Robust Linear Modelling of EEG data: the LIMO EEG plug-in

Cyril Pernet, PhD

Edinburgh Imaging &
Centre for Clinical Brain Sciences

Copy the LIMOEEG folder and move limo_egg to plugins

Overview

- What is LIMO EEG for?
- Plug-in overview
- Theory and practice

➢ 1st level robust GLM

➢ 2nd level robust stats

➢ Design brainstorming

Accounting for within and between subjects variance

WHAT IS IT FOR?

Full 'brain' analysis

- Traditionally compute averages per condition and do your statistics on peaks

→Peaks are NOT necessarily important events (they are likely to mark the end of a process)

→Increase type 1 FWER by choosing electrodes

→Averages don't account for trial variability

→Fixed effect can be biased

→Design flexibility

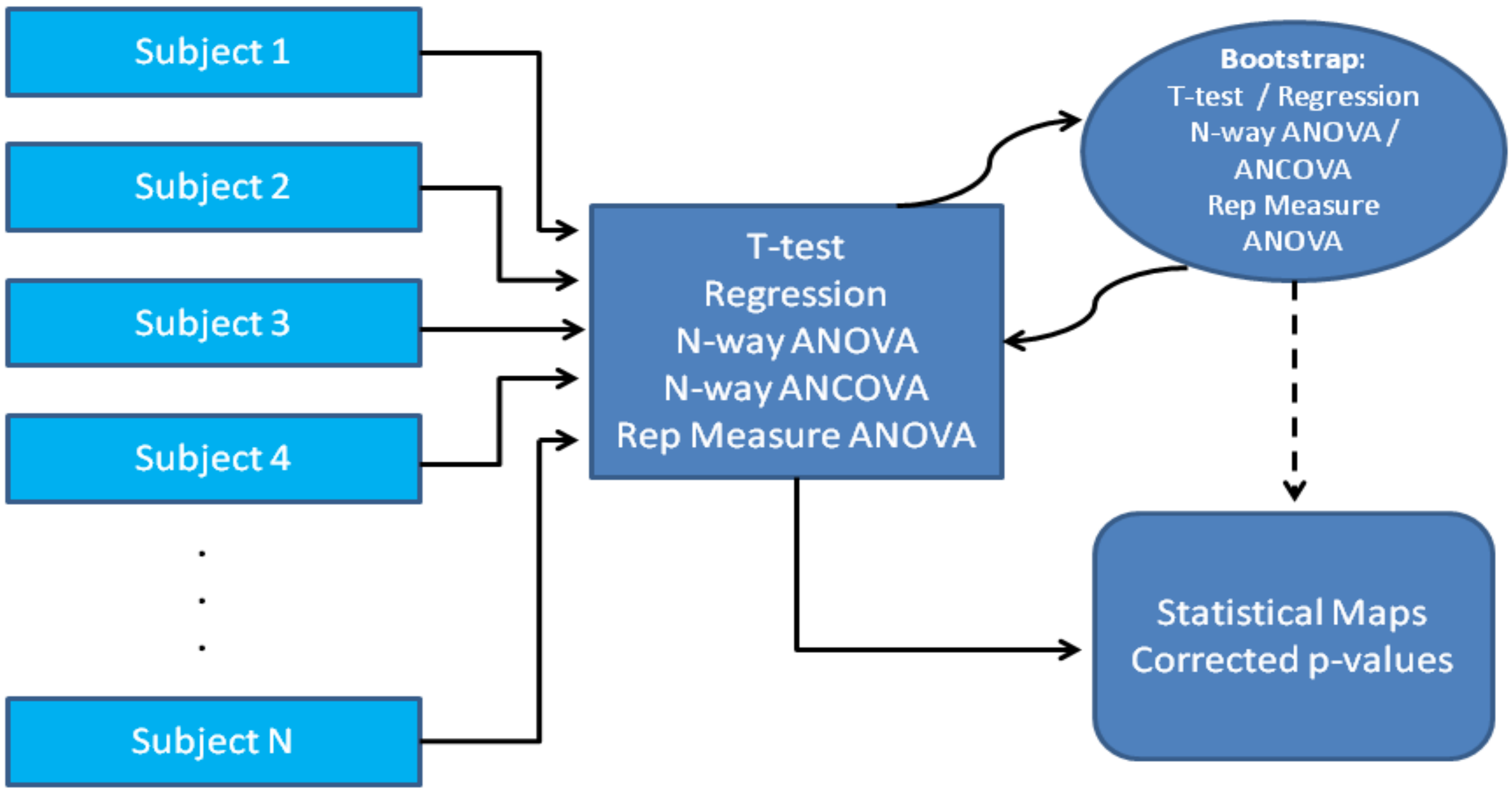
Rousselet & Pernet – It's time to up the Game Front. Psychol., 23 May 2011

Hierarchical Linear Model

<u>1st level analysis:</u>
GLM: Y=Xβ+ε
→1 β per column of X
(= within subject effects)

<u>2nd level analysis:</u>
Robust stats (Yuen t-tests, robust GLM, robust Hotelling T^2)

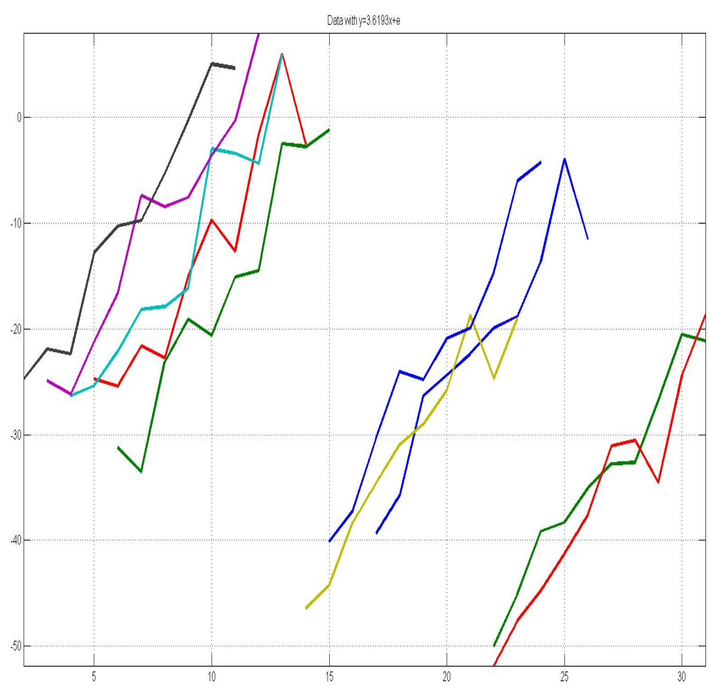
<u>Multiple Comnparison Correction:</u>
Max, Cluster-Mass, TFCE



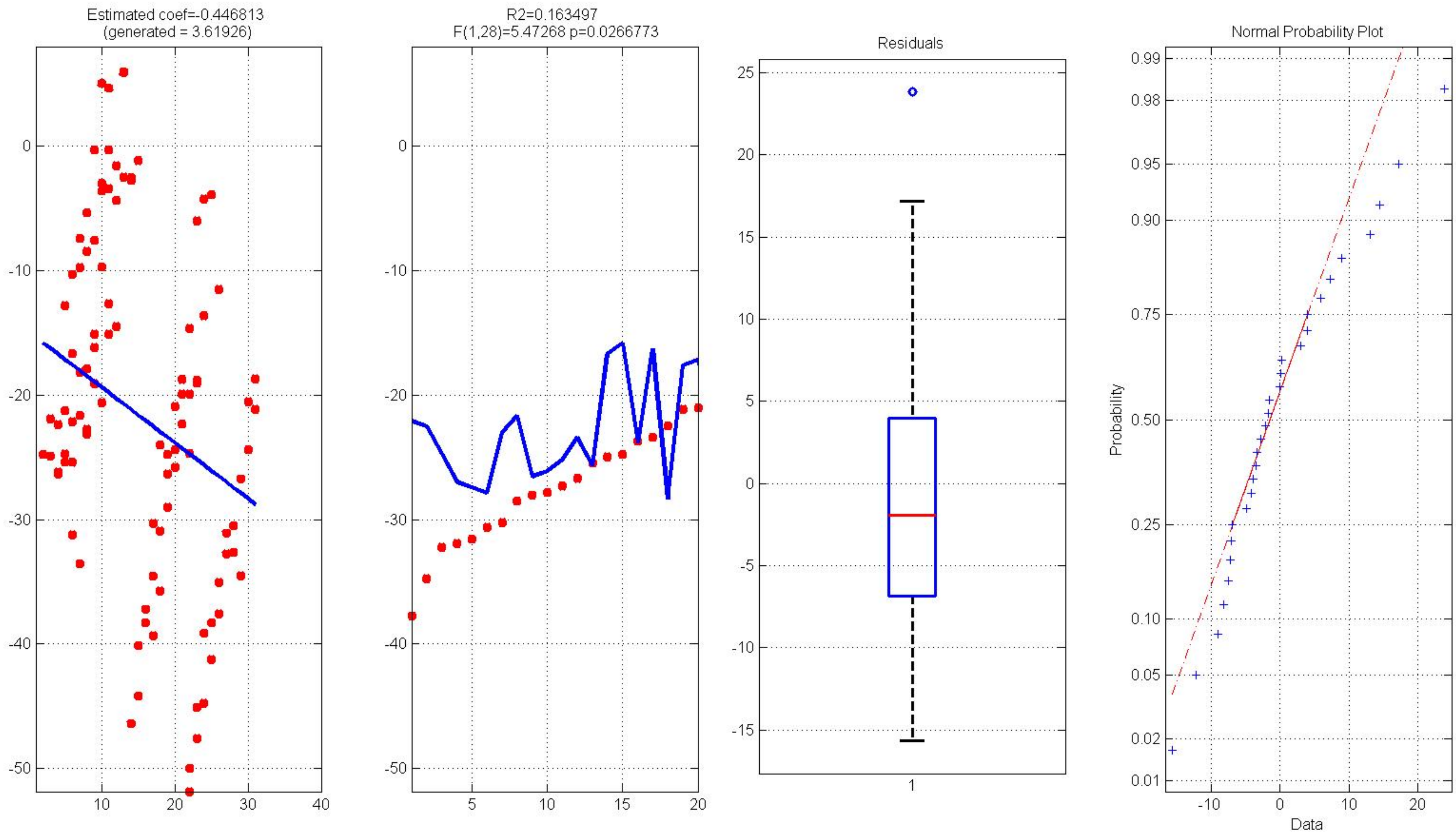
Random Effect Model

Model the data with fixed effects (the experimental conditions) and a random effect (subjects are allowed to have different overall values – considering subjects as a random variable)

<u>Example</u>: present stimuli from intensity -5 units to +5 units around the subject perceptual threshold and measure RT
→ Plot the data per intensity

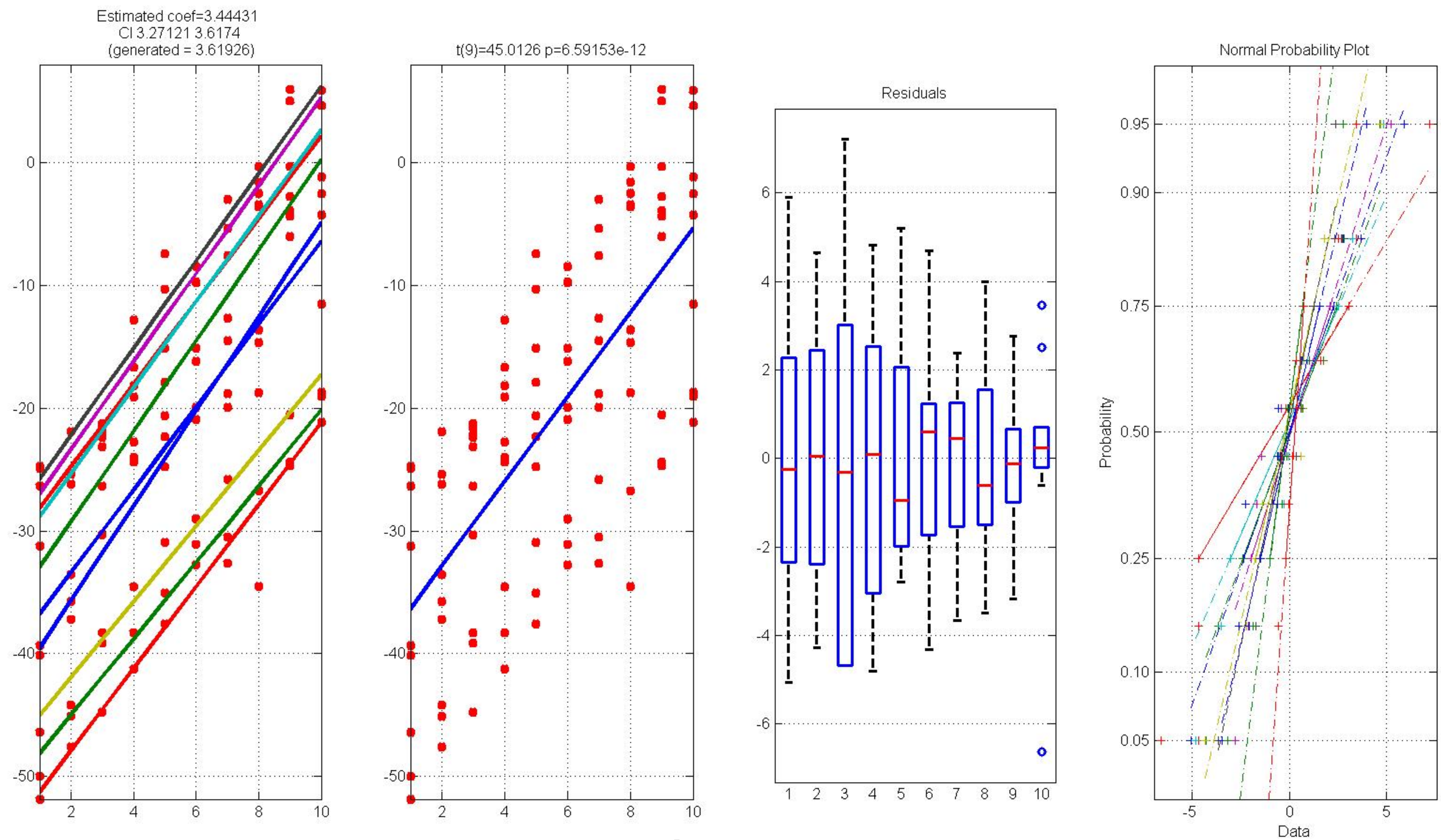


Random Effect Model



Fixed effect = average across subjects → negative correlation?

Random Effect Model

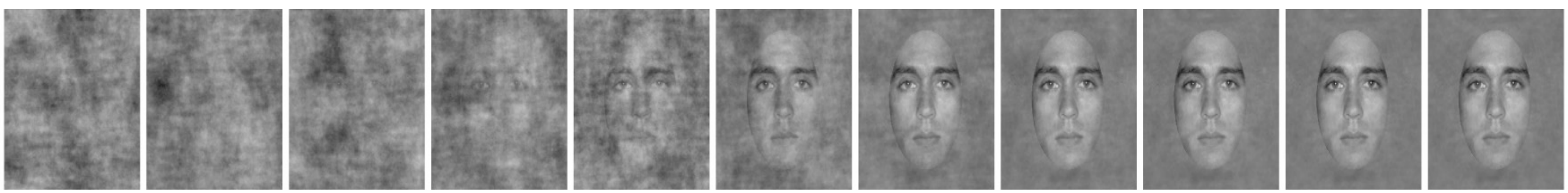


Fixed effect = average across subjects → negative correlation?
Mixed effect = effect per subject with variable (random) offsets

Design considerations

- ALLOWS YOU TO ANALYZE ANY (PAIRED / UNPAIRED) DESIGNS

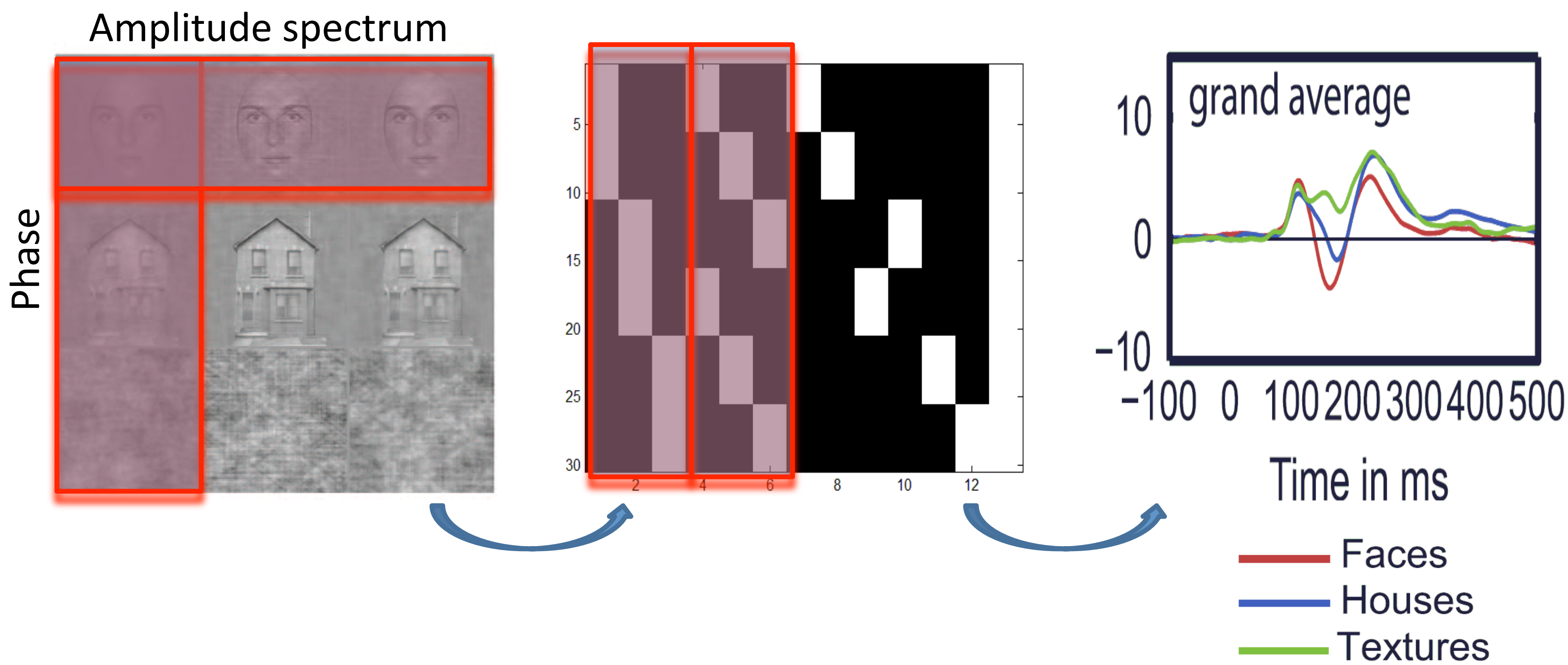
Illustration with a set of studies looking at the effect of stimulus phase information



Rousselet, Pernet, Bennet, Sekuler (2008). Face phase processing. *BMC Neuroscience 9:98*

Factorial Designs: N*N*N*…

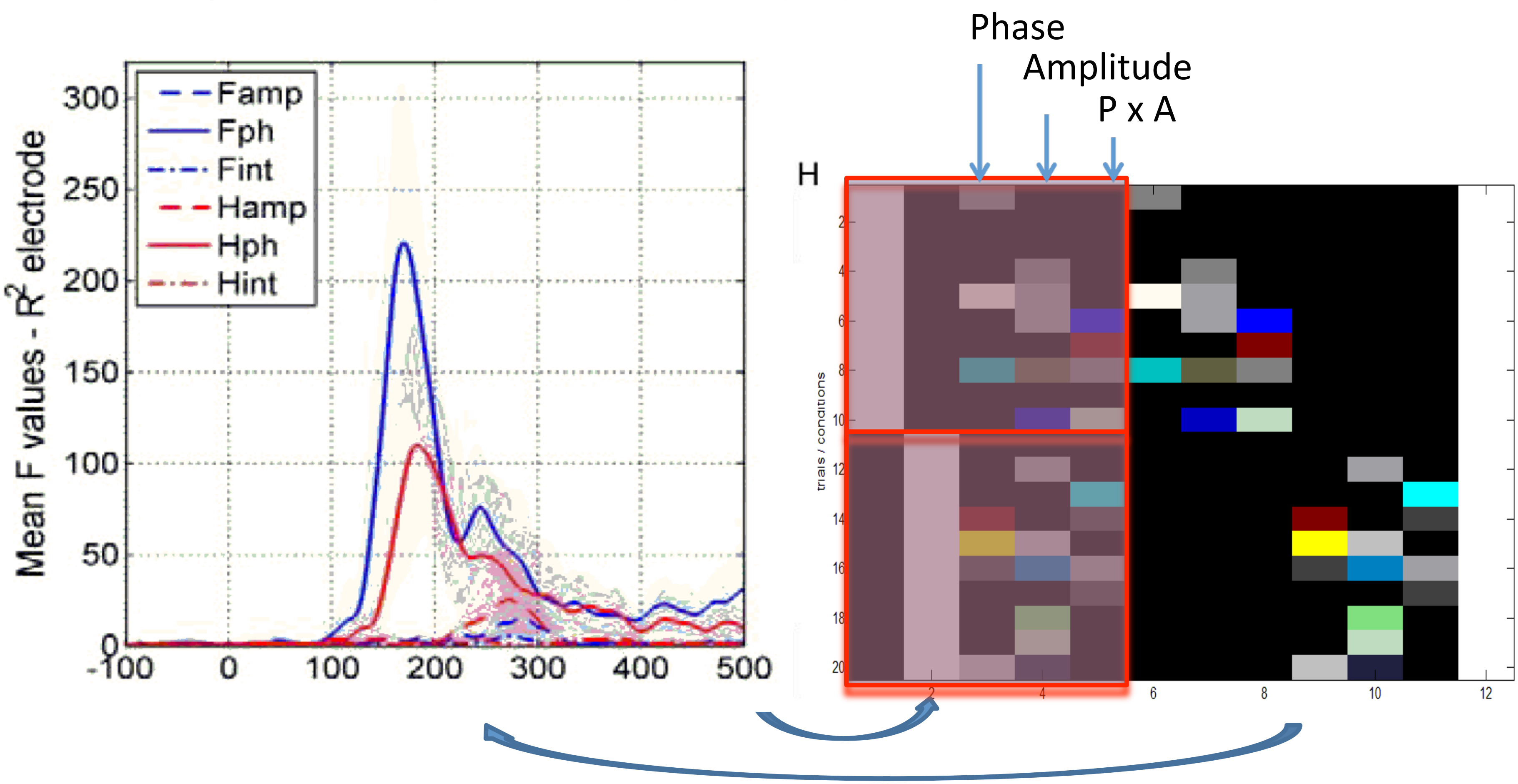
Categorical designs: Group level analyses of course but also Individual analyses with bootstrap



[Bienek, Pernet, Rousselet (2012). Phase vs Amplitude Spectrum. Journal of Vision 12(13), 1–24]()

Regression based designs

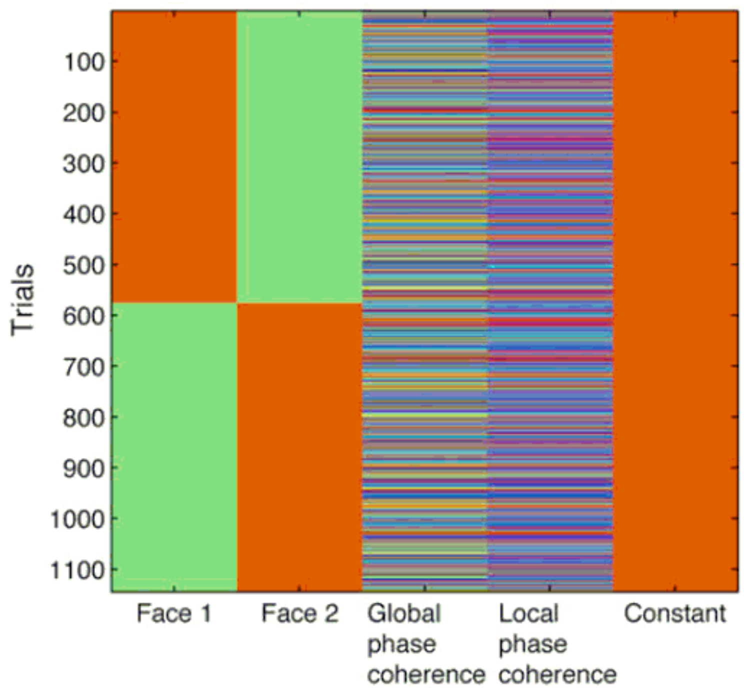
Mixed design: Control of low level physical properties

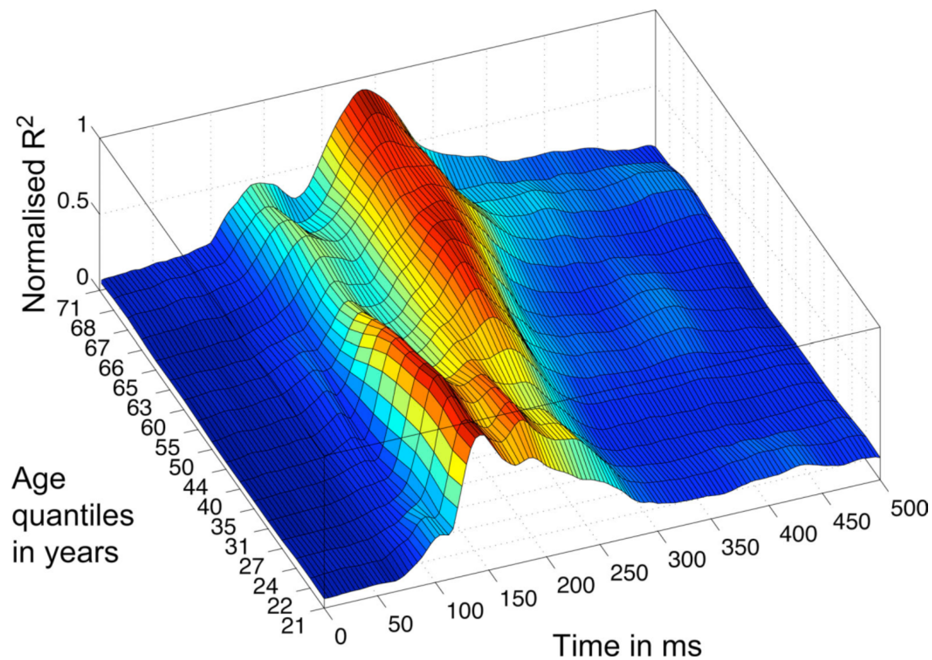


[Bienek, Pernet, Rousselet (2012). Phase vs Amplitude Spectrum. Journal of Vision 12(13), 1–24]

Regression based designs (2 levels)

Parametric designs:
study the effect of stimulus properties within subjects
effect of aging between subjects





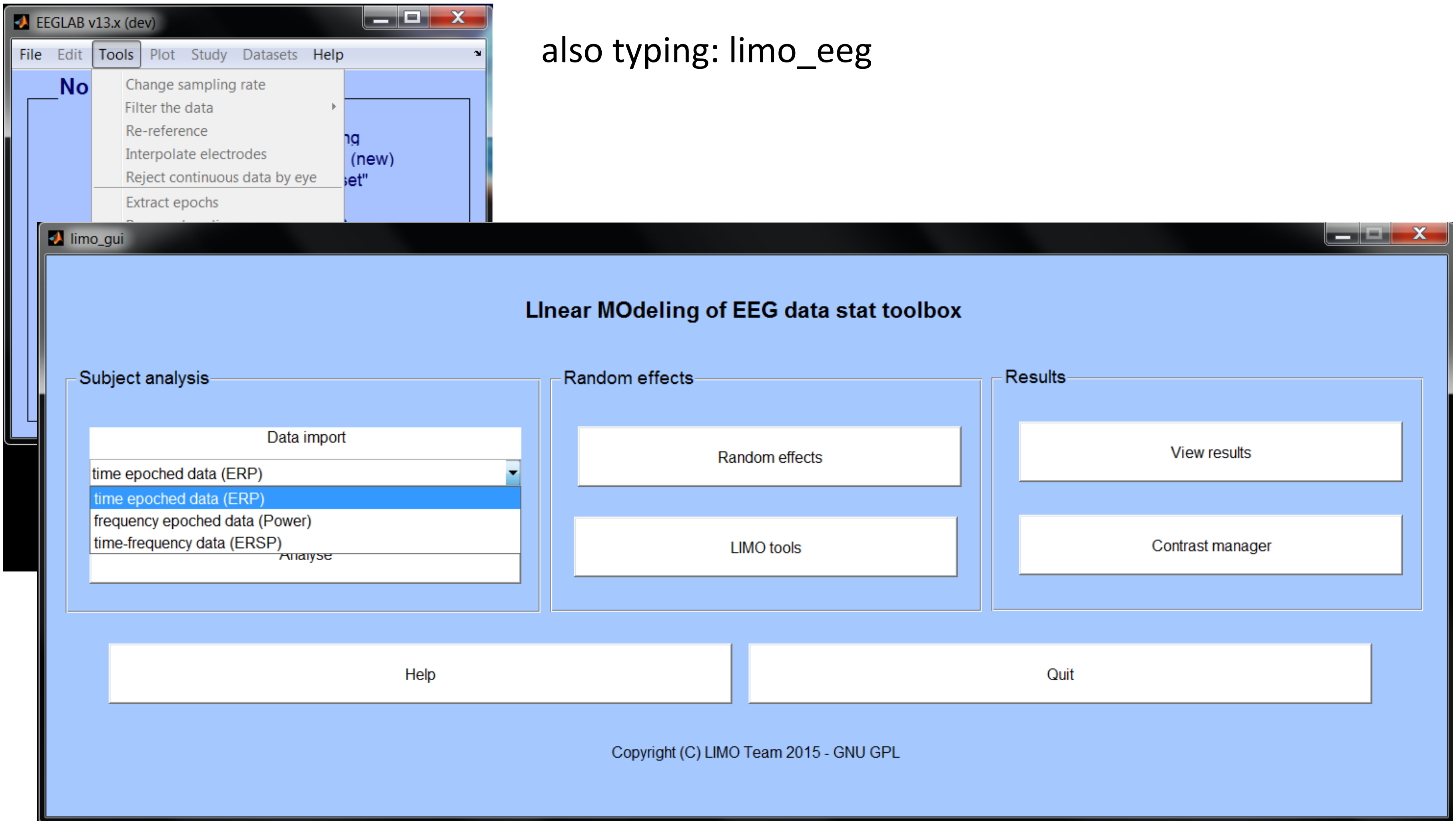
[Rousselet, Gaspar, Pernet, Husk, Bennett, Sekuler (2010). Aging and face perception. Front Psy]

Quick tour of the interface

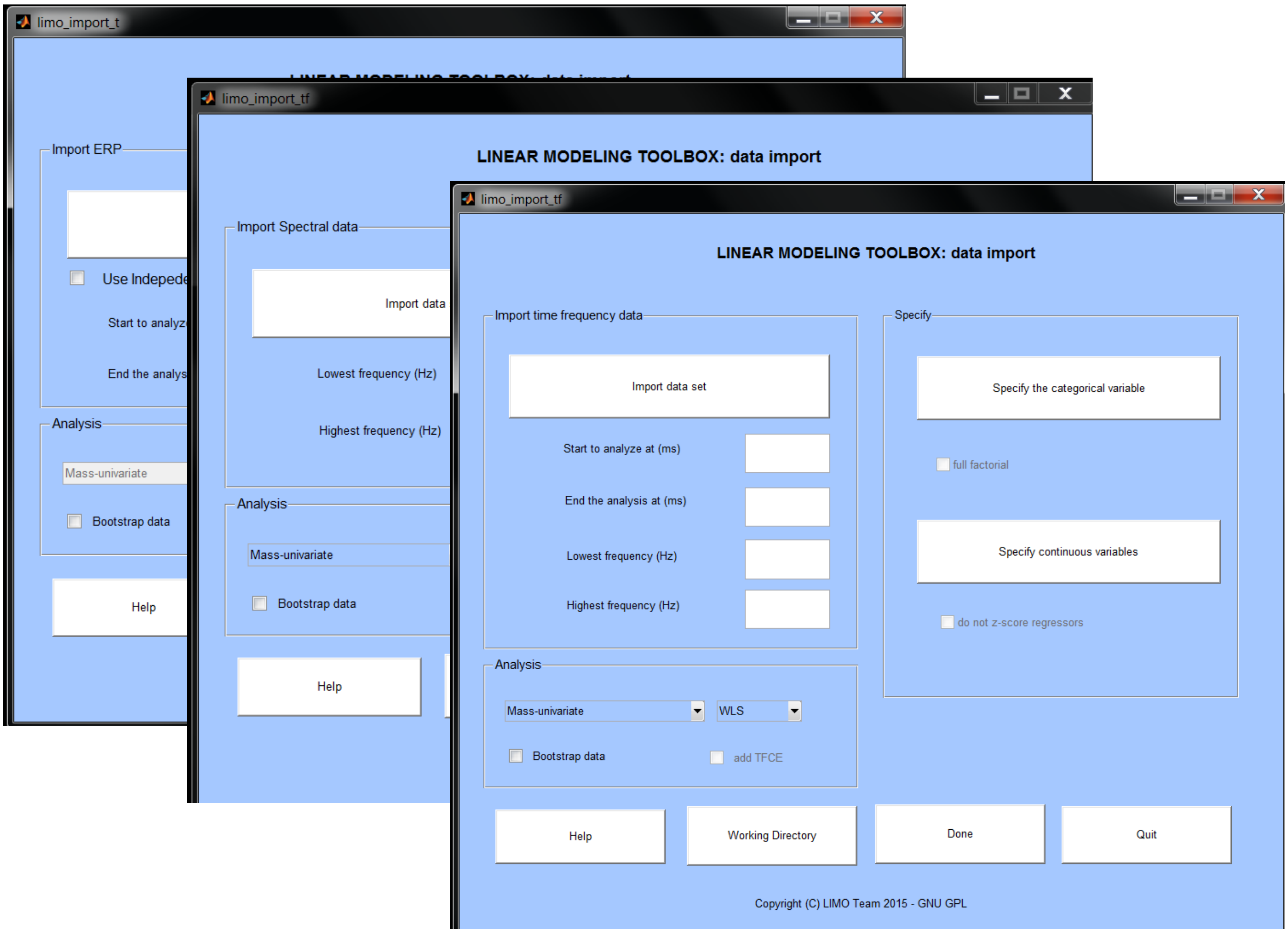
PLUG-IN OVERVIEW

Main GUI

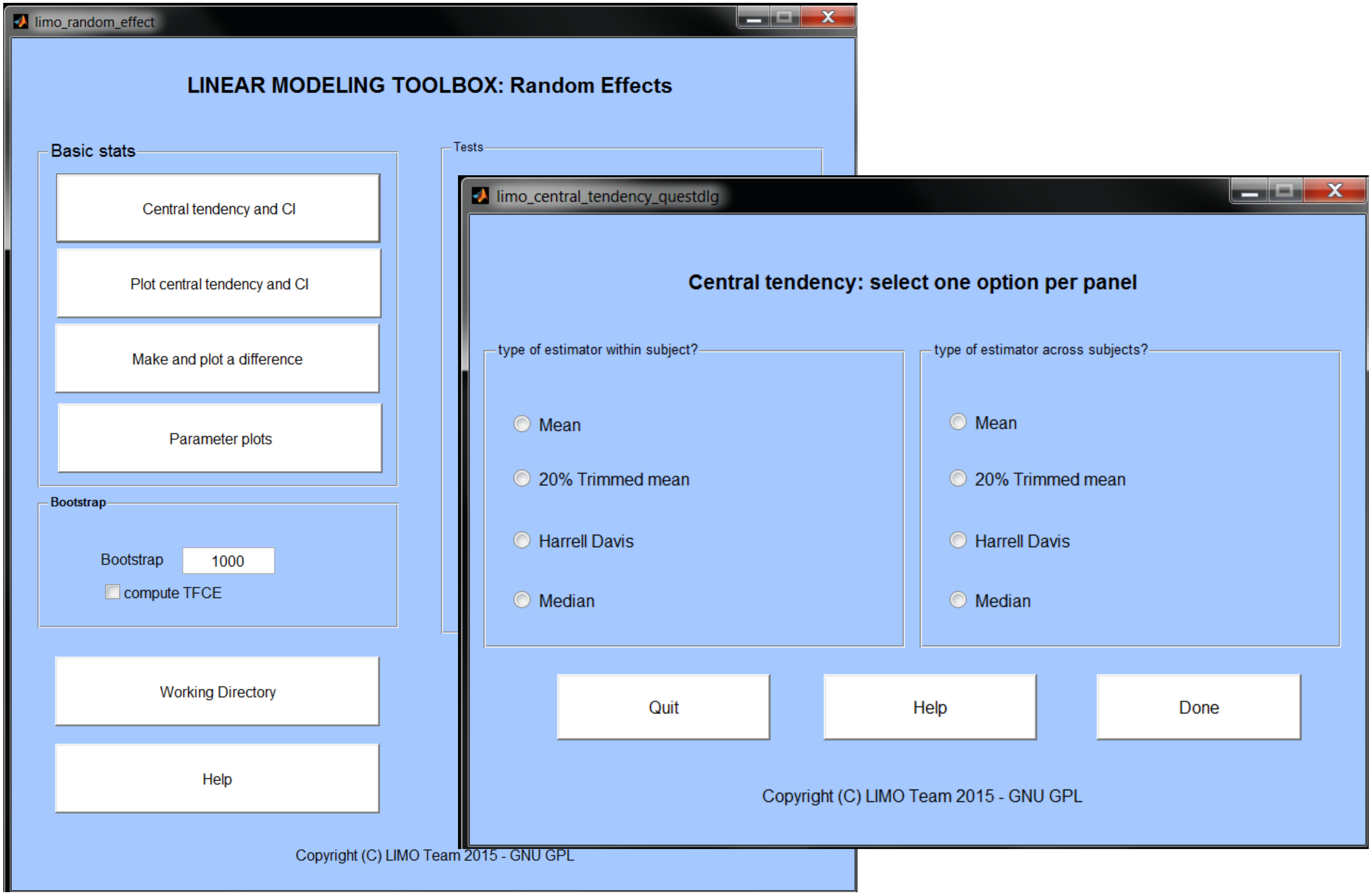
also typing: limo_eeg



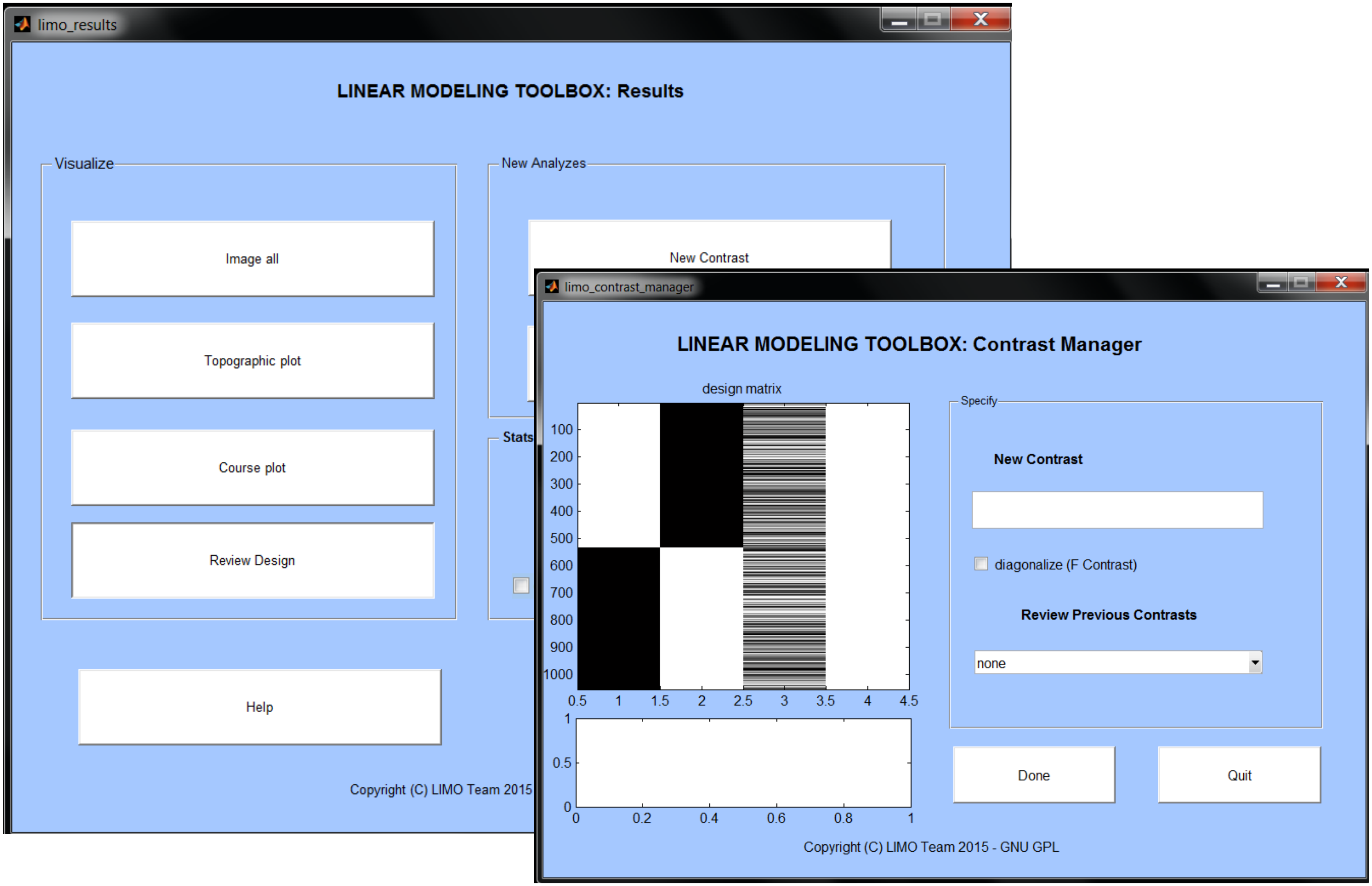
Import (1st level)



Random Effects (2nd level)



Results


limo_results
LINEAR MODELING TOOLBOX: Results
Visualize
Image all
Topographic plot
Course plot
Review Design
New Analyzes
New Contrast
Stats
Help
Copyright (C) LIMO Team 2015
limo_contrast_manager
LINEAR MODELING TOOLBOX: Contrast Manager
design matrix
Specify
New Contrast
diagonalize (F Contrast)
Review Previous Contrasts
none
Done
Quit
Copyright (C) LIMO Team 2015 - GNU GPL

GLM, robust stats, GLM, robust stats, GLM, robust stats, GLM, robust stat ..

THEORY AND PRACTICE

Linearity

- Means created by lines
- In maths it refers to equations or functions that satisfy 2 properties: additivity (also called superposition) and homogeneity of degree 1 (also called scaling)
- Additivity → $y = x1 + x2$ (output y is the sum of inputs xs)
- Scaling → $y = \beta\, x1$ (output y is proportional to input x)

http://en.wikipedia.org/wiki/Linear

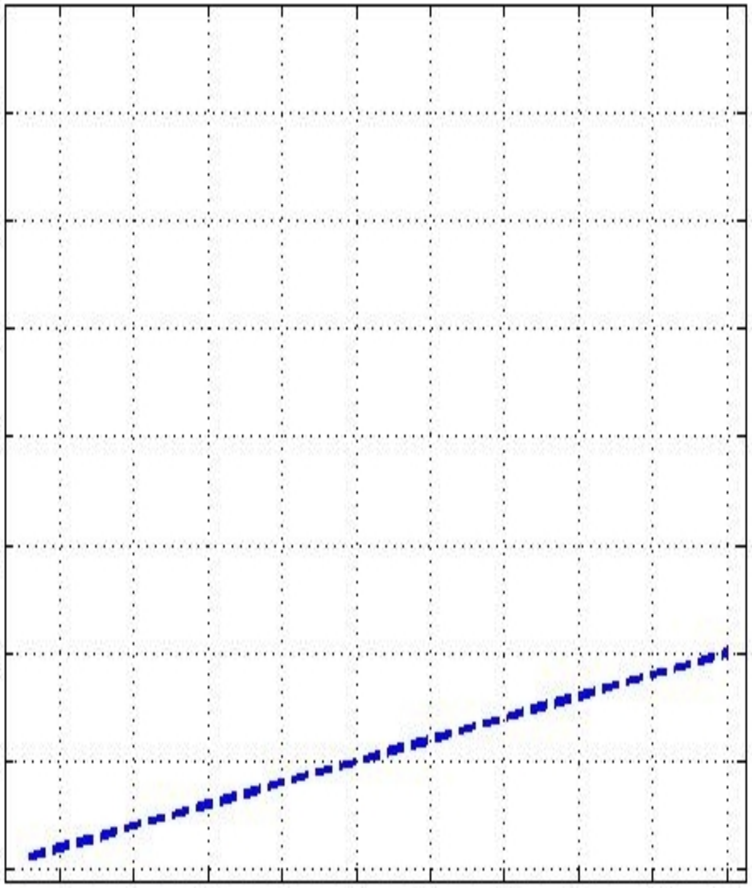
What is a linear model?

- An equation or a set of equations that models data and which corresponds geometrically to straight lines, planes, hyper-planes and satisfy the properties of additivity and scaling.

- Simple regression: $y = \beta 1x+\beta 2+\varepsilon$
- Multiple regression: $y = \beta 1x1+\beta 2x2+\beta 3+\varepsilon$
- One way ANOVA: $y = u+\alpha i+\varepsilon$
- Repeated measure ANOVA: $y=u+\alpha i+\varepsilon$
- ...

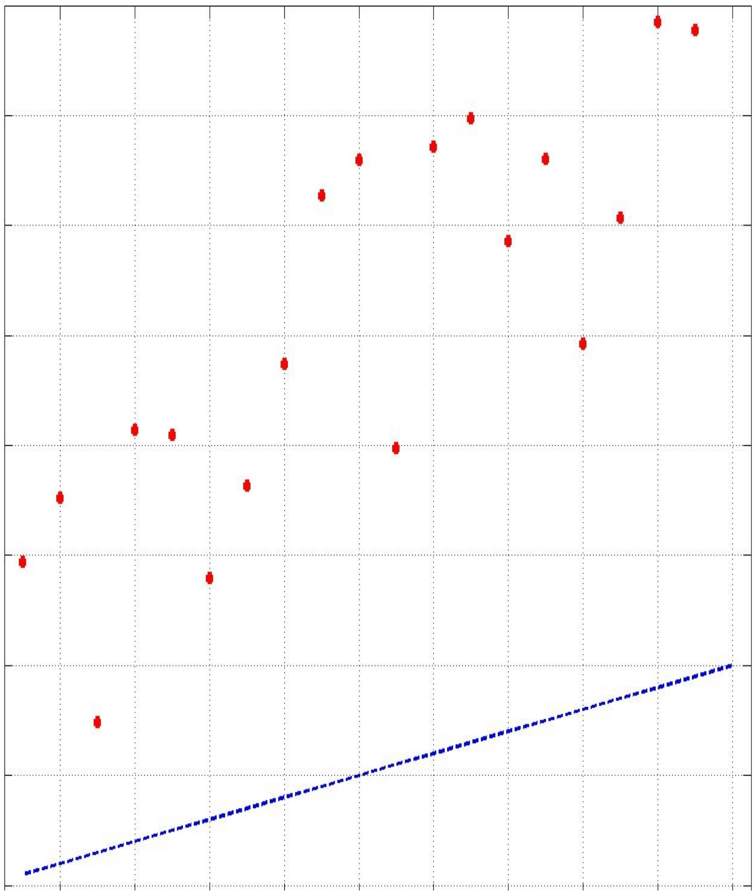
A regression is a linear model

- We have an experimental measure x (e.g. stimulus intensity from 0 to 20)



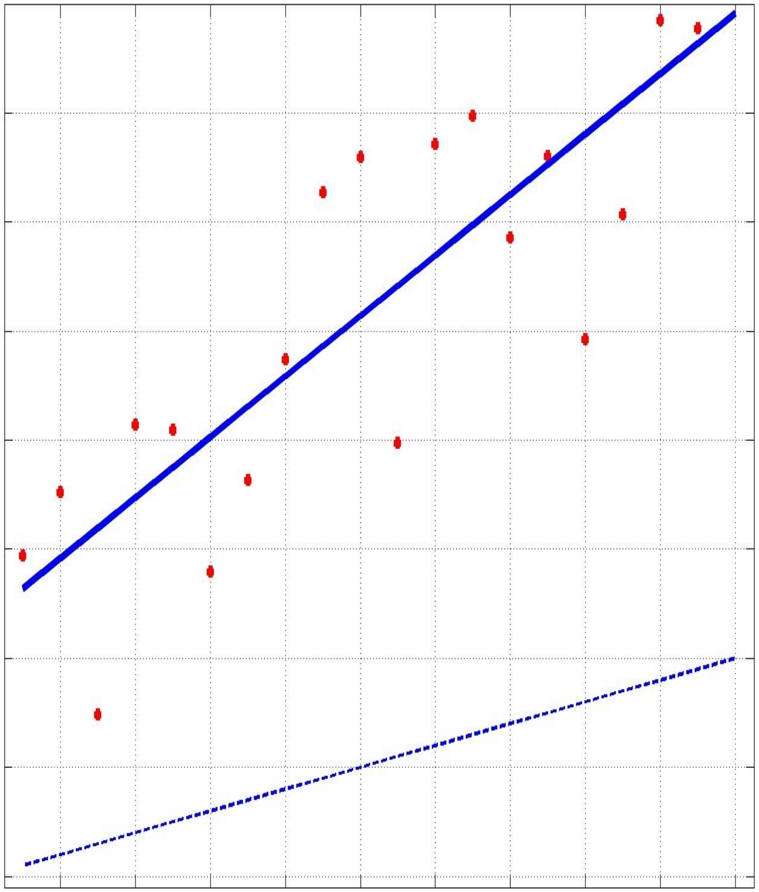
A regression is a linear model

- We have an experimental measure x (e.g. stimulus intensity from 0 to 20)
- We then do the expe and collect data y (e.g. RTs)



A regression is a linear model

- We have an experimental measure x (e.g. stimulus intensity from 0 to 20)
- We then do the expe and collect data y (e.g. RTs)
- Model: $y = \beta 1 x + \beta 2$
- Do some maths / run a software to find $\beta 1$ and $\beta 2$
- $y\text{^} = 2.7x + 23.6$

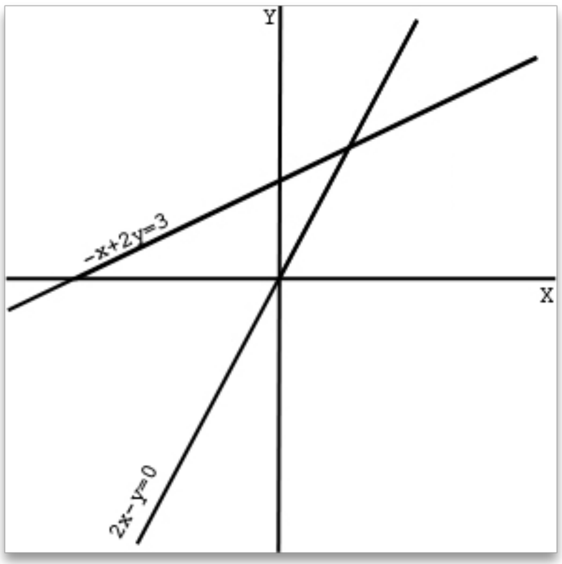


Linear algebra for regression

- Linear algebra has to do with solving linear systems, i.e. a set of linear equations
- For instance we have observations (y) for a stimulus characterized by its properties x_1 and x_2 such as $y = x_1 \beta_1 + x_2\beta2$

$$2\beta1 - \beta2 = 0$$
$$-\beta1 + 2\beta2 = 3$$

$$\beta1 = 1 \; ; \; \beta2 = 2$$



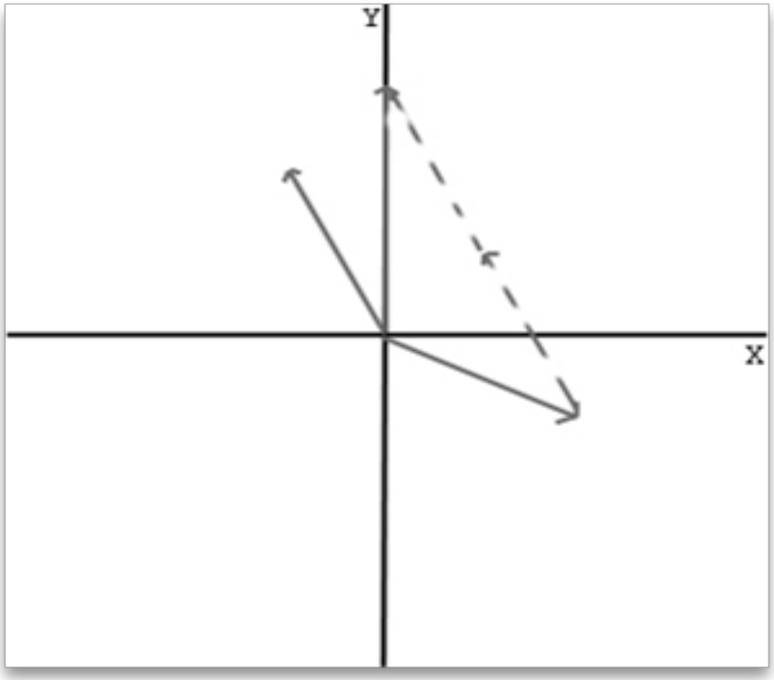
Linear algebra for regression

- With matrices, we change the perspective and try to combine columns instead of rows, i.e. we look for the coefficients with allow the linear combination of vectors

$$2\beta 1 - \beta 2 = 0$$
$$-\beta 1 + 2\beta 2 = 3$$

$$\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \left[\beta 1 \beta 2\right] = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$$

$$\beta 1 = 1 \; ; \; \beta 2 = 2$$



Linear algebra for ANOVA

- In text books we have y = u + xi + ε, that is to say the data (e.g. RT) = a constant term (grand mean u) + the effect of a treatment (xi) and the error term (ε)
- In a regression xi takes several values like e.g. [1:20]
- In an ANOVA xi is designed to represent groups using 1 and 0

Linear algebra for ANOVA

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

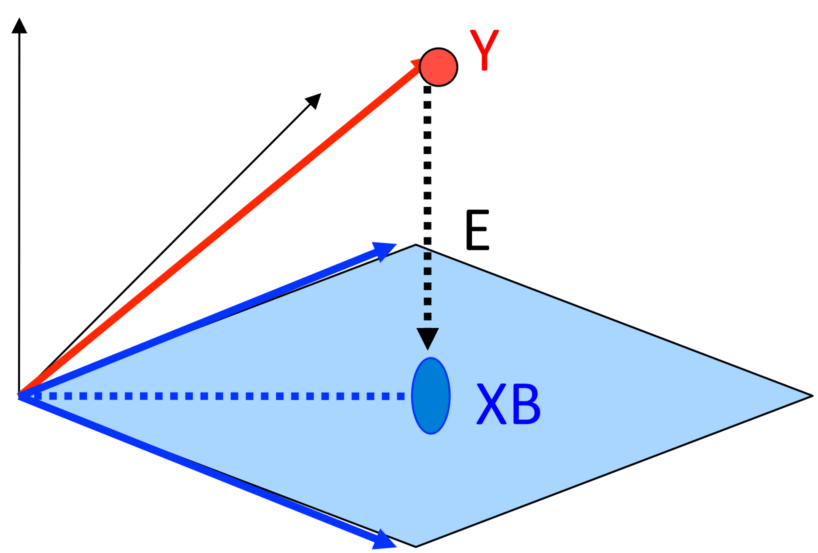
y(1..3)1= 1x1+0x2+0x3+0x4+c+e11
y(1..3)2= 0x1+1x2+0x3+0x4+c+e12
y(1..3)3= 0x1+0x2+1x3+0x4+c+e13
y(1..3)4= 0x1+0x2+0x3+1x4+c+e13

$$\begin{Bmatrix} 8 \\ 9 \\ 7 \\ 5 \\ 7 \\ 3 \\ 3 \\ 4 \\ 1 \\ 6 \\ 4 \\ 9 \end{Bmatrix} = \begin{Bmatrix} 1\,0\,0\,0\,1 \\ 1\,0\,0\,0\,1 \\ 1\,0\,0\,0\,1 \\ 0\,1\,0\,0\,1 \\ 0\,1\,0\,0\,1 \\ 0\,1\,0\,0\,1 \\ 0\,0\,1\,0\,1 \\ 0\,0\,1\,0\,1 \\ 0\,0\,1\,0\,1 \\ 0\,0\,0\,1\,1 \\ 0\,0\,0\,1\,1 \\ 0\,0\,0\,1\,1 \end{Bmatrix} * \begin{Bmatrix} \beta 1 \\ \beta 2 \\ \beta 3 \\ \beta 4 \\ c \end{Bmatrix} + \begin{Bmatrix} e1 \\ \vdots \\ e13 \end{Bmatrix}$$

→ This is like the multiple regression except that we have ones and zeros instead of 'real' values so we can solve the same way

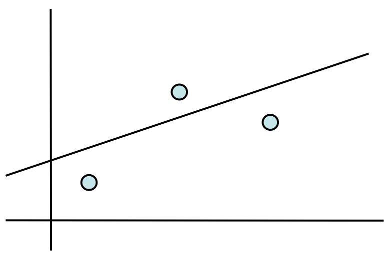
Linear Algebra, geometry and Statistics

- Y = 3 observations X = 2 regressors
- Y = XB+E → B = inv(X'X)X'Y → Y^=XB



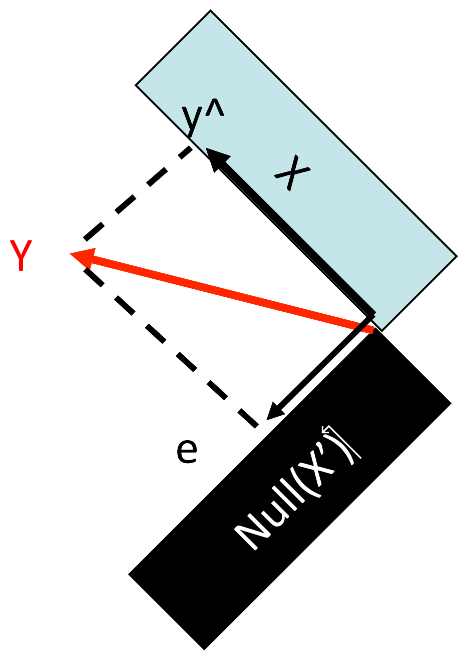
SS total = variance in Y
SS effect = variance in XB
SS error = variance in E
R2 = SS effect / SS total
F = SS effect/df / SS error/dfe

Linear Algebra, geometry and Statistics



$y = \beta x + c$

Projecting the points on the line at perpendicular angles minimizes the distance^2



Y = y^+e

P = X inv(X'X) X'

y^ = PY

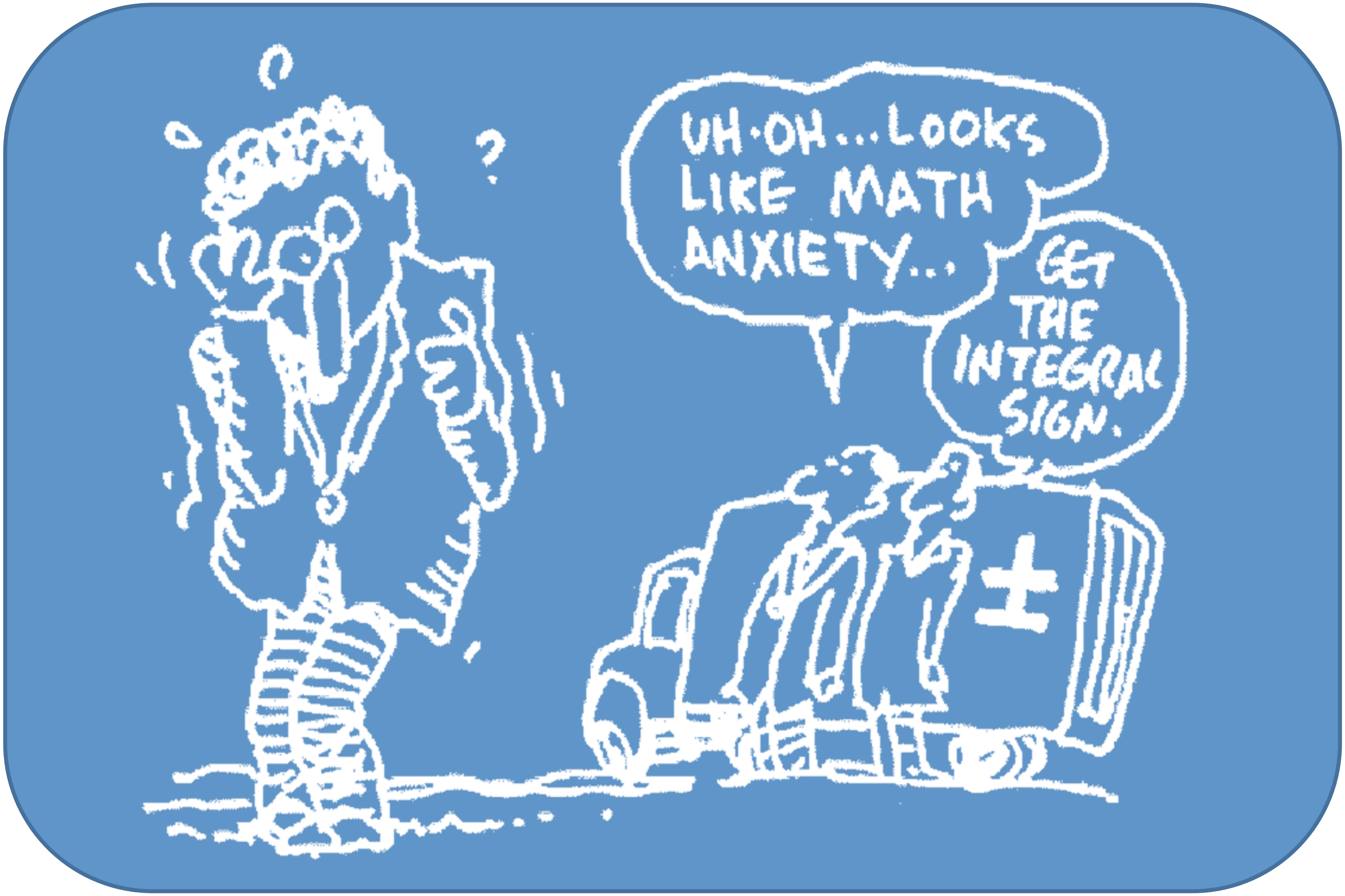
e = (I-P)Y

An 'effect' is defined by which part of X to test (i.e. project on a subspace)

```
R0 = I - (X0*pinv(X0));
P = R0 - R;
Effect  = (B'*X'*P*X*B);
```

Linear Algebra, geometry and Statistics

- Projections are great because we can now constrain Y^ to move along any combinations of the columns of X
- Say you now want to contrast gp1 vs gp2 in a ANOVA with 3 gp, do C = [1 -1 0 0]
- Compute B so we have XB based on the full model X then using P(C(X)) we project Y^ onto the constrained model (think doing a multiple regression gives different coef than multiple simple regression → project on different spaces)


UH-OH...LOOKS LIKE MATH ANXIETY...
GET THE INTEGRAL SIGN.

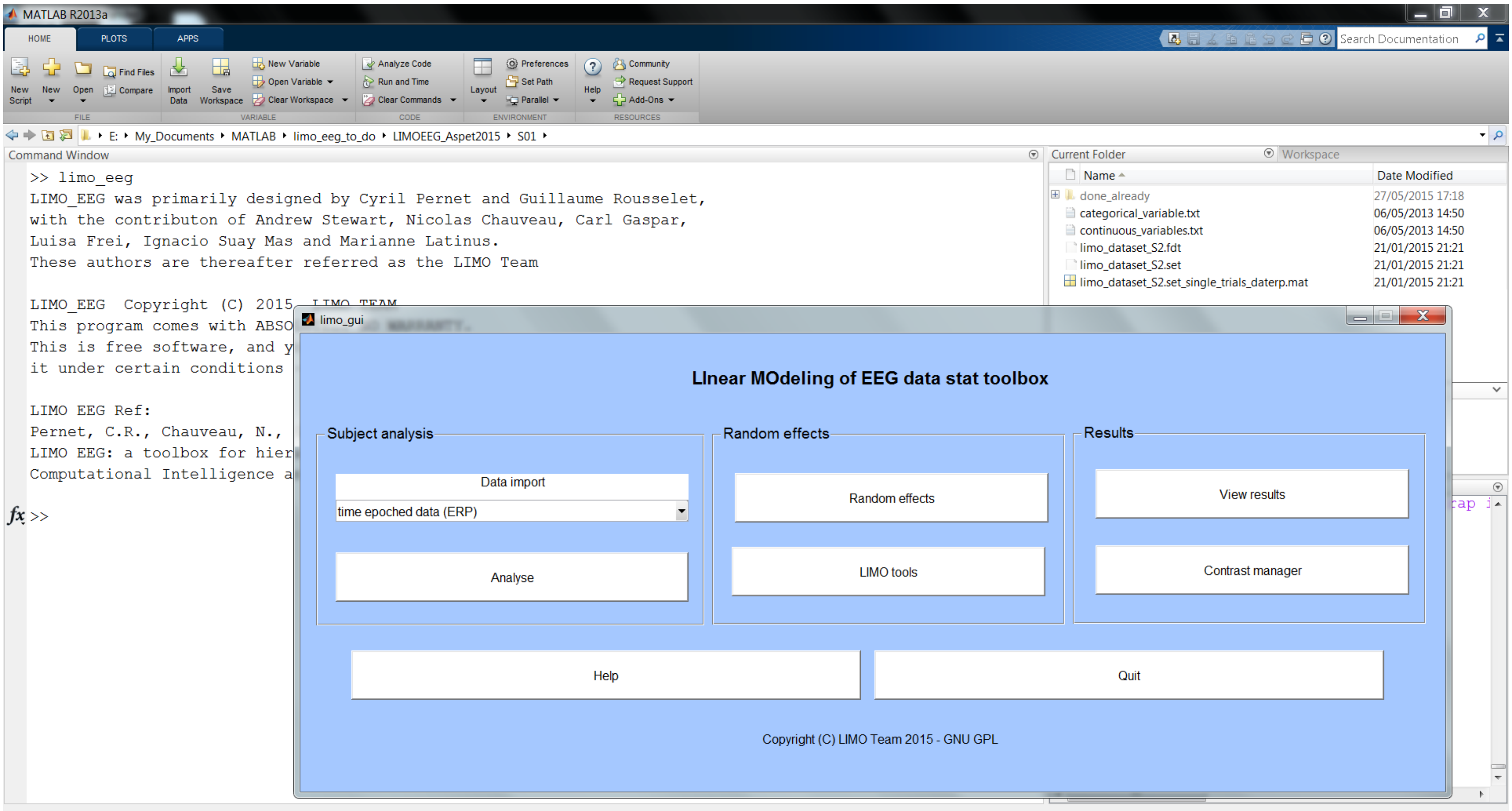
Let’s analyse one subject

- **Design**: 2 faces (cond1/cond2) + a continuous variable related to the phase information in the stimulus space (~noise)

- LIMO EEG – 1st level analysis

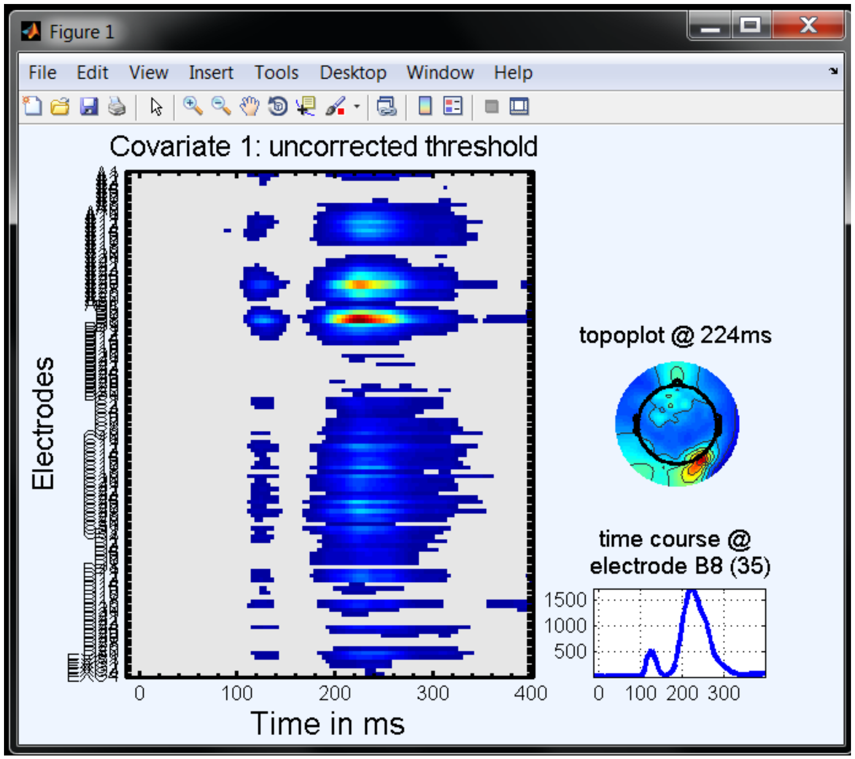
= make a parameter file per condition (like we would for erp)

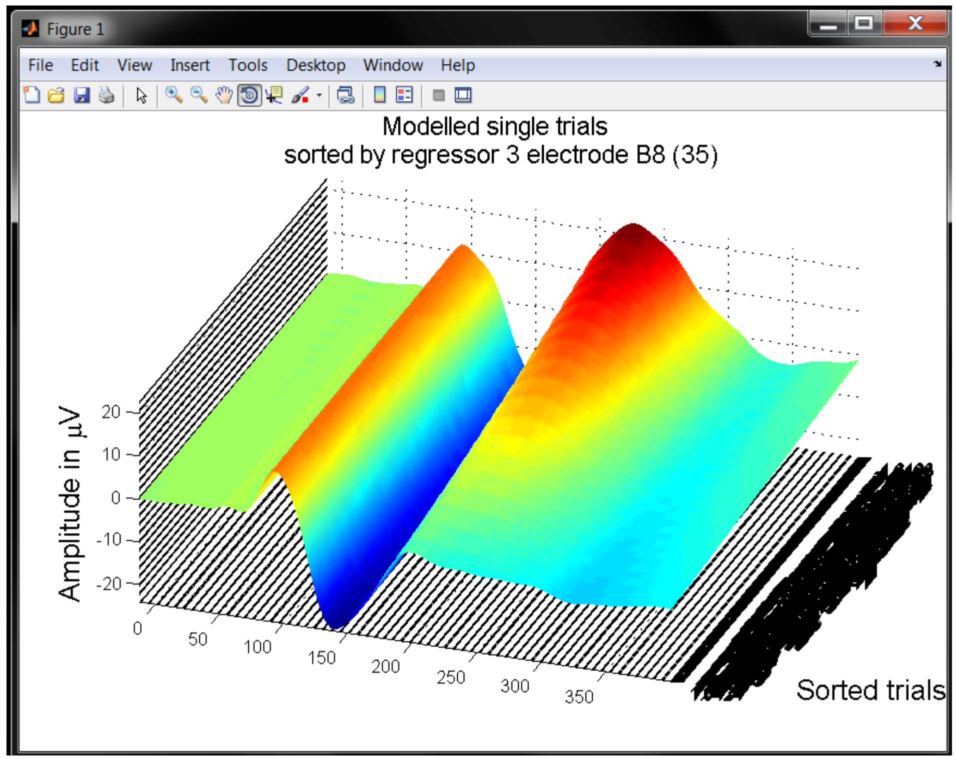
Let’s analyse one subject



What have we done: results

- Image all (R2, condition, covariate)
- Course plots, topoplots





Robust Statistics

WHY & HOW?

Issues with standard stats

- Standard stats are all instantiations of a GLM using an Ordinary Least Square solution → implies looking at the mean
- the breakdown point of an estimator is the proportion of incorrect observations (e.g. arbitrarily large observations) an estimator can handle before giving an incorrect
- For data x1 to xn – the mean has a bkdp of 0 because we can make the mean large changing any xi – the median has a bkdp of 50%

http://en.wikipedia.org/wiki/Robust_statistics

Yes but my data are Gaussian

- Are you sure?
- Micceri (1989). The Unicorn, The Normal Curve, and Other Improbable Creatures. Psych Bul. 105, 156-166
- If the data are Gaussian, the median, the trimmed mean is the same as the mean ! So no reason not to use alternative techniques.

- 1st level, uses weighted least square (weights down bad trials)
- 2nd level involves 20% trimmed mean (weights = 0 for bad subjects): t-tests, 1-way ANOVA, Repeated Measures ANOVA (soon)
- For regressions and N-way ANOVA/ANOVA we use an IRLS (all subjects have weights from 0 to 1)

http://en.wikipedia.org/wiki/Robust_statistics

Practical

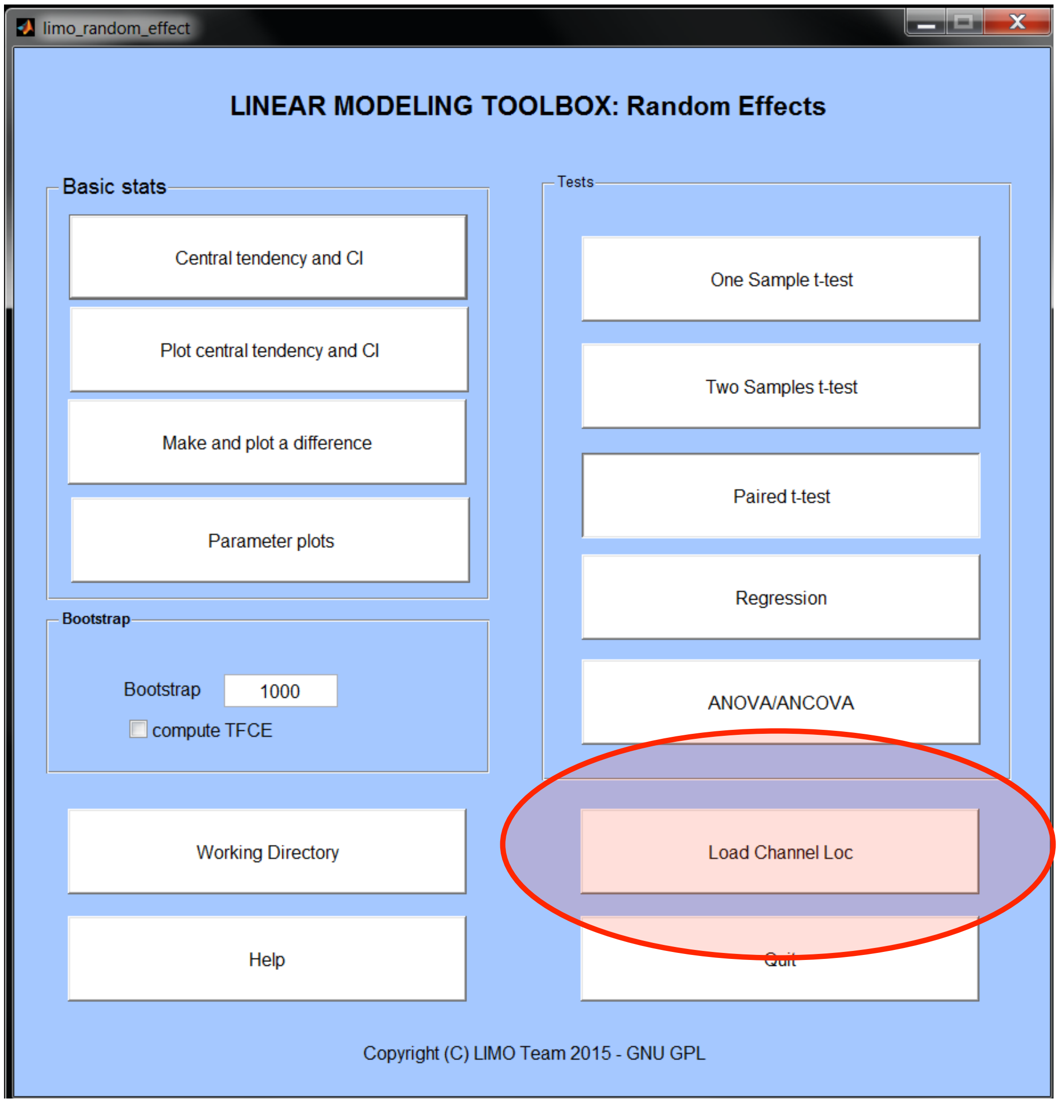
- One sample t-test on ‘noise’ regressor

→You can select files by hand, by it’s easier to build lists – right click/run sheffield_mklist.m

→This makes a list_of_Betas.txt we can use

→From the GUI, choose ‘Random Effect’

Practical

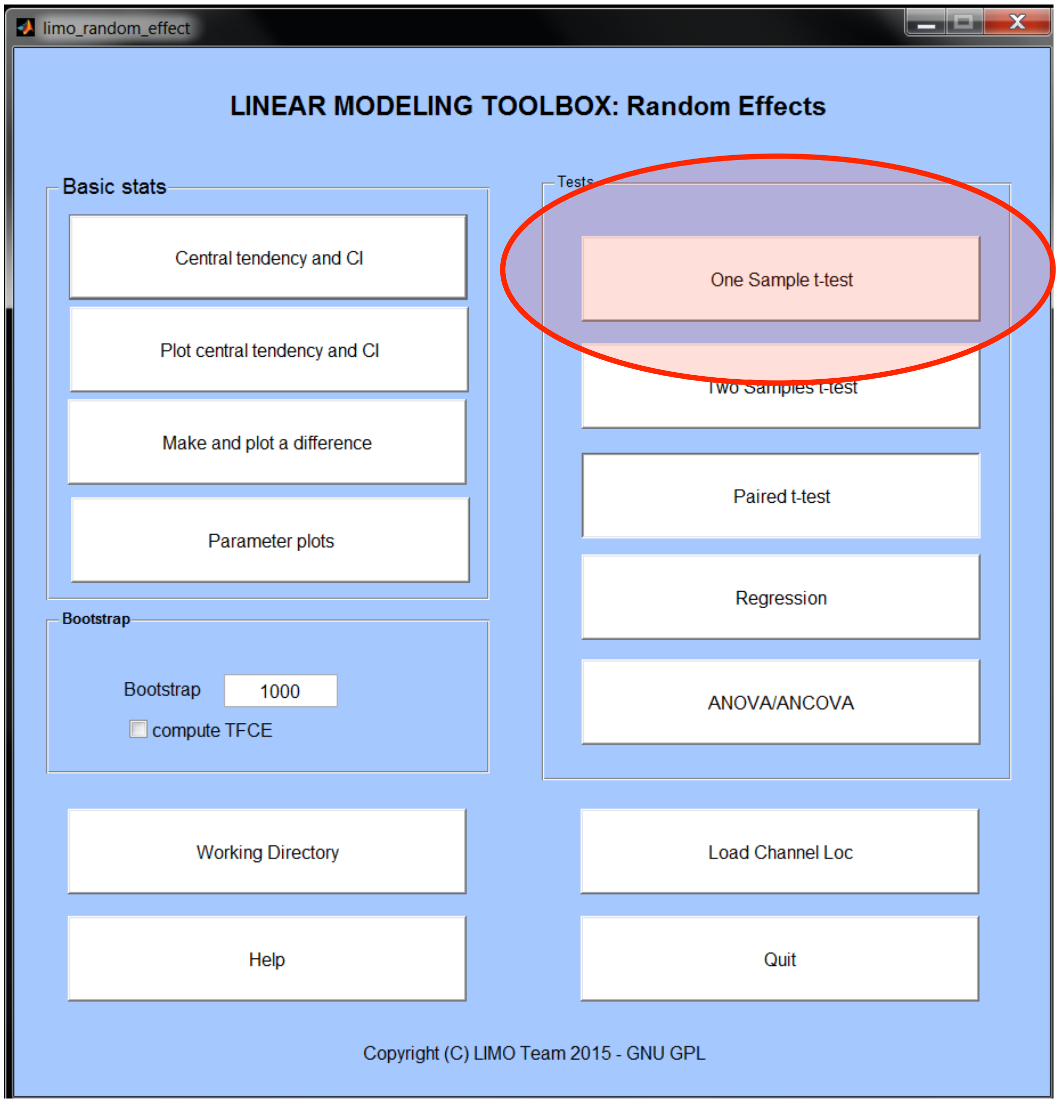


LIMO EEG expect you to build a template cap to use across subjects (LIMO TOOLS) because only valid electrodes are analysed per subject

= no interpolated values at the subject level
= LIMO EEG deals with missing data

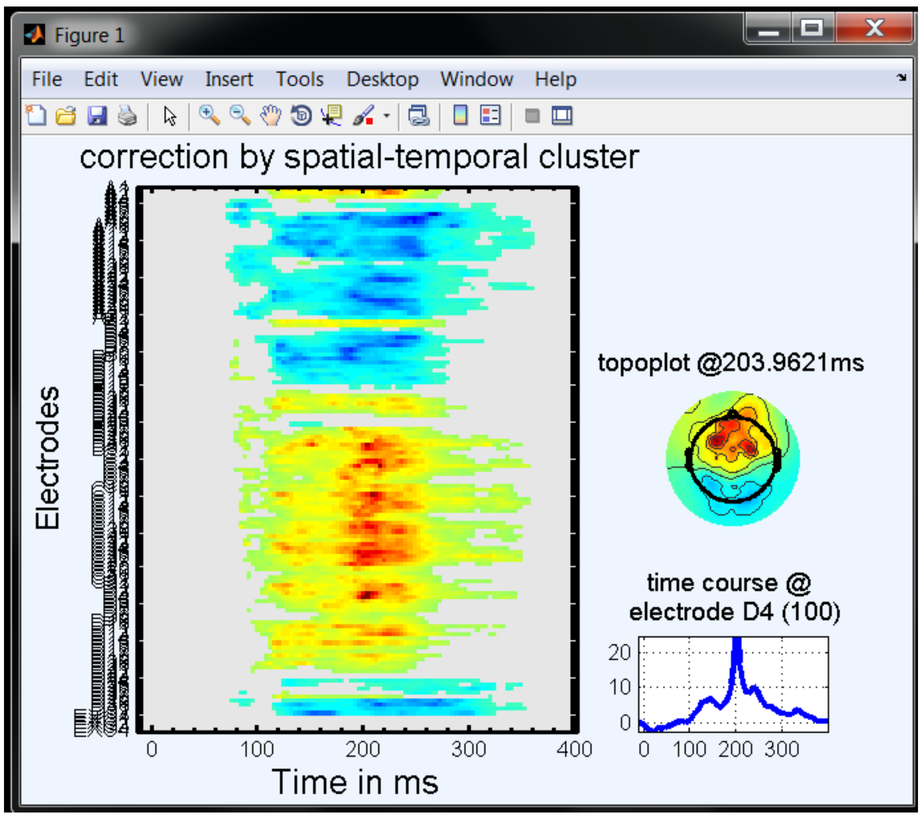
The expected channlocs is in the gp_effects directory

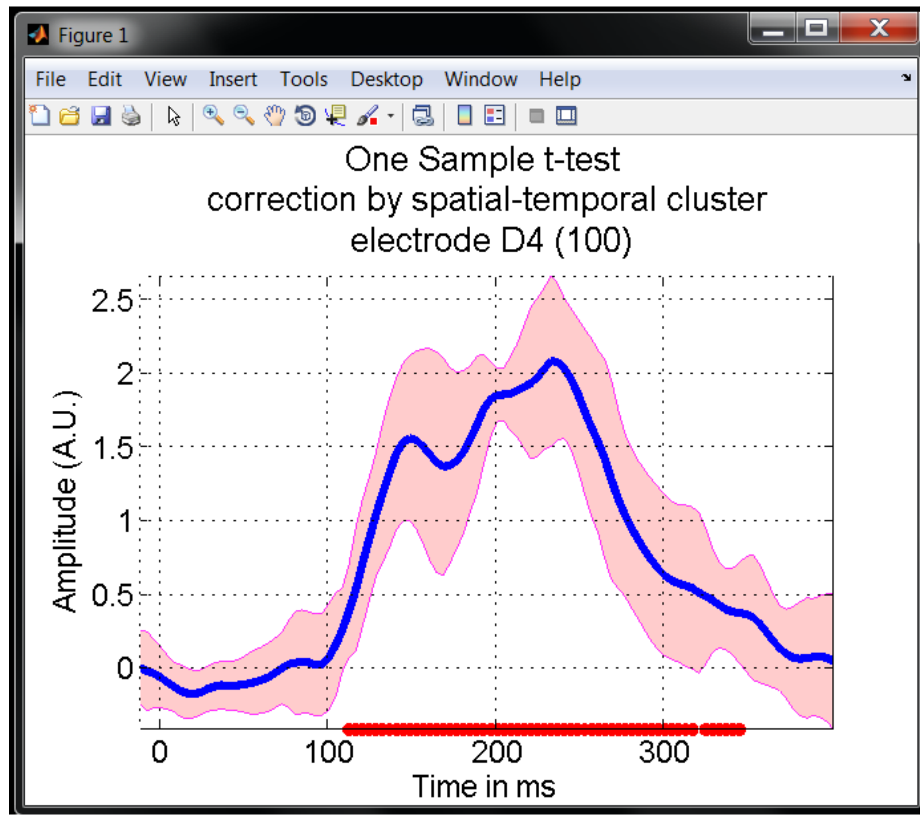
Practical



Compute a one-sample t-test on betas parameter 3 i.e. the Effect of stimulus phase information on ERP

Review gp level results





Design questions!

- Let's think how to analyse your data!
- Nb of conditions / covariates
- contrasts
- 1^{st} level covariates
- 2^{nd} level covariates

Design questions!

- Typical 2*2 design

→ 1st level vs 2nd level, where to model the interaction

- Testing the effect of covariate within or between conditions?