

Modelling Single-subject fMRI Data

Dr. Martyn McFarquhar

Division of Neuroscience & Experimental Psychology
The University of Manchester

Biennial Edinburgh SPM course, 2019

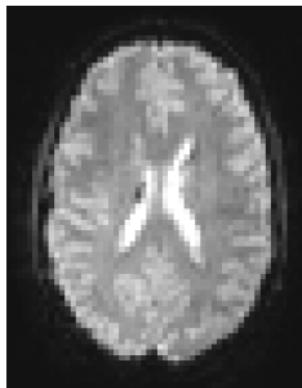
Contents

1. Principles of statistical modelling
2. The general linear model (GLM)
3. Issues with applying the GLM to fMRI data
 1. Low-frequency noise
 2. Autocorrelation in the time-series
 3. Sluggish Nature of the BOLD response
 4. Arbitrary scale of the BOLD signal
4. Contrasts and hypothesis tests



Principles of statistical modelling

Why do we want to model the data?



A volume of dimensions **50 x 60 x 50** will contain **150,000** voxels

If we had **TR = 2** and scanned for **10 minutes** we would collect **300 volumes**

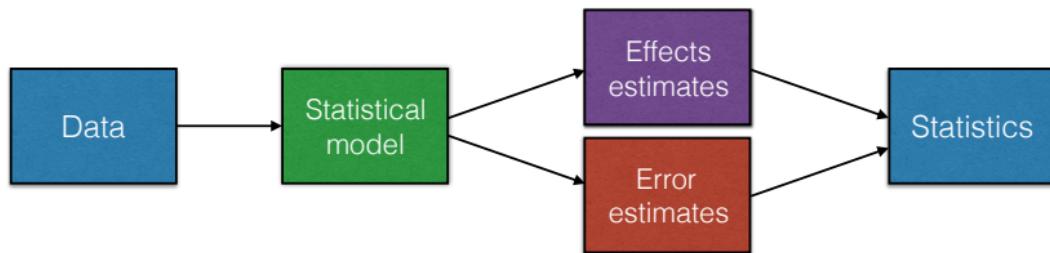
This would mean a total of **45,000,000** data-points

Modelling the data is a means of **condensing** all these values into a **smaller set** of values that provide information on the **magnitude of the experimental effects**

Principles of statistical modelling

Why do we want to model the data?

Modelling the data allows us to **separate** the **effects of interest** from the **error/noise**

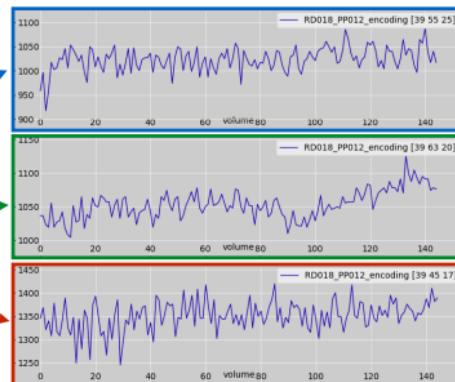
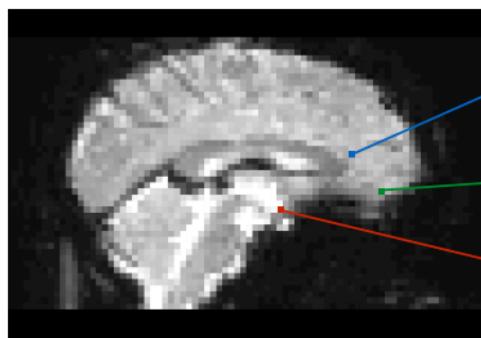


This allows for **statistical inference** about the **experimental task** performed in the scanner:

- Which regions are associated with different conditions?
- How does a single region differ across conditions?

Principles of statistical modelling

How do we model the data?



The **Statistical Parametric Mapping** approach is to fit the **same model** to each time-series **separately** — **mass-univariate**

The tool used to achieve this is the **General Linear Model**



The General Linear Model (GLM)

A general linear model for observation i from **one voxel** is

$$y_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_k x_{ik} + \epsilon_i$$

y_i is the i th value of the time series ($i = 1, \dots, n$)

x_{ij} is the i th value of predictor variable j ($j = 1, \dots, k$)

β_0, \dots, β_k are the model parameters

ϵ_i is the i th model error

This is just **multiple linear regression**

applied to the time-series from **one voxel**

$$y_i = \beta_0 + \sum_{j=1}^k \beta_j x_{ij} + \epsilon_i$$

The General Linear Model (GLM)

There are always n **regression equations** — one per value of y

$$y_1 = \beta_0 + \beta_1 x_{11} + \cdots + \beta_k x_{1k} + \varepsilon_1$$

$$y_2 = \beta_0 + \beta_1 x_{21} + \cdots + \beta_k x_{2k} + \varepsilon_2$$

⋮

$$y_n = \beta_0 + \beta_1 x_{n1} + \cdots + \beta_k x_{nk} + \varepsilon_n$$

Notice that across these equations the **values** of y , x and ε **change**,
but β_0 and $\beta_1 \dots \beta_k$ **stay the same**

This tells us that β_0 and $\beta_1 \dots \beta_k$ **provide a summary** across all the
data points — something **consistent** that is **separate** from the noise
in the data



The General Linear Model (GLM)

These n **regression equations** can be compactly written in **matrix** notation

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & \cdots & x_{1k} \\ 1 & x_{21} & \cdots & x_{2k} \\ 1 & x_{31} & \cdots & x_{3k} \\ \vdots & \vdots & & \vdots \\ 1 & x_{41} & \cdots & x_{4k} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \vdots \\ \varepsilon_4 \end{bmatrix}$$

This gives the standard expression for the GLM:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

The matrix of predictors (\mathbf{X}) is known as the **design matrix**



The General Linear Model (GLM)

Although regression is often thought of in terms of **continuous predictor variables**, it is possible to include **categorical predictor variables** through the use of **dummy variables**

Dummy variables:

- Take on values of **0** or **1** to indicate the **presence** or **absence** of a **categorical effect**
- Also known as **indicator variables**

Index	Category	Dummy 1	Dummy 2
1	1	1	0
2	1	1	0
3	2	0	1
4	2	0	1
5	3	0	0
6	3	0	0

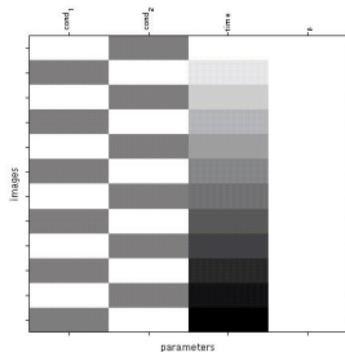
The General Linear Model (GLM)

The GLM can therefore contain a mixture of **continuous** and **categorical** predictor variables

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 5 \\ 1 & 1 & 0 & 7 \\ 1 & 0 & 1 & 3 \\ 1 & 0 & 1 & 2 \\ 1 & 0 & 0 & 4 \\ 1 & 0 & 0 & 8 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \end{bmatrix}$$

The General Linear Model (GLM)

The GLM can therefore contain a mixture of **continuous** and **categorical** predictor variables



In SPM the **design matrix** is **visualised** as a means of understanding the model you have specified

A **categorical predictor** is given as **blocks of colour** whereas a **continuous predictor** is given as a **gradient**



The General Linear Model (GLM)

Parameter estimation

No matter the forms of the predictor variables, the **unknown values** in the GLM must be **estimated** from the data

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} 1 + x_{11} + \dots + x_{1k} \\ 1 + x_{21} + \dots + x_{2k} \\ 1 + x_{31} + \dots + x_{3k} \\ 1 + x_{41} + \dots + x_{4k} \end{bmatrix} + \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \end{bmatrix}$$

Known values

We have measured values of the **response variable** and the **predictor variables**

Unknown values

We do not know the **parameter values** or the **errors** — they must be **estimated**



The General Linear Model (GLM)

Parameter estimation

Estimation of the β values is based on **least-squares**

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$$

The **residuals** come from subtraction of the **estimates** from the **data**

$$\boldsymbol{\epsilon} = \mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}$$

The **variance** comes from the sum-of-squares of the **residuals**, divided by the **error degrees of freedom**

$$\hat{\sigma}^2 = \frac{\boldsymbol{\epsilon}'\boldsymbol{\epsilon}}{n - k}$$

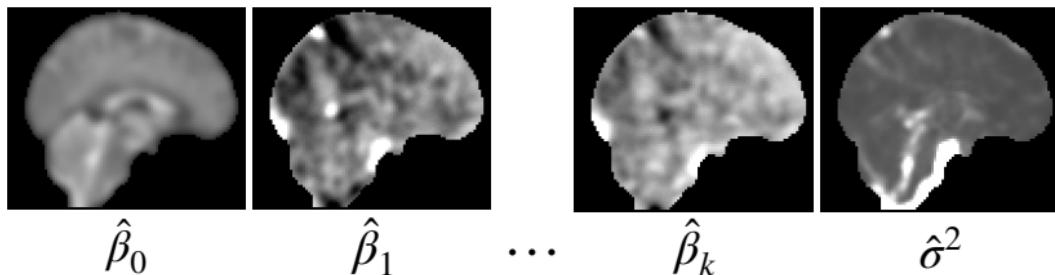


The General Linear Model (GLM)

Parameter estimation

Remember that the parameters are estimated **at every voxel**

The outputs from the **mass-univariate** GLM are **images** of the parameter estimates



The voxel values in each image are the **parameter estimates** resulting from fitting the model **at that voxel** ($\sim 150,000$ models)



The General Linear Model (GLM)

Parameter estimation

Tiny part of the SPM code (in **spm_spm.m**)

```
%--Weighted Least Squares estimation
%=====
beta      = xX.pKX*KWY;                      %--Parameter estimates
if any(cmask)
    res     = spm_sp('r',xX.xKXs,KWY);        %--Residuals
else
    res     = zeros(nScan,0);
end
ResSS     = sum(res.^2);                       %--Residual SSQ

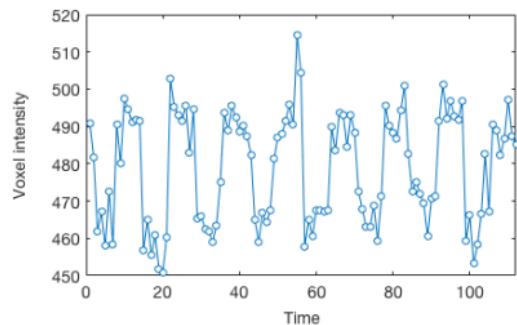
%--Write ResSS into ResMS (variance) file scaled by tr(RV)
%
c(cmask)   = ResSS / xX.trRV;
VResMS    = spm_data_write(VResMS, c, chunk);
```

The estimation can be done **very quickly**, despite the large number of models

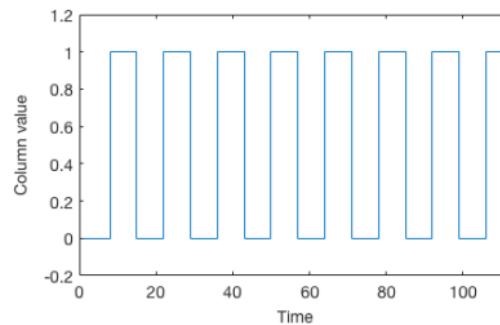
The General Linear Model (GLM)

What do the parameters tell us?

Example of a **finger tapping task** — 7 volumes of finger tapping and 7 volumes of rest



Time-series from **one voxel**
in the **motor cortex**

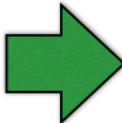
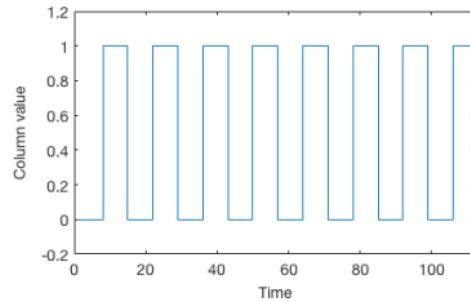


Dummy variable coding
the **off-on** block pattern



The General Linear Model (GLM)

What do the parameters tell us?



$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \\ y_8 \\ y_9 \\ y_{10} \\ y_{11} \\ y_{12} \\ y_{13} \\ y_{14} \\ y_{15} \\ y_{16} \\ y_{17} \\ y_{18} \\ y_{19} \\ y_{20} \\ y_{21} \\ \vdots \\ y_{112} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ \vdots \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \\ \varepsilon_7 \\ \varepsilon_8 \\ \varepsilon_9 \\ \varepsilon_{10} \\ \varepsilon_{11} \\ \varepsilon_{12} \\ \varepsilon_{13} \\ \varepsilon_{14} \\ \varepsilon_{15} \\ \varepsilon_{16} \\ \varepsilon_{16} \\ \varepsilon_{17} \\ \varepsilon_{18} \\ \varepsilon_{19} \\ \varepsilon_{20} \\ \varepsilon_{21} \\ \vdots \\ \varepsilon_{112} \end{bmatrix}$$

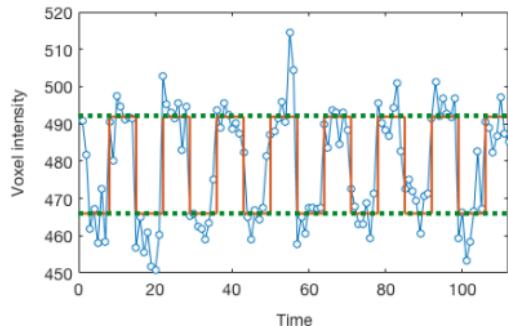
Under this parameterisation:

- β_0 will be the **average** signal during the **rest** periods — **off**
- β_1 will be the **change** in signal from **rest** to **tapping** — from **off** to **on**

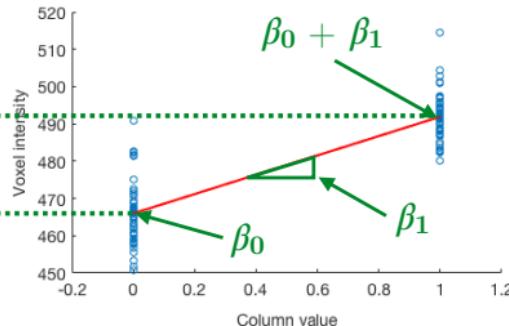


The General Linear Model (GLM)

What do the parameters tell us?



Timeseries perspective



Regression perspective

β_1 therefore tells us the **experimental effect of interest** — the **average change** in signal from **rest** to **finger tapping**



The General Linear Model (GLM)

What do we do with the parameter values?

We can see that the **parameter estimates** tell us something about the **magnitude** and **direction** of the **experimental effect**

For **statistical inference** on the parameters we would typically perform a **hypothesis test** on the parameter values

In the **finger tapping** example we would perform a test at each voxel to see whether the signal change from rest was **significantly different** to **0** — $H_0: \beta_1 = 0$

Before getting to hypothesis testing, there are some **issues** we need to resolve first...

Issues with applying the GLM to fMRI data

There are **several issues** with fMRI data that can cause problems for the GLM

The key problems are:

1. Low-frequency noise (signal drift)
2. Autocorrelation in the time-series
3. The sluggish nature of the BOLD signal
4. The arbitrary scaling of the BOLD signal

SPM has methods of dealing with all of these — some are **automatic** and some are controlled with **user options**

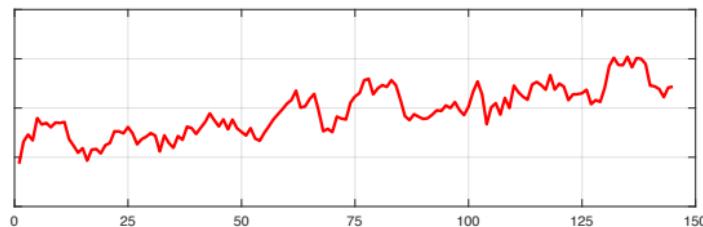
Even if they are **automatic** it is worthwhile knowing about them — SPM should not be seen as a **black box**



Issues with applying the GLM to fMRI data

1. Low-frequency noise

fMRI data are usually contaminated with **low-frequency noise**



An issue with the scanner — cannot prevent very easily — will **bias the parameter estimates**

It is important to design studies where the **frequency of experimental manipulations** does not fall within the **frequency range of the drift**

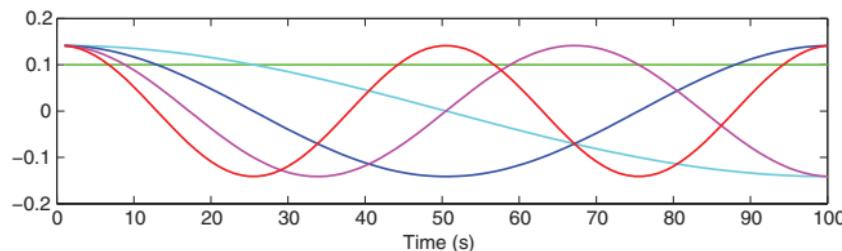


Issues with applying the GLM to fMRI data

1. Low-frequency noise — *solution*

SPM will **high-pass filter** the data to **remove** low-frequency signal components

This is done using a **discrete cosine transform** (DCT) basis set



Adding these cosines as columns in the **design matrix** has the effect of **filtering** the data up to the **highest frequency** cosine



Issues with applying the GLM to fMRI data

1. Low-frequency noise – *solution*

You won't actually see these columns in SPM

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boxed{\mathbf{X}_0\boldsymbol{\beta}_0} + \boldsymbol{\epsilon}$$

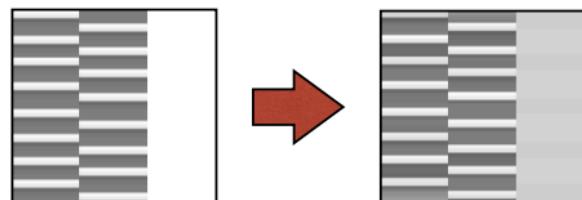
↓

$$\boxed{\mathbf{S}}\mathbf{y} = \mathbf{S}\mathbf{X}\boldsymbol{\beta} + \mathbf{S}\boldsymbol{\epsilon}$$

Residual forming matrix of the cosines **Cosines and their parameters**

$$\mathbf{S} = \mathbf{I} - \mathbf{X}_0\mathbf{X}_0^+$$

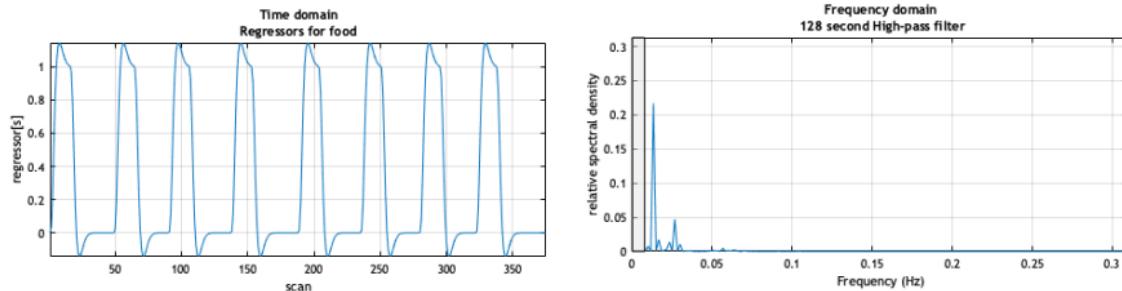
After **filtering** your **design matrix** will change



Issues with applying the GLM to fMRI data

1. Low-frequency noise — *solution*

The **Design Review** features in SPM can help you design a task where the **experimental manipulations** lie **outside** the range of the **high-pass filter**

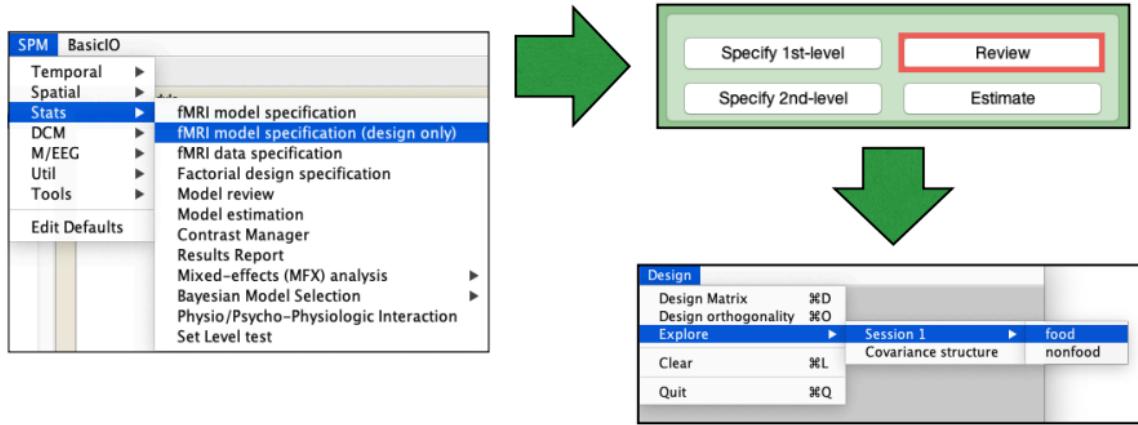


The highest power should lie **outside** of the **grey box** in the **frequency domain**

Issues with applying the GLM to fMRI data

1. Low-frequency noise — *solution*

The **Design Review** features in SPM can help you design a task where the **experimental manipulations** lie **outside** the range of the **high-pass filter**





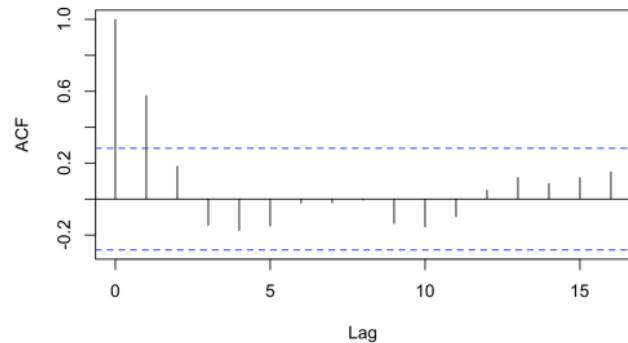
Issues with applying the GLM to fMRI data

2. Autocorrelation in the time-series

The **Gauss-Markov** theorem tells us that **least squares** produces **unbiased estimates** if the data are **not correlated**

This is a problem for fMRI because **time-series** data will be **autocorrelated**

This means values **close in time** will be **more similar** than values **further away**





Issues with applying the GLM to fMRI data

2. Autocorrelation in the time-series — *solution*

SPM uses an **AR(1)** noise model to **estimate** the correlation structure from the data

This is done using the **residuals** of an **initial model fit**

$$\epsilon_t = \rho \epsilon_{t-1} + \tau_t$$

$$\tau_t \sim \mathcal{N}(0, \sigma_\tau^2)$$

The outcome is a **covariance structure** where the **correlation** between time-points **decreases** the further away in time

$$\text{COV}(\epsilon_{t+n}, \epsilon_t) = \frac{\sigma_\tau^2}{1 - \rho^2} \rho^n \quad \forall \quad n \geq 0$$



Issues with applying the GLM to fMRI data

2. Autocorrelation in the time-series – *solution*

Once this **covariance structure** is estimated it is **remove** using a procedure known as **pre-whitening**

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \quad \boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{V})$$

$$\mathbf{V}^{-1} = \mathbf{W}\mathbf{W}' \quad \rightarrow \quad \mathbf{W}\mathbf{V}\mathbf{W}' = \mathbf{I}$$

$$\mathbf{W}\mathbf{y} = \mathbf{W}\mathbf{X}\boldsymbol{\beta} + \mathbf{W}\boldsymbol{\epsilon} \quad \mathbf{W}\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I})$$

This works because

$$\text{Cov}(\mathbf{W}\boldsymbol{\epsilon}) = \mathbf{W}\text{Cov}(\boldsymbol{\epsilon})\mathbf{W}' = \sigma^2 \mathbf{W}\mathbf{V}\mathbf{W}' = \sigma^2 \mathbf{I}$$



Issues with applying the GLM to fMRI data

2. Autocorrelation in the time-series — *solution*

There are some **limitations** to this approach:

1. The **AR(1)** model only uses **one-step** — may not be as accurate as **AR(p)** or **ARIMA** models
2. The **accuracy** of the **pre-whitening** depends on the **accuracy** of the estimate of **V**
3. In SPM, **V** is **not** estimated at each voxel — “significant” voxels are **pooled** and a **single estimate** of **V** is created

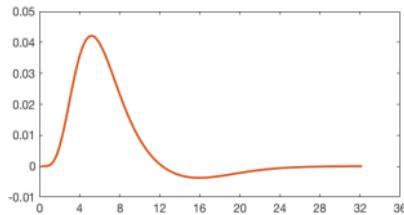
The **biggest drawback** to the SPM approach to correlation-correction is that a **simple correlation structure** is assumed to be **the same in every region of the brain**



Issues with applying the GLM to fMRI data

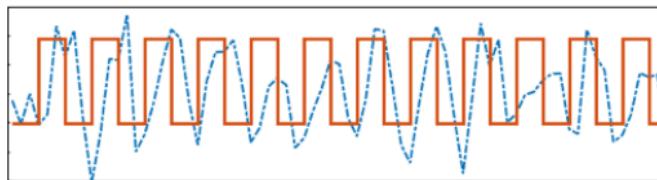
3. Sluggish nature of the **BOLD** signal

The **BOLD** signal takes ~6 seconds to reach a peak **after** a stimulus and does not return to baseline until ~20 seconds



A **dummy variable** representing our **experimental conditions** will not capture the **true shape** of the signal

This will lead to **larger errors** and a **less sensitive model**





Issues with applying the GLM to fMRI data

3. Sluggish nature of the BOLD signal — *solution*

A **more realistic** predicted signal can be created by **convolving** our **dummy variable** with a **model** of the BOLD response



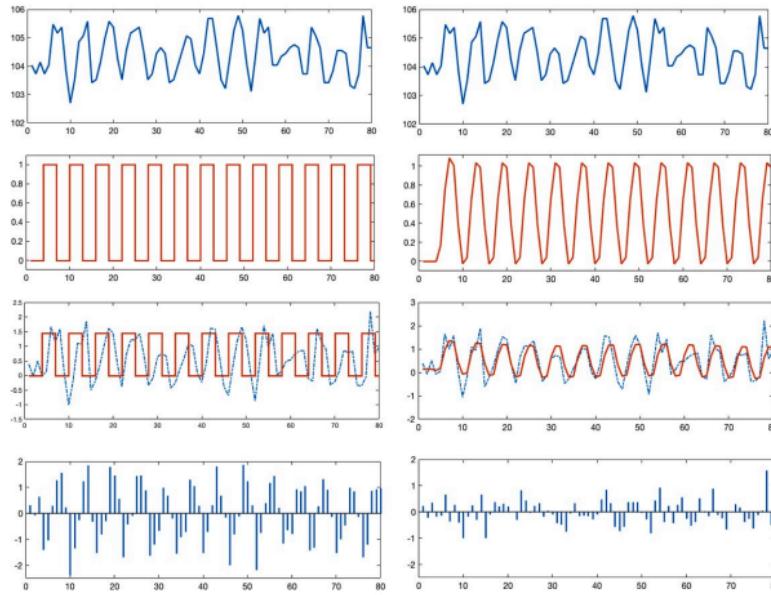
This model is known as the **hemodynamic response function** (HRF) — mixture of **two Gamma distributions**

Early studies used **deconvolution** methods to characterise the HRF (e.g. Friston *et al.*, 1994, Friston *et al.*, 1998)

The shape is **fixed** by SPM — only the **height** is **free to vary**

Issues with applying the GLM to fMRI data

3. Sluggish nature of the BOLD signal – *solution*



Raw signal

Predicted signal

Model fit

Error

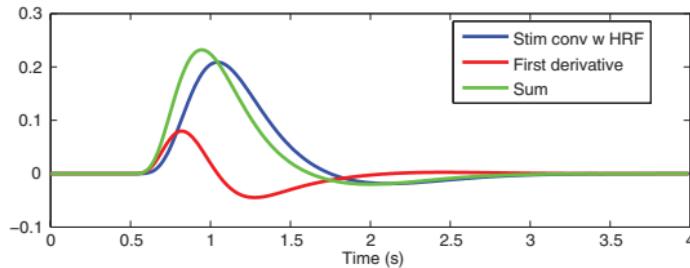


Issues with applying the GLM to fMRI data

3. Sluggish nature of the BOLD signal – *solution*

There are **alternatives** to the **canonical HRF** that aim to provide **more flexibility** at the expense of **more parameters**

A common alternative is the **HRF + derivatives** approach that includes **time** and **dispersion** derivatives to allow the **onset** and **width** of the response to vary



Issues with applying the GLM to fMRI data

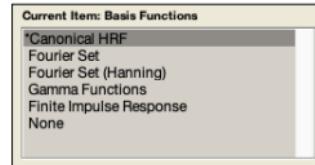
3. Sluggish nature of the BOLD signal – *solution*

There are **alternatives** to the **canonical HRF** that aim to provide **more flexibility** at the expense of **more parameters**

A common alternative is the **HRF + derivatives** approach that includes **time** and **dispersion** derivatives to allow the **onset** and **width** of the response to vary

Other approaches include:

- Finite Impulse Response (FIR)
- Fourier Set
- Gamma Functions



There can be better for **characterising** the **shape** of the response, but are **difficult** to use because of the **number of parameters**



Issues with applying the GLM to fMRI data

4. Arbitrary scaling of the BOLD signal

The magnitude of the BOLD signal is **arbitrary** and may differ from **subject** to **subject** and from **scanner** to **scanner**

The GLM **parameters** are on the scale of the data — β_1 gives the **change in signal** for a **unit change** in x_1

How can these parameter values be **meaningfully compared between subjects** if their value is governed by the **arbitrary scale of the signal?**



Issues with applying the GLM to fMRI data

4. Arbitrary scaling of the BOLD signal — *solution*

SPM will **scale** the data such that the **grand mean** (over voxels and volumes) is equal to 100 — **grand mean scaling** — placing all subjects on the **same scale**

Note that this **does not** mean the parameters can be interpreted as **percentage signal change**

frontiers in
NEUROSCIENCE

REVIEW ARTICLE
published: 21 January 2014
doi: 10.3389/fnins.2014.00001



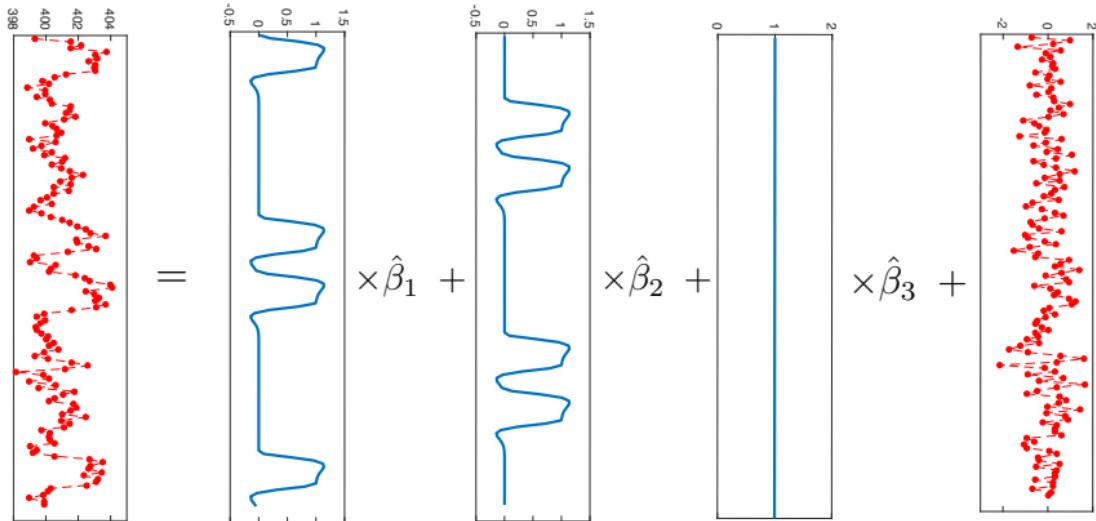
Misconceptions in the use of the General Linear Model applied to functional MRI: a tutorial for junior neuro-imagers

Cyril R. Pernet *

Brain Research Imaging Centre, Imaging Sciences, University of Edinburgh, Edinburgh, UK



The adapted GLM

Scaled, filtered and
whitened data**A** \mathbf{Y}^* = $\mathbf{X}^* \boldsymbol{\beta}$ **B**

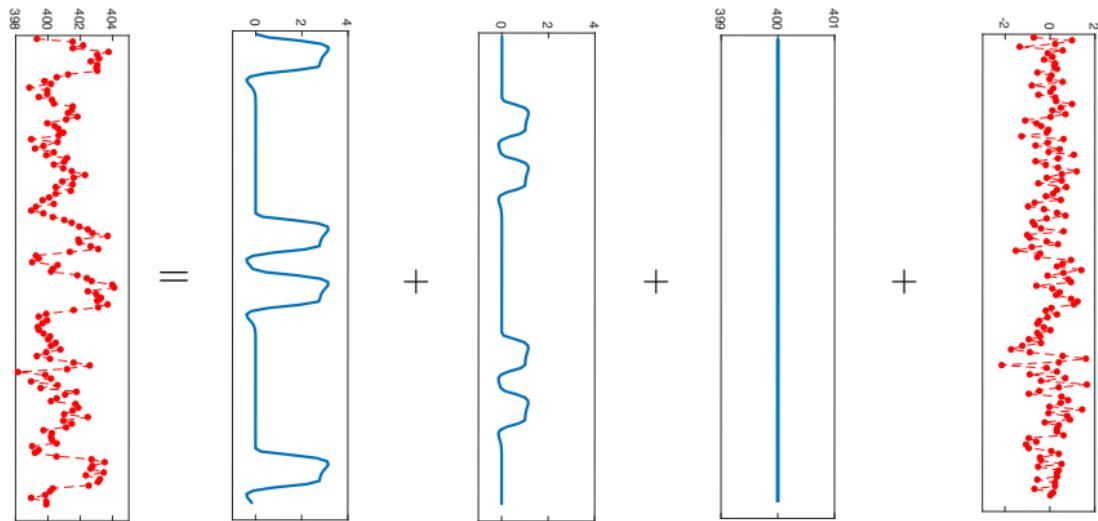
Constant

Error

+ ϵ^*



The adapted GLM

Scaled, filtered and
whitened data**A****B**

Constant

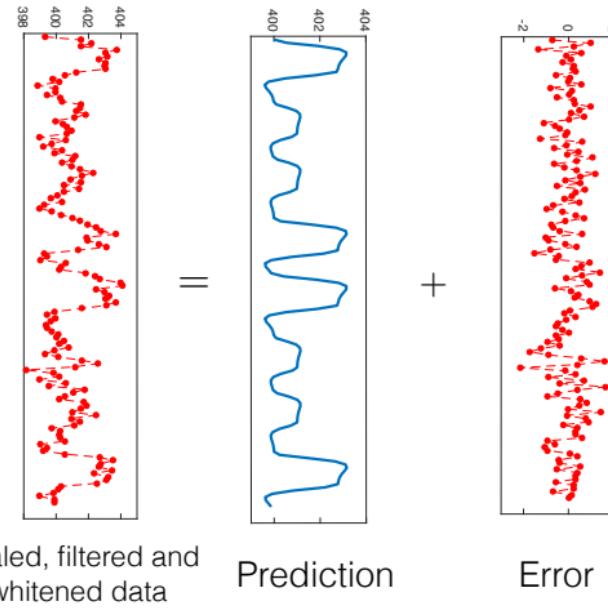
Error

$$\mathbf{Y}^* =$$

$$\mathbf{X}^* \beta$$

$$+ \epsilon^*$$

The adapted GLM



$$\mathbf{Y}^* = \mathbf{X}^* \boldsymbol{\beta} + \boldsymbol{\epsilon}^*$$



Contrasts and hypothesis tests

The **GLM** is essentially **multiple regression** in matrix form, inclusive of both continuous and categorical predictor variables

The GLM is **not suitable** for **raw** fMRI data, however, we can use it after **whitening**, **filtering** and **scaling** the data

The GLM **parameters** are **estimated** from the data and provide **numeric summaries** of the **experimental effects**

How do we use these **summaries** to reach **conclusions** about our experiment?

We need to ask **questions** about our model using the **language** of **contrasts** and **hypothesis tests**



Contrasts and hypothesis tests

What is a contrast?

A **linear combination** of the parameter estimates $\mathbf{c} = \mathbf{L}\hat{\beta}$

Where \mathbf{L} is a matrix/vector of **weights**

$$\hat{\beta}_1 - \hat{\beta}_2 = [1 \quad -1] \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$$

In SPM, we are always testing **null hypotheses** of the form

$$\mathcal{H}_0 : \mathbf{L}\hat{\beta} = 0$$

Some **linear combination** (*average, difference etc.*) of the parameter estimates is **equal to zero** (in the population)

Contrasts and hypothesis tests

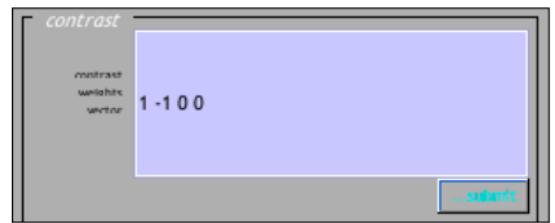
This is a very **flexible** system

$$\mathcal{H}_0 : [1 \quad 0] \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = 0 \quad \beta_1 \text{ is equal to } 0$$

$$\mathcal{H}_0 : [1 \quad -1] \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = 0 \quad \text{the } \mathbf{difference} \text{ between } \beta_1 \text{ and } \beta_2 \text{ is } 0$$

$$\mathcal{H}_0 : [0.5 \quad 0.5] \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = 0 \quad \text{the } \mathbf{average} \text{ of } \beta_1 \text{ and } \beta_2 \text{ is } 0$$

The **weights** are entered into SPM to test the hypothesis defined by the contrast

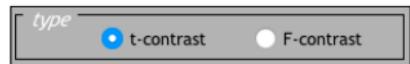




Contrasts and hypothesis tests

t-contrasts and **F**-contrasts

SPM uses two kinds of contrasts: *t* and *F*



A *t*-contrast will form a ***t*-statistic** at **every voxel** ($\nu = 1 \dots V$)

$$t_\nu = \frac{\mathbf{L}\hat{\beta}_\nu}{\sqrt{\sigma_\nu^2 \mathbf{L}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{L}'}}$$

A *t*-contrast is defined by

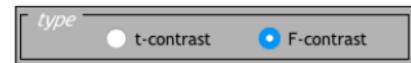
- An **L** matrix with a **single row**
- Hypothesis testing using a *t*-statistic
- **One-tailed** *p*-values for testing **directional hypotheses**
 - e.g. [1 -1] and [-1 1] will give **different results**



Contrasts and hypothesis tests

***t*-contrasts and *F*-contrasts**

SPM uses two kinds of contrasts: *t* and *F*



An *F*-contrast will form an ***F*-statistic** at **every voxel** ($v = 1 \dots V$)

$$F_v = \frac{(\mathbf{L}\hat{\beta}_v)'(\mathbf{L}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{L}')^{-1}(\mathbf{L}\hat{\beta}_v)}{r\sigma_v^2}$$

An *F*-contrast is defined by

- An **L** matrix with **multiple rows**
- Hypothesis testing using an *F*-statistic
- **Two-tailed** *p*-values for testing **non-directional hypotheses**
 - e.g. [1 -1] and [-1 1] will give the **same result**

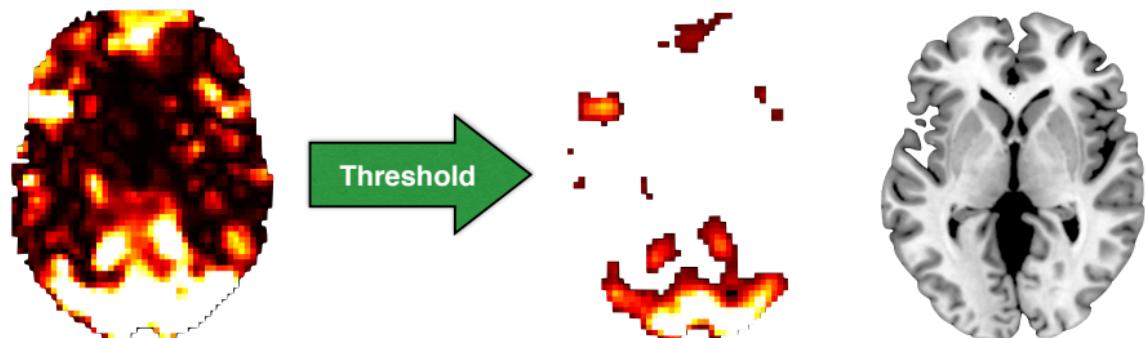


Contrasts and hypothesis tests

The Statistical Parametric Map

Whether a t -contrast or F -contrast, SPM will calculate the test statistic using the contrast weights **at each voxel**

This creates a **Statistical Parametric Map** — an image of test statistics



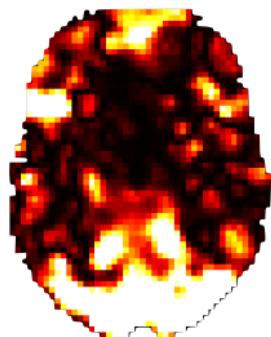


Contrasts and hypothesis tests

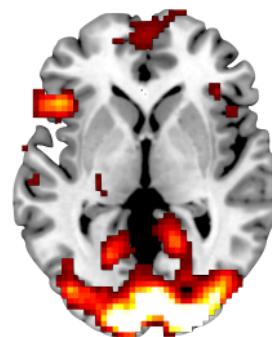
The Statistical Parametric Map

Whether a t -contrast or F -contrast, SPM will calculate the test statistic using the contrast weights **at each voxel**

This creates a **Statistical Parametric Map** — an image of test statistics



Threshold





Summary of the SPM approach

The tool we use for modelling our data is the **General Linear Model** — multiple regression in matrix form

To make the GLM suitable for fMRI we have to **filter**, **whiten** and **scale** the data as well as **convolving** our predictions with the **HRF**

The parameters of the GLM tells us about the **magnitude** and **direction** of the experimental effects — we ask questions by using **contrasts** of the parameter estimates at **each voxel**

These contrasts are converted to **test statistics** (t or F) to form a **Statistical Parametric Map**

These maps are then **thresholded** to indicate **brain regions** where the experimental effects are deemed “**significant**” — how this is done is coming up next!

References

Brett M, Penny W & Kiebel S (2007). Parametric Procedures. In *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Academic Press.

Friston KJ, Jezzard P & Turner R (1994). Analysis of functional MRI time series. *Human Brain Mapping*, 1, 1-19.

Friston KJ, Fletcher P, Josephs O, Holmes A, Rugg MD & Turner R (1998). Event-related fMRI: Characterizing differential responses. *NeuroImage*, 7, 30-40.

Kiebel SJ & Holmes AP (2007). The General Linear Model. In *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Academic Press.

Poline JB & Brett M (2012). The general linear model and fMRI: does love last forever? *NeuroImage*, 62(2), 871-80.