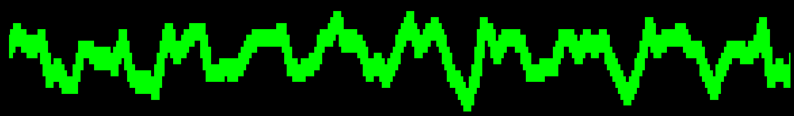
# First-level fMRI modeling

Jeanette Mumford
University of Wisconsin - Madison

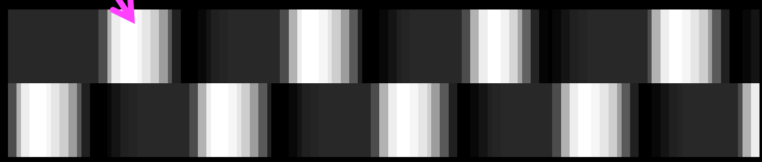
# Quick GLM overview

Data: BOLD
time series



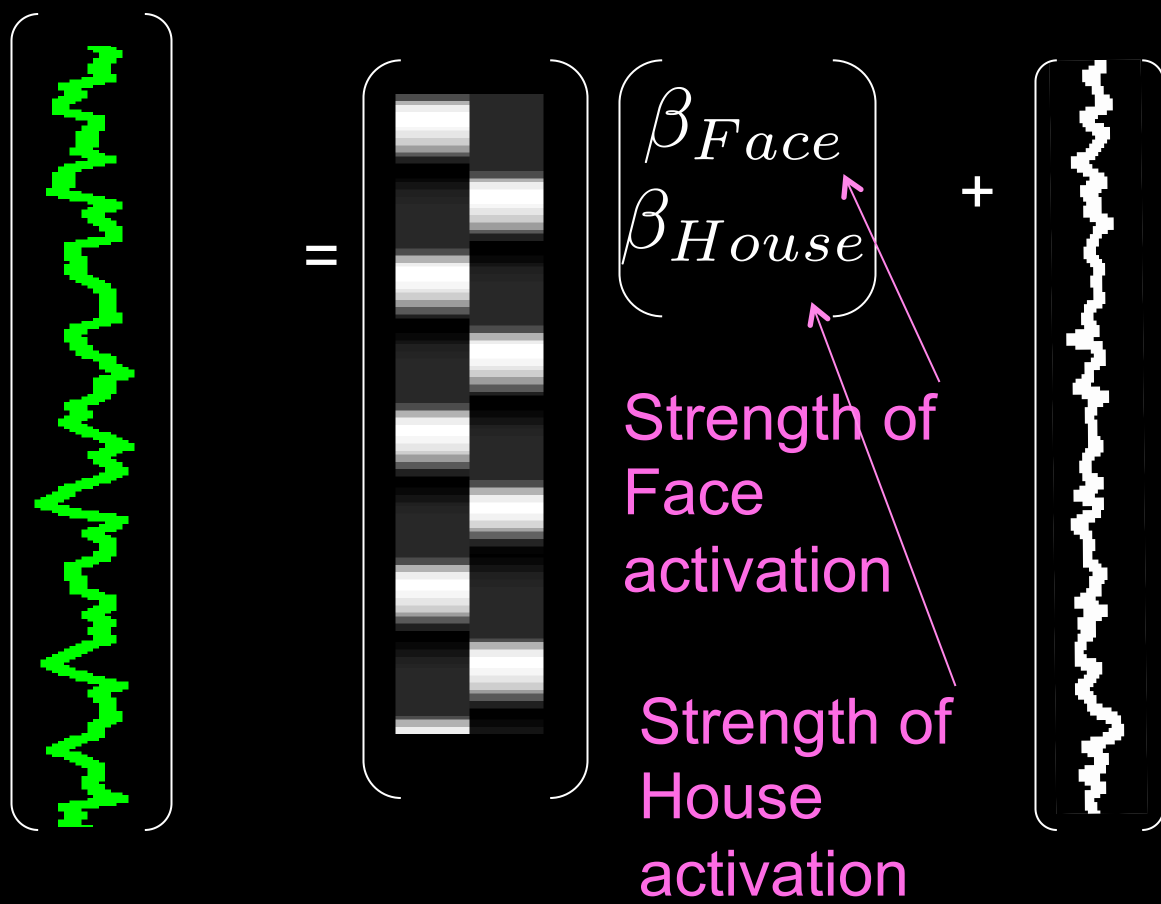
Model: Expected
BOLD response
for 2 blocked
tasks

When Houses occur

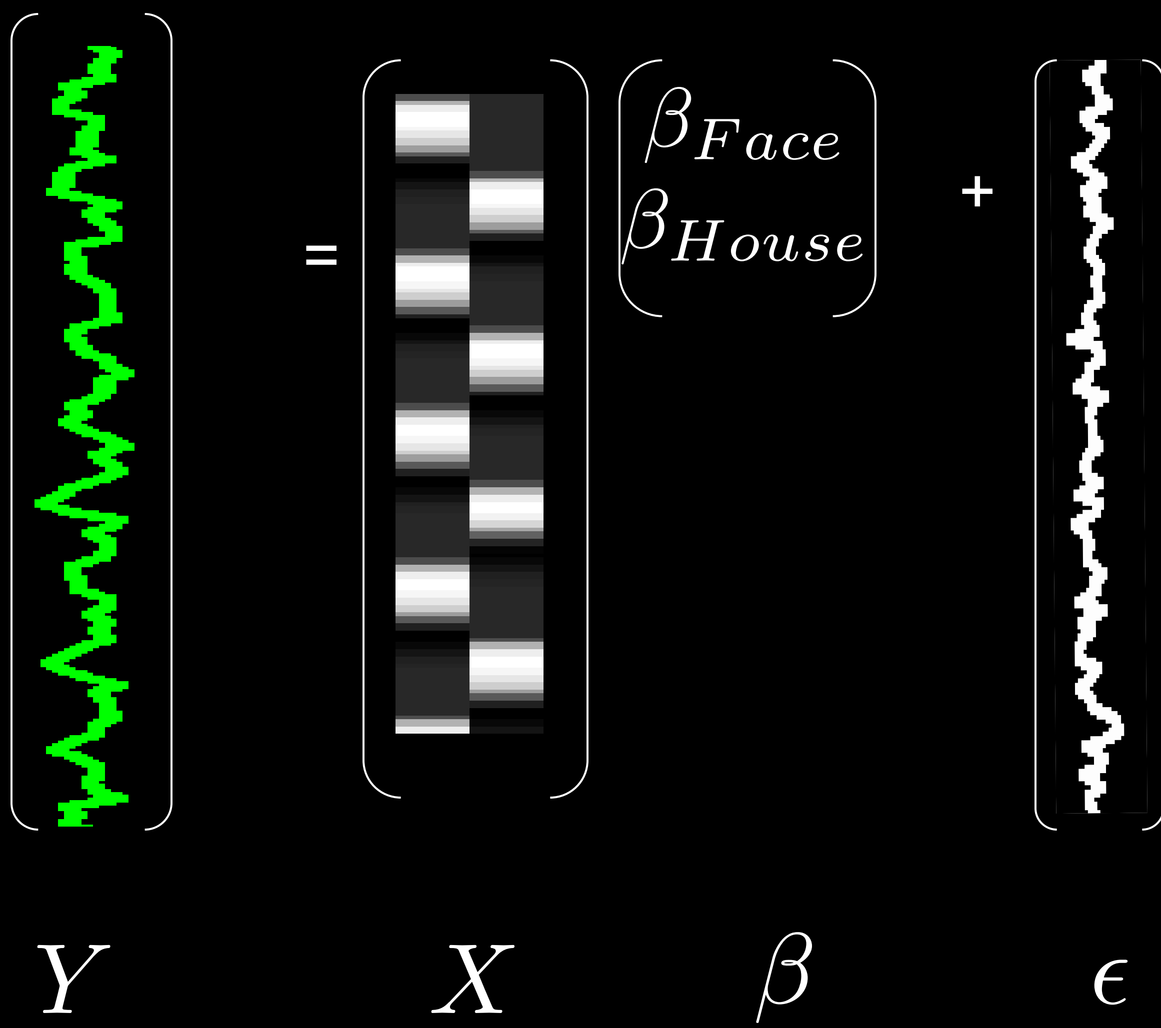


When Faces occur

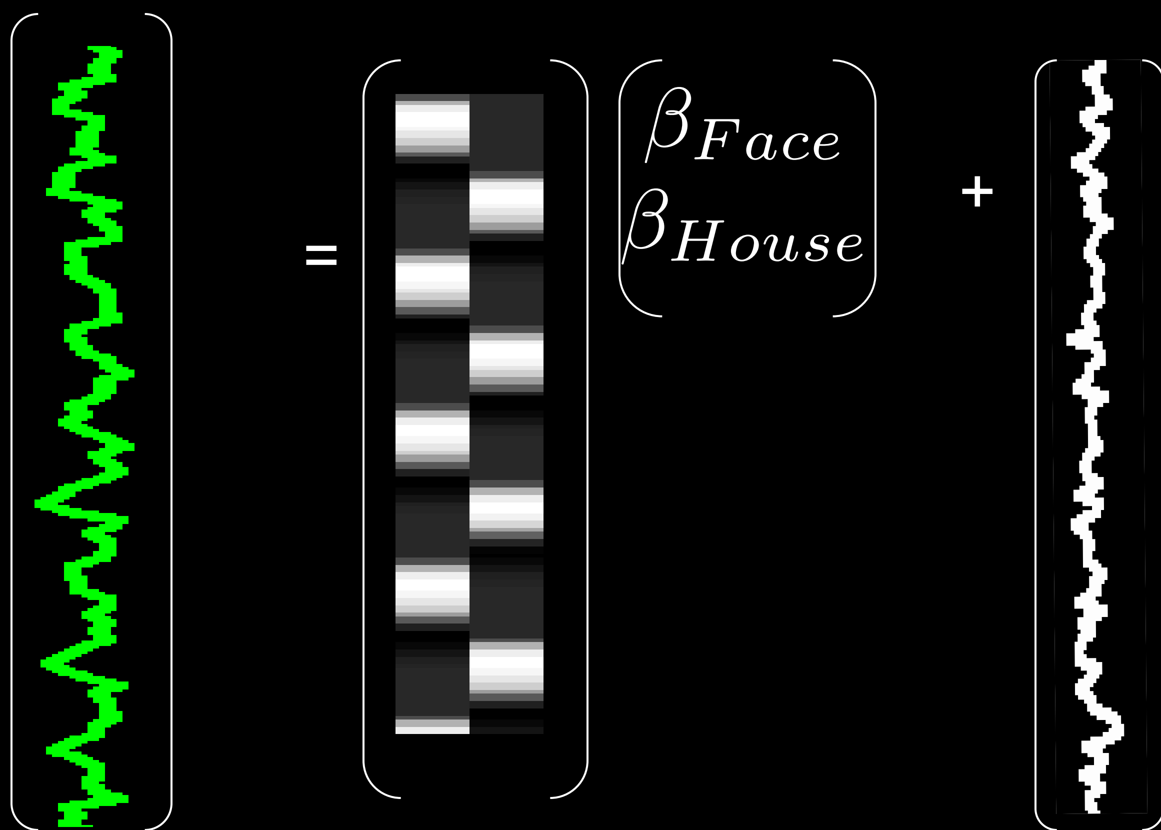
# Quick GLM overview



# Quick GLM overview



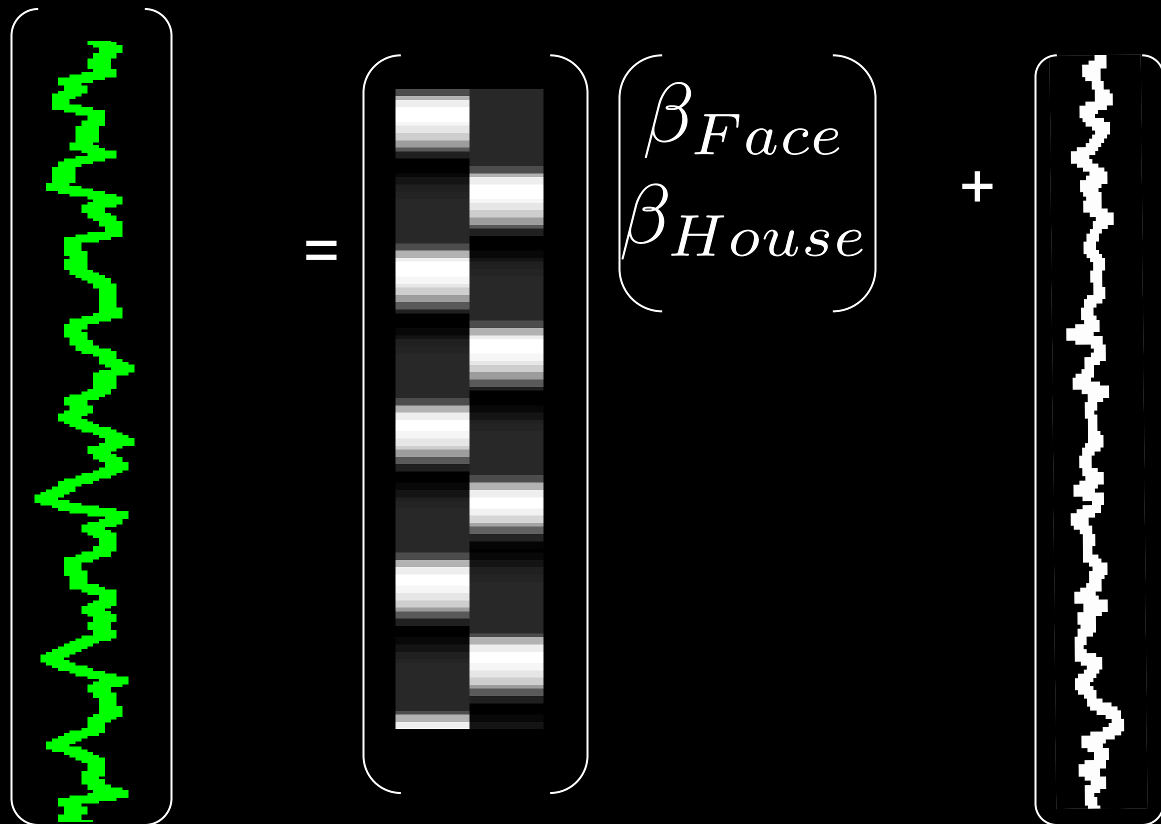
# Quick GLM overview




How do we find the FFA?

$$\beta_{Face} > \beta_{House} \rightarrow \beta_{Face} - \beta_{House} > 0$$

# Quick GLM overview




$$\begin{bmatrix} 1 & -1 \end{bmatrix} \begin{bmatrix} \beta_{Face} \\ \beta_{House} \end{bmatrix} > 0$$

# What do we need to know?

- What is the general structure of a t-statistic?

- What are the assumptions we make with the linear model (Gauss Markov)?

covariance is the same across time (homoskedesicity)
&& independence of observations

- What is the residual?

# T-statistic structure

$$\frac{\text{contrast you are estimating}}{\sqrt{\text{estimated variance of contrast you are estimating}}}$$

$$= \frac{\text{contrast you are estimating}}{\sqrt{(\text{variance contribution from model})(\text{residual variance})}}$$

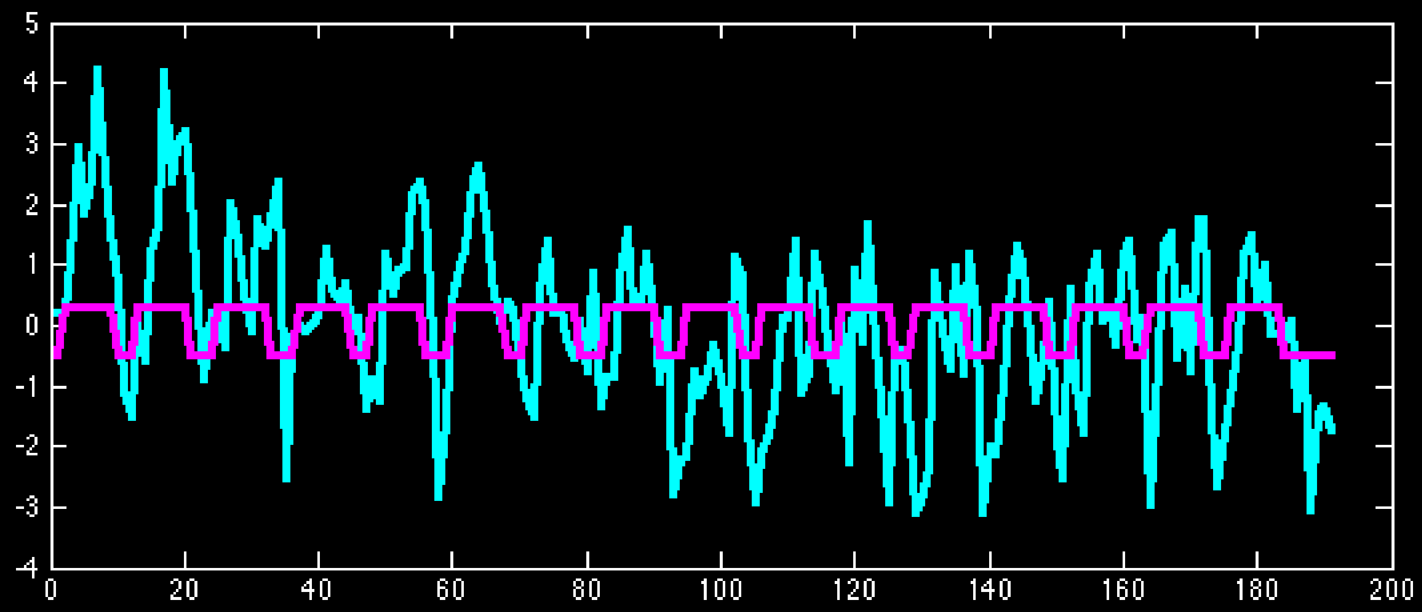
$$= \frac{c(X'X)^{-1}X'Y}{\sqrt{c(X'X)^{-1}c'\hat{\sigma}}}$$

In a new experiment the numerator & the residual variance are unknown. But, can run an efficiency calculation to estimate size of the t stat by focusing on the final component, the variance contribution from the model.

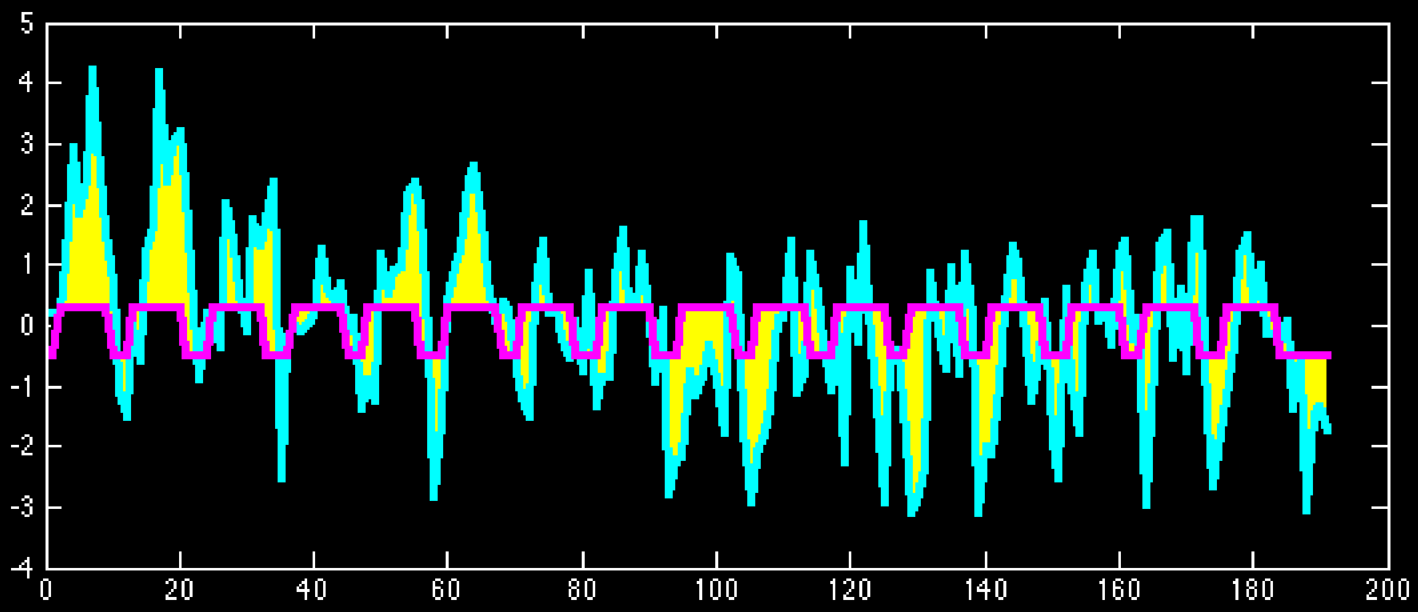
# Gauss Markov Theorem

- If the errors
  - Have mean 0
  - Are uncorrelated
  - Have the same variance
- The least squares estimates are unbiased and have minimum variance among all unbiased estimators!

# Goal

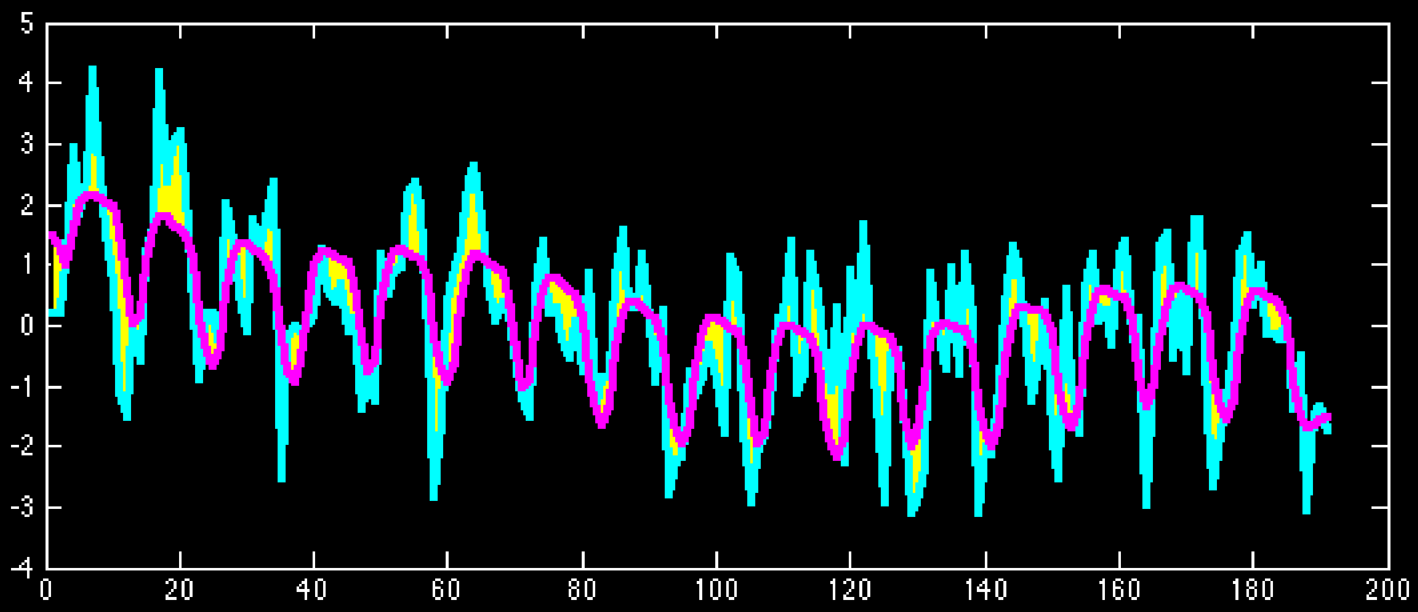


# Goal



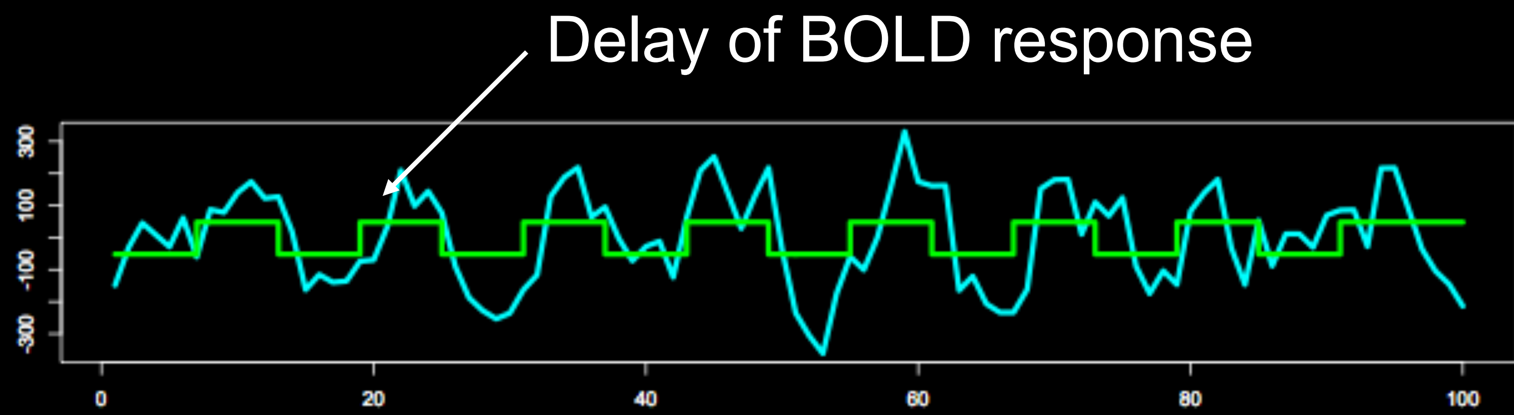

Residual shown here in yellow

# Goal



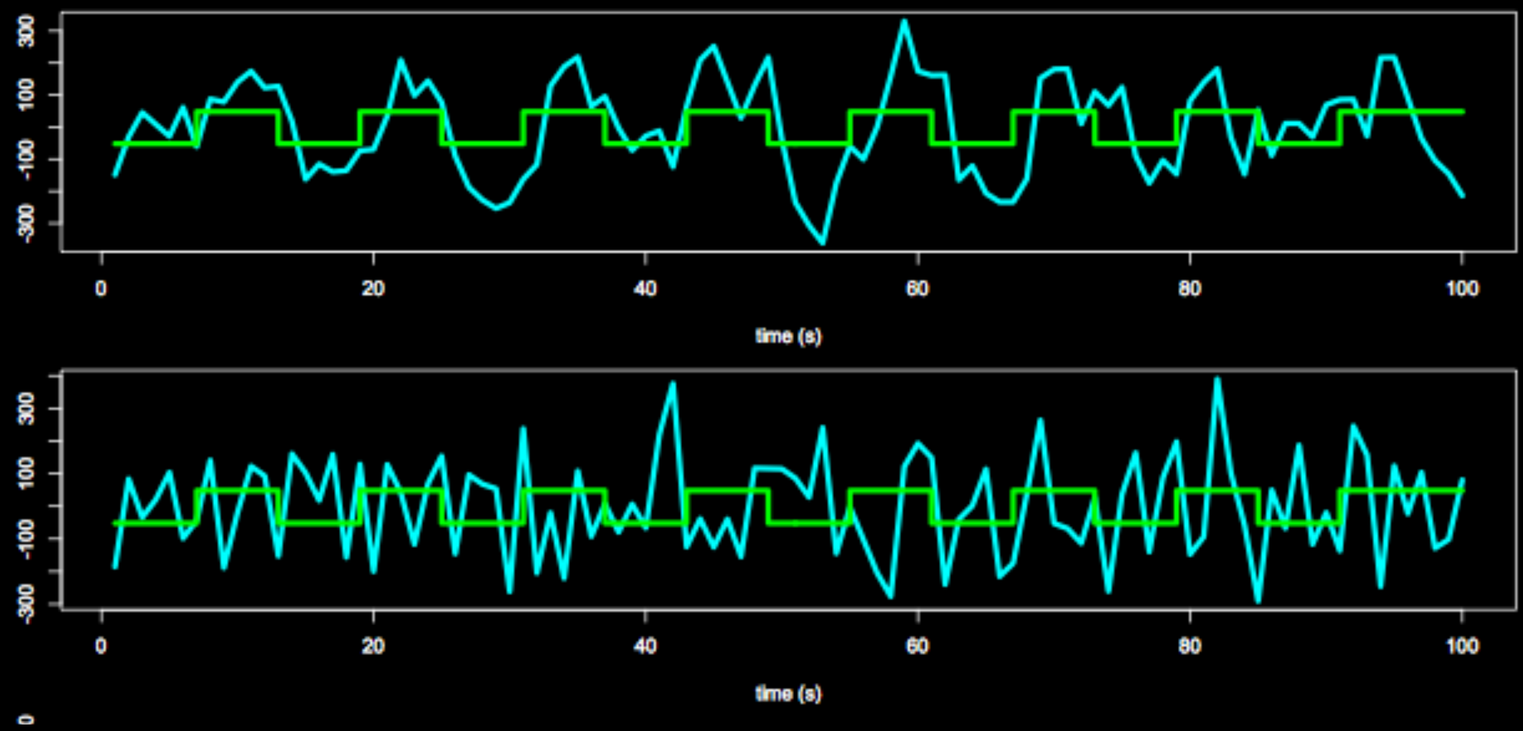

Who wants smaller residuals??
You do!!

Voxel with signal



Voxel
with
signal

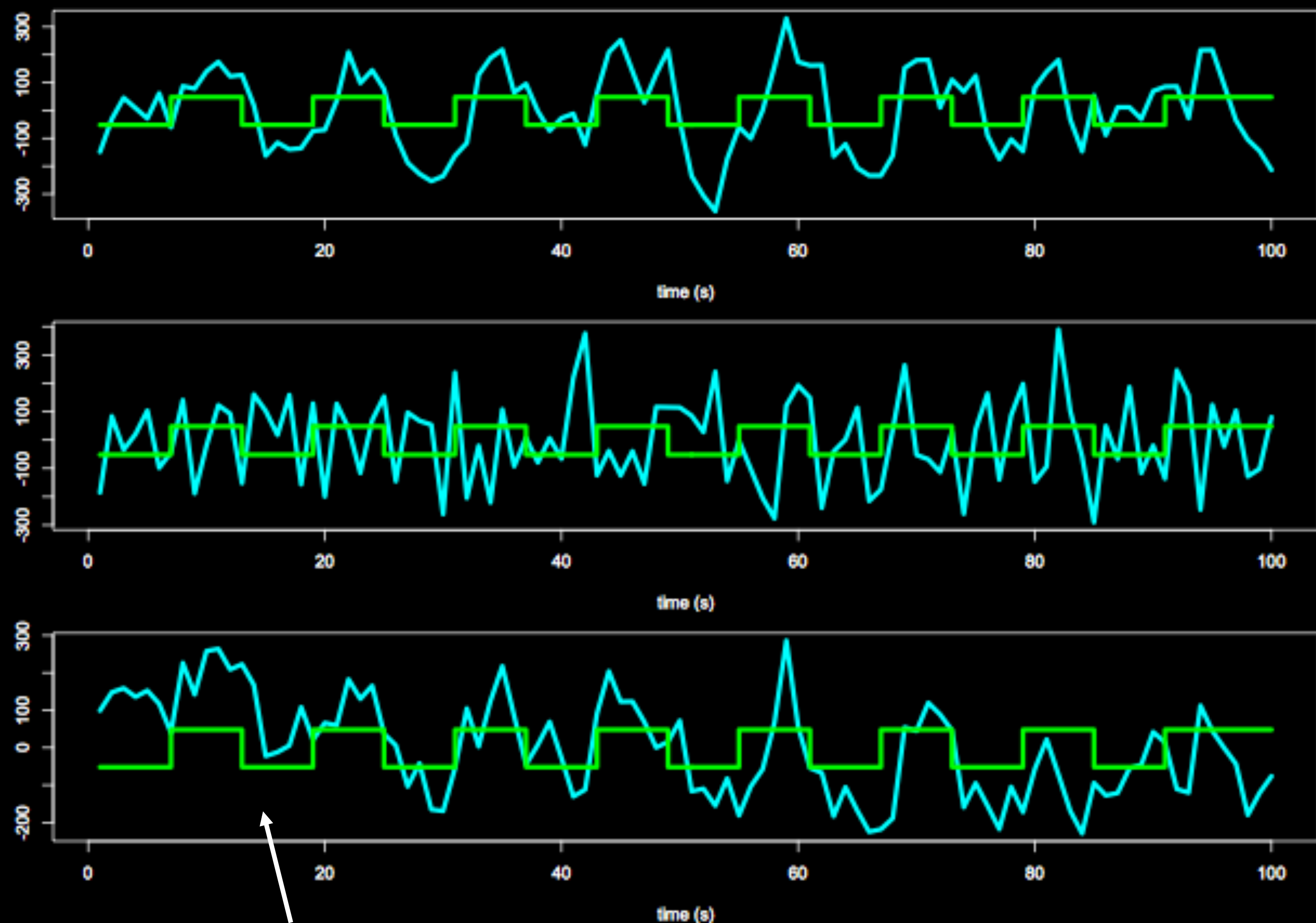
Voxel
no
signal



Voxel with signal

Voxel no signal

Voxel signal and drift

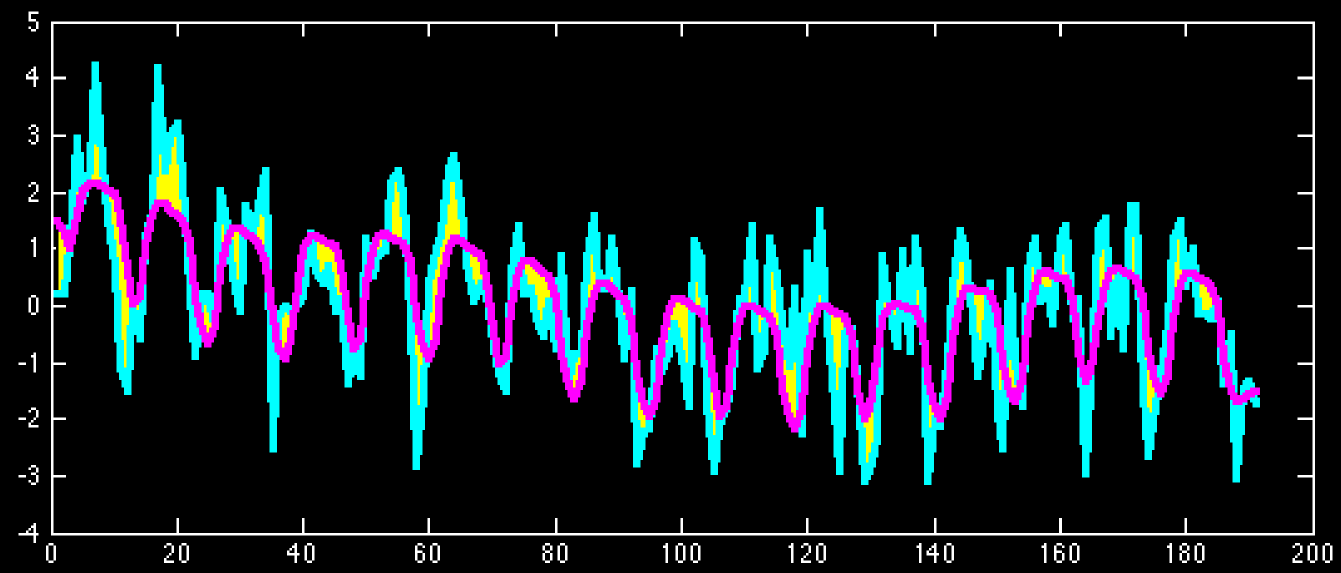



Starts off high

# BOLD issues

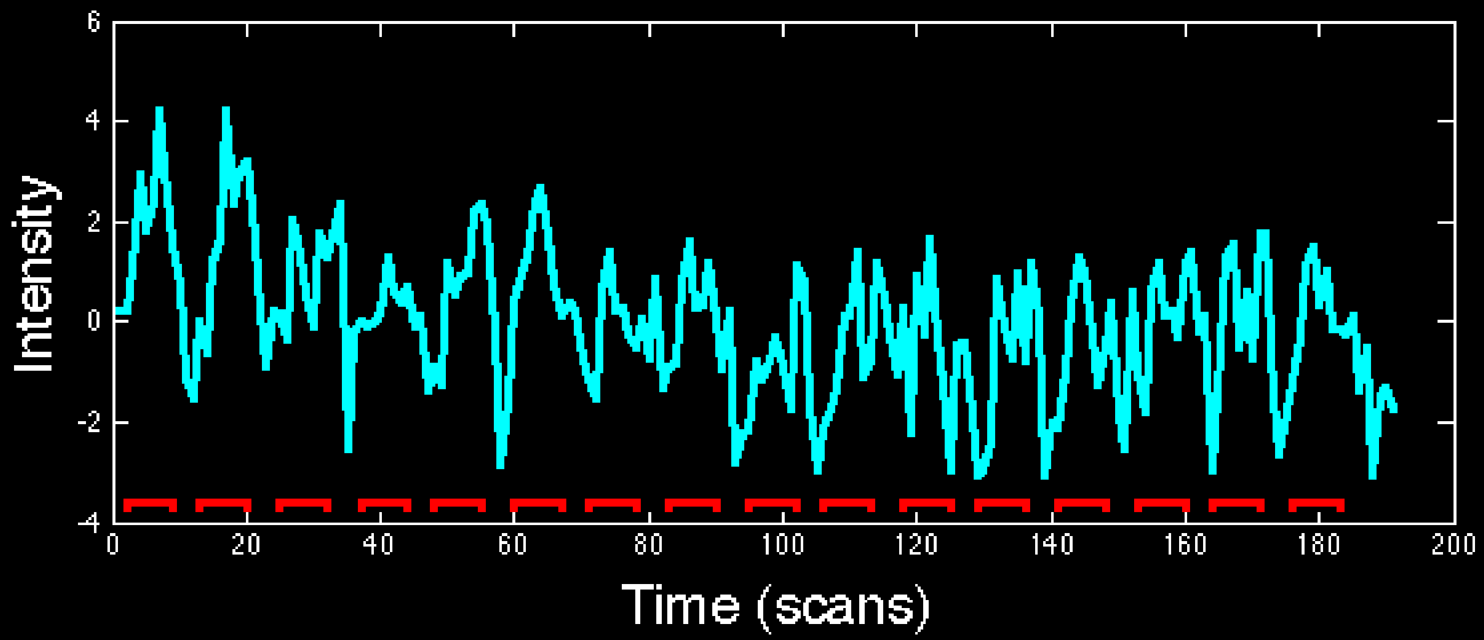
- BOLD response is delayed
  - Convolution
  - FIR modeling    finite impulse response modeling
- BOLD time series suffer from low frequency noise
  - Highpass filtering
  - Prewhitening
- Scaling the data
  - Grand mean scaling
  - Intensity normalization

# How to make a good model




Big residuals ⇒Big variance⇒Small t stat

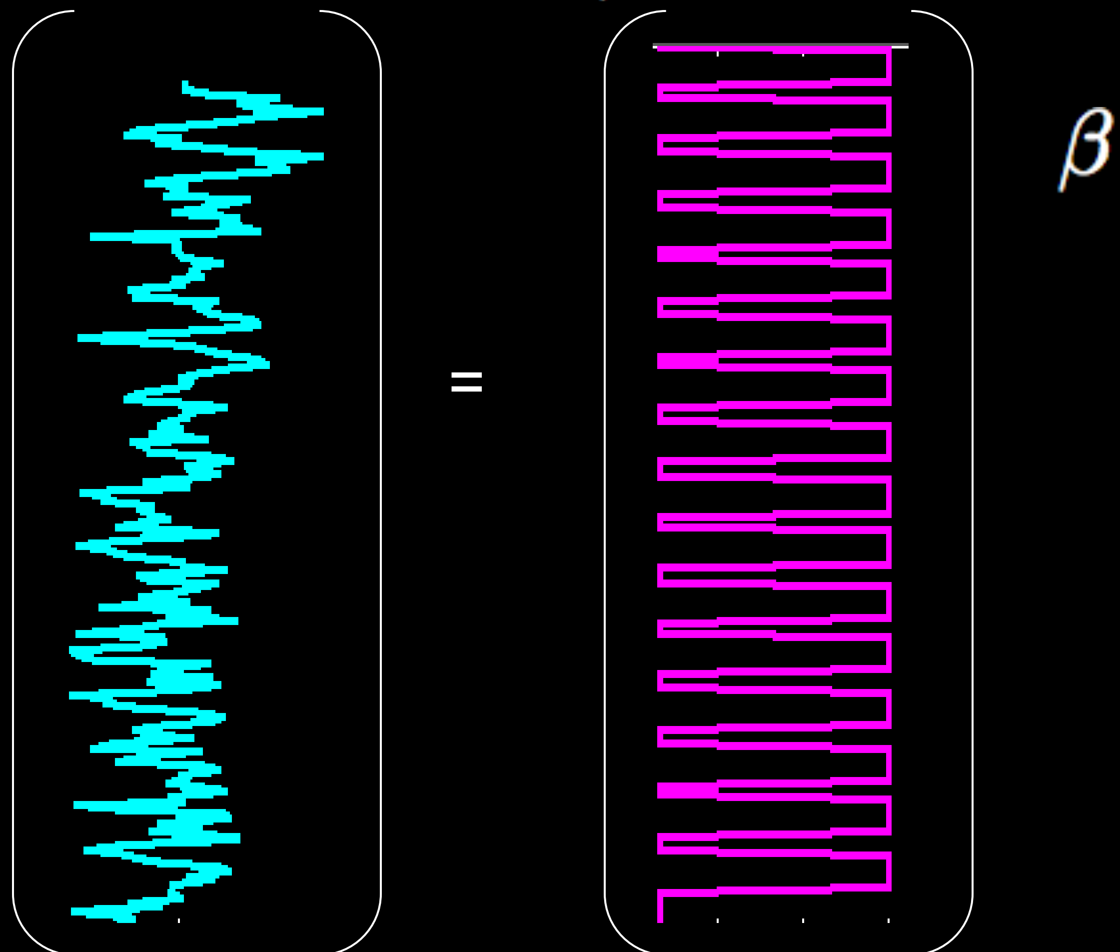
# Understanding the data



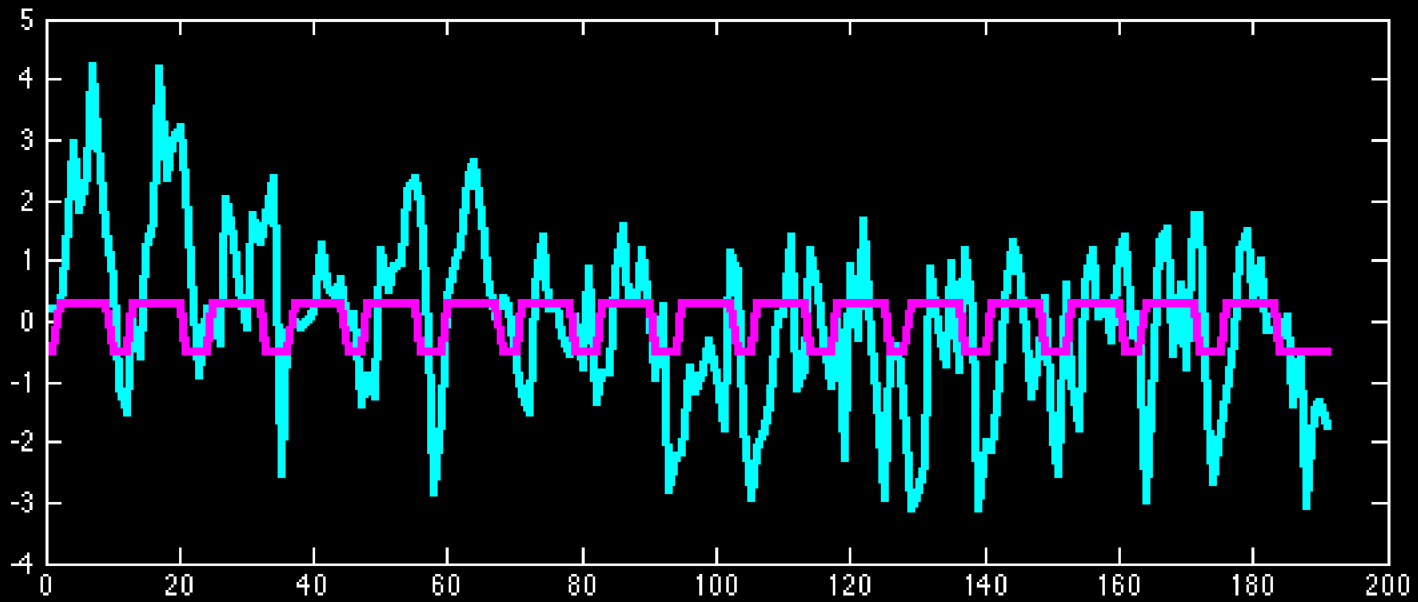

- Time series drifts down in beginning
- BOLD response is delayed

# Simplest Model

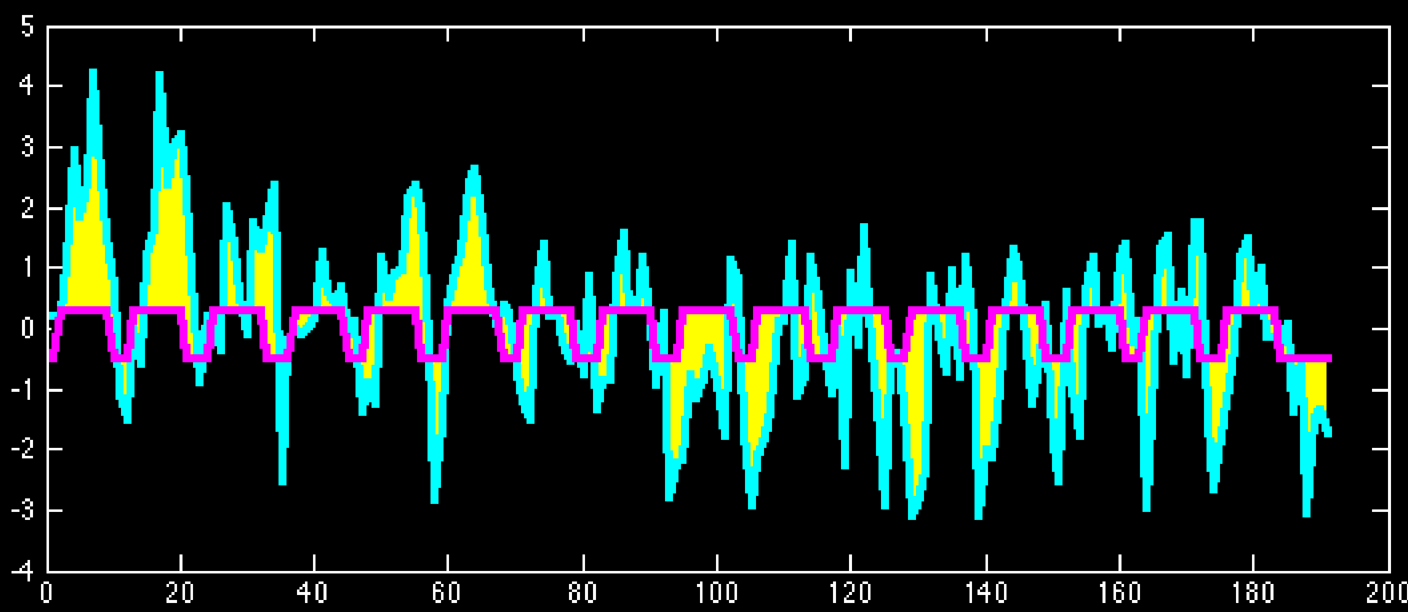
$$Y = X\beta$$



# Simplest Model



# Simplest Model




$$t = \frac{c(X'X)^{-1}X'Y}{\hat{\sigma}\sqrt{c(X'X)^{-1}c'}}$$

$$t = \frac{0.3}{1.41 \times 0.06} = 3.55$$

the green component is fixed b/c it is dependent upon the study design

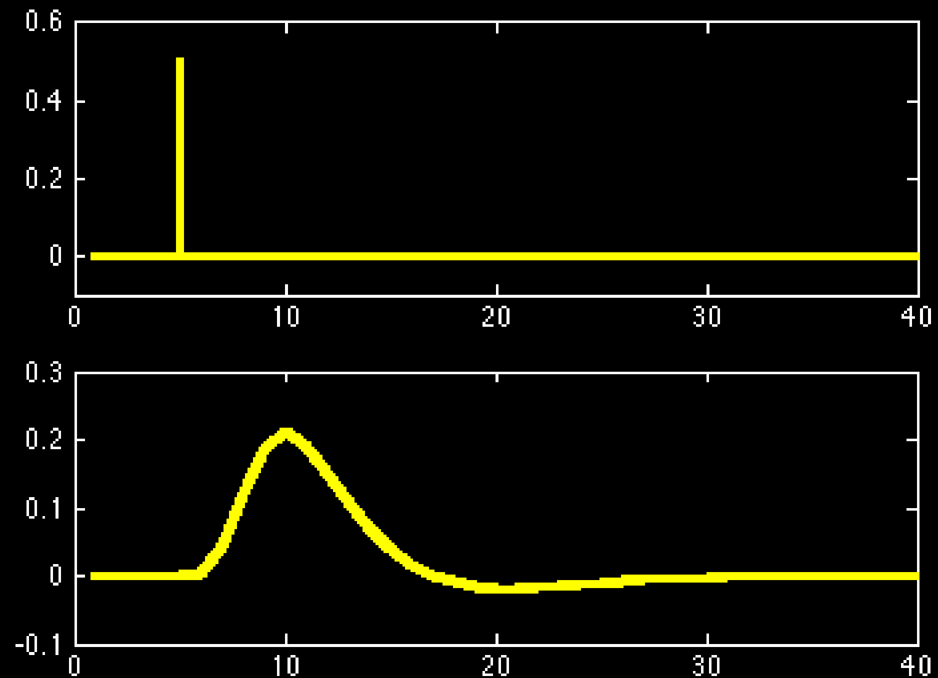
# Modeling the delay

- Hemodynamic response function
  - Real data was used to find good models for the hemodynamic response

Stimulus

HRF

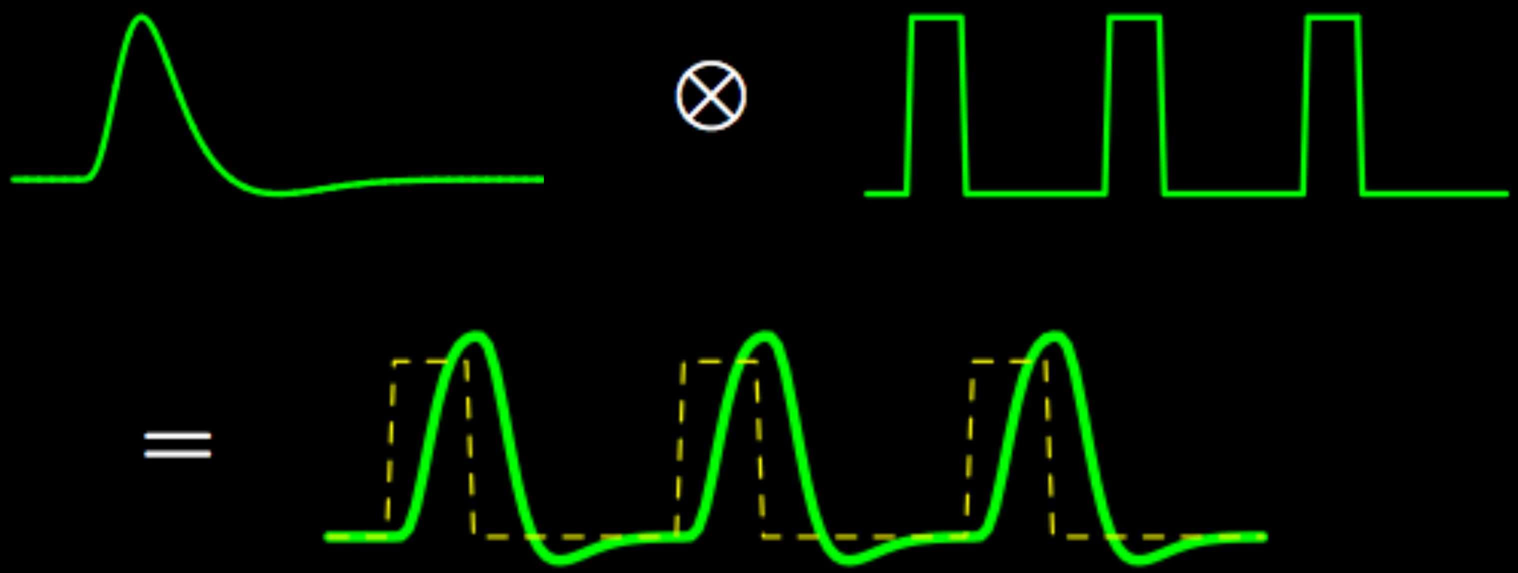
(double gamma)




but FSL's default is a single gamma?

# Convolution

- Combine HRF and expected neural response



Typically model derivative of convolved HRF to adjust for small differences in onset (<1s)

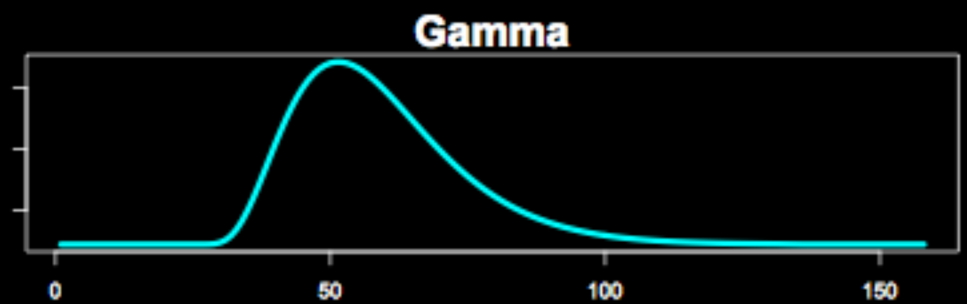
# Different HRF's

Too symmetric

# Different HRF's

Too symmetric

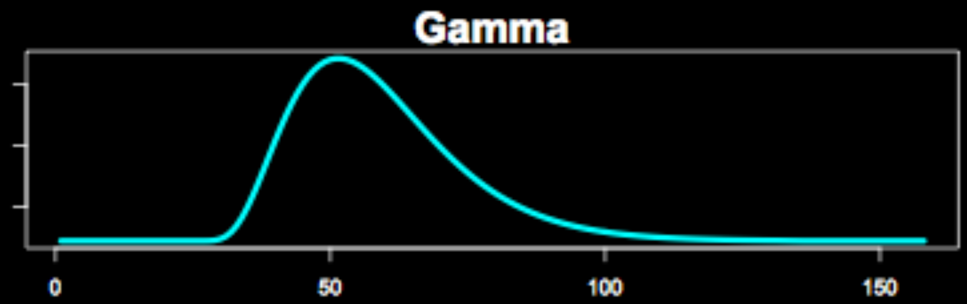

Basic shape okay, but no post stimulus undershoot



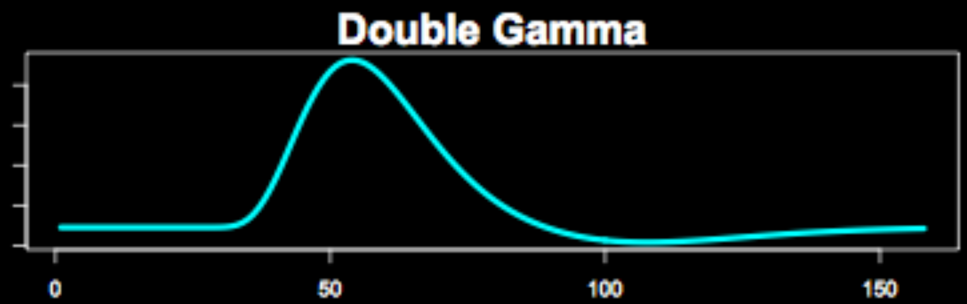
# Different HRF' s

Too symmetric


Basic shape okay, but no post stimulus undershoot

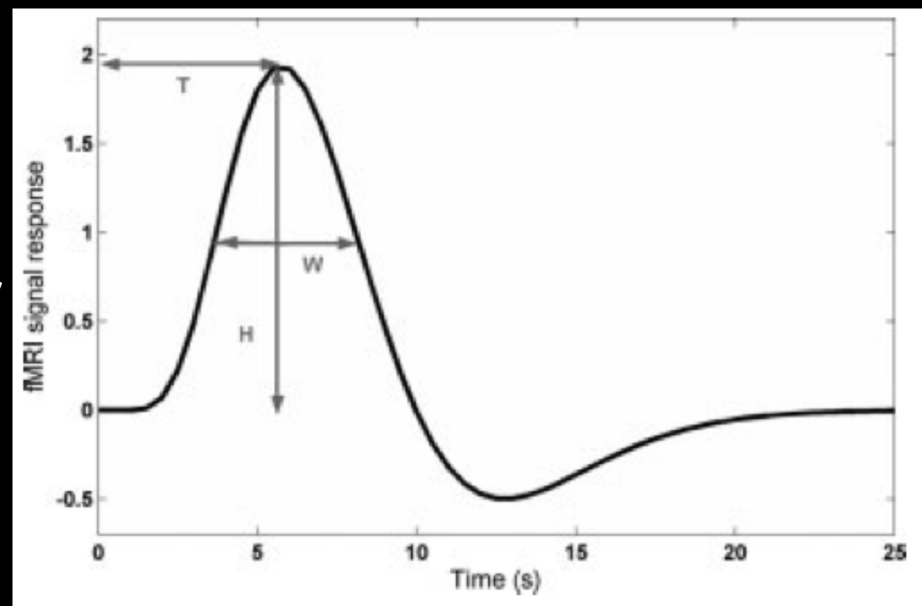



Includes post stimulus undershoot



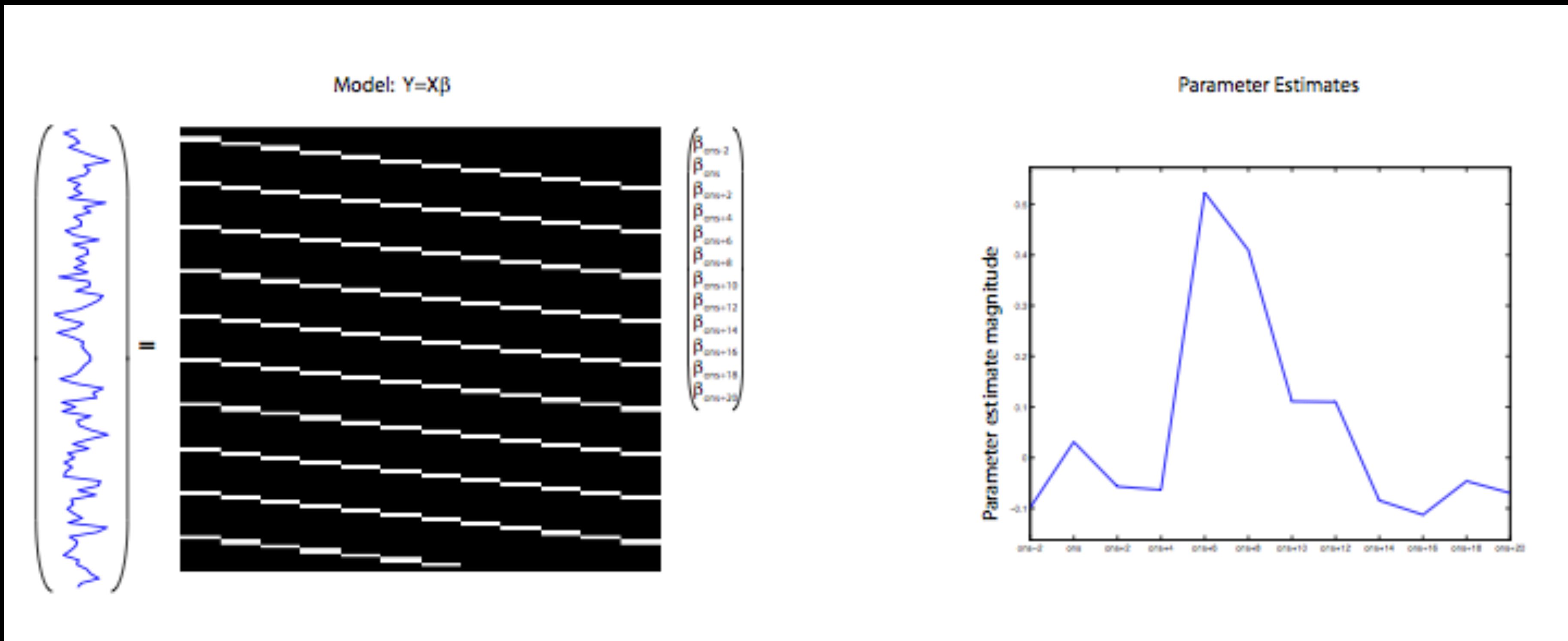
# Assumptions of canonical HRF

- The width, height and delay are correct
- Lindquist & Wager (2007)
  - Fit H/W/T separately
  - Works okay-ish



# Finite impulse response model

- FIR
  - Make no assumption about the shape of the HRF

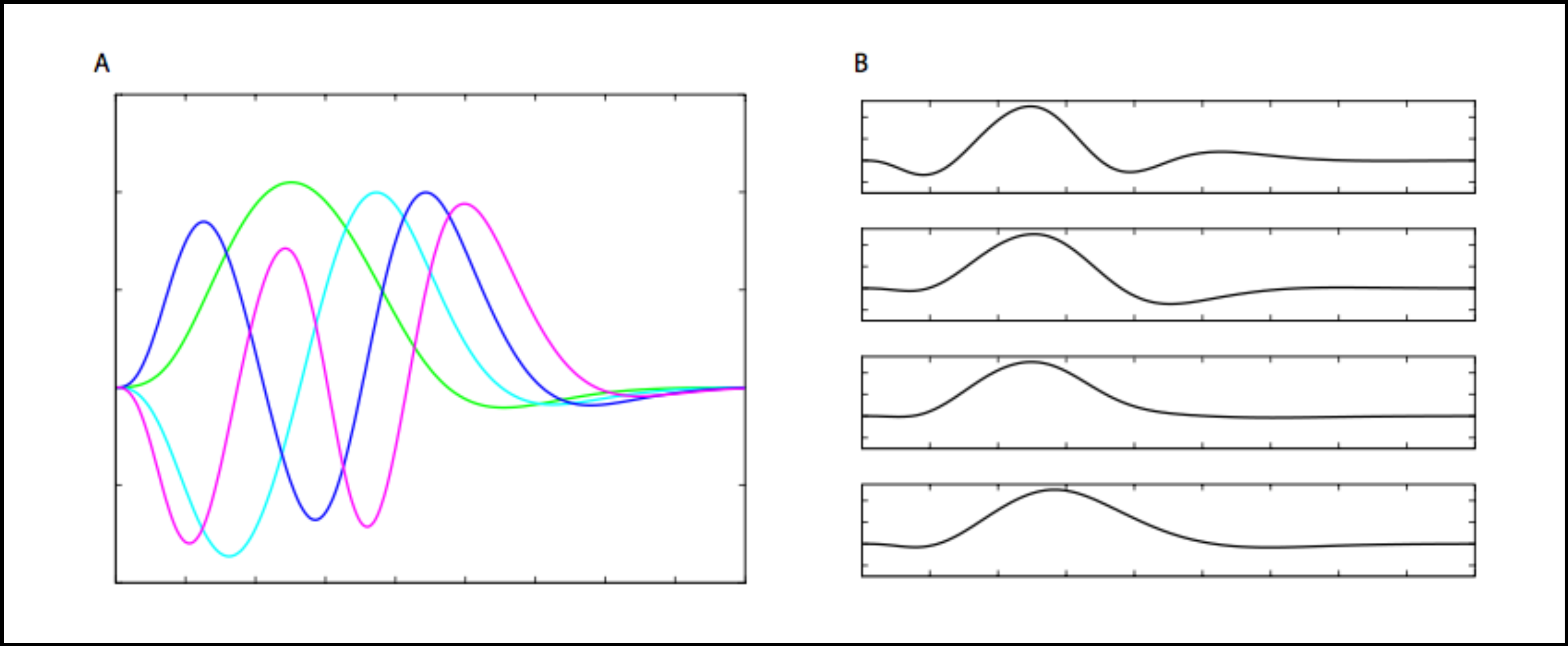


# Constrained basis set

- Lower the number of regressors in the model by using a basis set

- Constrained to shapes that are reasonable for HRF shapes

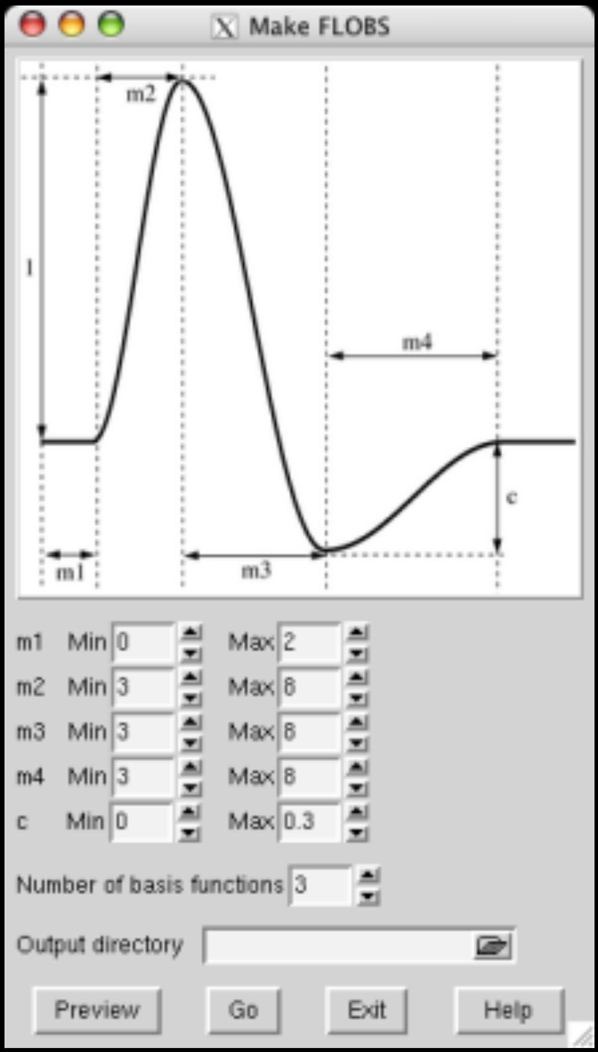
# Constrained basis set



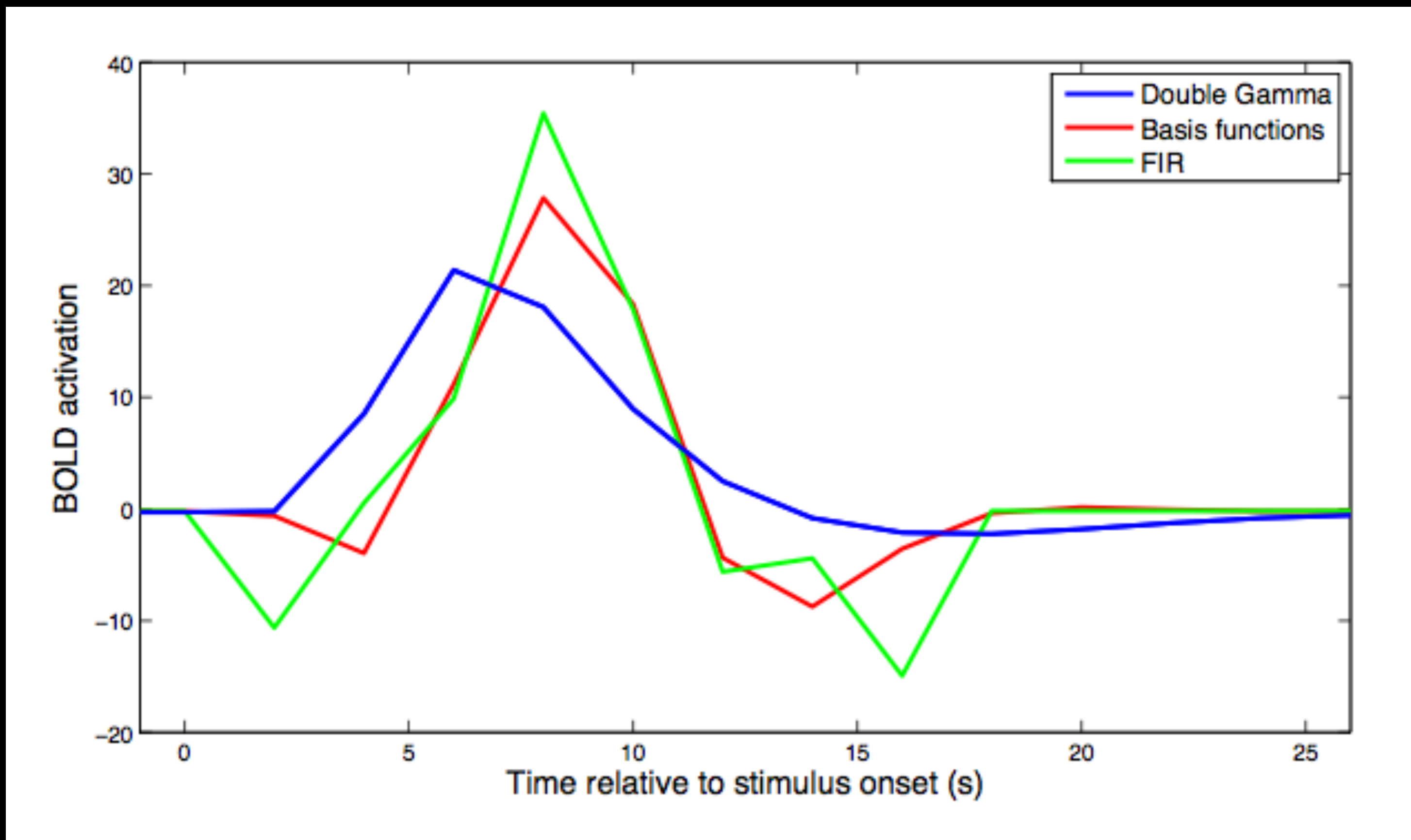

Basis set HRF possibilities

# FLOBS

- fMRIB Linear Optimal Basis Sets
  - Generates a set of basis sets to model signal
  - Specify ranges for different portions of the hrf



# Comparison



Double Gamma
Basis functions
FIR
BOLD activation
Time relative to stimulus onset (s)
40
30
20
10
0
−10
−20
0
5
10
15
20
25

# More thoughts about canonical HRF

- Advantages:
  - Simpler analysis
  - Easily interpretable outcome
  - Simplifies group analysis
- Disadvantages
  - Biased if canonical HRF is incorrect

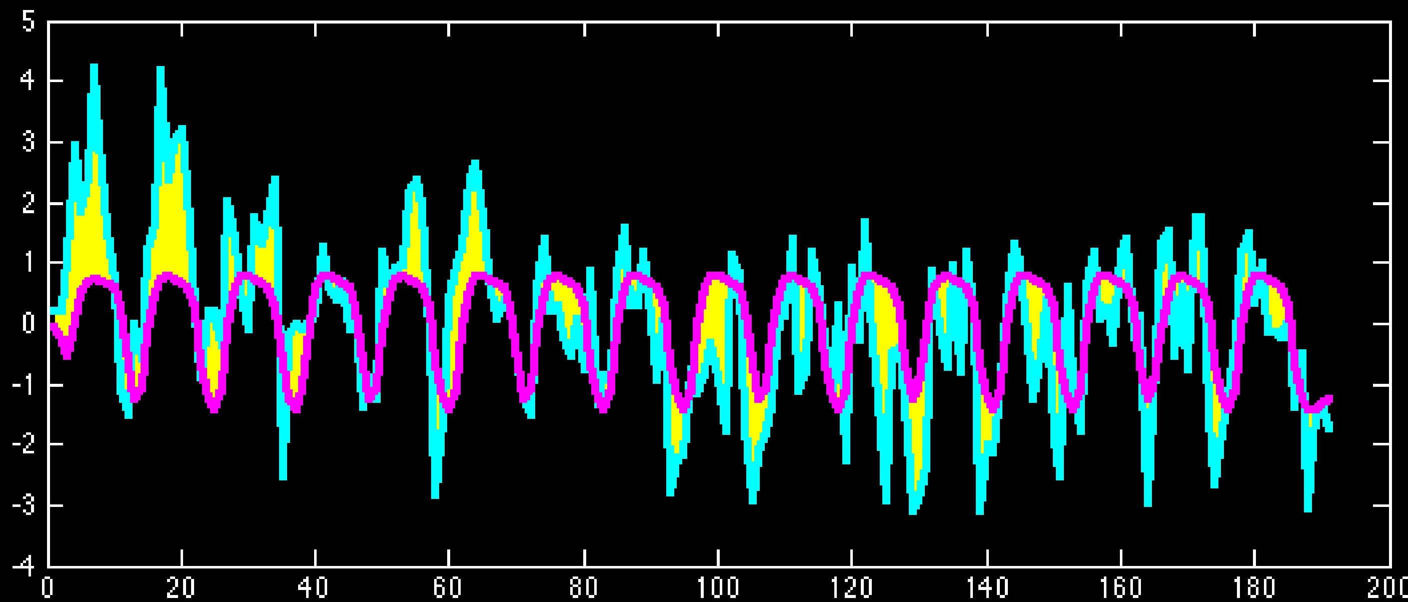
# Unbiased basis sets

- Advantages
  - Not biased towards a particular shape
  - Allows testing of hypotheses about specific HRF parameters
- Disadvantages
  - Less powerful
  - Makes group analysis more difficult
  - Tend to overfit the data (i.e., fit noise)

# We can make a design matrix!

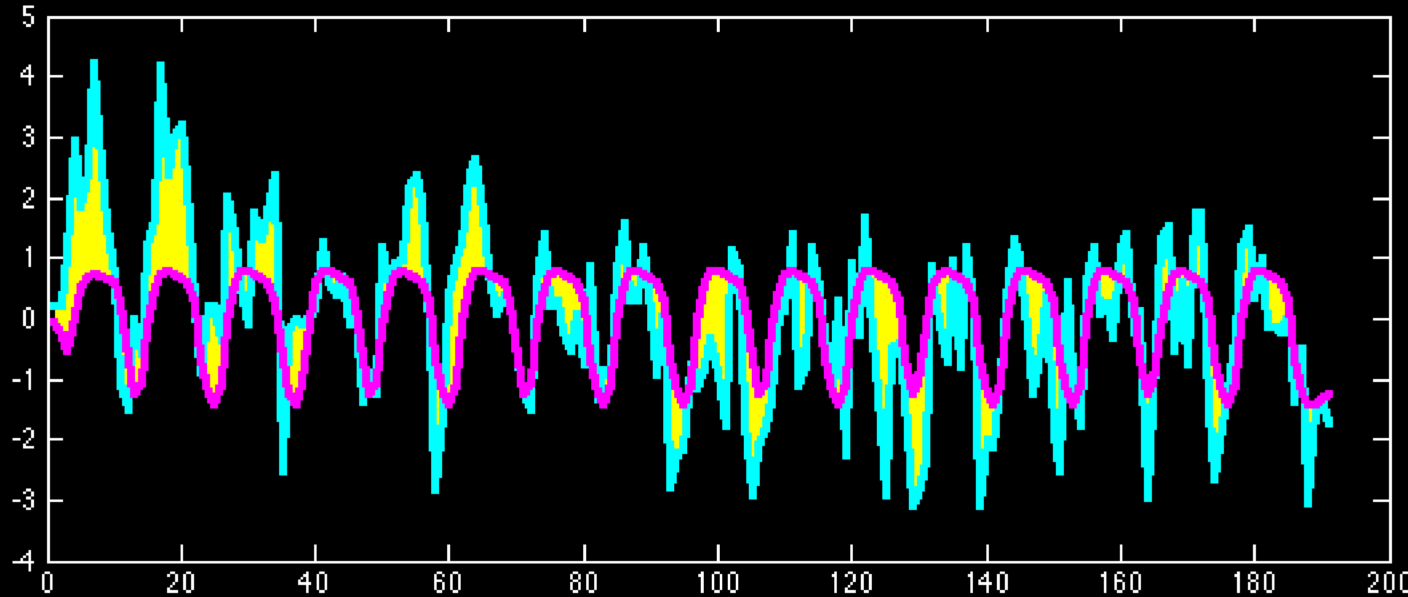
- Start with task blocks or delta functions
- Convolve
- Estimate the GLM and carry out hypothesis!

# Convolved Boxcar




$$t = \frac{0.66}{1.22 \times 0.06} = 9.02$$

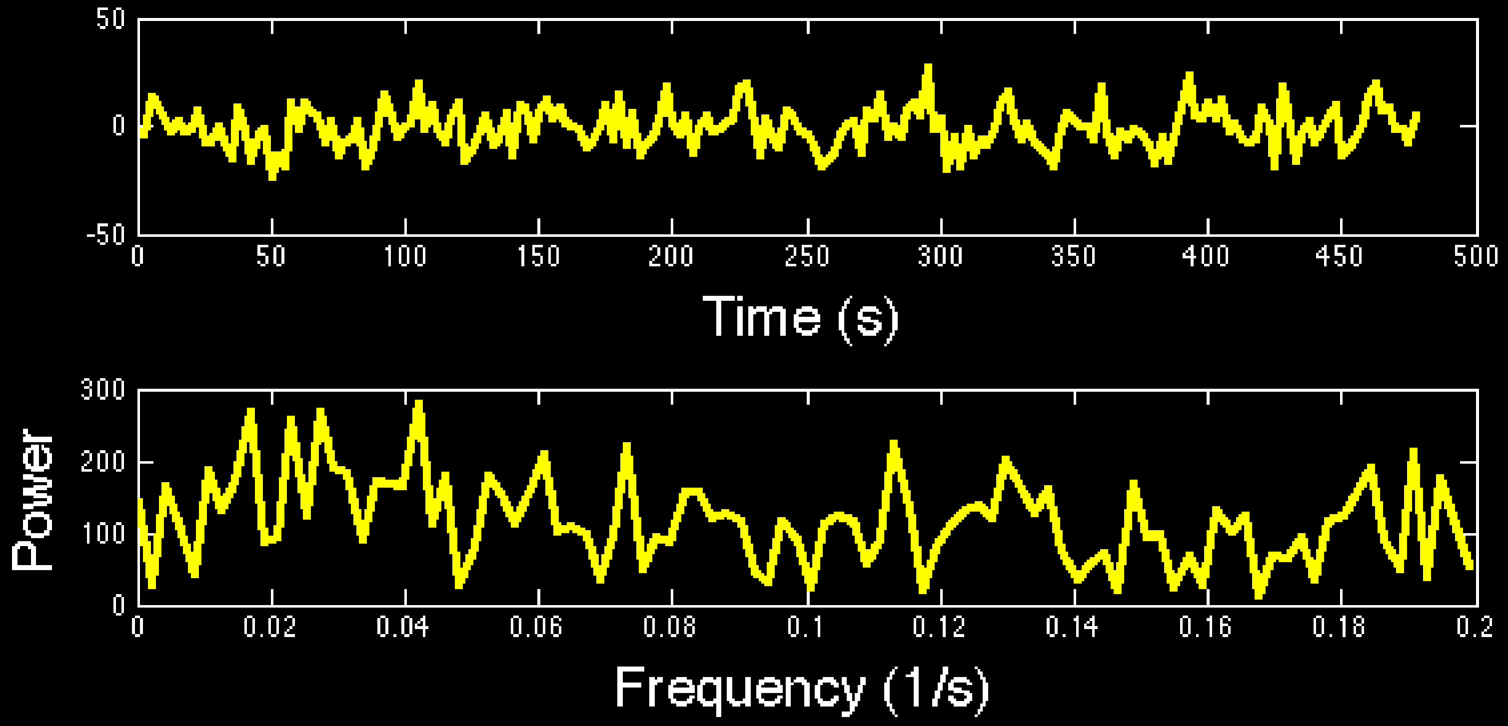
# Convolved Boxcar




$$t = \frac{0.66}{1.22 \times 0.06} = 9.02$$

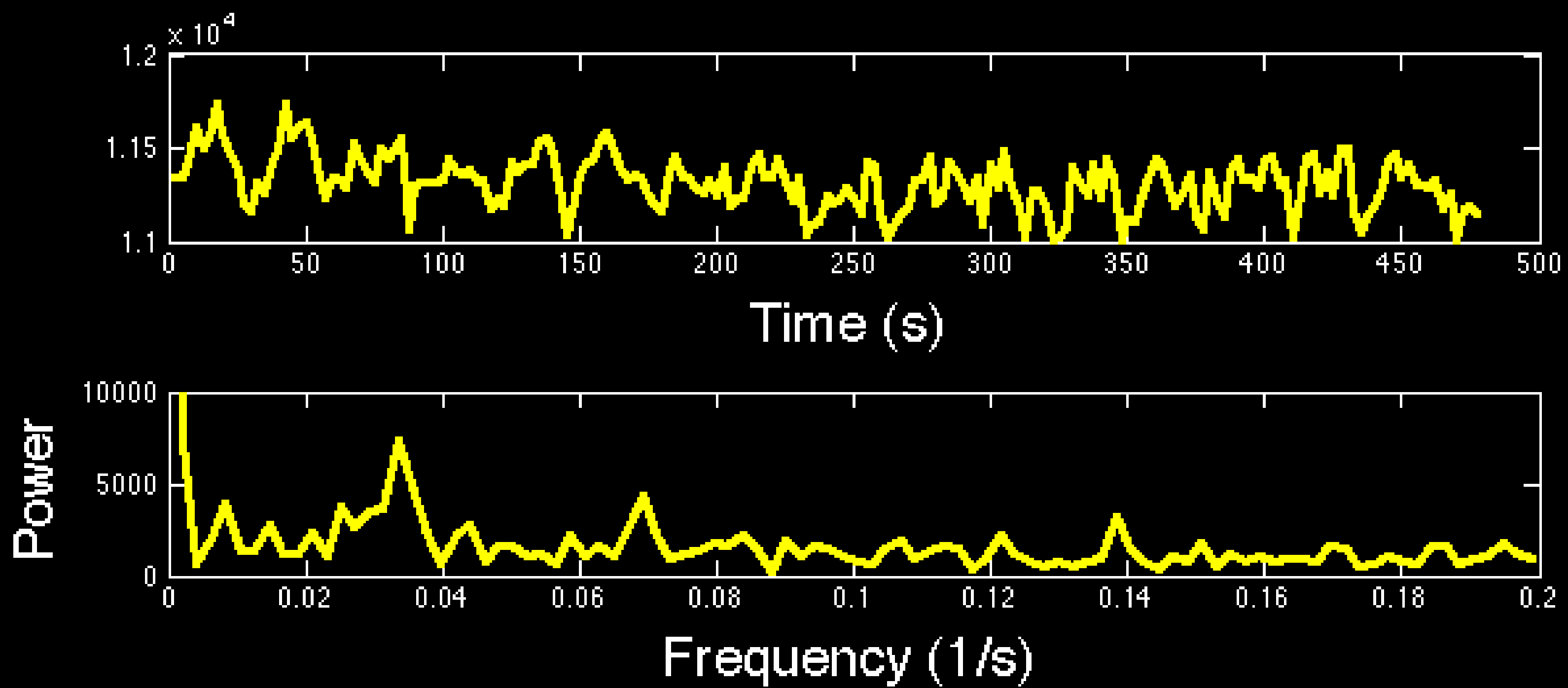
# The Noise

- White noise
  - All frequencies have similar power
  - Not a problem for OLS



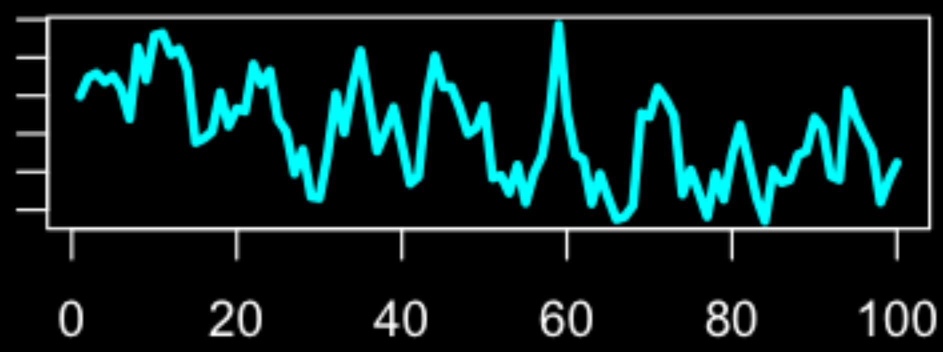
# More Noise

- Colored noise
  - Has structure
  - OLS needs help!

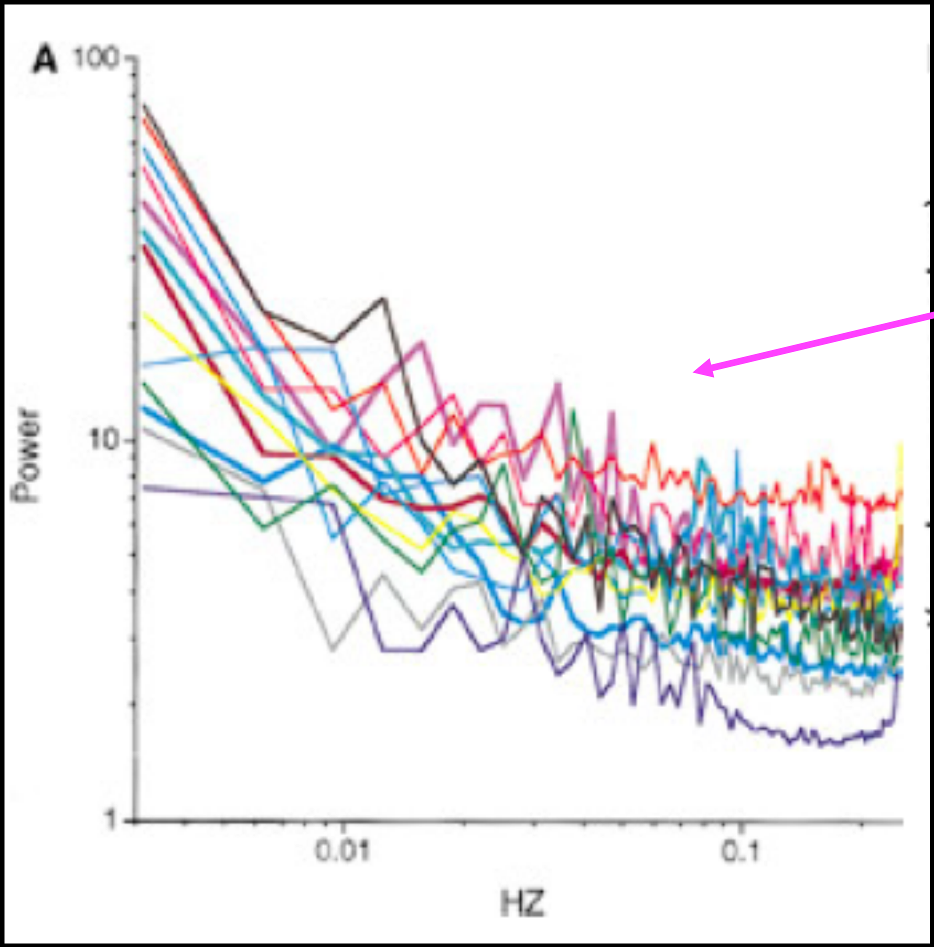


# What about the drift?

- Sources
  - Head motion
  - Cardiac noise
  - Respiratory noise
  - Scanner noise



# What the noise looks like

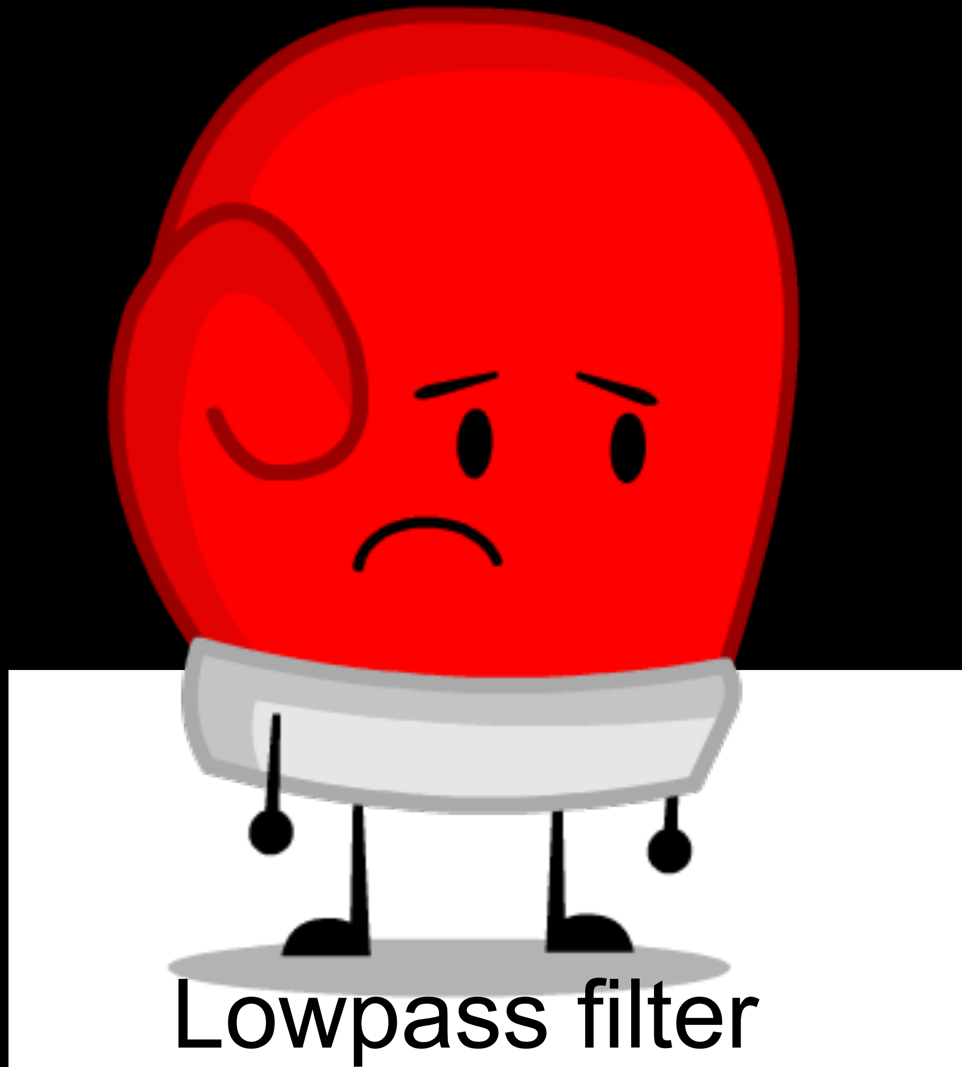



1/f
structure

Power spectra of noise data
(Zarahn, Aguirre, D' Esposito, NI, 1997)

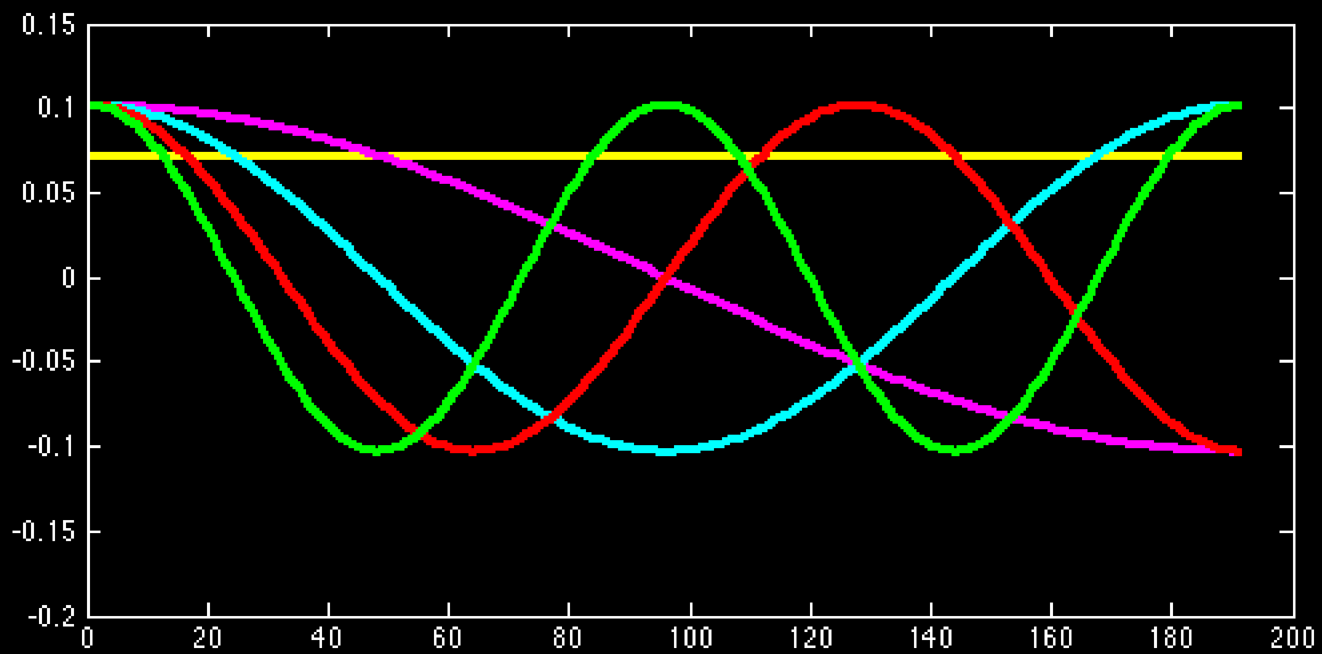
# The 1-2 punch

# Typically not used



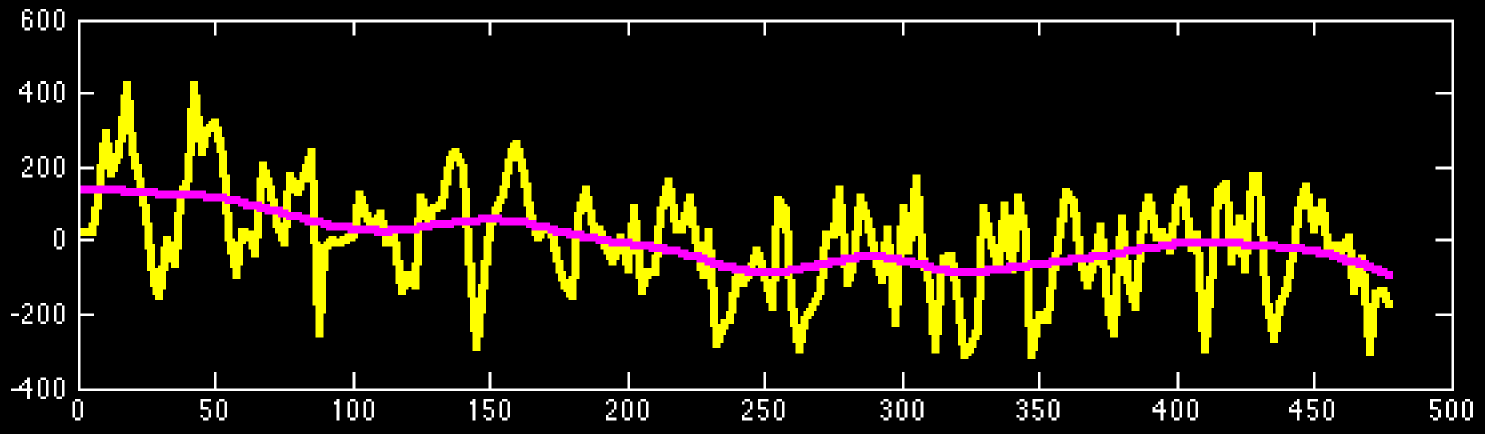
# Highpass filtering

- Remove low frequency noise
  - SPM: Adds a discrete cosine transform basis set to design matrix



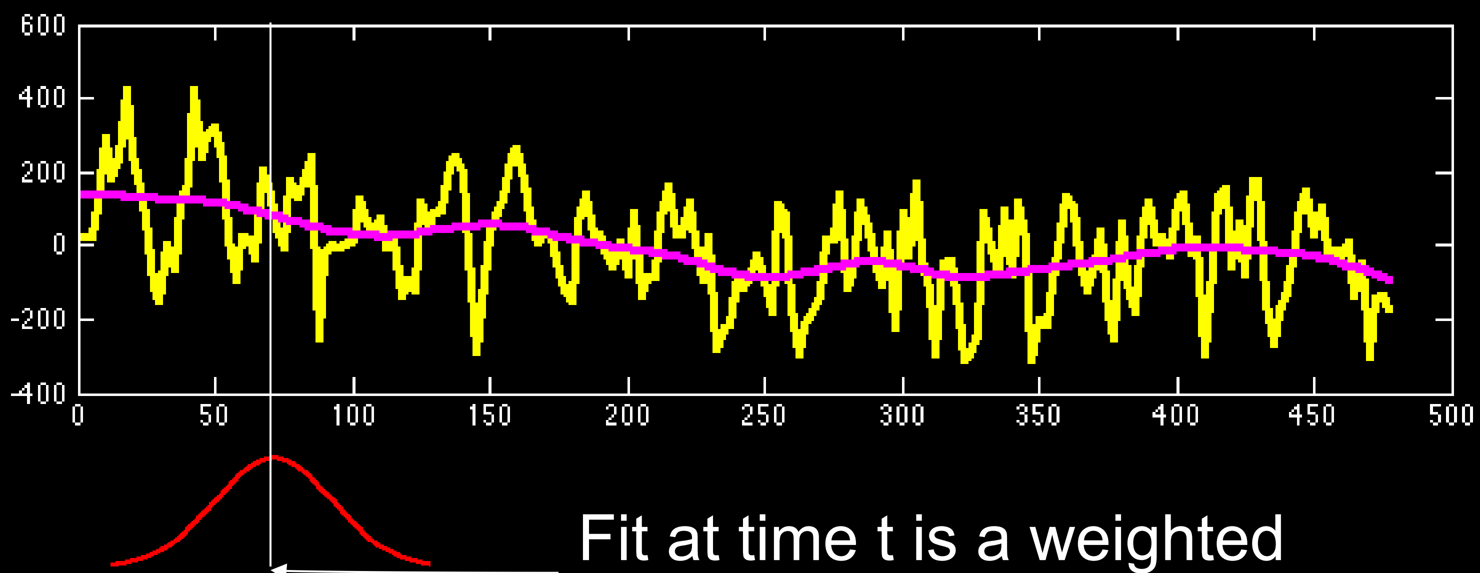
# Highpass filtering

- FSL: Gaussian-weighted running line smoother
  - Step 1: Fit a Gaussian weighted running line



# Highpass filtering

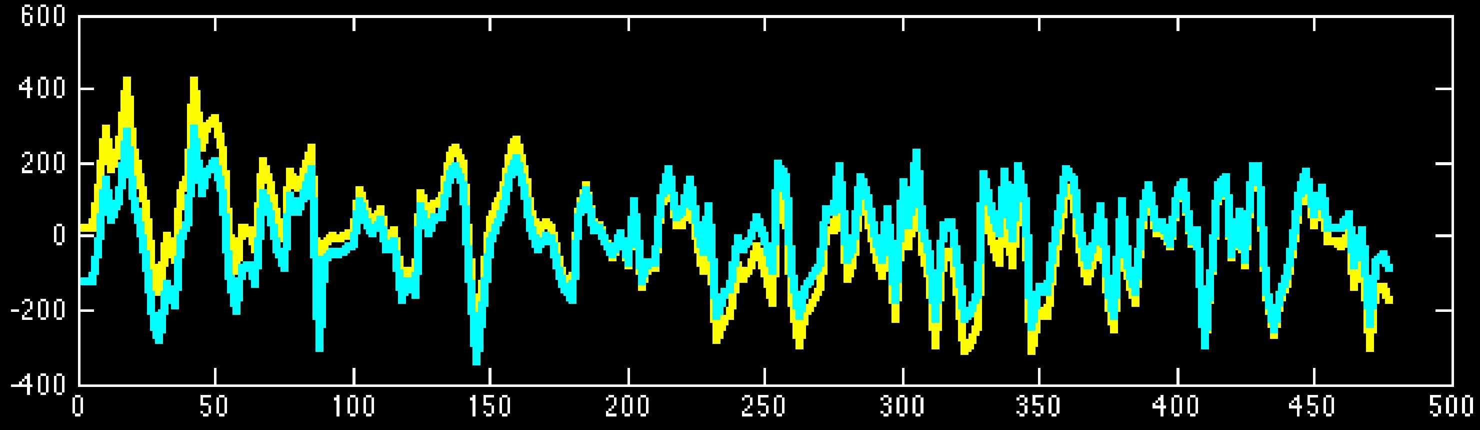
- FSL: Gaussian-weighted running line smoother
  - Step 1: Fit a Gaussian weighted running line



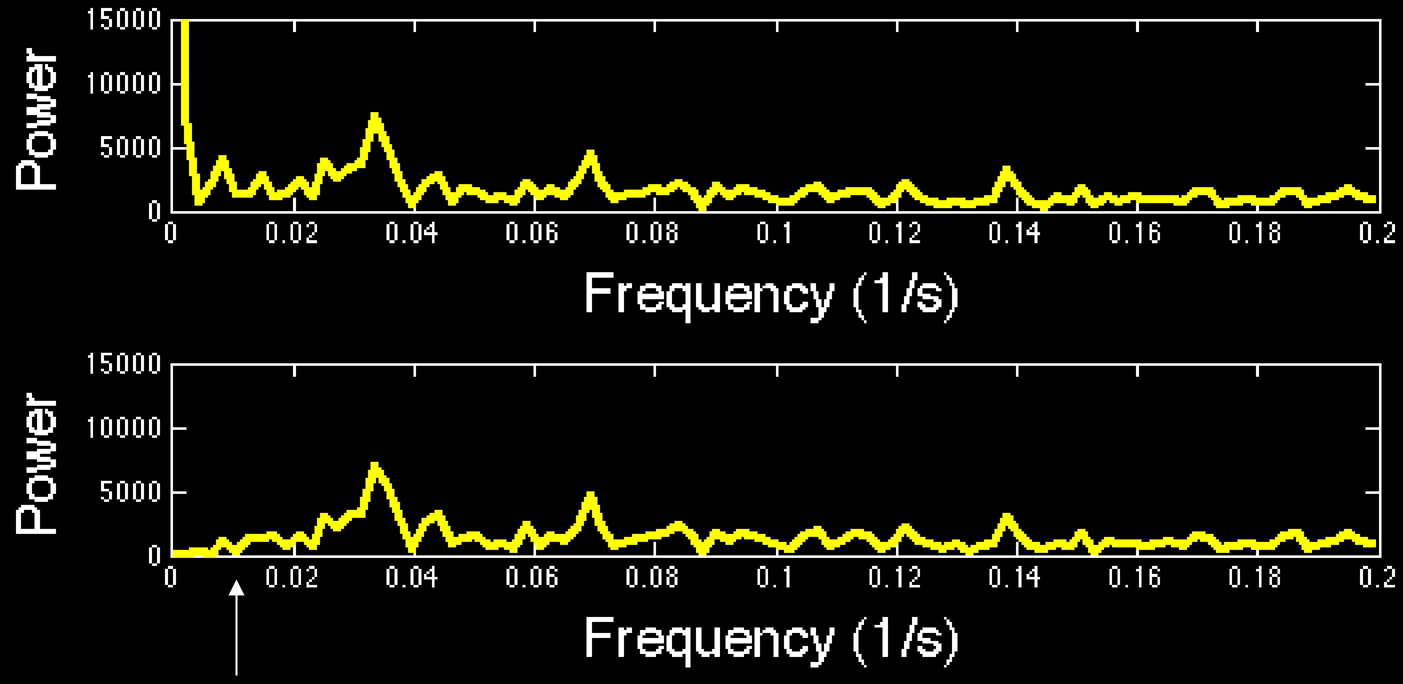

Fit at time t is a weighted average of data around t

# Highpass filtering

– Step 2: Subtract Gaussian weighted running line fit



# Highpass filtering




Filter below .01 Hz

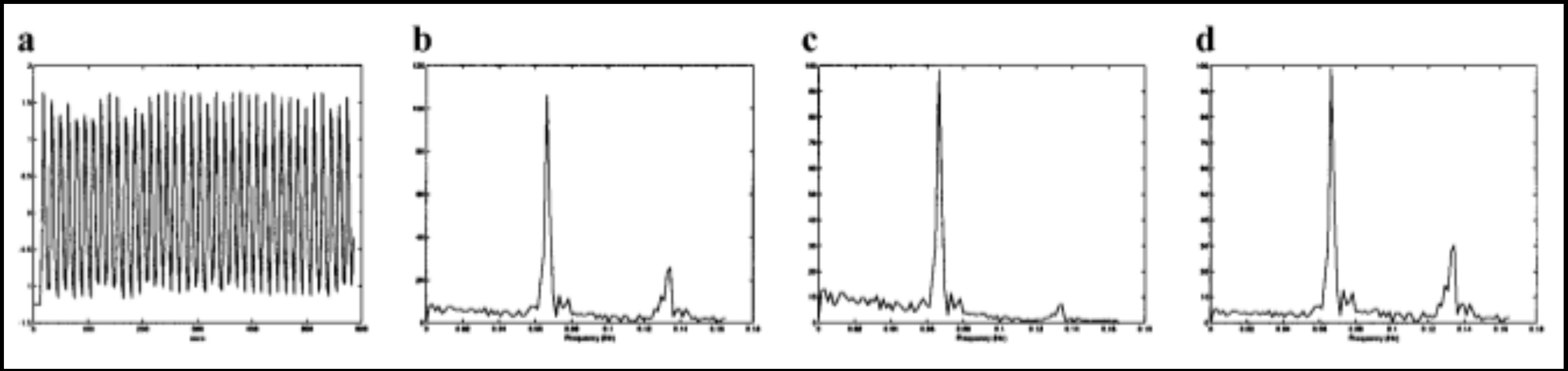
# Filter cutoff

- High, but not higher than paradigm frequency
  - Look at power spectrum of your design and base cutoff on that
  - Block design:  Longer than 1 task cycle… usually twice the task cycle
  - Event related design: Larger than 66 s (based on the power spectrum of a canonical HRF of a single response)

In FSL, need to apply a high pass filter to both the data and the design matrix. In SPM it is automatically applied to both.

# Highpass vs lowpass filtering

- What does it do to the signal??




Signal

Power spectrum

Lowpass filter

Highpass filter

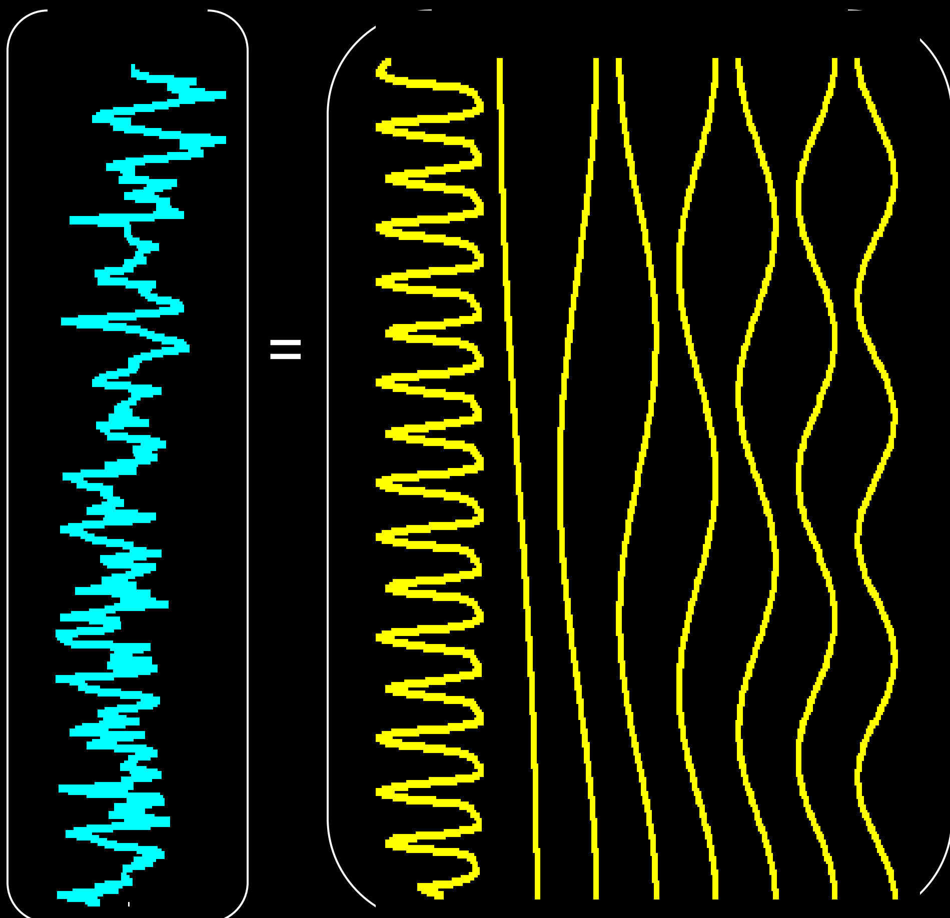
Woolrich et al, NI 2001

# Filtering Conclusions

- Highpass filter
  - Yes!
  - Just don’t remove your signal
- Lowpass filter
  - No (typically)
  - Tends to remove signal

# Model with HP filter

$$Y = X\beta$$

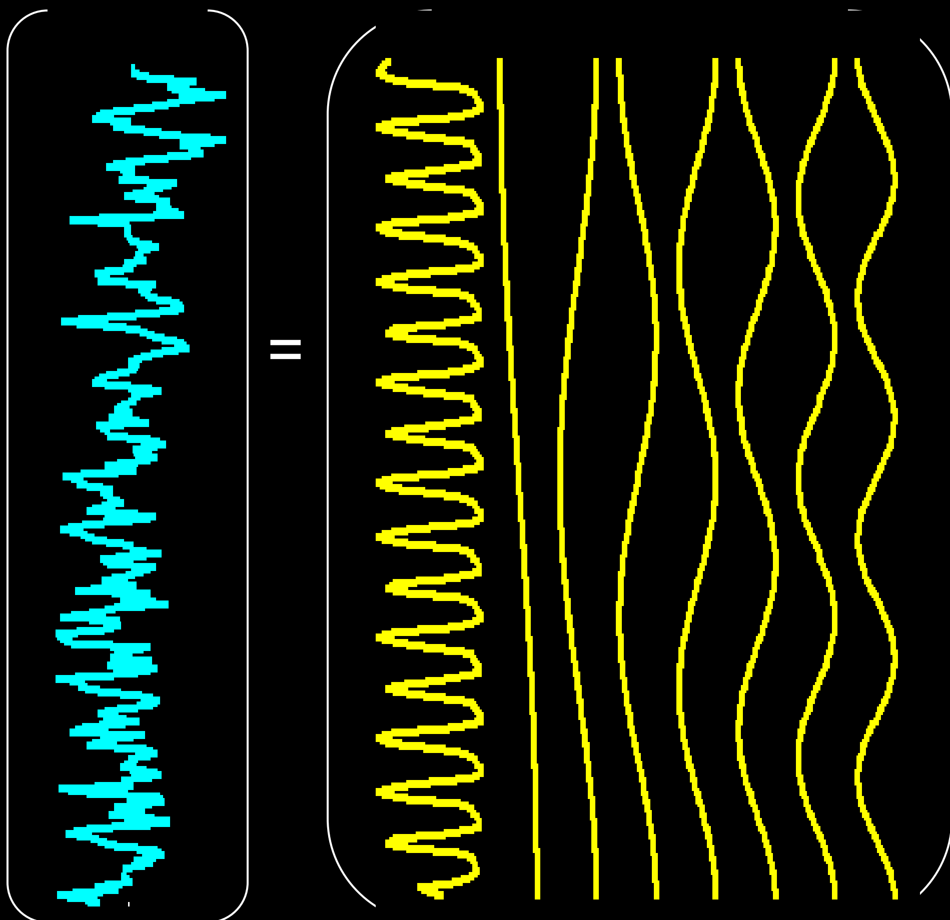


$$\begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \\ \beta_6 \\ \beta_7 \\ \beta_8 \end{pmatrix}$$

Parameter of interest

# Model with HP filter

$$Y = X\beta$$




$$\begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \\ \beta_6 \\ \beta_7 \\ \beta_8 \end{pmatrix}$$

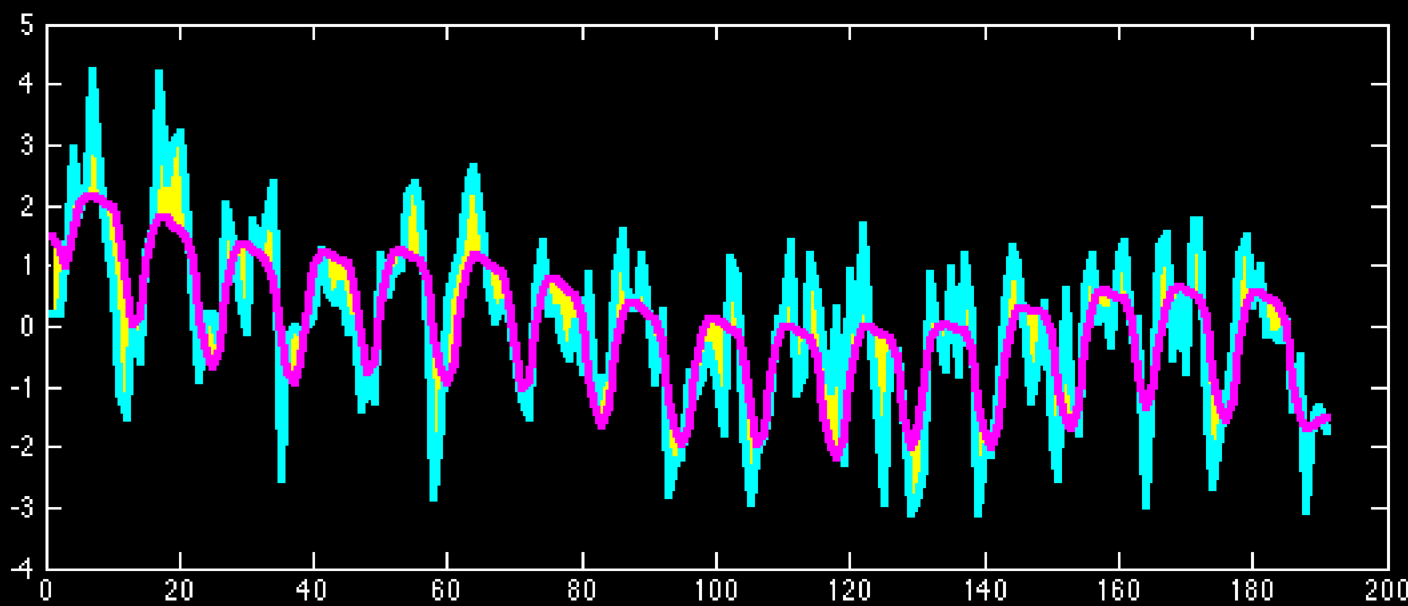
Parameter of interest

$$H_0 : c\beta = 0$$

Use contrast

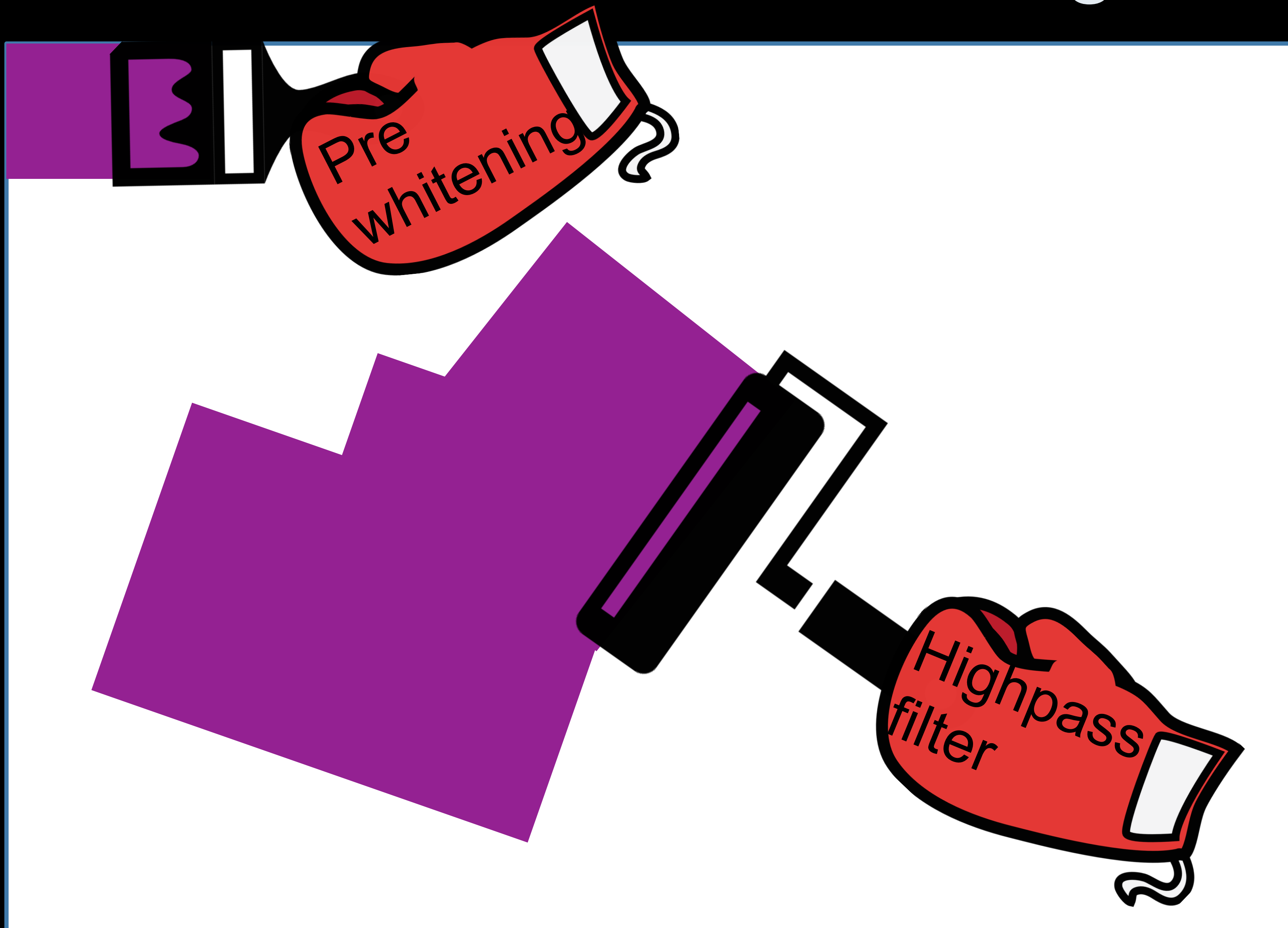
c=(1 0 0 0 0 0 0 0 )

# Convolution & HP filter




$$t = \frac{0.64}{1.04 \times 0.06} = 10.26$$

# Punch 2: Prewhitening



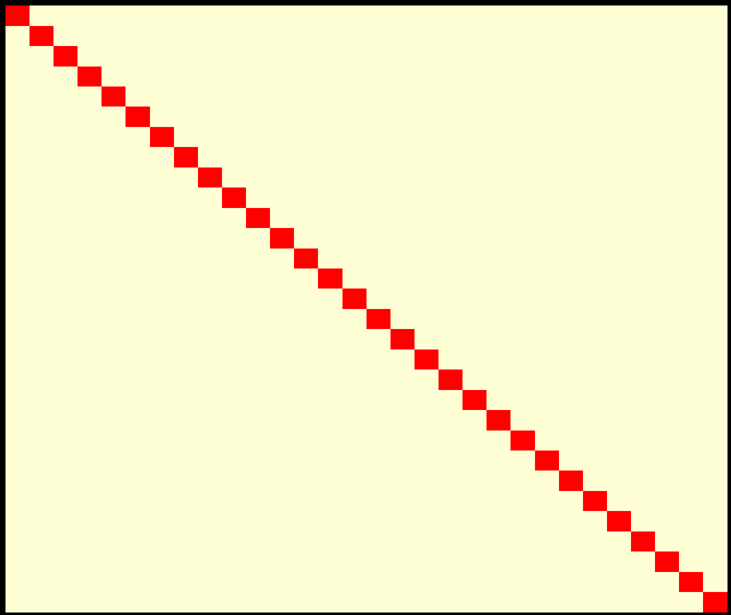
# Prewhitening

- Remember Gauss Markov?
  - If our errors are distributed with mean 0, constant variance and not temporally autocorrelated then our estimates are unbiased and have the smallest variance of all unbiased estimators.
  - Uh oh,

# Prewhitening

$$Y = X\beta + \epsilon$$

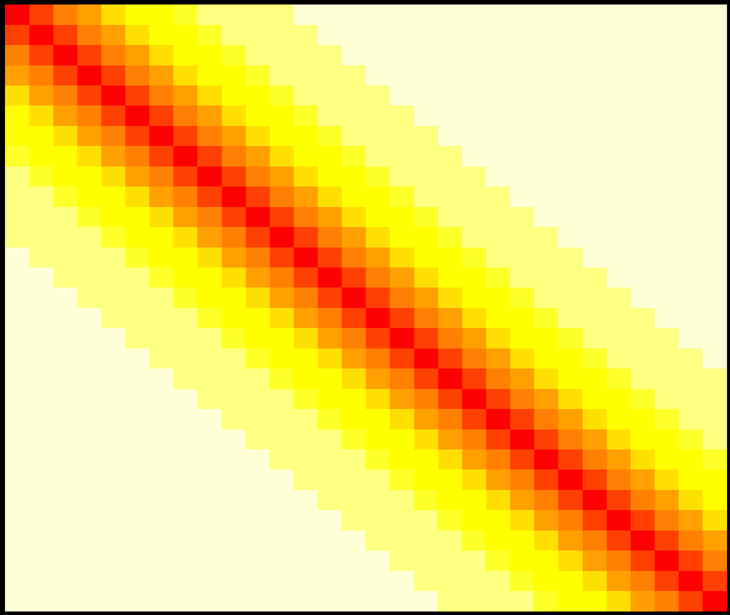
$$\mathrm{Cov}(\epsilon) =$$



# Prewhitening

$$Y = X\beta + \epsilon$$

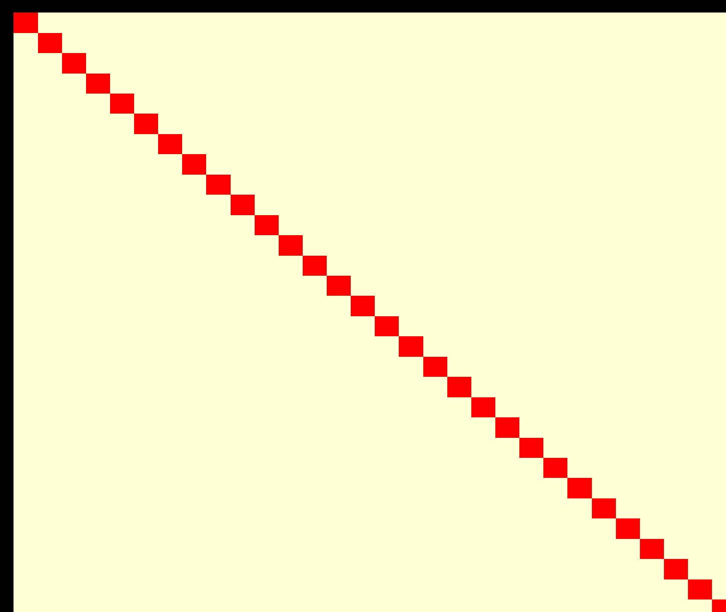
$$\mathrm{Cov}(\epsilon) =$$



# Prewhitening

- Matrix magic
  - Find a matrix K such that

$$\mathrm{Cov}(K\epsilon) =$$



SPM uses a global (not voxel by voxel) estimate to find matrix K. FSL provides a voxelwise estimate (tho it uses neighboring voxels to inform assessments)

# Prewhitening

- Estimate this GLM

$$KY = KX\beta + K\epsilon$$

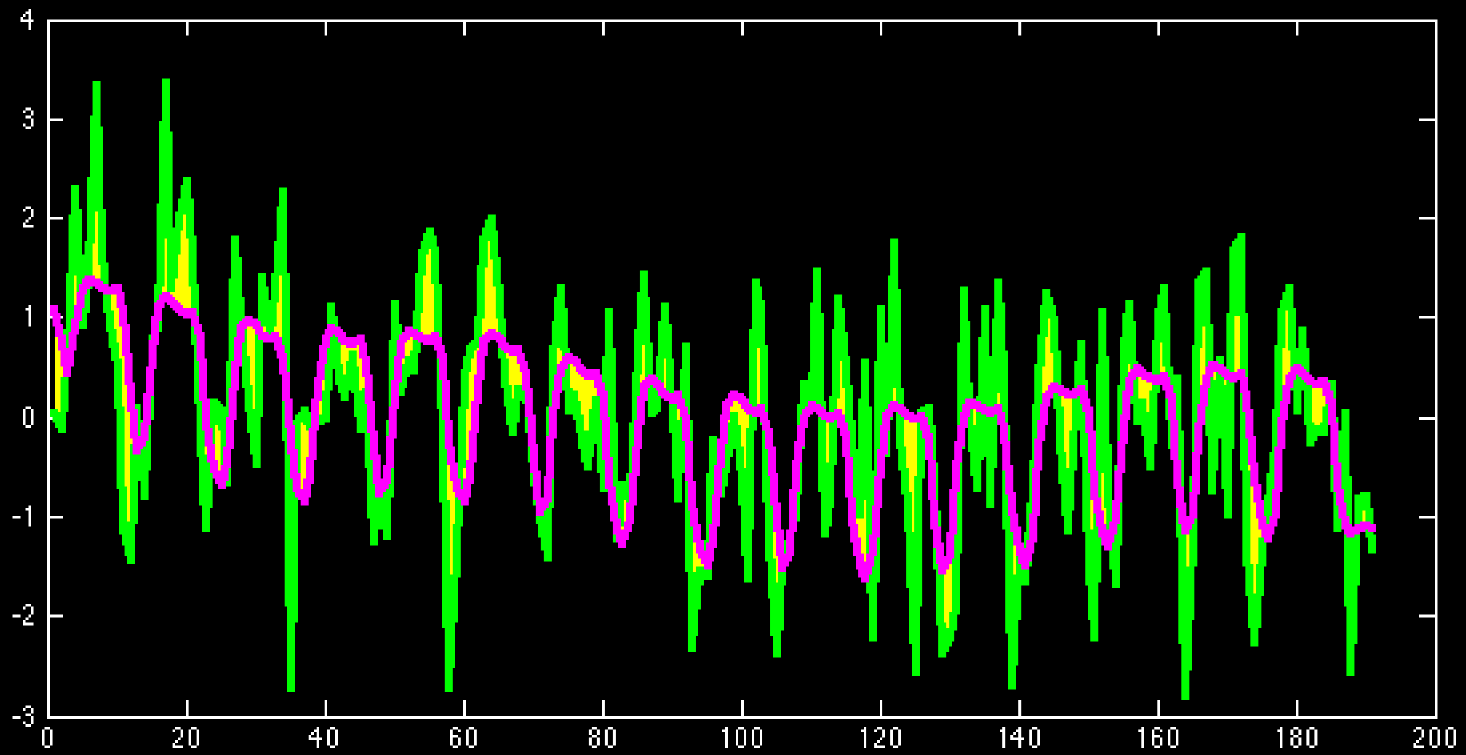
# Prewhitening

- Step 1: Fit the linear regression ignoring temporal autocorrelation
- Step 2: Use residuals from first regression to estimate temporal autocorrelation to obtain K
- Step 3: Create prewhitened model and estimate in usual way

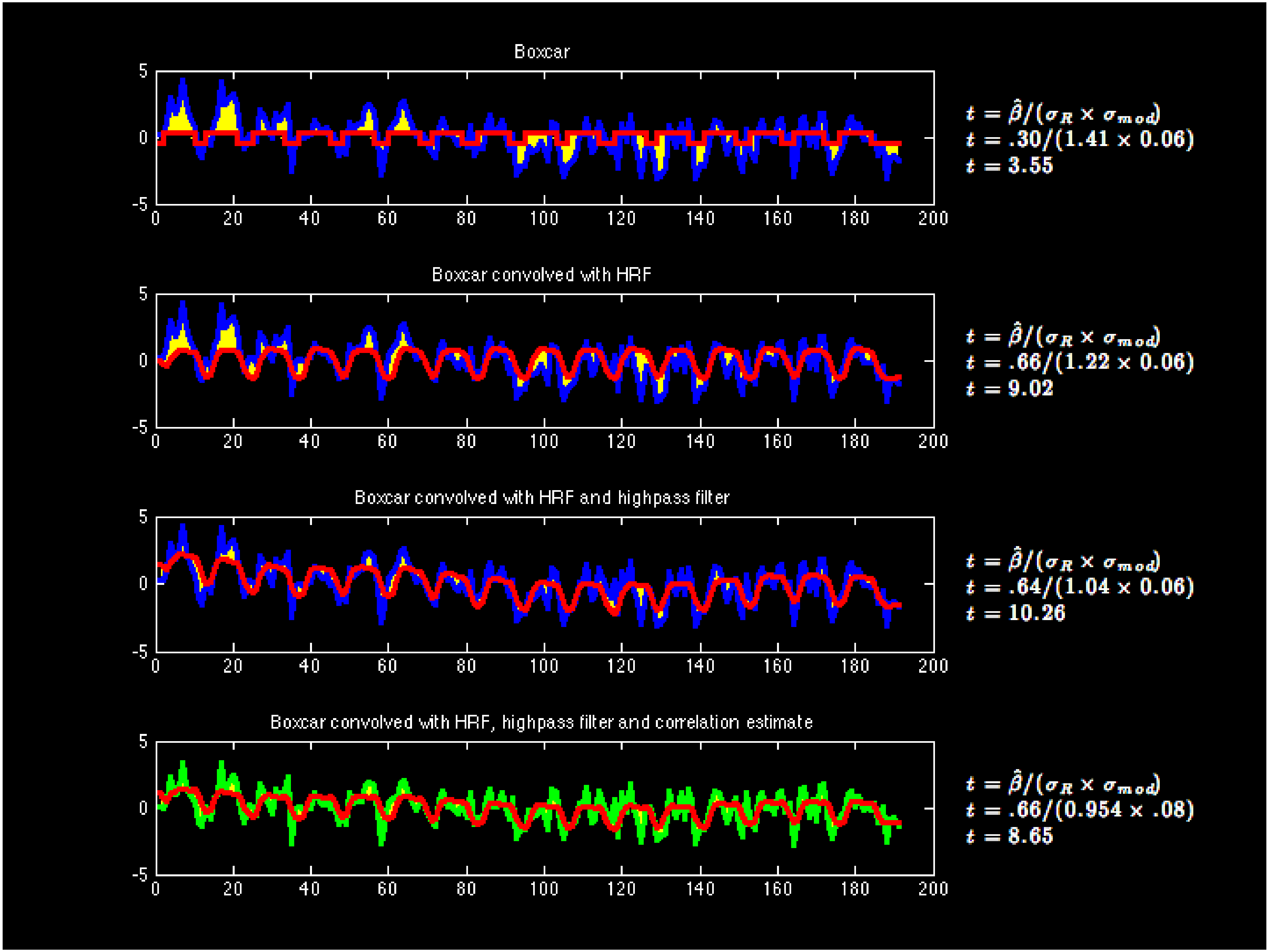
# Estimating V

- There's a bias problem….
  - SPM uses a global covariance estimate to help with this
  - FSL uses a local estimate, but smoothes it

# Convolution, HP filter, Whitening




$$t = \frac{0.66}{0.954 \times 0.08} = 8.65$$



Boxcar
$t = \hat{\beta}/(\sigma_R \times \sigma_{mod})$
$t = .30/(1.41 \times 0.06)$
$t = 3.55$
Boxcar convolved with HRF
$t = \hat{\beta}/(\sigma_R \times \sigma_{mod})$
$t = .66/(1.22 \times 0.06)$
$t = 9.02$
Boxcar convolved with HRF and highpass filter
$t = \hat{\beta}/(\sigma_R \times \sigma_{mod})$
$t = .64/(1.04 \times 0.06)$
$t = 10.26$
Boxcar convolved with HRF, highpass filter and correlation estimate
$t = \hat{\beta}/(\sigma_R \times \sigma_{mod})$
$t = .66/(0.954 \times .08)$
$t = 8.65$

# Scaling

- Grand Mean Scaling (good)
  - Removes intersession variance
  - Allows us to combine data across subjects
  - Whole 4D data set is scaled by a single number
  - Automatically done in software packages

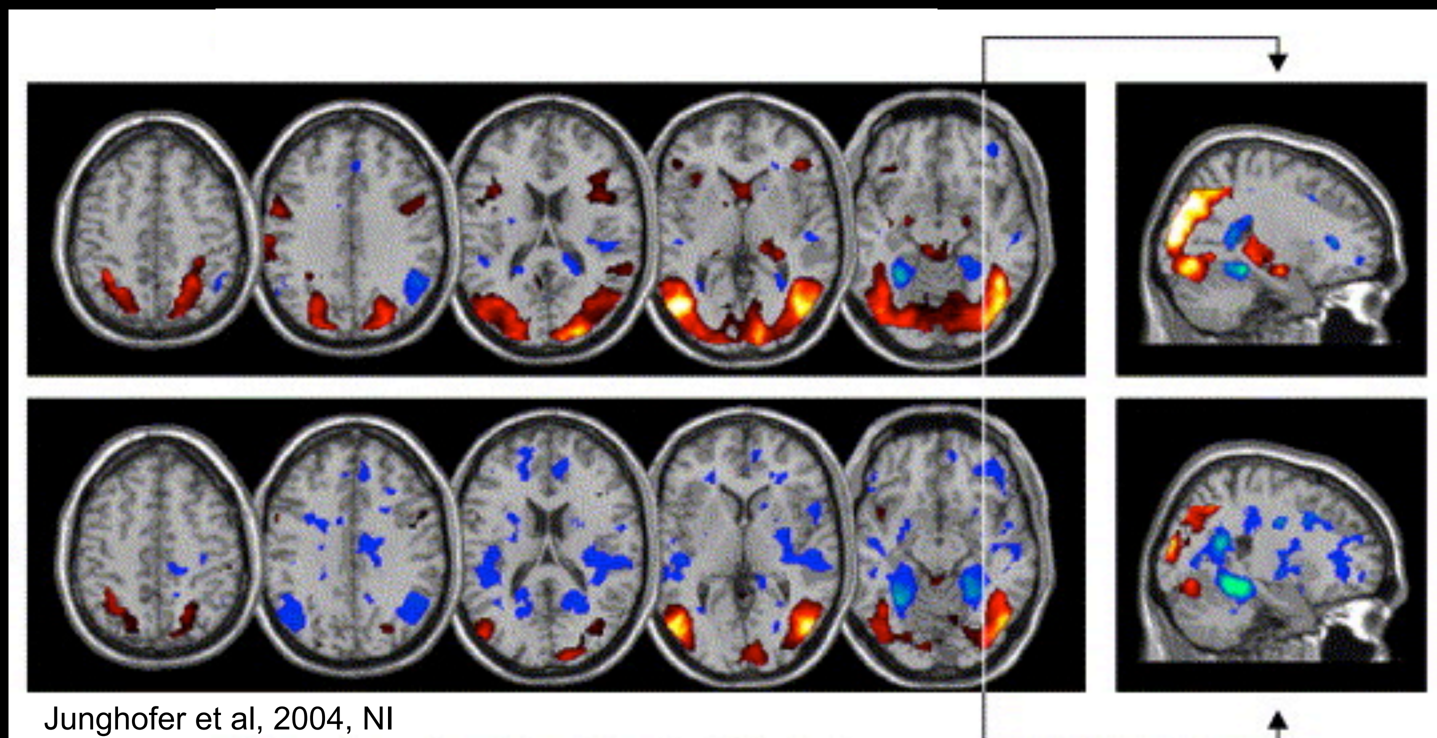
# Scaling

- Proportional scaling aka Intensity normalization (not good)
  - Forces each volume of 4D dataset to have the same mean
  - Also done by modeling the global signal
  - Idea is to remove background activity

# Intensity Normalizaton

without

with



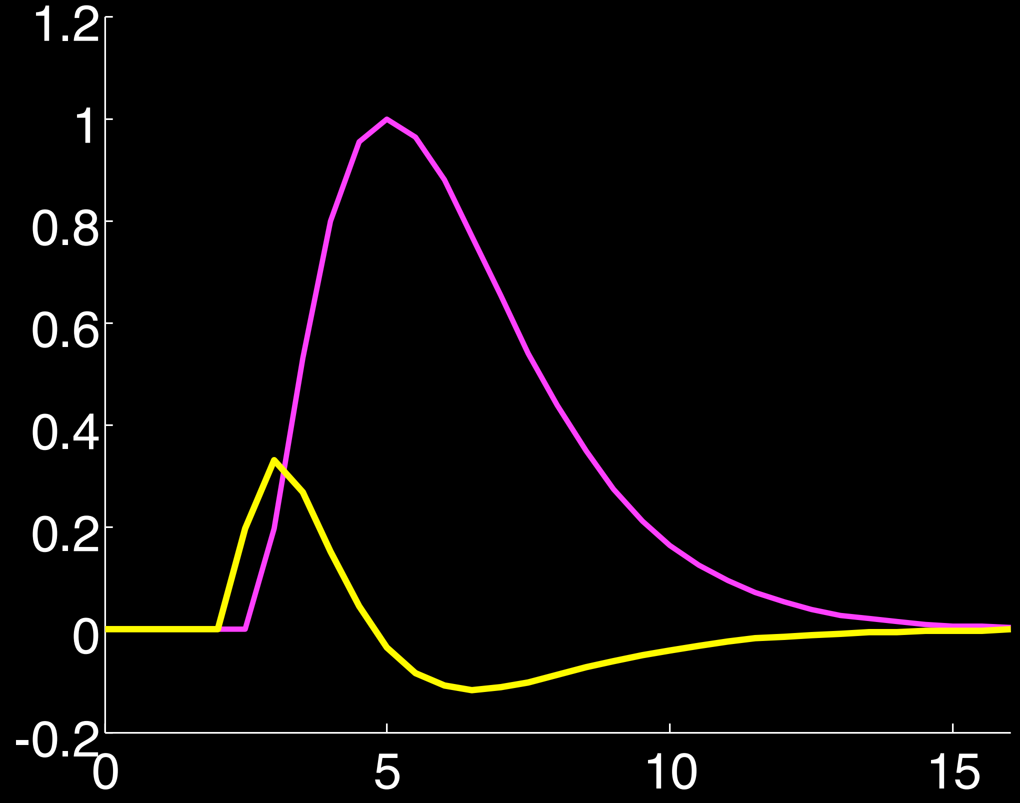
Junghofer et al, 2004, NI

Signal is lost and negative activation artifacts

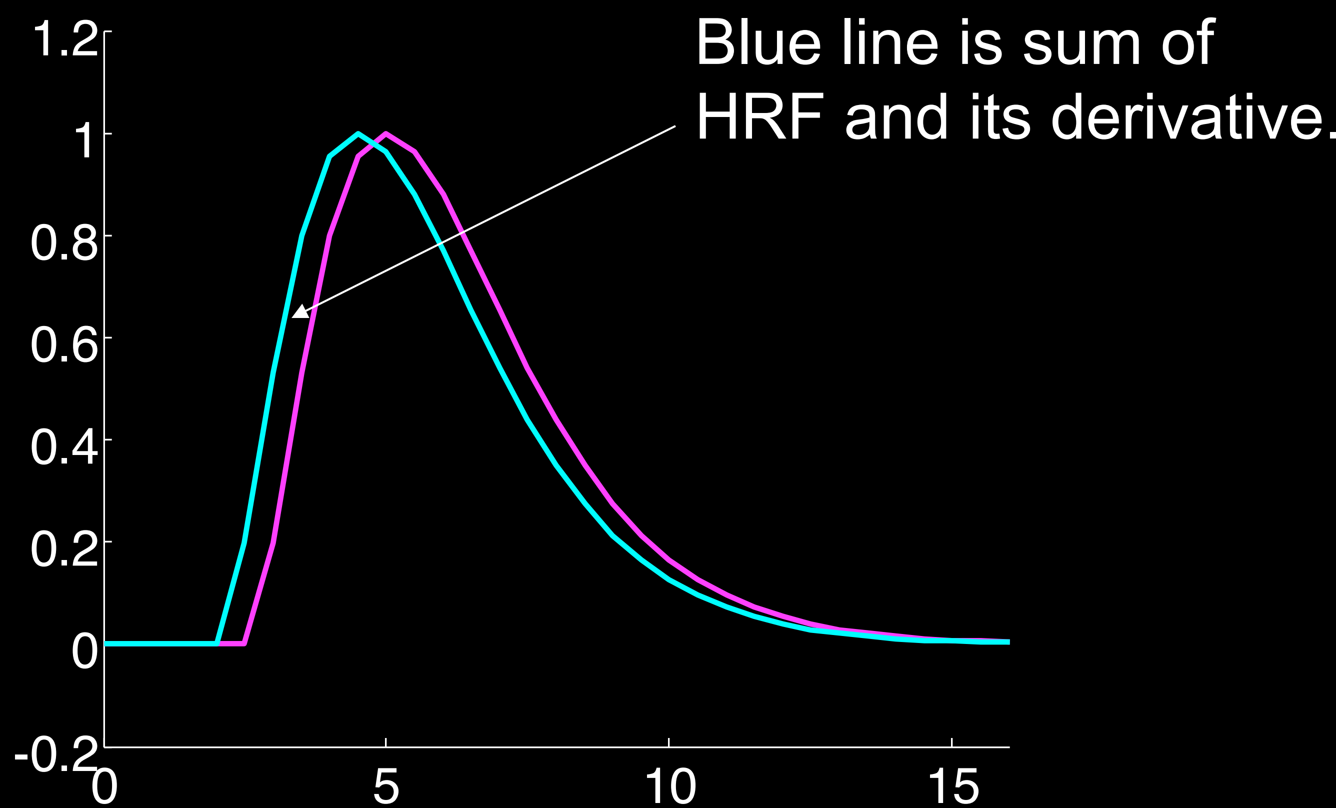
# Other modeling considerations

- Adding the derivative of the HRF
- Adding motion parameters to the model

# Model HRF & Derivative



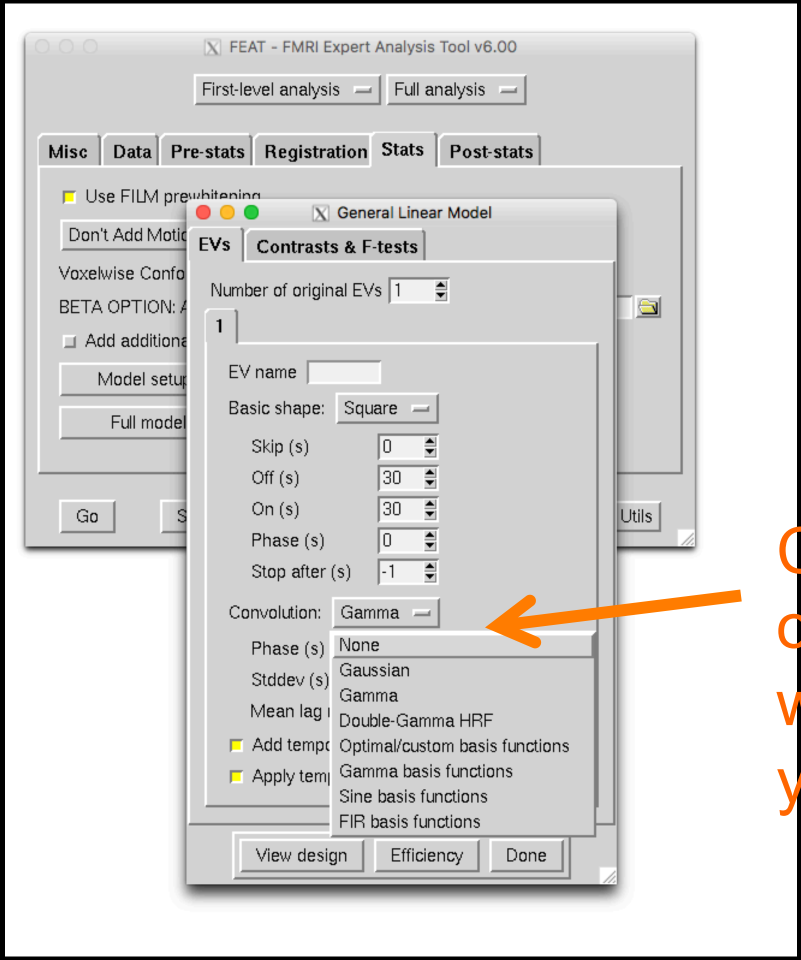
# “Shifted” HRF


Blue line is sum of
HRF and its derivative.
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
0
5
10
15

# Temporal derivative

- We model the derivative, but don't study inferences of it
  - Lindquist, et al (NI, 2008) suggest this is a bad idea…may lead to bias

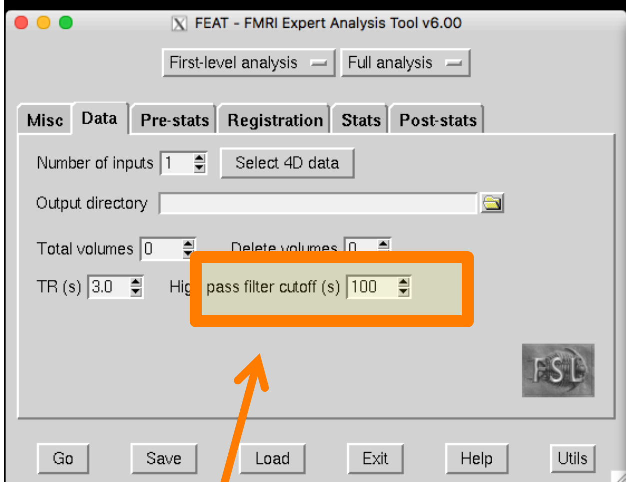
# Where are these options in the Feat GUI?



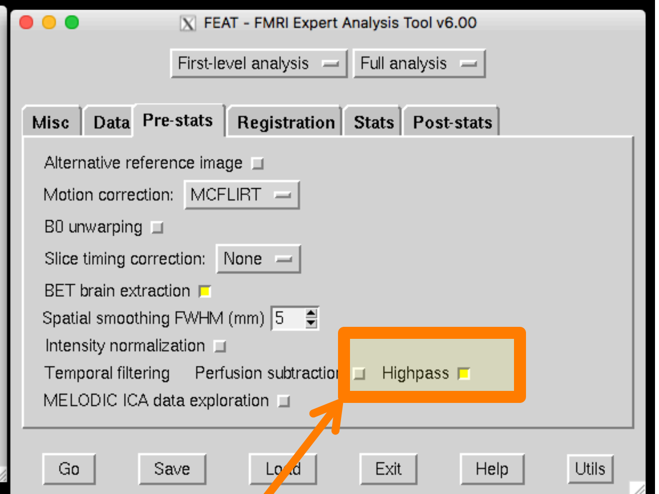

Convolution option . . . which one should you use?

# Where are these options in the Feat GUI?

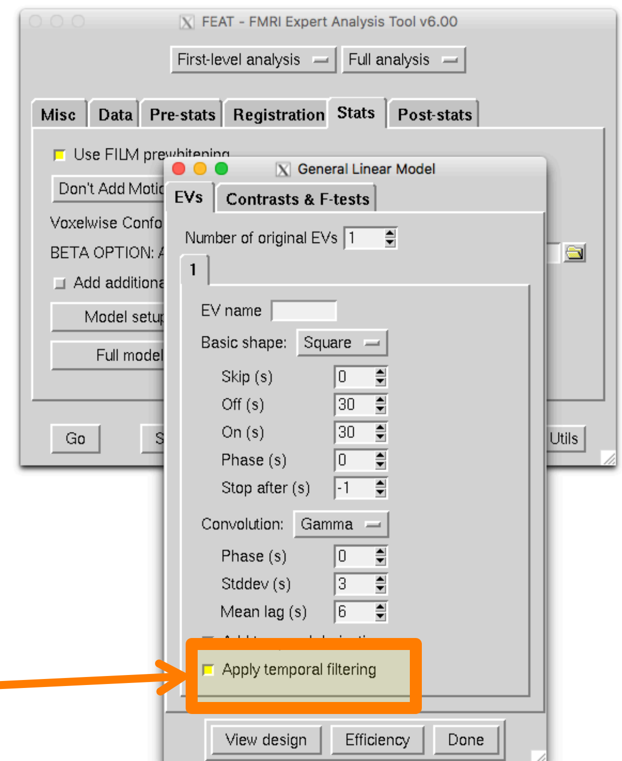
- Highpass filter (3 places!!)




Sets cutoff for data and design

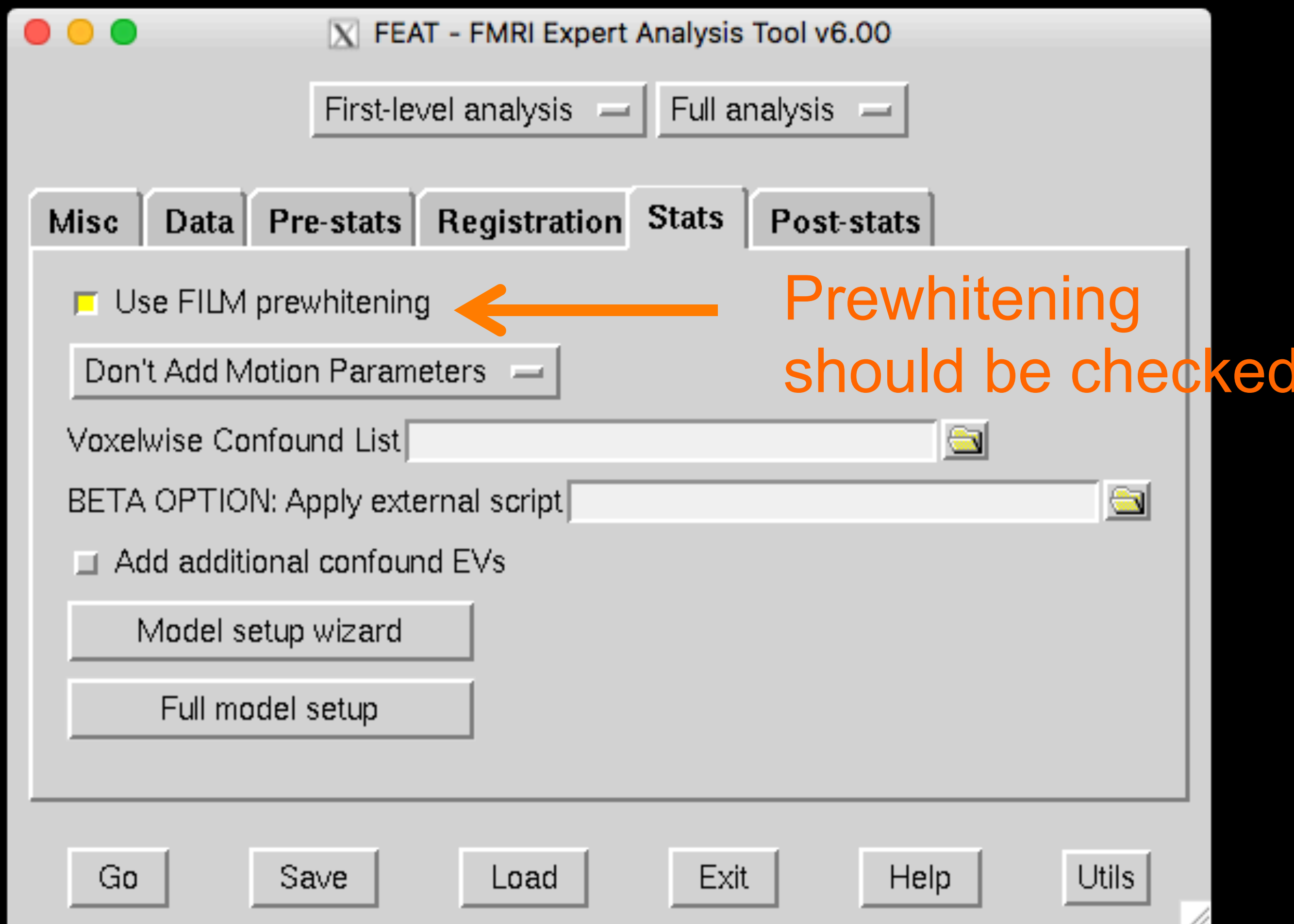



Turns it on for *data*

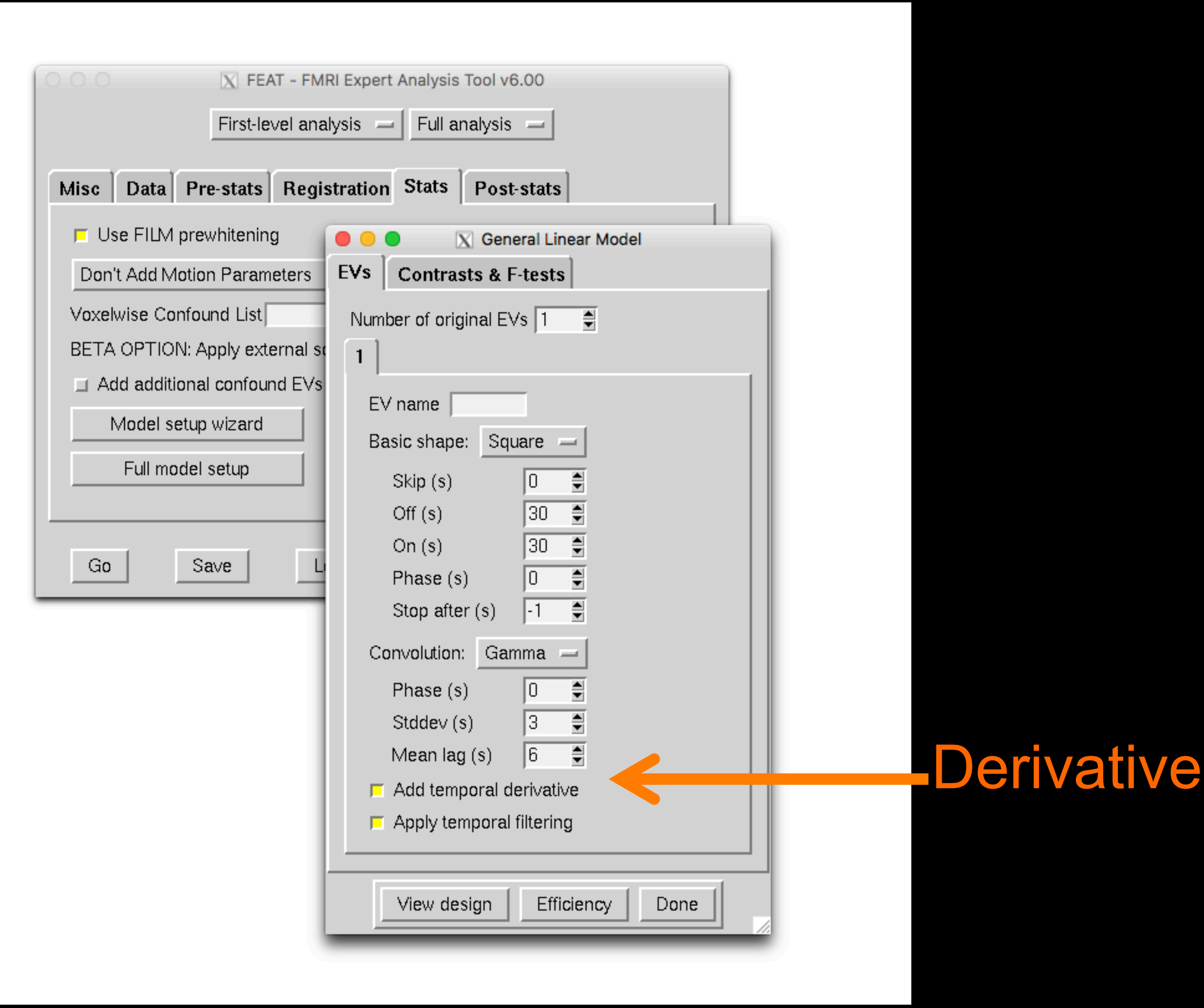



Turns it on for *design*

# Where are these options in the Feat GUI?

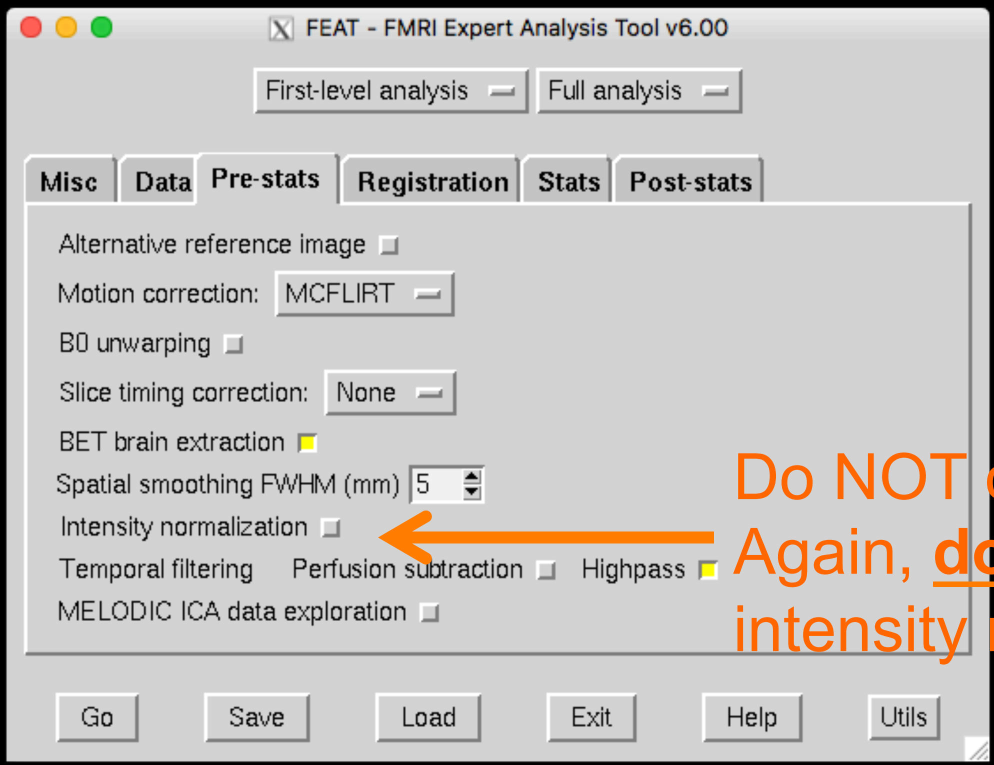


# Where are these options in the Feat GUI?



# Where are these options in the Feat GUI?

- Intensity normalization




Do NOT check this. Again, **<u>don't use</u>** intensity normalization

# Questions?