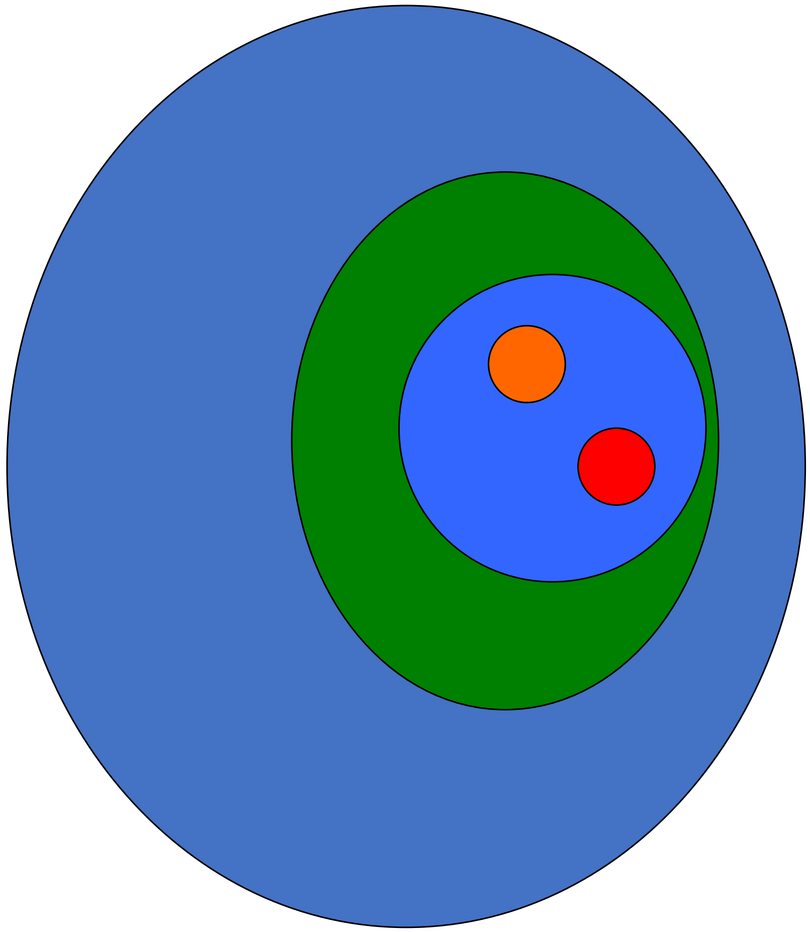
GLM in EEGLAB/LIMO

Arnaud Delorme
(with slides from C. Pernet)

The GLM Family



T-tests

Simple regression

ANOVA

Multiple regression

General linear model

- Mixed effects/hierarchical
- Timeseries models (e.g., autoregressive)
- Robust regression
- Penalized regression (LASSO, Ridge)

Generalized linear models

- Non-normal errors
- Binary/categorical outcomes (logistic regression)

One-step solution

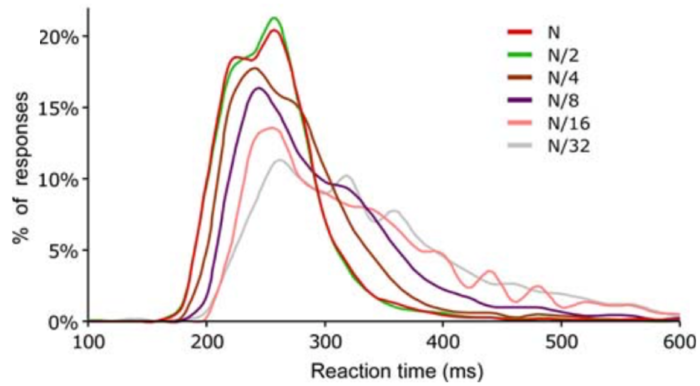
Iterative solutions (e.g., IWLS)

Tor Wager's slide

A regression is a linear model

Varying factor: Contrast of image

Outcome: Reaction time

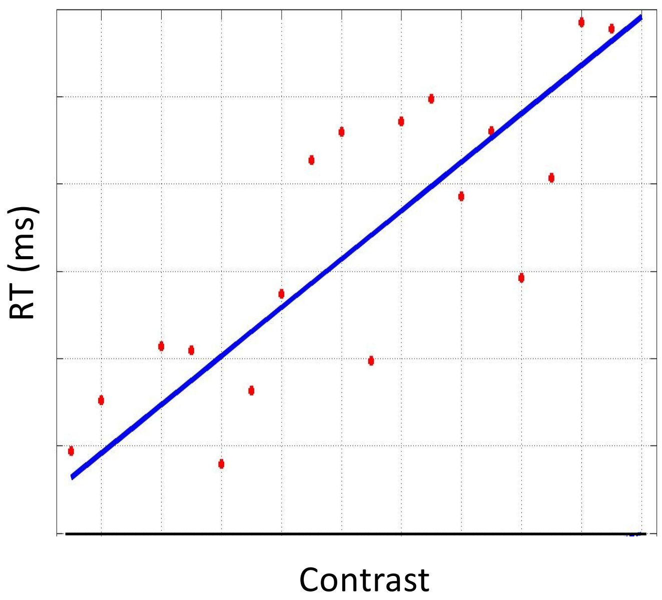


Mace, M., Delorme, A., Richard, G., Fabre-Thorpe, M. (2010) Spotting animals in natural scenes: efficiency of humans and monkeys at very low contrasts. *Animal Cognition*, 13(3):405-18.

A regression is a linear model

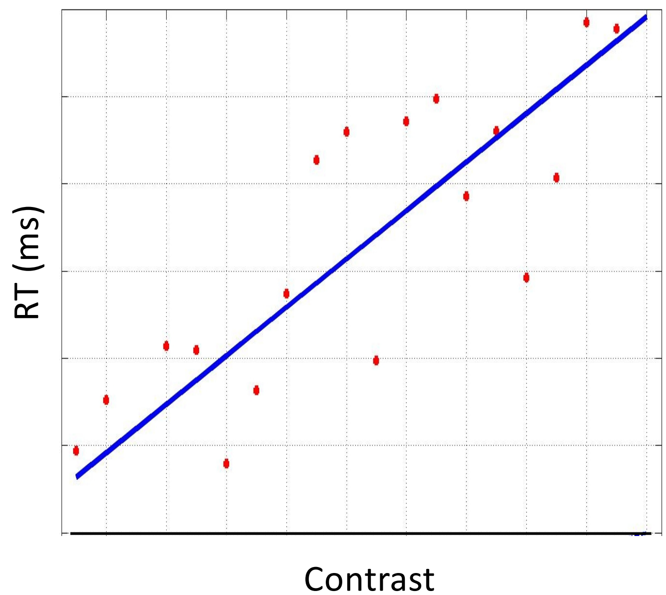
- We have an experimental measure x (e.g. contrast)

- We then do the expe and collect data RT (e.g. reaction time)



A regression is a linear model

- We have an experimental measure x (e.g. contrast)
- We then do the expe and collect data RT (e.g. reaction time)
- Model: RT = $\beta_0 + x\beta_1 + \varepsilon$



A regression is a linear model

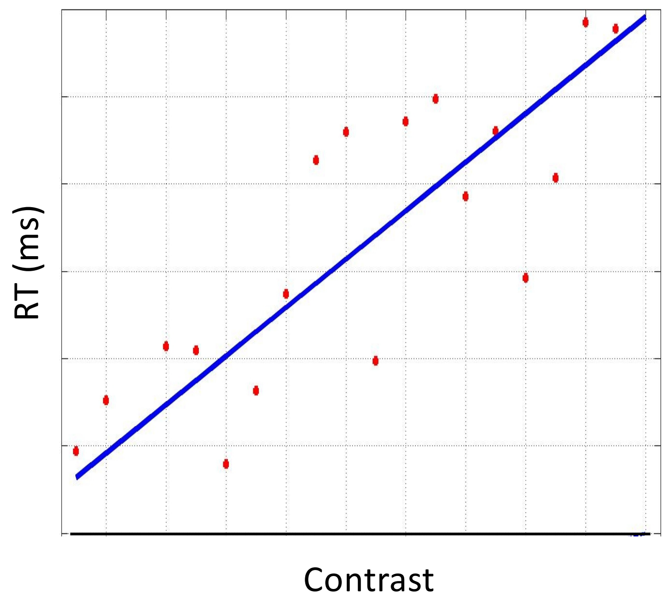
- We have an experimental measure x (e.g. contrast)

- We then do the expe and collect data RT (e.g. reaction time)

- Model: $RT = \beta_0 + x\beta_1 + \varepsilon$

- Do some maths / run a software to find β_1 and β_0

- $\hat{RT} = 23.6 + 2.7x$



A regression is a linear model

For each trial

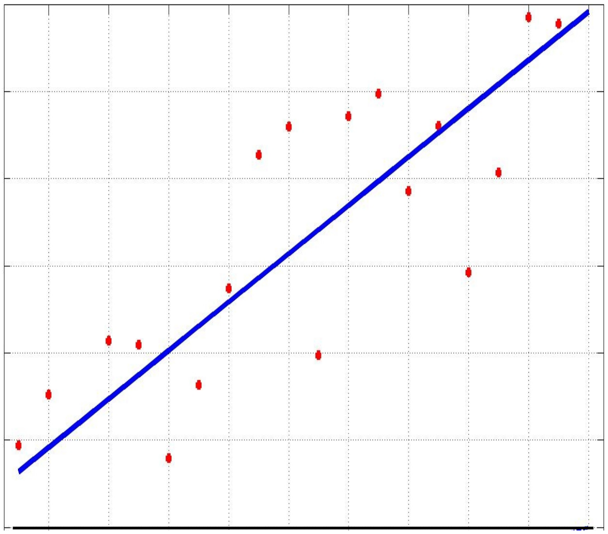
Contrast level

$RT_1 = \beta_0 + 10*\beta_1 + \varepsilon_1$

$RT_2 = \beta_0 + 5*\beta_1 + \varepsilon_2$

$RT_3 = \beta_0 + 7*\beta_1 + \varepsilon_3$

…



A regression is a linear model

For each trial

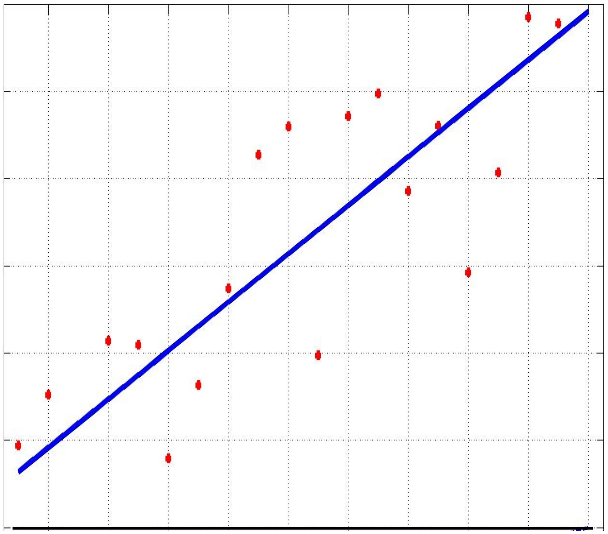
Contrast level

$RT_1 = \beta_0 + 10*\beta_1 + \varepsilon_1$

$RT_2 = \beta_0 + 5*\beta_1 + \varepsilon_2$

$RT_3 = \beta_0 + 7*\beta_1 + \varepsilon_3$

…



To test for significance compare the original regression model $RT_i = \beta_0 + c_i*\beta_1 + \varepsilon_i$ with the simplified model $RT_i = \beta_0 + \varepsilon_i$

Test if 0 included in confidence interval

Compare these errors

An ANOVA is a linear model

Varying factor: Type of image

Outcome: Reaction time (go/no-go)

Fishes **Birds** **Reptiles**

No-animal

Delorme, A., Richard, G., Fabre-Thorpe, M. (2010). Key visual features for rapid categorization of animals in natural scenes. *Frontier in psychology*, 1:21

$$RT_{i,j} = \beta_0 + \beta_i + \varepsilon_{i,j}$$

that is to say the data (e.g. RT) = a constant term (grand mean β_0) + the effect of a treatment (β_1 for fishes 1 and β_2, β_3 for birds and reptiles) and the error term ($\varepsilon_{i,j}$)

$$RT_{i,j} = \beta_0 + \beta_i + \varepsilon_{i,j}$$

that is to say the data (e.g. RT) = a constant term (grand mean β_0) + the effect of a treatment (β_1 for fishes 1 and β_2, β_3 for birds and reptiles) and the error term ($\varepsilon_{i,j}$)

For trial 4 (for example first trial of birds) we have

$$RT_{2,1} = \beta_0 + 0*\beta_1 + 1*\beta_2 + 0*\beta_3 + \varepsilon_{2,1}$$

$$RT_{i,j} = \beta_0 + \beta_i + \varepsilon_{i,j}$$

that is to say the data (e.g. RT) = a constant term (grand mean β_0) + the effect of a treatment (β_1 for fishes 1 and β_2, β_3 for birds and reptiles) and the error term ($\varepsilon_{i,j}$)

For trial 4 (for example first trial of birds) we have

$$RT_{2,1} = \beta_0 + 0*\beta_1 + 1*\beta_2 + 0*\beta_3 + \varepsilon_{2,1}$$

For trial 13 (for example second trial of birds) we have

$$RT_{2,2} = \beta_0 + 0*\beta_1 + 1*\beta_2 + 0*\beta_3 + \varepsilon_{2,2}$$

$$RT_{i,j} = \beta_0 + \beta_i + \varepsilon_{i,j}$$

that is to say the data (e.g. RT) = a constant term (grand mean β_0) + the effect of a treatment (β_1 for fishes 1 and β_2, β_3 for birds and reptiles) and the error term ($\varepsilon_{i,j}$)

For trial 4 (for example first trial of birds) we have

$$RT_{2,1} = \beta_0 + 0*\beta_1 + 1*\beta_2 + 0*\beta_3 + \varepsilon_{2,1}$$

For trial 13 (for example second trial of birds) we have

$$RT_{2,2} = \beta_0 + 0*\beta_1 + 1*\beta_2 + 0*\beta_3 + \varepsilon_{2,2}$$

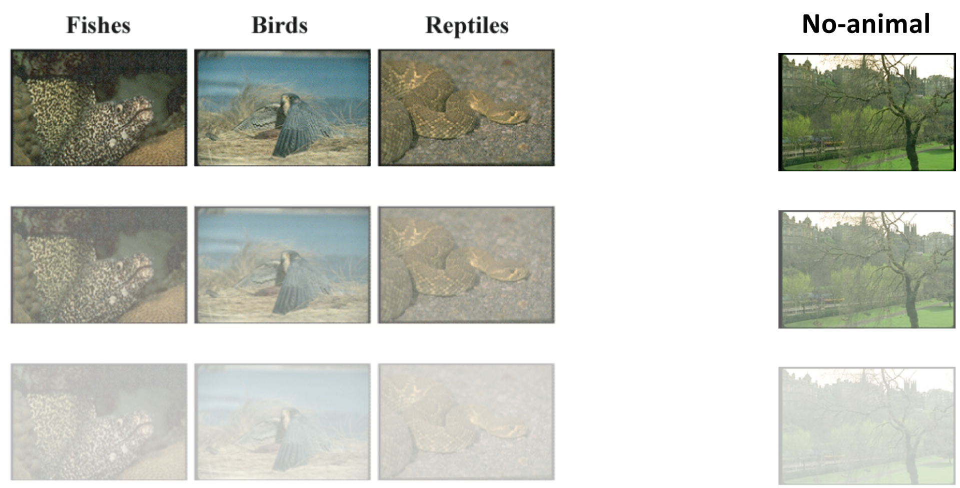
Statistics: if there is an effect of treatment then error of the simplified model $RT_{i,j} = \beta_0 + \varepsilon_{i,j}$ should be lower than the original model $RT_{i,j} = \beta_0 + \beta_i + \varepsilon_{i,j}$

Compare these errors

This is a GLM that is also equivalent to running an ANOVA

The GLM can do both
a Regression and an ANOVA (ANCOVA)

Varying factor: Type of image **AND** contrast
Outcome: Reaction time (go/no-go)



For example, for trial (first bird with contrast $c_{2,1}$) we have

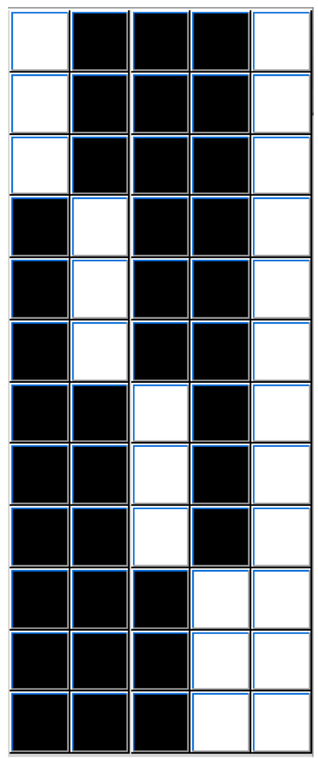
$$RT_{2,1} = \beta_0 + \underbrace{0*\beta_1 + 1*\beta_2 + 0*\beta_3 + 0*\beta_3}_{\text{Categorical var. ANOVA}} + \underbrace{c_{2,1}*\beta_4}_{\text{Continous var. REGRESSION}} + \varepsilon_{2,1}$$

The design matrix

Y	Gp
8	1
9	1
7	1
5	2
7	2
3	2
3	3
4	3
1	3
6	4
4	4
9	4

Design matrix

G_1 G_2 G_3 G_4 C



y(1..3) = 1xβ1+0xβ2+0xβ3+0xβ4+c+error

y(4..6) = 0xβ1+1xβ2+0xβ3+0xβ4+c+error

y(7..9) = 0xβ1+0xβ2+1xβ3+0xβ4+c+error

y(10..12) = 0xβ1+0xβ2+0xβ3+1xβ4+c+error

...

$$\begin{Bmatrix} 8 \\ 9 \\ 7 \\ 5 \\ 7 \\ 3 \\ 3 \\ 4 \\ 1 \\ 6 \\ 4 \\ 9 \end{Bmatrix} = \begin{Bmatrix} 10001 \\ 10001 \\ 10001 \\ 01001 \\ 01001 \\ 01001 \\ 00101 \\ 00101 \\ 00101 \\ 00011 \\ 00011 \\ 00011 \end{Bmatrix} * \begin{Bmatrix} \beta 1 \\ \beta 2 \\ \beta 3 \\ \beta 4 \\ c \end{Bmatrix} + \begin{Bmatrix} e1 \\ \vdots \\ e13 \end{Bmatrix}$$

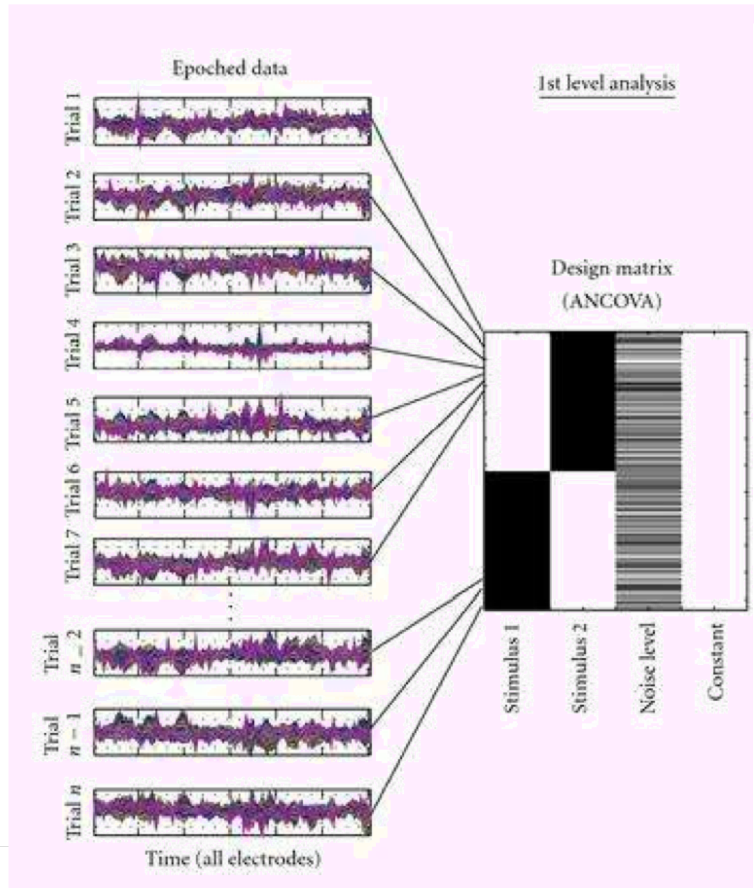
Measures

Model/ Design matrix

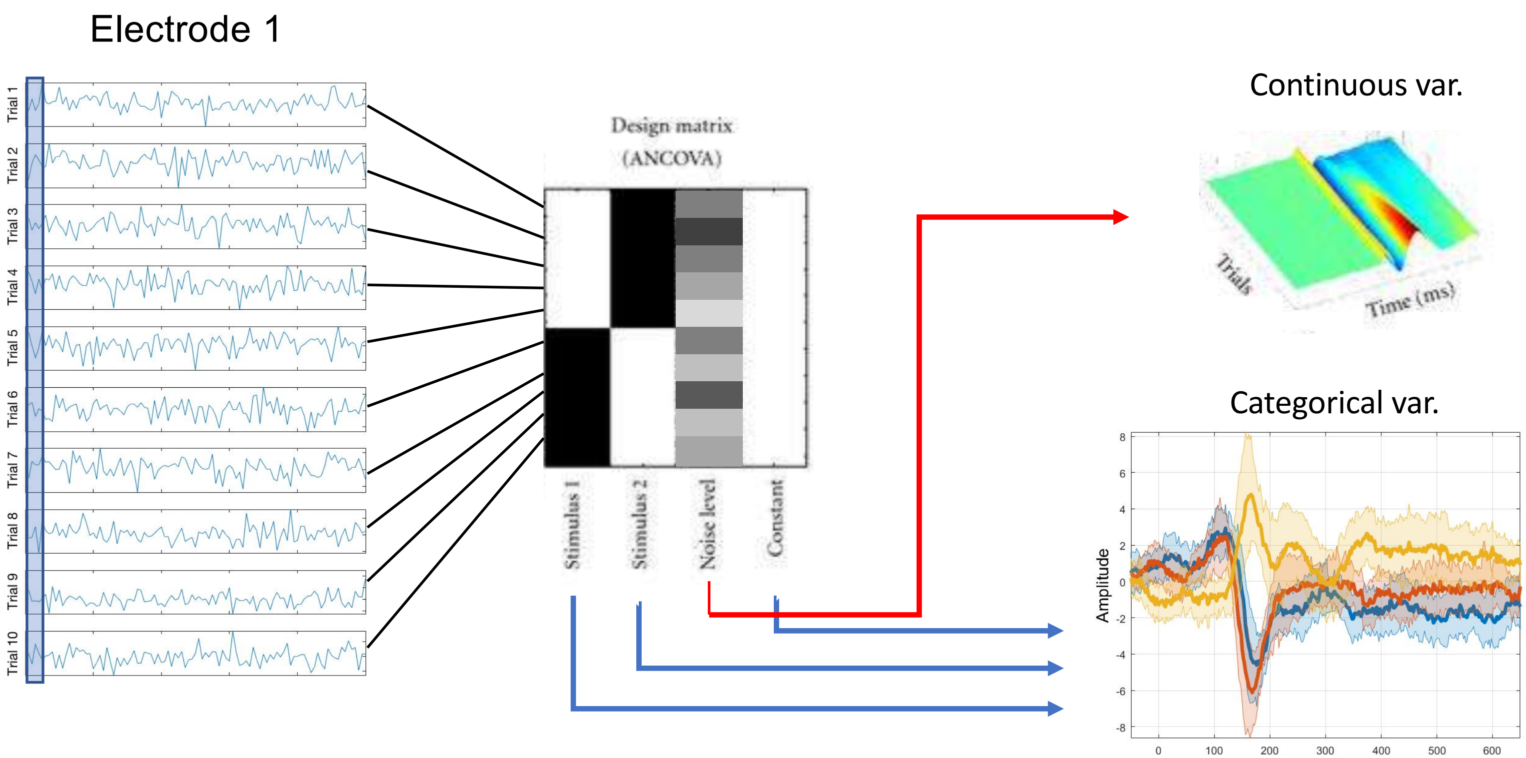
Unknown

Errors

Linear Modeling of EEG data



Linear Modeling of EEG data: level 1

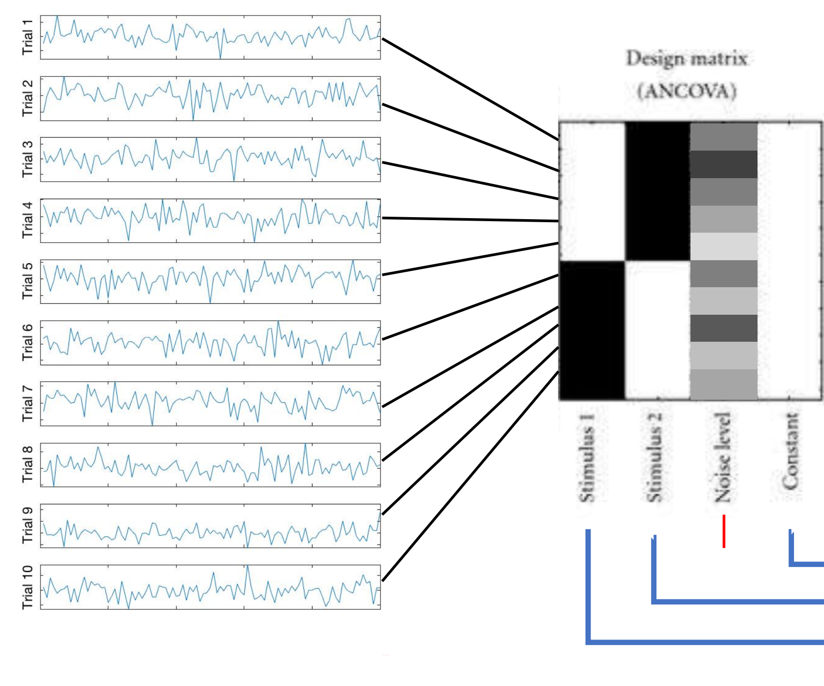


GLM: ordinary least square (OLS) versus weighted least square (WLS)

Significance: bootstrap trials to get confidence interval of beta parameters

Linear Modeling of EEG data: level 1

Electrode 1

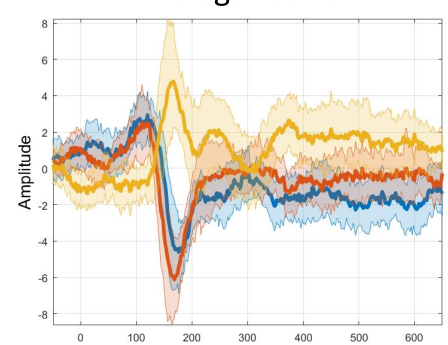


Hypotheses:

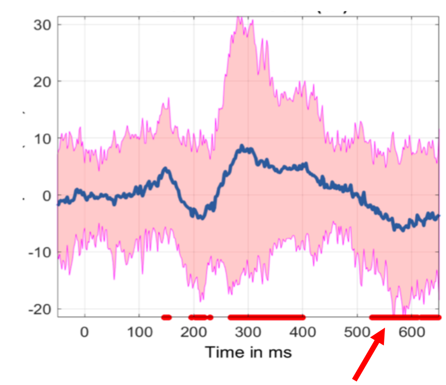
1. Effect of stimulus 1 -> is beta 1 significant
   (0 outside of beta1 confidence interval)

2. Difference between stimulus 1 and 2 (faces vs house)
   -> is beta 1 minus beta 2 significant
   (are the confidence intervals overlapping)

Categorical var.



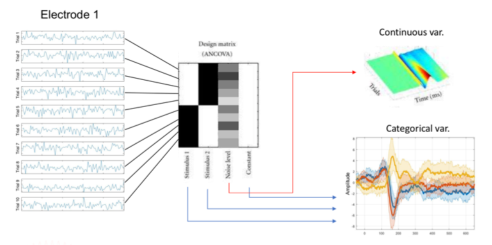
Electrode difference
Between conditions



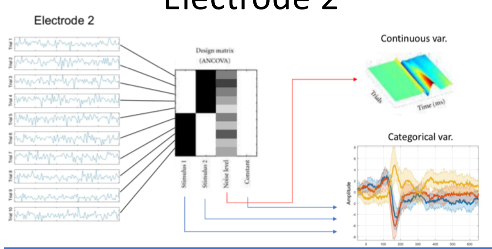
Significance based on beta params.

Linear Modeling of EEG data: 1st level

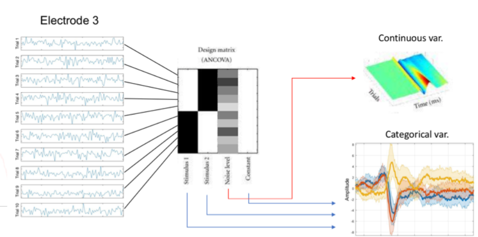
Electrode 1



Electrode 2

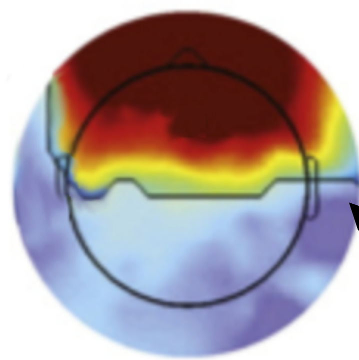


Electrode 3

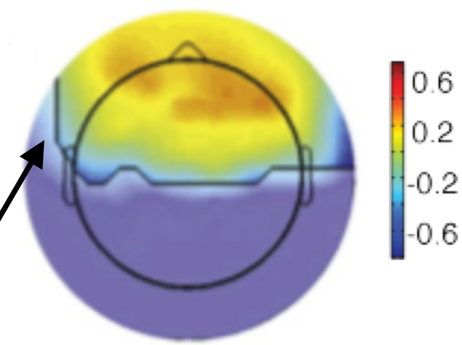


...

Scalp topography of **beta difference** at a given latency

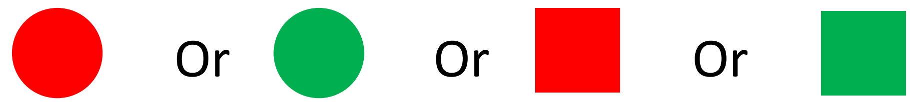


It is possible to plot the **potential difference** between condition at a given latency and assess significance using the beta difference

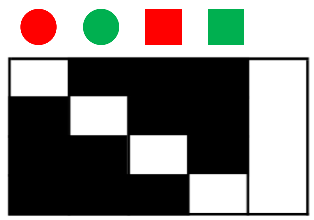


Limit of the regions masked for significance

Stimuli



1. Interaction design
(EEGLAB default)



Use beta as direct input into repeated measure ANOVA 2nd level to compute main effect and interaction effect (no need to build contrasts)

www.nature.com/scientificdata

SCIENTIFIC DATA

OPEN

SUBJECT CATEGORIES

» Electroencephalography
-EEG

» Brain imaging

A multi-subject, multi-modal human neuroimaging dataset

Daniel G. Wakeman[1,2] & Richard N. Henson[2]

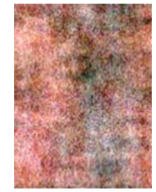
• *Scientific Data* **2**, Article number: 150001 (2015)
• doi:10.1038/sdata.2015.1
https://www.nature.com/articles/sdata20151

The Data

Famous

Unfamiliar

Scambled



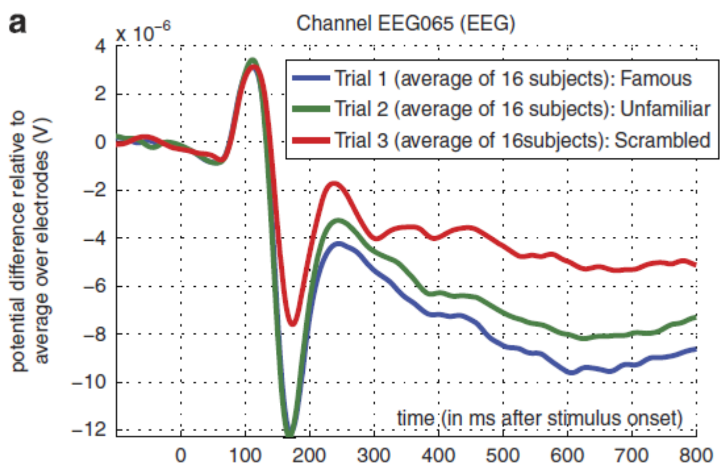
- 3 types of stimuli: Famous faces, Non-famous faces, Scrambled faces
- 3 levels of repetition: 1st time, 2nd time (right after), 3rd time (delayed)

→Priming experiment with a possible interaction with the type of stimuli.

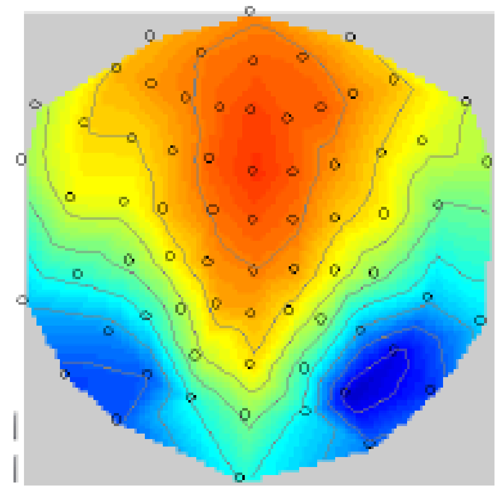
We need the conditions computed per subject (1st level) and then do the repeated measure ANOVA to test main effects and interactions.

What are we going to do?

- 1 – Replicate Henson et al. – faces vs. scrambled



Topography 170 ms



- 2 – learn about HLM, robust statistics and multiple comparison corrections

Preprocessing in EEGLAB

- Step 1. Raw data importation
- Step 2: Downsample the data
- Step 3: High-pass filter the data
- Step 4: Remove strong line noise
- Step 5: Detect and reject bad channels
- Step 6: Re-reference the scalp-channel data to average reference
- Step 7: Extract epochs centered on Famous, Unfamiliar, and Scrambled face presentations
- Step 8: Further clean the data by rejecting noisy epochs
- Step 9: Perform ICA decomposition
- Step 10: Select independent components
- Step 11: Fit equivalent current dipole models to components

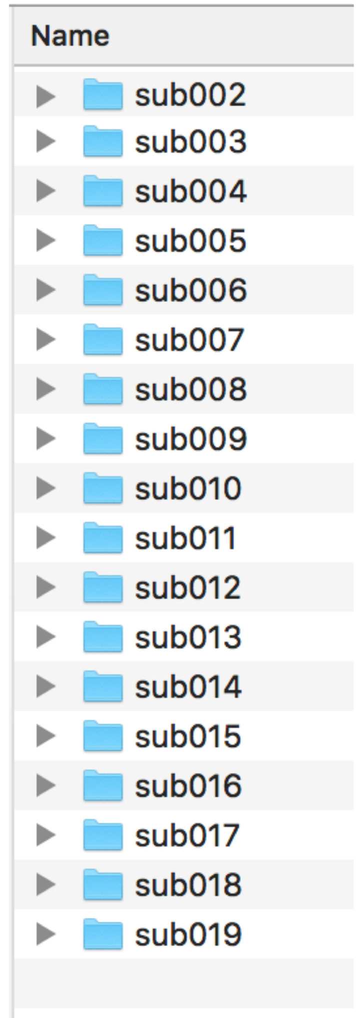
Assessing Event-Related EEG Brain Dynamics Using EEGLAB

Scott Makeig, Ramon Martinez-Cancino, Makoto Miyakoshi, Zeynep Akalin Acar, Luca Pion-Tonachini, John Iversen, Cyril Pernet, Arnaud Delorme

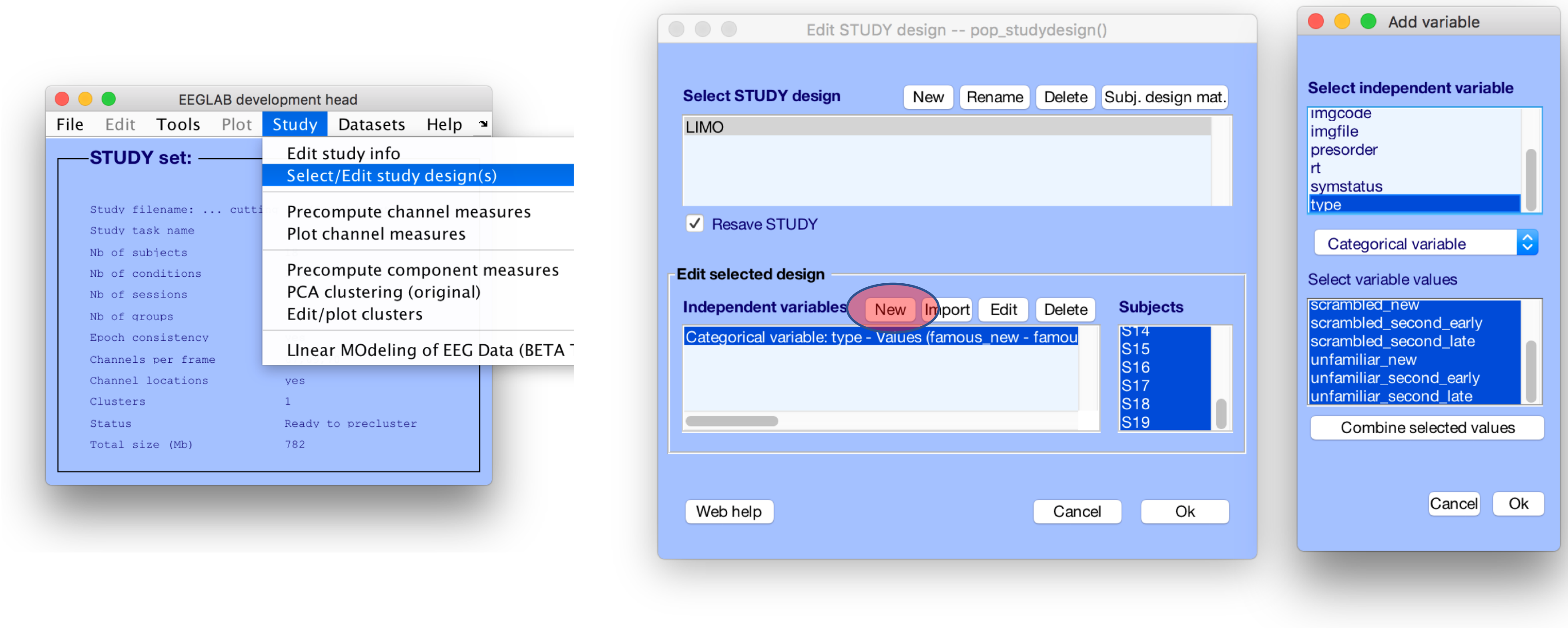
In preparation for a special issue of Frontiers in Neuroimaging methods

Let's get started

- Open Matlab
- Start eeglab
- Move to the folder containing the data

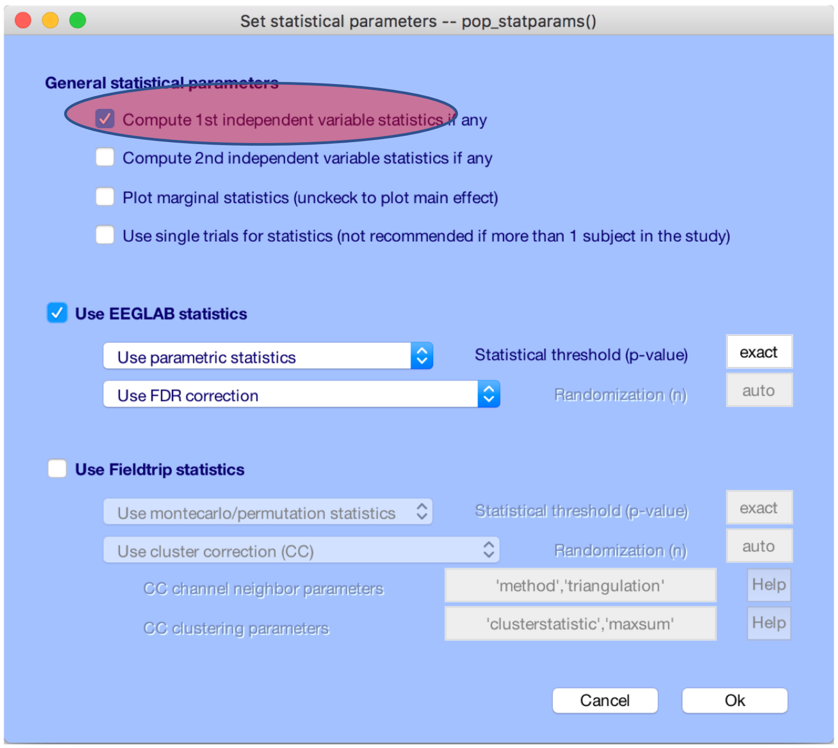


Create study designs

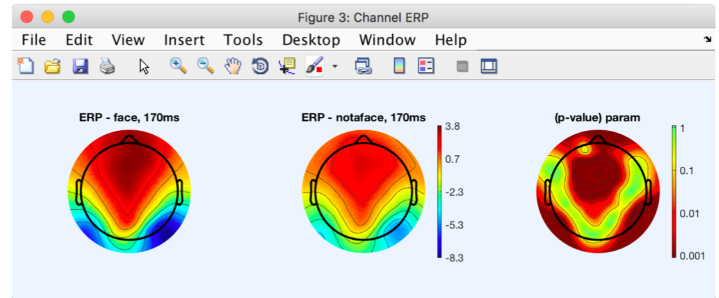


Here, we pick the 'type' and select all 9 conditions
(events tagged during preprocessing appear here)

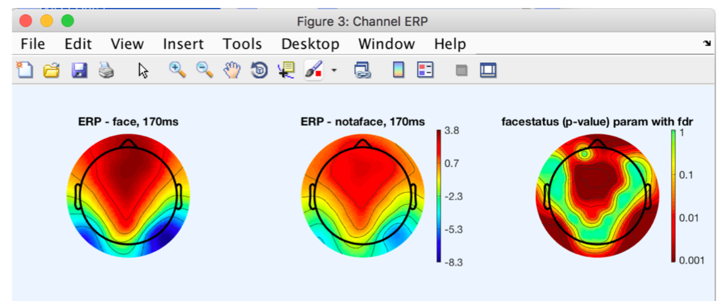
Standard EEGLAB statistics



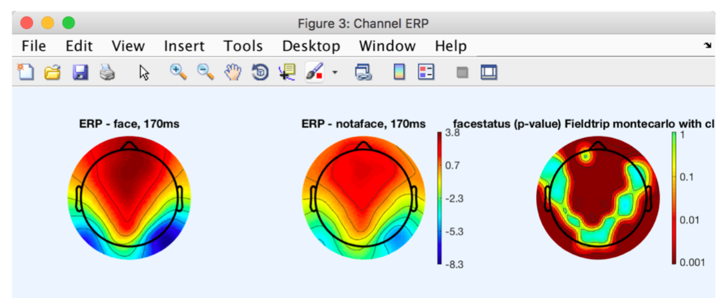
Uncorrected



FDR corrected



Cluster corrected



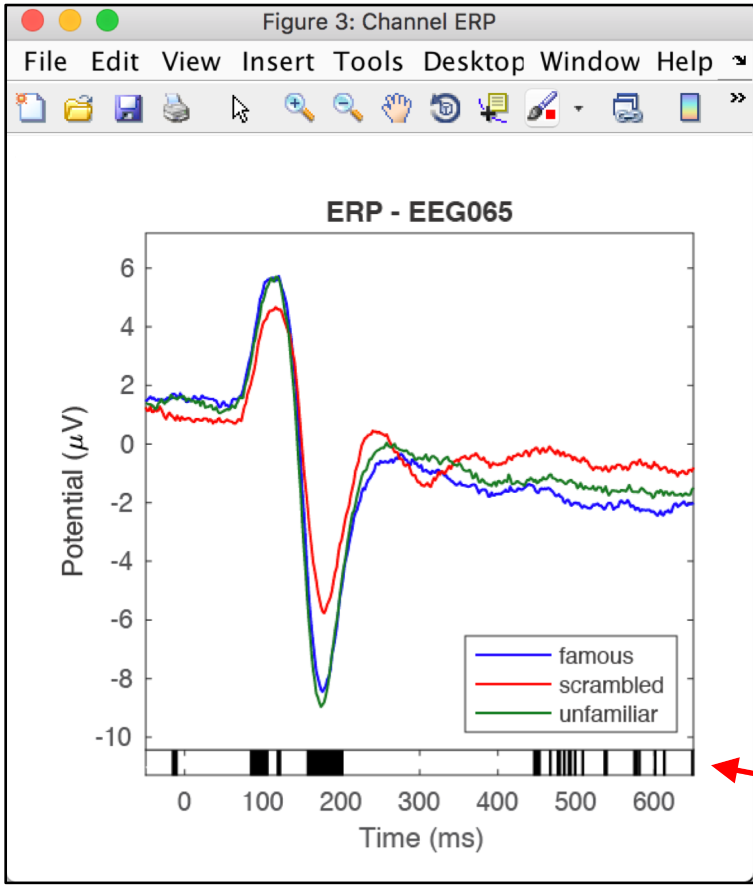
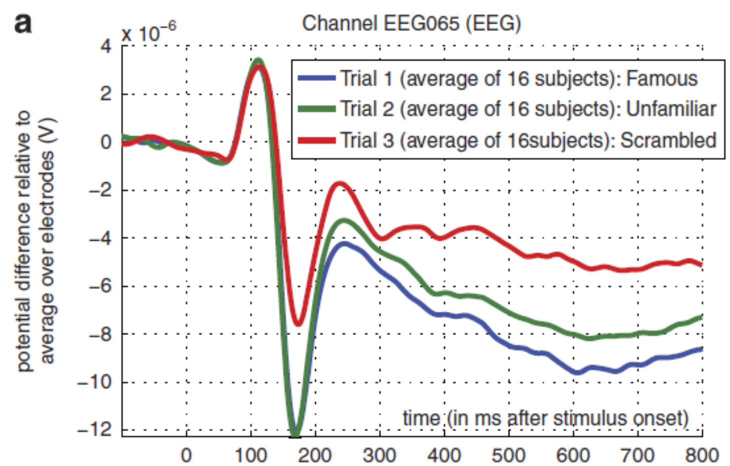


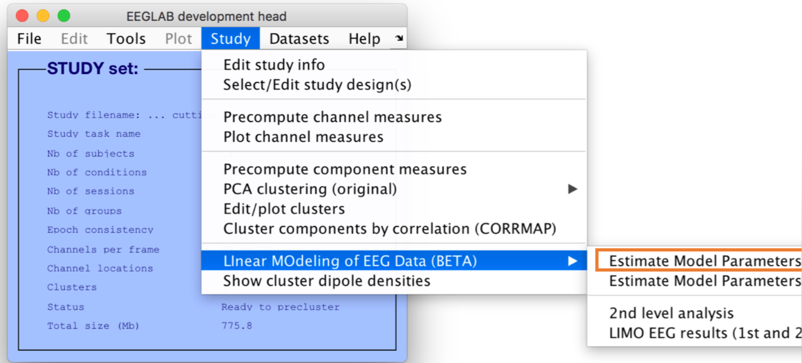
Figure 3: Channel ERP
File Edit View Insert Tools Desktop Window Help
ERP - EEG065
Potential (μV)
Time (ms)
famous
scrambled
unfamiliar

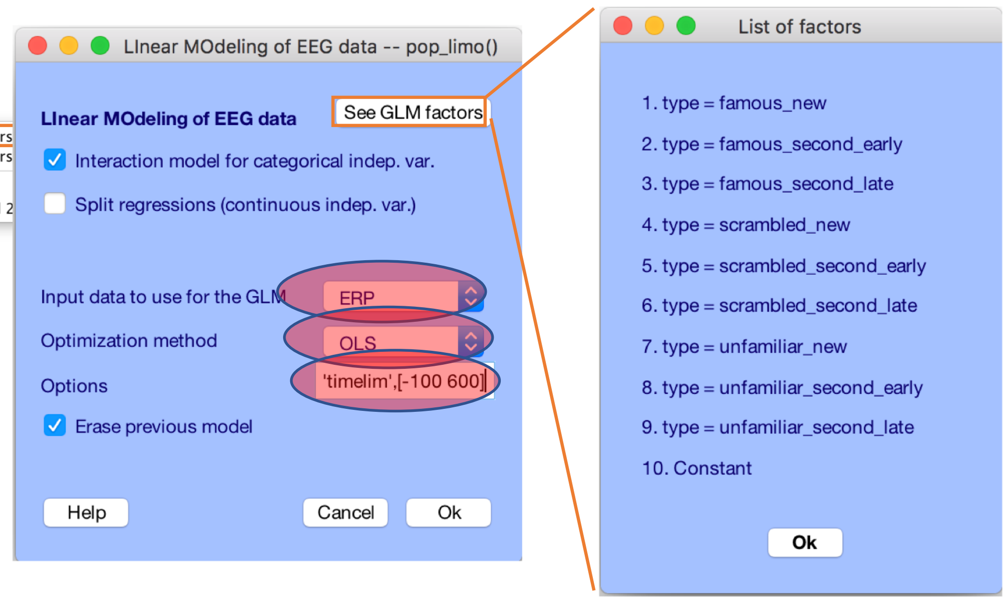
p<0.001



a
Channel EEG065 (EEG)
x 10^{-6}
potential difference relative to average over electrodes (V)
Trial 1 (average of 16 subjects): Famous
Trial 2 (average of 16 subjects): Unfamiliar
Trial 3 (average of 16subjects): Scrambled
time (in ms after stimulus onset)

Estimate model parameters





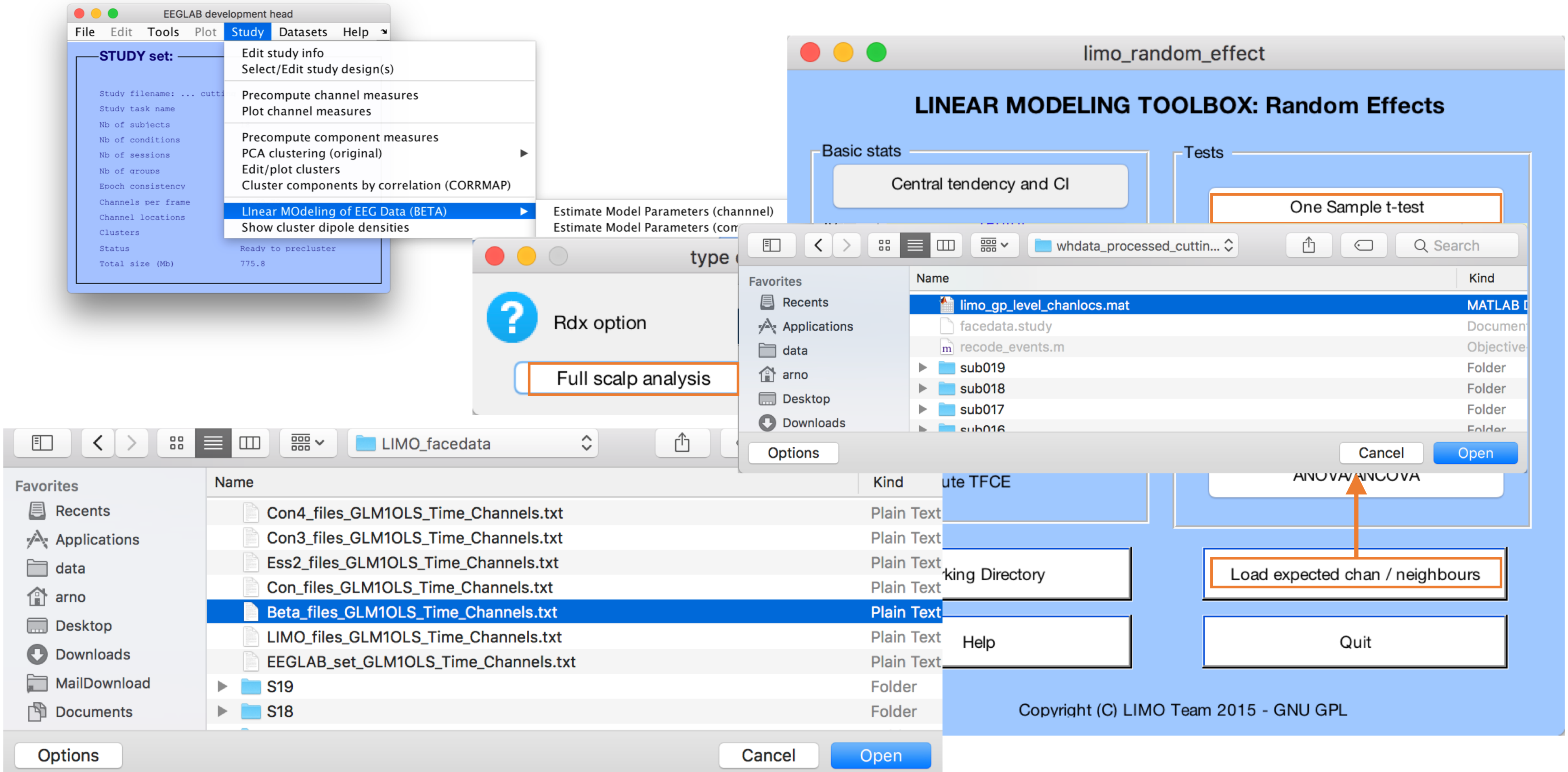
All designs are pooled in the GLM

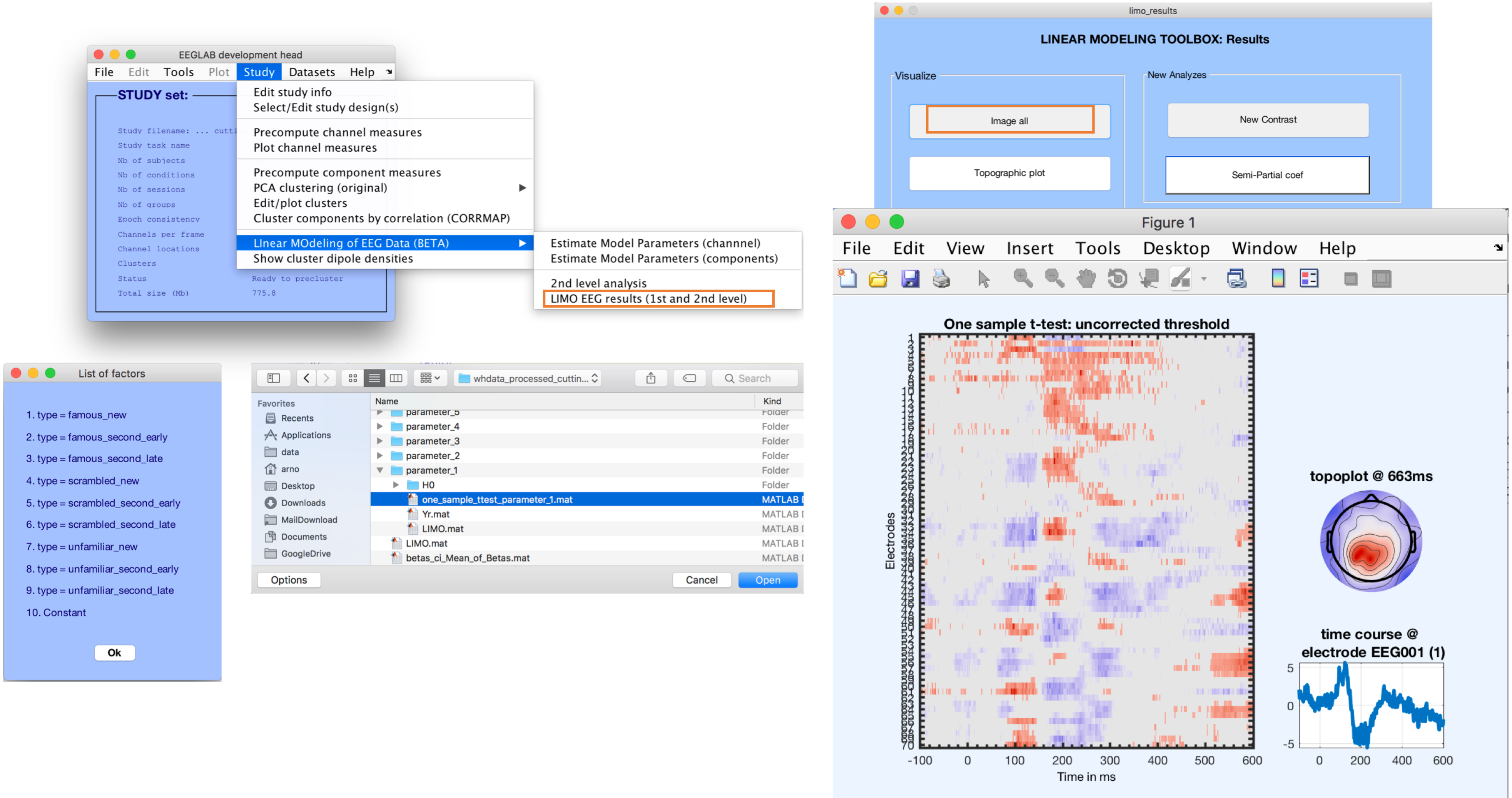
Estimate Model Parameter

Have generated single trials, specified the model, we now do the stats

→ Restrict ‘timelim’ [-50 650]

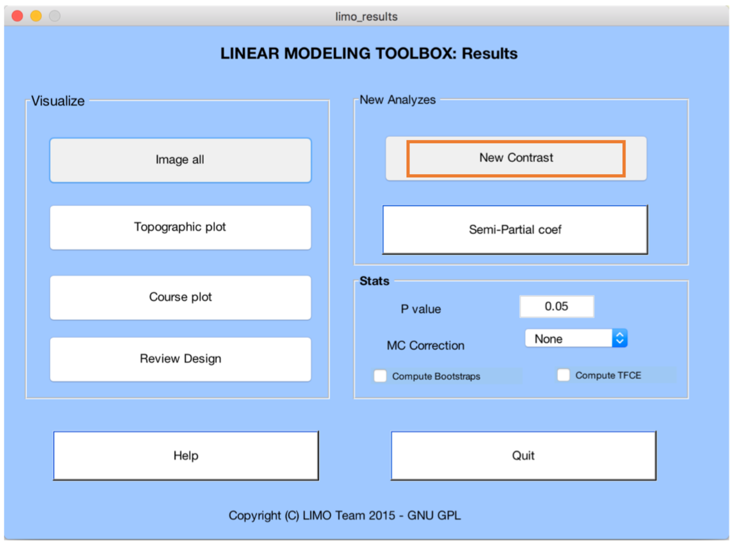
Are Beta significant?

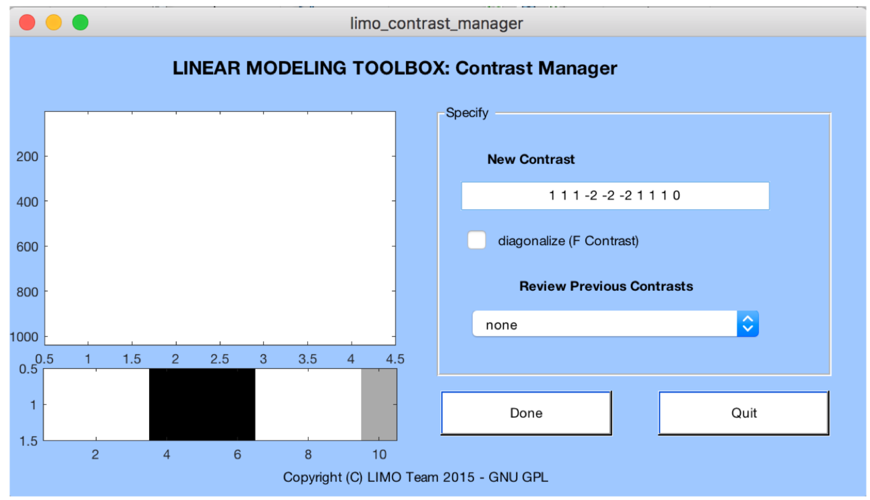




EEGLAB development head
File Edit Tools Plot Study Datasets Help
STUDY set:
Edit study info
Select/Edit study design(s)
Precompute channel measures
Plot channel measures
Precompute component measures
PCA clustering (original)
Edit/plot clusters
Cluster components by correlation (CORRMAP)
Linear MOdeling of EEG Data (BETA)
Show cluster dipole densities
Estimate Model Parameters (channnel)
Estimate Model Parameters (components)
2nd level analysis
LIMO EEG results (1st and 2nd level)
Ready to precluster
775.8
limo_results
LINEAR MODELING TOOLBOX: Results
Visualize
New Analyzes
Image all
New Contrast
Topographic plot
Semi-Partial coef
List of factors
1. type = famous_new
2. type = famous_second_early
3. type = famous_second_late
4. type = scrambled_new
5. type = scrambled_second_early
6. type = scrambled_second_late
7. type = unfamiliar_new
8. type = unfamiliar_second_early
9. type = unfamiliar_second_late
10. Constant
Ok
whdata_processed_cuttin...
Search
Favorites
Recents
Applications
data
arno
Desktop
Downloads
MailDownload
Documents
GoogleDrive
Name
Kind
parameter_4
parameter_3
parameter_2
parameter_1
H0
one_sample_ttest_parameter_1.mat
Yr.mat
LIMO.mat
LIMO.mat
betas_ci_Mean_of_Betas.mat
Folder
MATLAB
Options
Cancel
Open
Figure 1
File Edit View Insert Tools Desktop Window Help
One sample t-test: uncorrected threshold
Electrodes
-100 0 100 200 300 400 500 600
Time in ms
topoplot @ 663ms
time course @
electrode EEG001 (1)
5
0
-5
0 200 400 600

Grouping betas and differences between conditions





Faces vs non-faces	Famous	Scrambled	Unfamiliar	ANOVA (famous/scambled/unfamiliar)			ANOVA (new/early/late)		
1	1	0	0	1	0	0	1	0	0
1	1	0	0	1	0	0	0	1	0
1	1	0	0	1	0	0	0	0	1
-2	0	1	0	0	1	0	1	0	0
-2	0	1	0	0	1	0	0	1	0
-2	0	1	0	0	1	0	0	0	1
1	0	0	1	0	0	1	1	0	0
1	0	0	1	0	0	1	0	1	0
1	0	0	1	0	0	1	0	0	1
0	0	0	0	0	0	0	0	0	0

ANOVA (famous/scambled/unfamiliar)