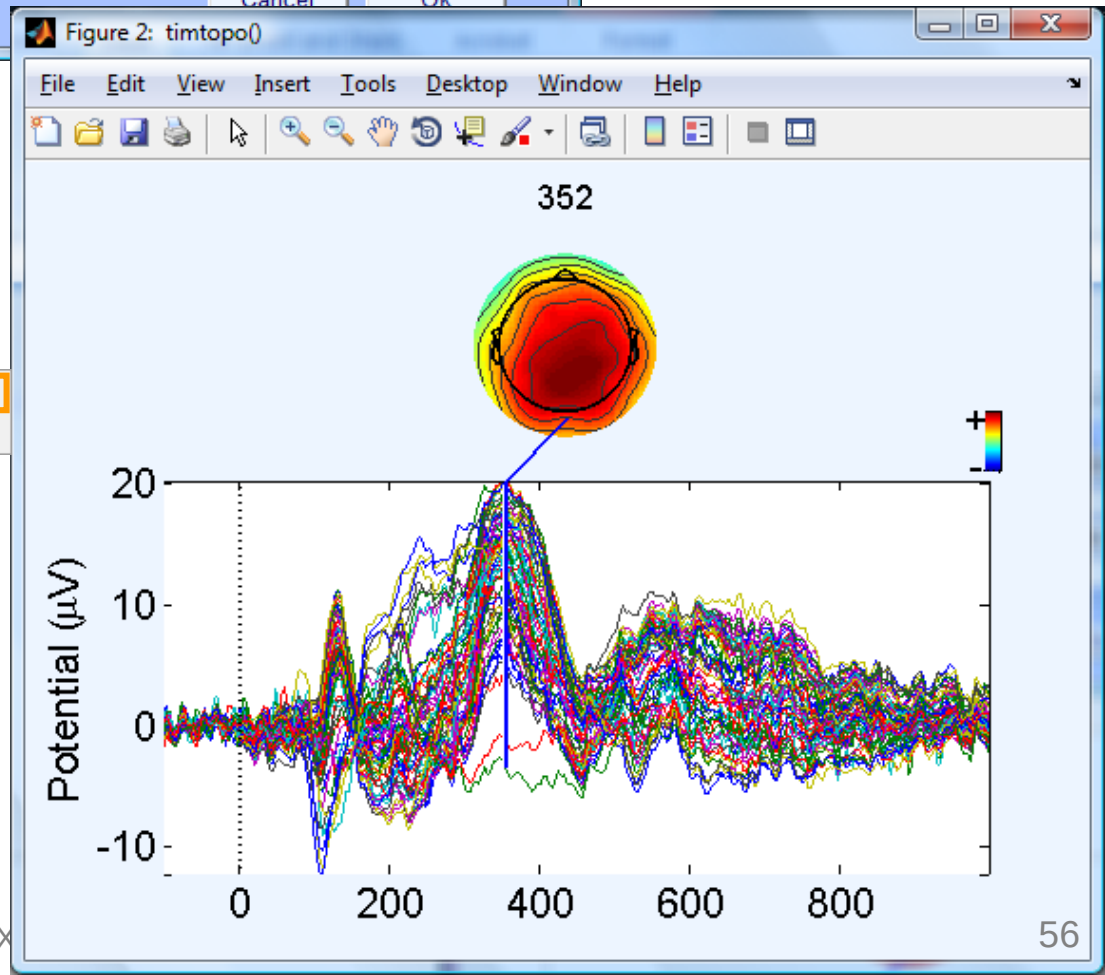


Visualize ERP scalp distribution

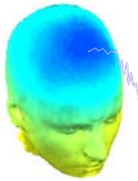


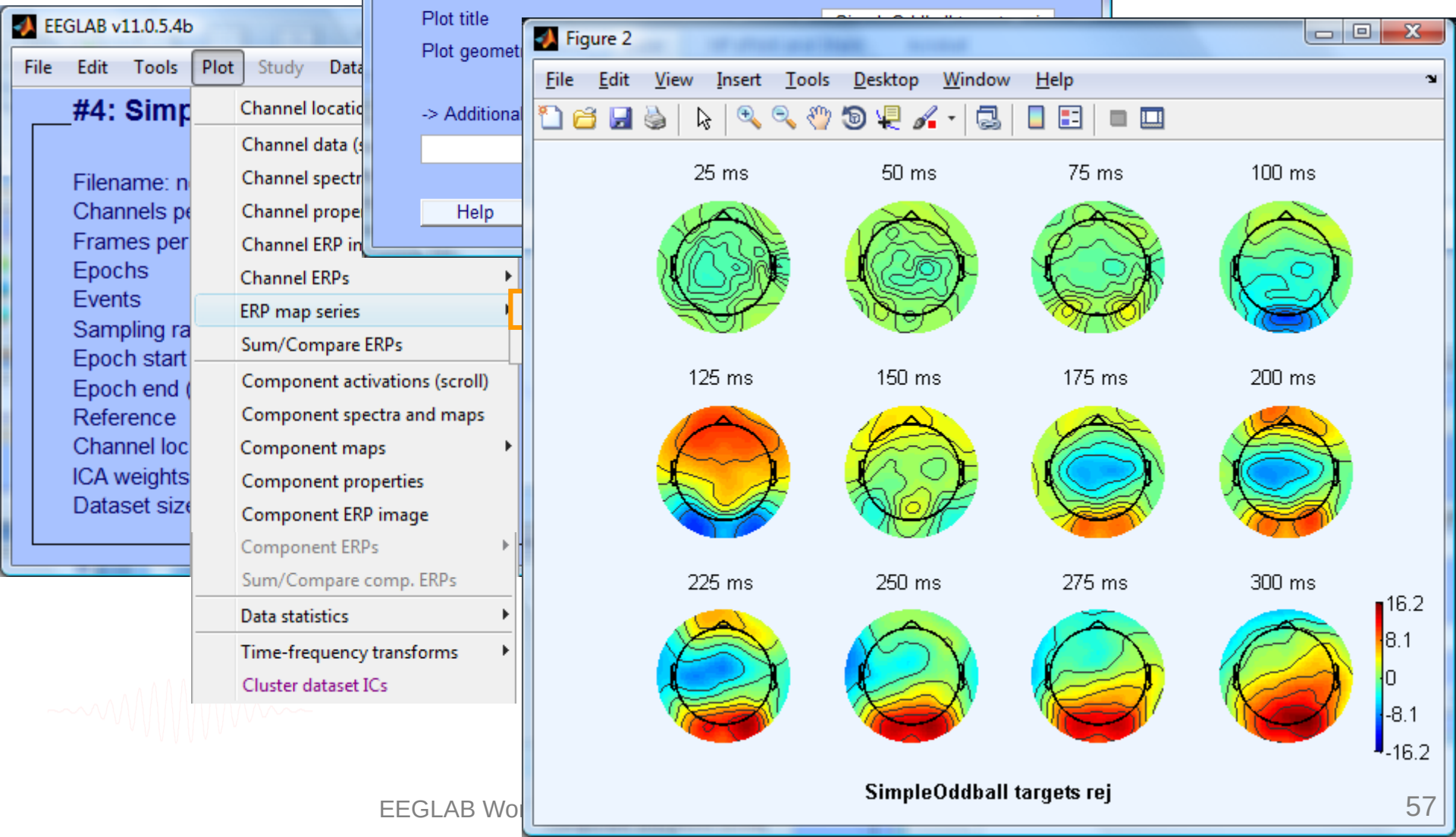




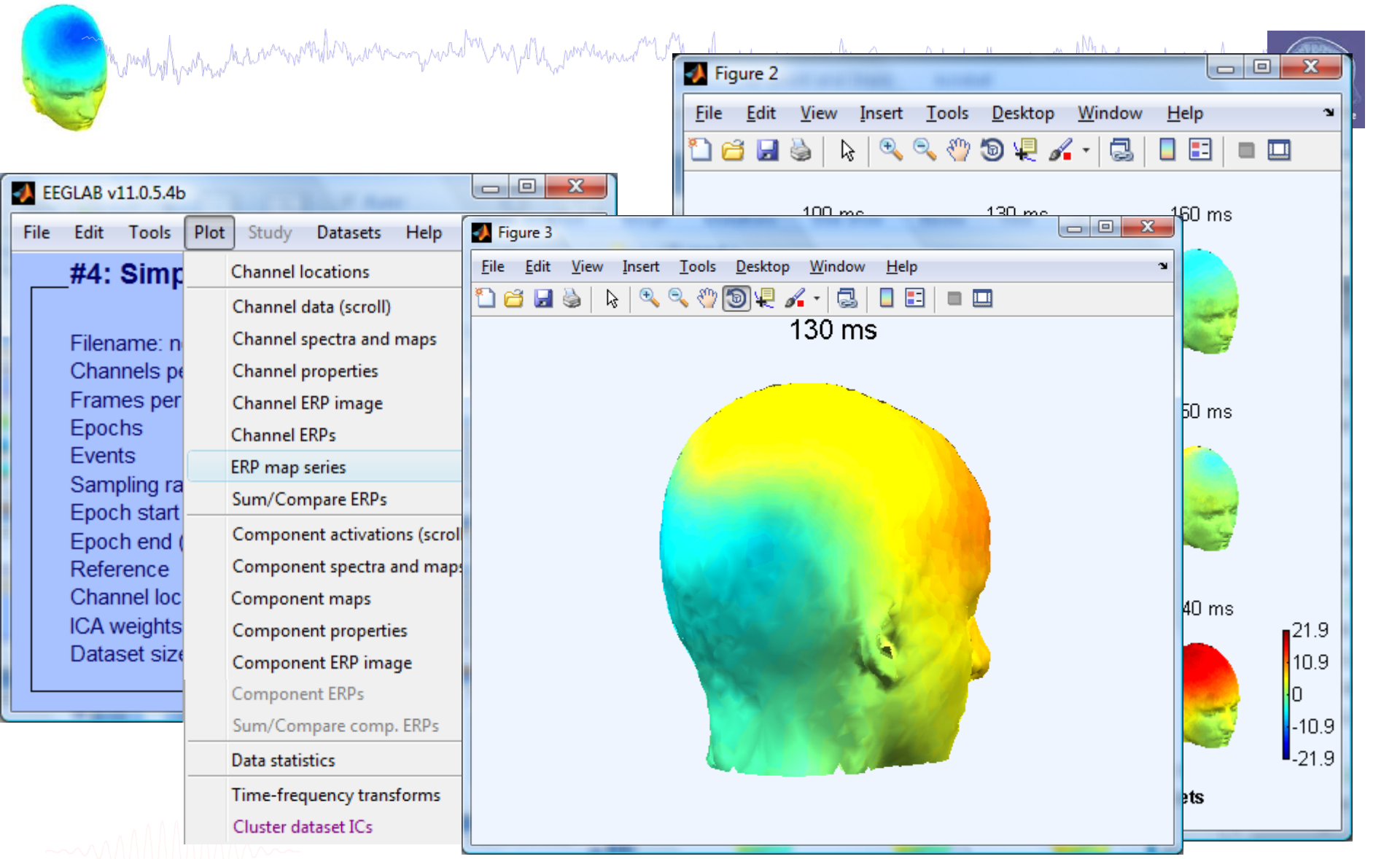


Visualize channel ERPs in 2D

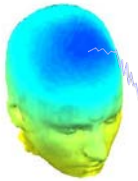




Visualize channel ERPs in 3D



Pre-processing pipeline



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

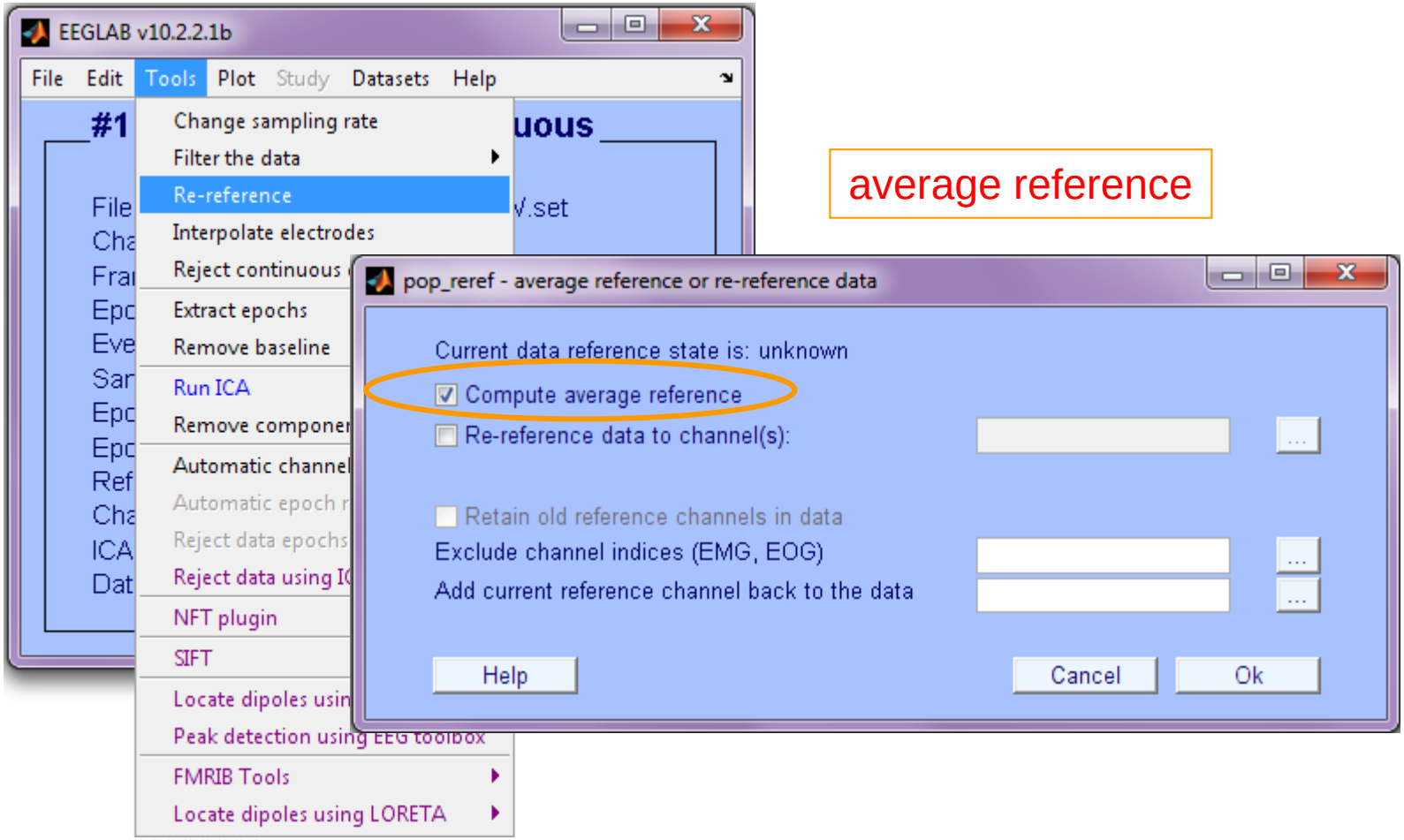
Examine raw data

Reject large artifact time points

Re-reference/ down-sample (if necessary)

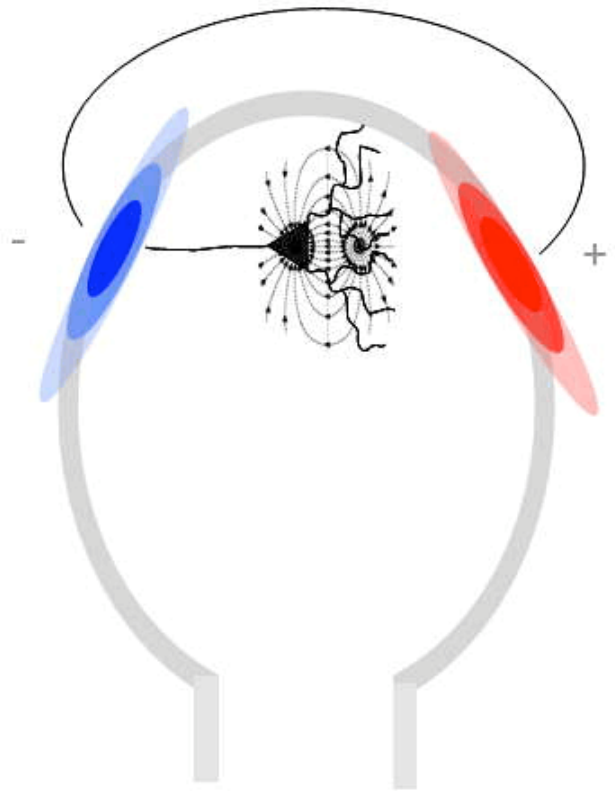
Run ICA

Re-reference data (if necessary/desired)



average reference

On Average Referencing

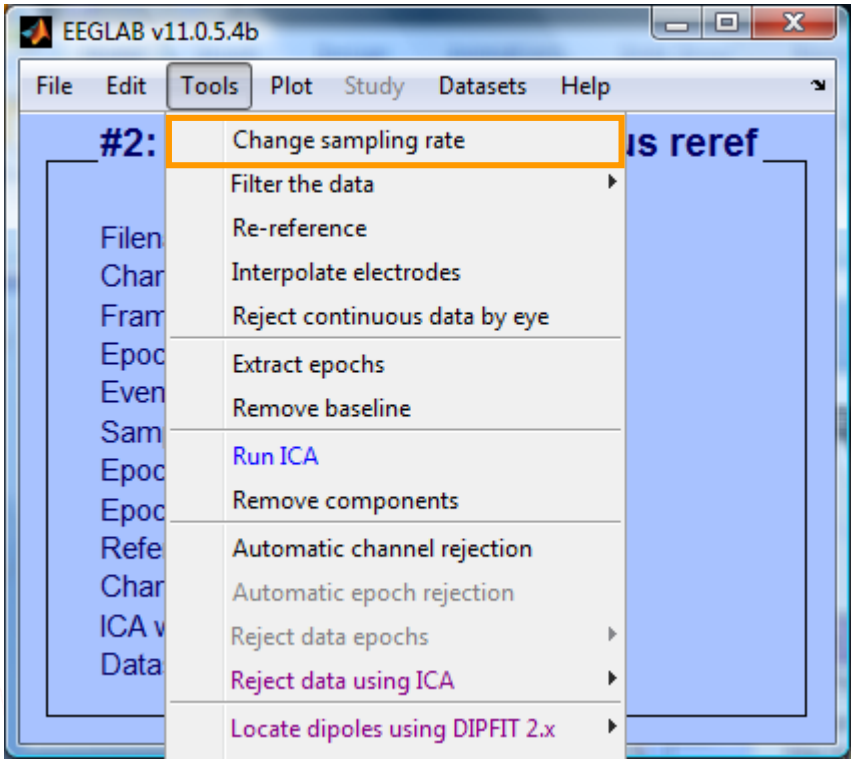


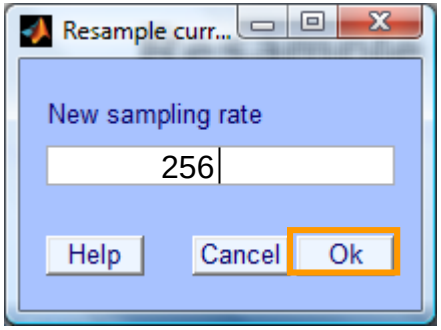
- In theory, positive and negative current across entire head should balance—no net current source or sink: Average referencing enforces this.
- ICA is invariant to re-referencing, except for
  - Effect of rank deficiency
  - DC difference.
- Average referencing reduces data rank by 1~~, so you must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

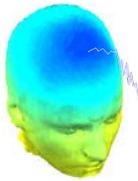
https://sccn.ucsd.edu/wiki/Makoto's_preprocessing_pipeline#Re-reference_the_data_to_average

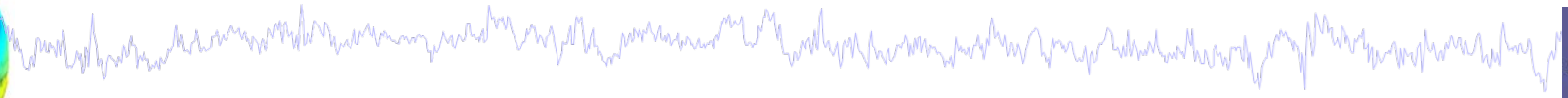
Resample data (if desired)

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





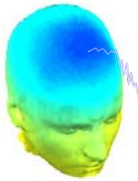




Swartz Center for Computational Neuroscience

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