Hands-On Session: Regression Analysis

File: afni05_regression.pdf

SSCC/NIMH/NIH/HHS

Overview

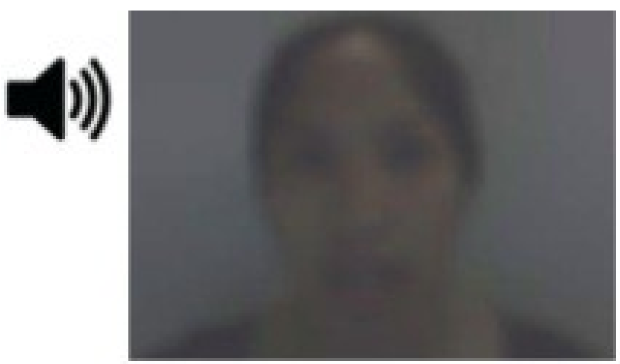
- What we have learned so far
  - Use data viewer ‘afni’ interactively
  - Model HRF with a fixed-shape basis function (*e.g.*, Gamma variate)
    - Assume the brain responds with the **same shape**
      - Across subjects and any activated regions
    - Differ in magnitude: β (and its significance) is focus
- What we will do in this hands-on session
  - Data pre-processing overview for time series regression analysis
  - Basic concepts
    - Regressors, design matrix, and confounding effects
    - Statistical significance testing in regression analysis
  - Navigation with GUI ‘**afni**’
    - Spot check for the original data
    - Statistic thresholding with data viewer ‘**afni**’ (two-sided vs. one-sided with t)
    - Model performance (visual check of curve fitting and test via full F or R^2)

FMRI Regression Analysis

- Voxel-wise regression model: $y = X\beta+\varepsilon$
  - y: signal (time series) at a voxel – different across voxels
  - X: explanatory (independent) variables (regressors) – same across voxels
  - β: regression coefficients (response strength) – different across voxels
  - ε: residuals (anything we can't account for) – different across voxels
- Regressors in design matrix $X = [x_1, x_2, \ldots, x_k]$
  - Regressors of interest: hemodynamic responses (HDR)
  - Regressors of no interest: drift effect (polynomials), head motion, *etc*.
- Association between stimulus and BOLD signal: HDR/HRF
  - Fixed shape regardless of subjects, brain regions, stimuli: regression
  - No assumption about the HDR shape: deconvolution + regression
  - Middle ground: regression
- Residuals
  - White noise: OLS – 3dDeconvolve
  - Serially correlated: ARMA(1,1)+REML – 3dREMLfit

A Case Study

- Speech Perception Task: Subjects were presented with audiovisual speech presented in a predominantly auditory or predominantly visual modality.
- A digital video system was used to capture auditory and visual speech from a female speaker.
- 2 types of stimulus conditions:



(1) **Auditory-Reliable**

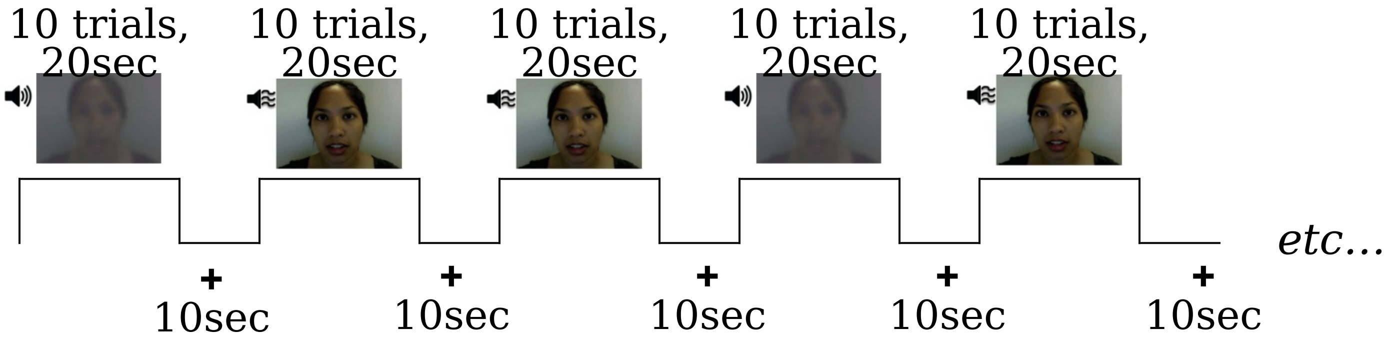
Example: Subjects can clearly *hear* the word "cat," but the video of a woman mouthing the word is degraded.

(2) **Visual-Reliable**

Example: Subjects can clearly *see* the video of a woman mouthing the word "cat," but the audio of the word is degraded.

Experiment Design

- 3 runs in a scanning session
- Each run consisted of randomized 10 blocks:
  - 5 blocks contained Auditory-Reliable (*Arel*) stimuli, and
  - 5 blocks contained Visual-Reliable (*Vrel*) stimuli
- Each block contained 10 trials of *Arel* OR *Vrel* stimuli
  - Each block lasted for 20s (1s for stimulus presentation, followed by a 1s inter-stimulus interval)
- Each baseline block consisted of a 10s fixation point



Data Collected

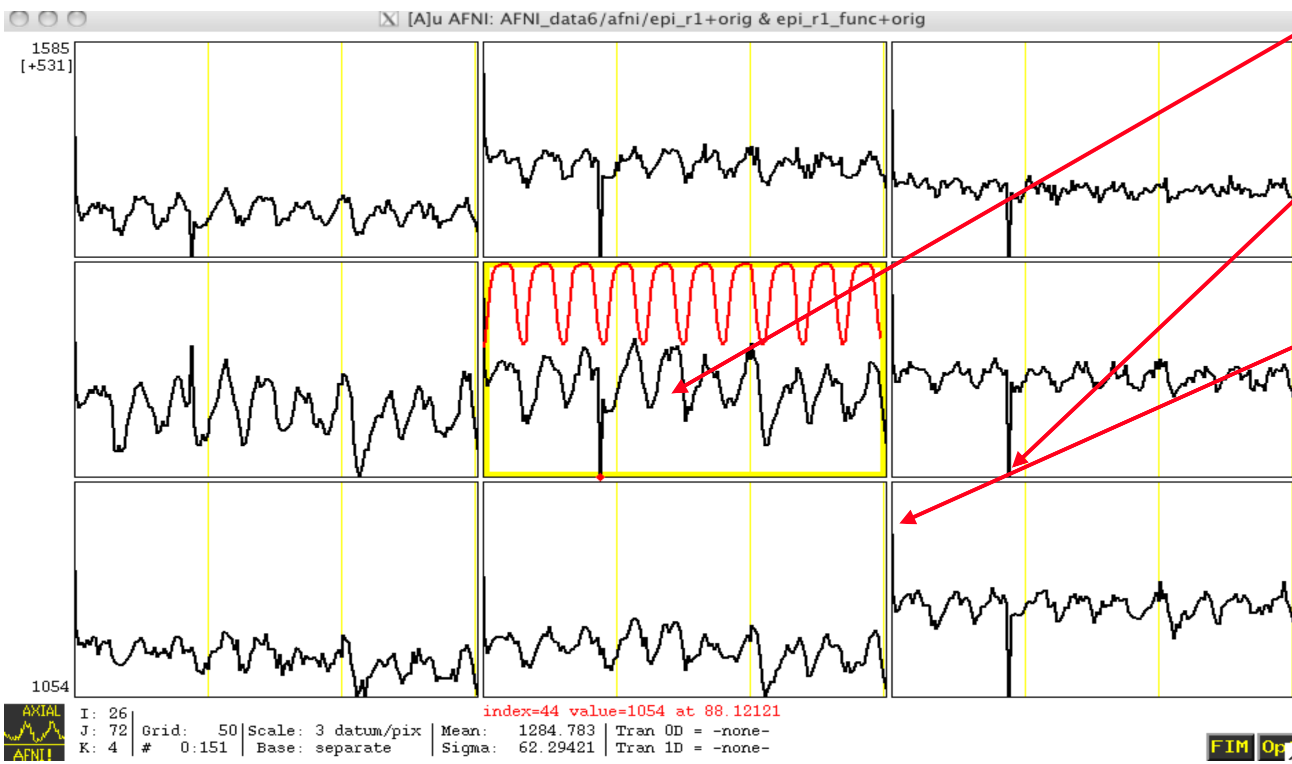
- 2 anatomical datasets for each subject from a 3T
  - 175 sagittal slices
  - voxel dimensions = 1.0 × 0.938 × 0.938 mm^3

- 3 time series (EPI) datasets for each subject
  - 33 axial slices × 152 volumes (TRs) per run
  - TR = 2s; voxel dimensions = 2.75 × 2.75 × 3.0 mm^3

- Sample size, n = 10 (all right-handed subjects)
  - 10 is far too few for most modern experiments

Data Quality Check

- To look at the data: type **cd AFNI_data6/afni**, then **afni**
- **Switch Underlay** to dataset **epi_r1**
  - Then **Axial Image** and **Graph**
  - **FIM→Pick Ideal** ; then click **afni/epi_r1_ideal.1D** ; then **Set**
  - Right-click in image, **Jump to (ijk)**, then **26 72 4**, then **Set**



- Data clearly has activity in sync with reference
  - 20s blocks
- Data also has a big spike at 89s
  - Head motion
- Spike at $t = 0$
- Some tricks with keyboard
  - a: automatic scaling
  - v: video mode
  - m/M: voxel matrix sizing on Graph window

Preparing Data for Analysis

- Following preparatory steps are common (e.g., afni_proc.py):
  - Outliers: **3dToutcount**, **3dDespike**
  - Temporal alignment or slice timing correction (sequential/interleaved): **3dTshift**
  - EPI Image/volume registration (head motion correction): **3dvolreg**
  - EPI to anatomy registration: **align_epi_anat.py**
  - Spatial normalization (standard space conversion): **@auto_tlrc, auto_warp.py**
  - Blurring/smoothing: **3dmerge, 3dBlurToFWHM, 3dBlurInMask**
  - Masking: **3dAutomask**
  - Temporal mean scaling: **3dTstat** and **3dcalc**

- Not all steps are necessary or desirable in any given case

Regression Analysis

- Regression model: $y = X\beta + \varepsilon$
- Run script by typing **tcsh rall_regress** (takes a few minutes)

```
3dDeconvolve -input rall_vr+orig -polort 1                              \
    -concat '1D: 0 150 300'                                             \
    -num_stimts 8                                                       \
    -stim_times 1 stim_AV1_vis.txt 'BLOCK(20,1)' -stim_label 1 Vrel    \
    -stim_times 2 stim_AV2_aud.txt 'BLOCK(20,1)' -stim_label 2 Arel    \
    -stim_file 3 motion.1D'[0]' -stim_base 3 -stim_label 3 roll        \
    -stim_file 4 motion.1D'[1]' -stim_base 4 -stim_label 4 pitch       \
    -stim_file 5 motion.1D'[2]' -stim_base 5 -stim_label 5 yaw         \
    -stim_file 6 motion.1D'[3]' -stim_base 6 -stim_label 6 dS          \
    -stim_file 7 motion.1D'[4]' -stim_base 7 -stim_label 7 dL          \
    -stim_file 8 motion.1D'[5]' -stim_base 8 -stim_label 8 dP          \
    -gltsym 'SYM: Vrel -Arel' -glt_label 1 V-A                          \
    -tout -x1D rall_X.xmat.1D -xjpeg rall_X.jpg                         \
    -fitts rall_fitts -bucket rall_func                                 \
    -jobs 2
```

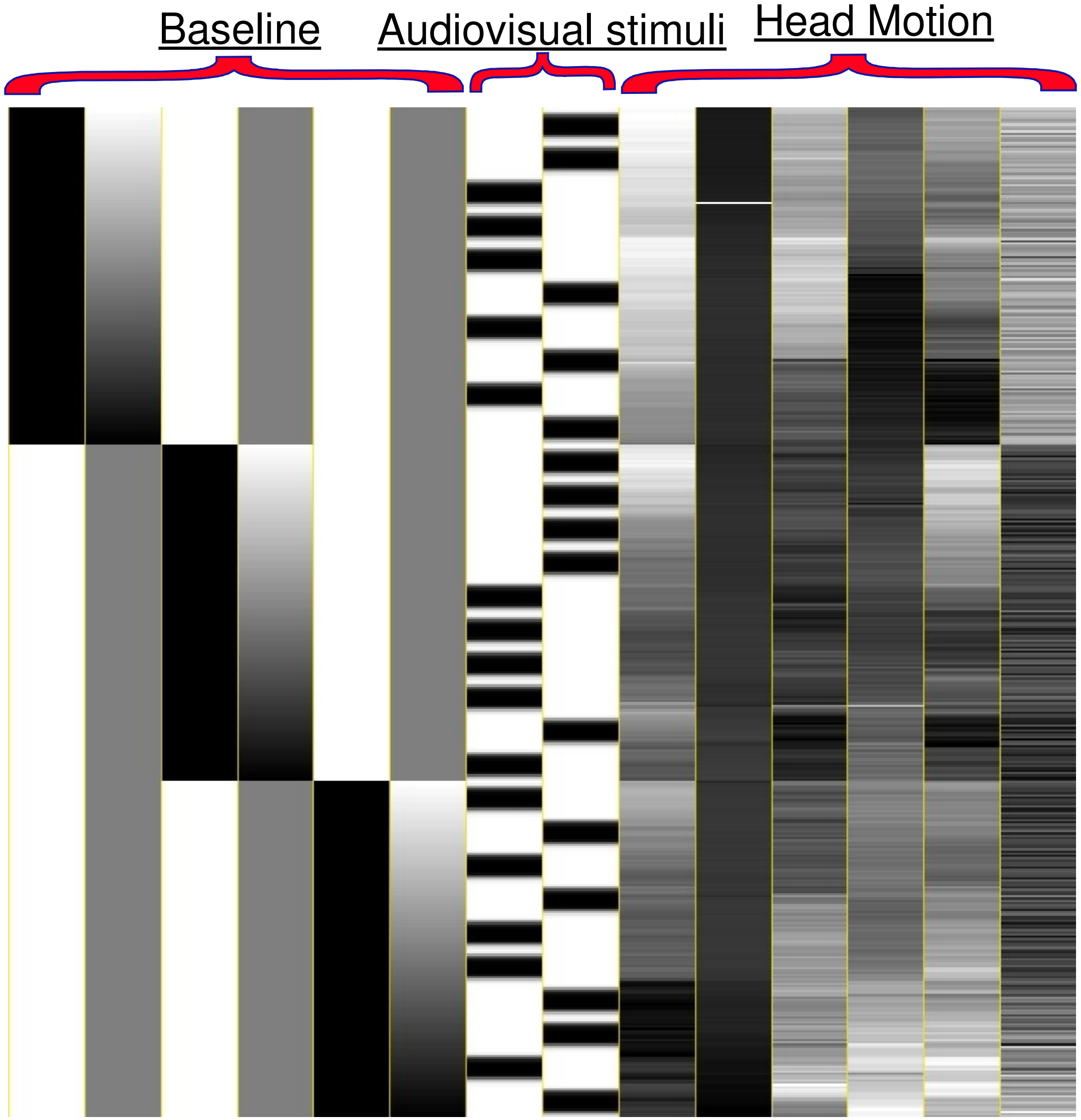
- 2 audiovisual stimulus classes were given using **`-stim_times`**
- **Important to include motion parameters as regressors**?
  - May remove the confounding effects due to motion artifacts
  - 6 motion parameters as covariates via **`-stim_file`** + **`-stim_base`**
  - **`motion.1D`** generated from **`3dvolreg`** with the **`-1Dfile`** option
  - Test the significance of head motion parameters
    - Add **`-bout`** or remove **`-stim_base`**
    - Use **`-gltsym 'SYM: roll \ pitch \yaw \dS \dL \dP'`**

Modeling Serial Correlation in the Residuals

- Temporal correlation exists in the residuals of the time series regression model
- Within-subject variability (or statistical value) would get deflated (or inflated) if temporal correlation is not accounted for in the model
- Better correct for the temporal correlation if bringing both effect size and within-subject variability to group analysis
- ARMA(1, 1) assumed in **3dREMLfit**
- Script automatically generated by 3dDeconvolve (may use –x1D_stop)
  - ⋆ File **rall_func.REML_cmd** under **AFNI_data6/afni**
  - ⋆ Run it by typing ***tcsh -x rall_func.REML_cmd***

```
3dREMLfit -matrix rall_X.xmat.1D -input rall_vr+orig \
 -tout -Rbuck rall_func_REML -Rvar rall_func_REMLvar \
 -Rfitts rall_fitts_REML -verb
```

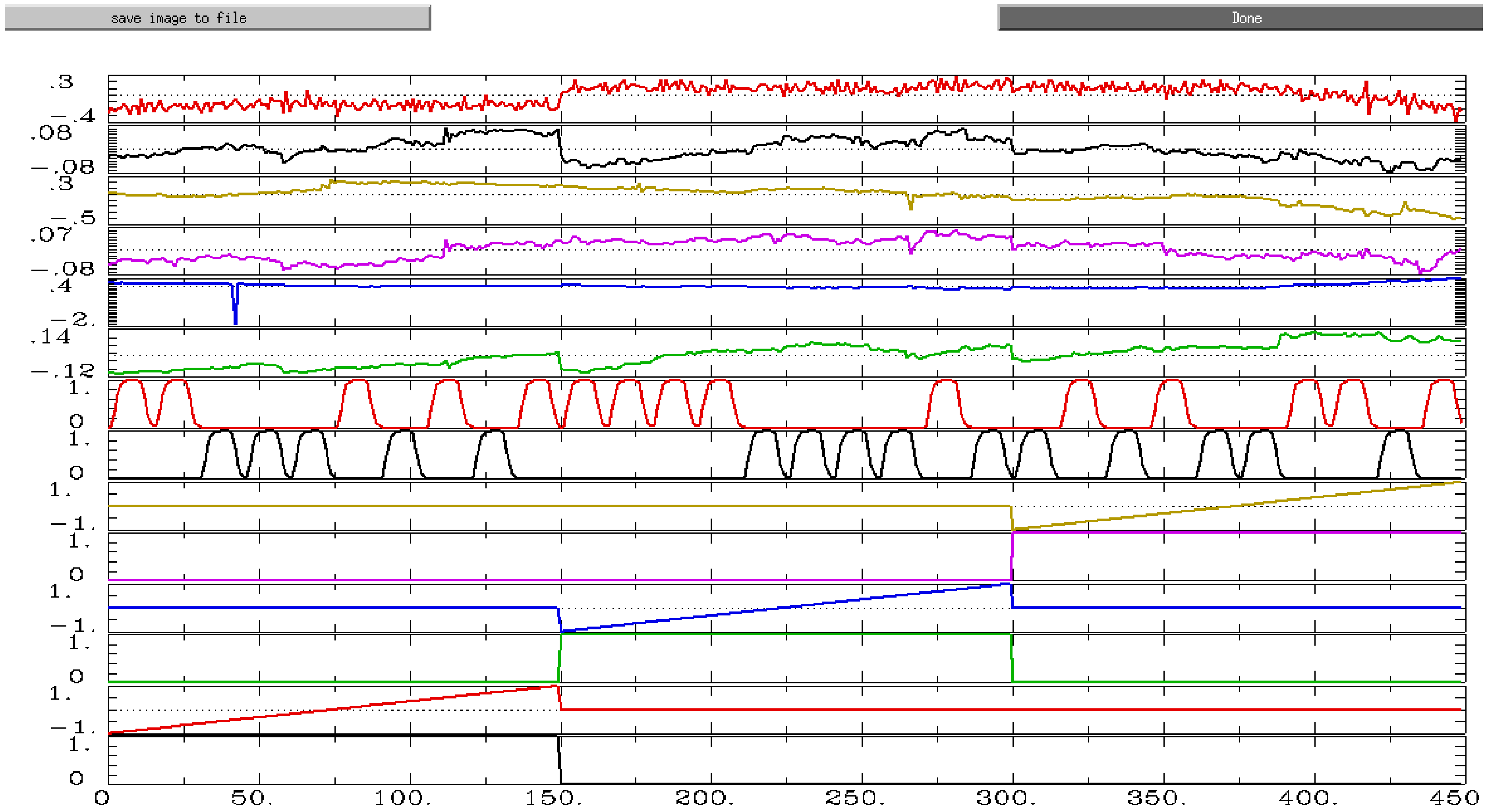
Regressor Matrix X for This Script (via -xjpeg)



aiv rall_xmat.jpg

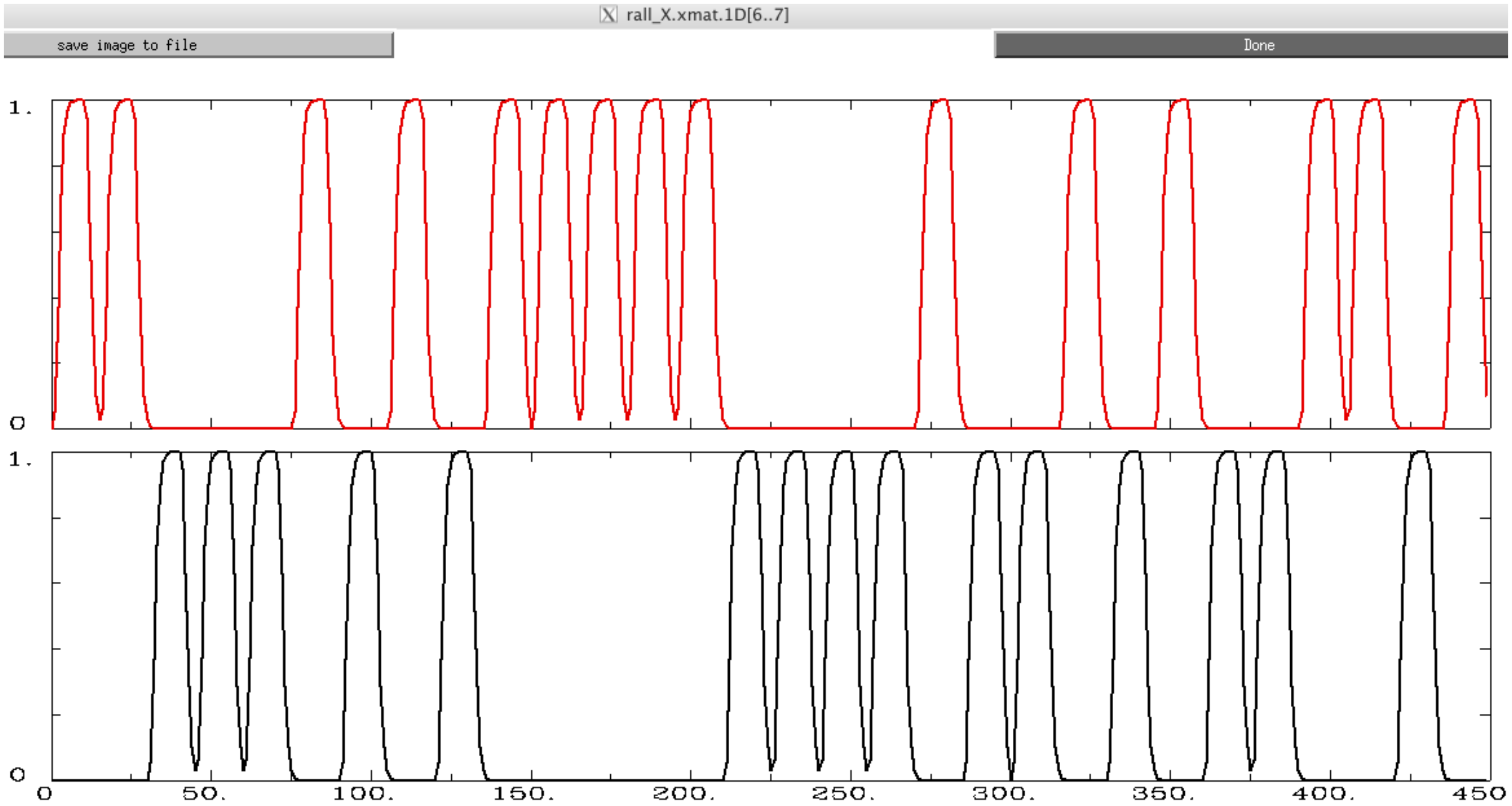
- 6 drift effect regressors
  - linear baseline
  - 3 runs times 2 params/run
- 2 regressors of interest
- 6 head motion regressors
  - 3 rotations and 3 shifts

Showing All Regressors (via -x1D)



All regressors: **1dplot -sepscl rall_X.mat.1D**

Plotting Regressors of Interest



Regressors of Interest: **1dplot rall_X.mat.1D'[6..7]'**

Options in 3dDeconvolve - 1

`-concat '1D: 0 150 300'`

- "File" that indicates where distinct imaging runs start inside the input file
  - Numbers are the time (TR) indexes inside the dataset file for start of runs
  - These time points are considered as discontinuities in the model
  - In this case, a text format .1D file put directly on the command line
    - Could also be a filename, if you want to store that data externally

`-num_stimts 8`

- 2 audiovisual stimuli (+6 motion), thus 2 **`-stim_times`** below
- Times given in the **`-stim_times`** files are *local* to the start of each run

`-stim_times 1 stim_AV1_vis.txt 'BLOCK(20,1)' -stim_label 1 Vrel`

- Content of **`stim_AV1_vis.txt`**

  60 90 120 180 240

  120 150 180 210 270

  0 60 120 150 240

  - ★ Each of 3 lines specifies start time in seconds for stimuli within the run

Options in 3dDeconvolve - 2

`-gltsym 'SYM: Vrel -Arel' -glt_label 1 V-A`

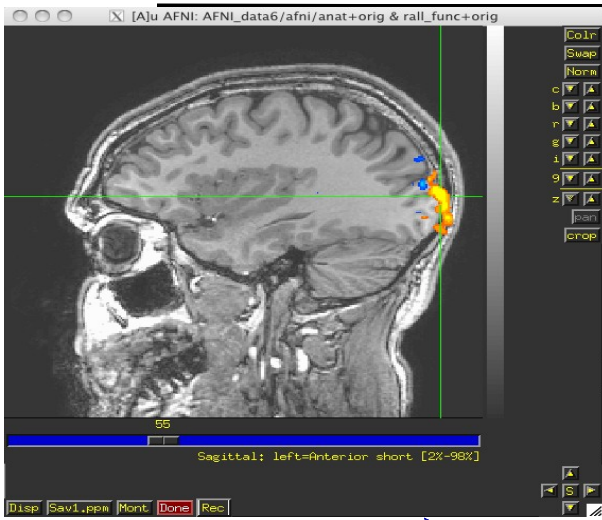
- **GLT**s: **G**eneral **L**inear **T**ests
- **3dDeconvolve** provides test statistics for each regressor separately, but to test combinations of the β weights in each voxel, we need **`-gltsym`** option
- Example above tests the difference between the β weights for the **`Virual-reliable`** and the **`Audio-reliable`** responses
  - **`SYM:`** means symbolic input is on command line
    - Otherwise inputs will be read from a file
  - Symbolic names for each regressor taken from **`-stim_label`** options
  - Stimulus label can be preceded by **`+`** or **`-`** to indicate sign to use in combination of β weights
  - Leave space after each label!
- Goal is to test a linear combination of the β weights
  - Null hypothesis $\beta_{Vrel} = \beta_{Arel}$
  - e.g., does **`Vrel`** get different response from **`Arel`**?
- What do **`'SYM: 0.5*Vrel +0.5*Arel'`** and **`'SYM: Vrel \ Arel'`** test?

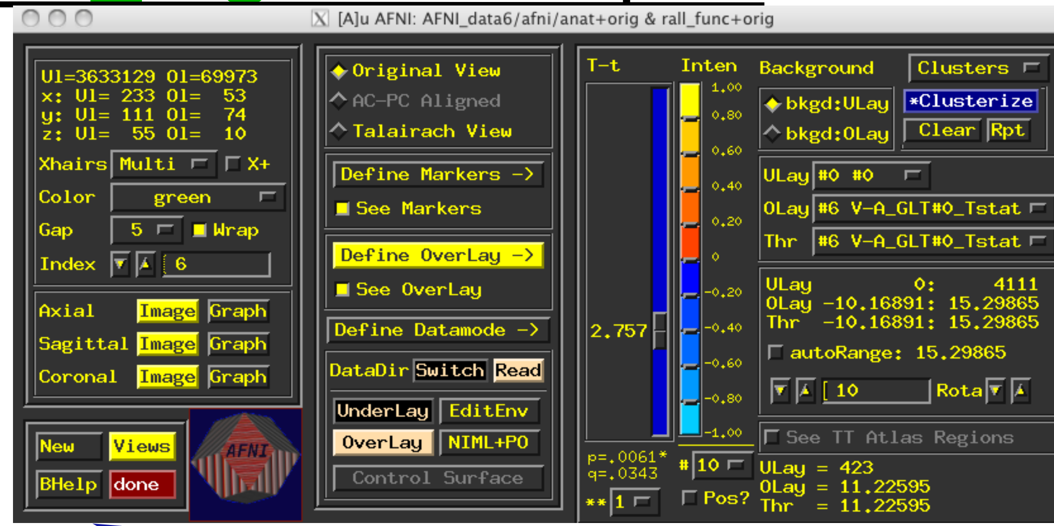
Options in 3dDeconvolve - 3

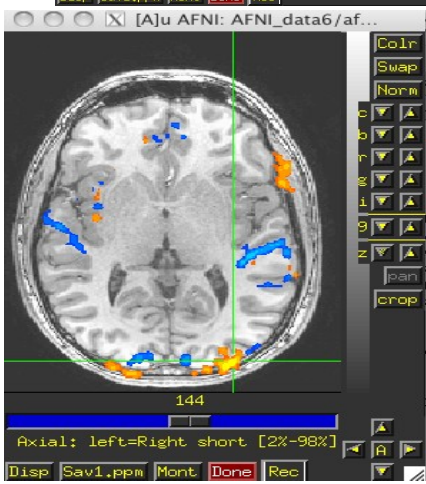
`-fout -tout` = output both *F*- and *t*-statistics for each stimulus class (**`-fout`**) and stimulus coefficient (**`-tout`**) — but not for the baseline coefficients (use **`–bout`** for baseline)

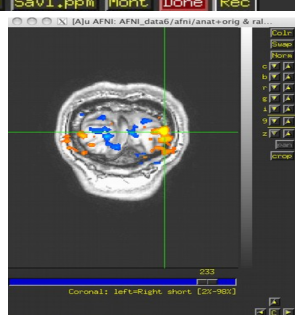
- The full model statistic is an *F*-statistic that shows how well all the regressors of interest explain the variability in the voxel time series data
  - Compared to how well *just* the baseline model time series fit the data times (in this example, we have 12 baseline regressor columns in the matrix — 6 for the linear drift, plus 6 for motion regressors)
  - $F = [SSE(r) - SSE(f)] / df(\mathrm{n}) \div [SSE(f) / df(\mathrm{d})]$
- The individual stimulus classes also will get individual *F*- (if **`–fout`** added) and/or *t*-statistics indicating the significance of their individual *incremental* contributions to the data time series fit
  - If DF=1 (e.g., F for a single regressor), ***t*** is equivalent to *F*: $t(n) = F^2(1, n)$

Results of **rall_regress** Script



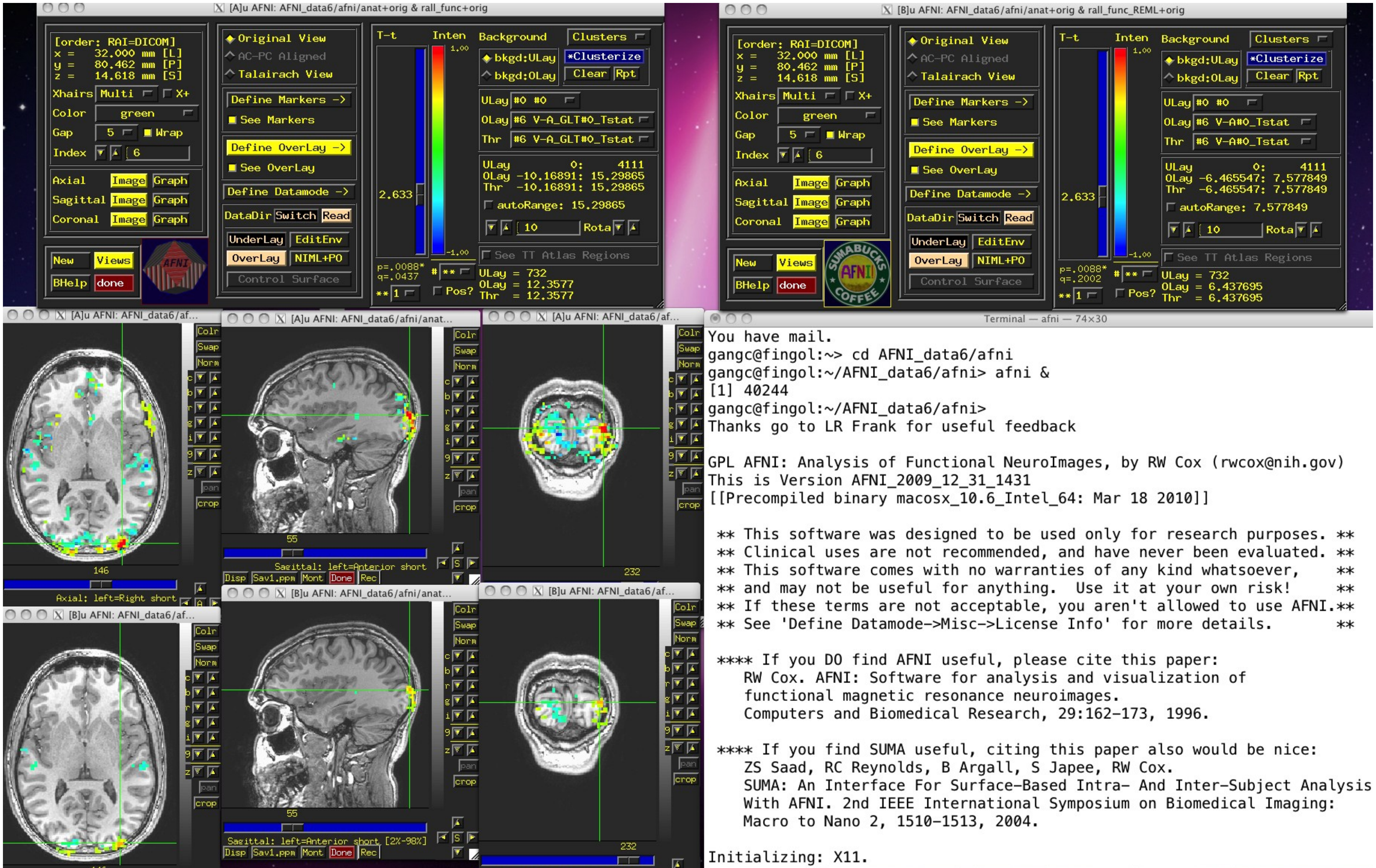






- Images showing results from third GLT contrast: **VrelvsArel**
- Menu showing labels from **3dDeconvolve**
- Play with these results yourself*!*

Compare 3dDeconvolve and 3dREMLfit



Group Analysis: will be carried out on β or **GLT** coef (+*t*-value) from single-subject analysis