Dynamic causal modeling for fMRI

Methods and Models for fMRI, HS 2016

Jakob Heinzle
Structural, functional & effective connectivity

anatomical/structural connectivity
- presence of physical connections
- DWI, tractography, tracer studies (monkeys)

functional connectivity
- statistical dependency between regional time series
- correlations, ICA

effective connectivity
- causal (directed) influences between neuronal populations
- DCM
Dynamic causal modelling (DCM) for fMRI

- DCM framework was introduced in 2003 for fMRI by Karl Friston, Lee Harrison and Will Penny (NeuroImage 19:1273-1302)
- part of the SPM software package
- Allows to do an effective connectivity analysis
From functional segregation to functional integration

localizing brain activity:  
**functional segregation**

« Where, in the brain, did my experimental manipulation have an effect? »

**u₁**  
**u₁ X u₂**

effective connectivity analysis:  
**functional integration**

« How did my experimental manipulation propagate through the network? »

« Where, in the brain, did my experimental manipulation have an effect? »
Introductory example: Attention to motion

Assess site of attention modulation during visual processing in fMRI paradigm reported by Büchel and Friston.

How can the classical results be explained mechanistically:
- Photic $\rightarrow$ V1
- + Motion $\rightarrow$ V5
- + Attention $\rightarrow$ V5 + parietal cortex (SPC)

Friston et al. (2003) *NeuroImage*
Bayesian model selection

\[ \ln p(y|m) \]

models marginal likelihood

V1 → V5

stim

m_1

Modulation By attention

PPC

m_2

Modulation By attention

PPC

m_3

Modulation By attention

PPC

m_4

Modulation By attention

PPC

→ \( p(y|m_1) \)

→ \( p(y|m_2) \)

→ \( p(y|m_3) \)

→ \( p(y|m_4) \)

Stephan et al. 2008, *NeuroImage*
Parameter inference

We will run this example in the tutorial

Stephan et al. 2008, NeuroImage
Generative model

1. enforces mechanistic thinking: how could the data have been caused?
2. generate synthetic data (observations) by sampling from the prior – can model explain certain phenomena at all?
3. inference about model structure: formal approach to disambiguating mechanisms \( \rightarrow p(m|y) \)
4. inference about parameters \( \rightarrow p(\theta|y) \)
DCM approach to effective connectivity

A simple model of a neural network … … described as a dynamical system … … causes the data (BOLD signal).

\[ \dot{x} = f(x, u, \theta_x) \]
\[ y = g(x, \theta_y) \]

Let the system run with input \((u)\) and parameters \((\theta_x, \theta_y)\), and you will get a BOLD signal time course \(y\) that you can compare to the measured data.
Approximating $f(x, u, \theta_x)$

\[ \frac{dx}{dt} = f(x, u) \approx f(x_0, 0) + \frac{\partial f}{\partial x} x + \frac{\partial f}{\partial u} u + \frac{\partial^2 f}{\partial x \partial u} xu + \frac{\partial^2 f}{\partial x^2} x^2 + \cdots \]

Bi-linear model

Non-linear model

Modulation By attention

stim V1 V5 PPC

stim V1 V5 PPC
Approximating $f(x, u, \theta_x)$

$$\frac{dx}{dt} = f(x, u) \approx f(x_0, 0) + Ax + Cu + B xu + D x^2 + \ldots$$
The neural model – summary.

Parameter sets...

A - fixed connectivity
B - modulation of connectivity
C - weight of driving inputs
D - weight of non-linear terms

... determine dynamics!
The neural equations – bilinear model

\[ \frac{dx}{dt} = \left( A + \sum_{i=1}^{m} u_i B^{(i)} \right) x + Cu \]

Parameters A, B and C define connectivity!
The neural equations – non-linear model

\[
\frac{dx}{dt} = \left( A + \sum_{i=1}^{m} u_i B^{(i)} + \sum_{j=1}^{n} x_j D^{(j)} \right) x + Cu
\]

Parameters A, B, C and D define connectivity!
DCM parameters = rate constant

\[ a_{11} \]

\[ \frac{dx_1}{dt} = a_{11} x_1 \]

\[ x_1(t) = x_1(0) \exp(a_{11} t) \]

If \( A \rightarrow B \) is 0.10 \( \text{s}^{-1} \) this means that, per unit time, the increase in activity in \( B \) corresponds to 10% of the current activity in \( A \)

\[ \tau = \ln 2 / a_{11} \]

Decay function

If \( A \rightarrow B \) is 0.10 \( \text{s}^{-1} \) this means that, per unit time, the increase in activity in \( B \) corresponds to 10% of the current activity in \( A \)
Bilinear neural state equation

\[
\dot{x} = \left( A + \sum_{j=1}^{m} u_j B^{(j)} \right) x + Cu
\]

State changes

External modulatory inputs

Current state

External driving inputs

\[\theta = \{ A, B^1, \ldots, B^m, C \}\]

Fixed connectivity weights

Weights (strength) of connectivity modulation

Weights for direct inputs
The problem of the hemodynamic response

↑ neural activity
↓
↑ blood flow
↓
↑ oxyhemoglobin
↓
↑ T2*
↓
↑ MR signal

Rest

oxy-Hb
↓
deoxygenated Hb

Activity

Brief Stimulus

Peak

Undershoot

Initial dip

Source, Huettel et al, 2004, fMRI (Book)
From neural activity to the BOLD signal

Local hemodynamic state equation

\[ \dot{s} = x - \kappa s - \gamma (f - 1) \]
\[ \dot{f} = s \]

Balloon model

\[ \tau \dot{v} = f - v^{1/\alpha} \]
\[ \tau \dot{q} = f E(f, E_0)/E_0 - v^{1/\alpha} q/v \]

Changes in volume \( (v) \) and dHb \( (q) \)

BOLD signal change equation

\[ y = \frac{\Delta S}{S_0} \approx V_0 \left[ k_1 (1 - q) + k_2 \left( 1 - \frac{q}{v} \right) + k_3 (1 - v) \right] \]

cf. Simulations in Lecture 1
The hemodynamic model in DCM

- 6 hemodynamic parameters:
  
  \[ \theta^h = \{ \kappa, \gamma, \tau, \alpha, \rho, \varepsilon \} \]

  important for model fitting, but of no interest for statistical inference

- Computed separately for each area (like the neural parameters)
  → region-specific HRFs!

Friston et al. 2000, *NeuroImage*
Stephan et al. 2007, *NeuroImage*
The hemodynamic model in DCM – role of $\varepsilon$

Neural state equation
\[
\frac{dx}{dt} = \left( A + \sum_{j=1}^{m} u_j B^{(j)} \right) x + Cu
\]

Hemodynamic state equations
\[
\dot{s} = x - \kappa s - \gamma (f - 1)
\]

Flow induction ($\tau$CBF)
\[
\dot{f} = s
\]

Balloon model
\[
\tau v = f - v^{\frac{1}{\alpha}}
\]
\[
\tau q = f E(f,E_0)/E_0 - v^{\frac{1}{\alpha}}q/v
\]

BOLD signal change equation
\[
\lambda(q,v) = \frac{\Delta S}{S_0} \approx V_0 \left[ k_1 (1 - q) + k_2 \left( 1 - \frac{q}{v} \right) + k_3 (1 - v) \right]
\]

\[
k_1 = 4.3 \beta E_0 \tau E
\]
\[
k_2 = \gamma E_0 \tau E
\]
\[
k_3 = 1 - \varepsilon
\]

Stephan et al. 2007, NeuroImage
Hemodynamic forward models are important for connectivity analyses of fMRI data.

Summary – the full model

Hemodynamic model

\[ \tau_i(t) \] and \[ q_i(t) \]

Neuronal states

\[ x_i(t) \]

BOLD signal change equation

\[ y_i(t) \]

BOLD signal

\[ y(t) \]

Inputs

\[ x_i(t) \], \[ \tau_i(t) \], \[ q_i(t) \]

Input

\[ s_i(t) \], \[ f_i(t) \]

Hidden states

\[ y_i(t) \]

time [s]
Summary – the full model

Inputs

Neuronal states \( x_i(t) \)

Hemodynamic model \( \nu_i(t) \) and \( q_i(t) \)

BOLD signal change equation \( y_i(t) \)

BOLD signal \( y(t) \)

Inputs

\( x_i(t) \)

\( s_i(t) \)

\( f_i(t) \)

\( \nu_i(t) \)

\( q_i(t) \)

Hidden states

Time [s]
Summary – parameters of interest

Inputs

- Neuronal states $x_i(t)$
- Hemodynamic model $v_i(t)$ and $q_i(t)$
- BOLD signal change equation $y_i(t)$
- BOLD signal $y(t)$

Connection weights: A, B, C and D

Hemodynamic parameters: $\tau$, $\kappa$, $\varepsilon$
Example traces 1: Single node

\[
\begin{align*}
\dot{x} &= Ax + u_2 B^{(2)} x + C u_1 \\
\begin{bmatrix}
\dot{x}_1 \\
\dot{x}_2
\end{bmatrix} &= 
\begin{bmatrix}
\sigma & 0 \\
0 & \sigma
\end{bmatrix} \cdot \begin{bmatrix}
x_1 \\
x_2
\end{bmatrix} + u_2 \begin{bmatrix}
0 & 0 \\
0 & 0
\end{bmatrix} \cdot \begin{bmatrix}
x_1 \\
x_2
\end{bmatrix} + \begin{bmatrix}
c_{11} & 0 \\
0 & 0
\end{bmatrix} \cdot \begin{bmatrix}
u_1 \\
u_2
\end{bmatrix}
\end{align*}
\]
Example traces 2: Connected nodes

\[ \dot{x} = Ax + u_2 B^{(2)} x + C u_1 \]

\[
\begin{bmatrix}
\dot{x}_1 \\
\dot{x}_2
\end{bmatrix} = 
\begin{bmatrix}
\sigma & 0 \\
a_{12} & \sigma
\end{bmatrix} \cdot 
\begin{bmatrix}
x_1 \\
x_2
\end{bmatrix} + u_2 \begin{bmatrix} 0 \\
0
\end{bmatrix} \cdot 
\begin{bmatrix}
x_1 \\
x_2
\end{bmatrix} + \begin{bmatrix}
c_{11} \\
0
\end{bmatrix} \cdot 
\begin{bmatrix}
u_1 \\
u_2
\end{bmatrix}
\]

Diagram showing connected nodes and stimuli with inputs and neuronal activity.
Example traces 3: Modulation of connection

\[
\dot{x} = Ax + u_2 B^{(2)} x + Cu_1
\]

\[
\begin{bmatrix}
\dot{x}_1 \\
\dot{x}_2
\end{bmatrix} =
\begin{bmatrix}
\sigma & 0 \\
ar_{12} & \sigma
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2
\end{bmatrix} + u_2
\begin{bmatrix}
0 \\
b_{12}^{(2)} & 0
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2
\end{bmatrix} +
\begin{bmatrix}
c_{11} & 0 \\
0 & 0
\end{bmatrix}
\begin{bmatrix}
u_1 \\
u_2
\end{bmatrix}
\]
Example traces 4: Modulation of self-connection

\[
\begin{align*}
\dot{x} &= Ax + u_2 B^{(2)} x + C u_1 \\
\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} &= \begin{bmatrix} \sigma & 0 \\ a_{12} & \sigma \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + u_2 \begin{bmatrix} 0 & 0 \\ 0 & b_{22}^{(2)} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} c_{11} & 0 \\ 0 & 0 \end{bmatrix} u_1
\end{align*}
\]
Bayesian inference – forward and inverse model

forward problem

\[ p(y|\mathcal{G}, m) \]

likelihood

posterior distribution

\[ p(\mathcal{G}|y, m) \]

inverse problem
Bayesian inference

- **Likelihood:** \( p(y|\theta,m) \)
- **Prior:** \( p(\theta|m) \)
- **Bayes rule:** \( p(\theta|y,m) = \frac{p(y|\theta,m)p(\theta|m)}{p(y|m)} \)
Dynamical systems in Bayes

Assume data is normally distributed around the prediction from the dynamical model:

\[ p(y(t)|\theta, m) = \mathcal{N}(y(t), \theta_\sigma) \]

Dynamical model defines the likelihood!
Combining likelihood (data) and priors

Likelihood = Probability of data
- Derived from dynamical system
  - Gaussian noise

Priors (constraints):
- Hemodynamic parameters
  - empirical
- Neural parameters
  - self connections: principled
  - other parameters (inputs, connections): shrinkage

\[ p(\theta|y,m) = \frac{p(y|\theta,m) \cdot p(\theta|m)}{p(y|m)} \]
The likelihood of the data

\[ p(y|\theta, m) = \prod_t p(y(t)|\theta, m) \]
Type role and impact of priors

• Types of priors:
  ✓ Explicit priors on model parameters (e.g., connection strengths)
  ✓ Implicit priors on model functional form (e.g., system dynamics)
  ✓ Choice of “interesting” data features (e.g., regional timeseries vs ICA components)

• Role of priors (on model parameters):
  ✓ Resolving the ill-posedness of the inverse problem
  ✓ Avoiding overfitting (cf. generalization error)

• Impact of priors:
  ✓ On parameter posterior distributions (cf. “shrinkage to the mean” effect)
  ✓ On model evidence (cf. “Occam’s razor”)
  ✓ On free-energy landscape (cf. Laplace approximation)
Model estimation: running the machinery

- **Goal:** Find posterior of parameters $p(\theta|y, m)$ that maximises model evidence $p(y|m)$ given data and priors
- This is often not possible analytically $\rightarrow$ approximate methods are used

Variational Bayes
Eduardo Aponte

MCMC algorithms
Eduardo Aponte

BMS and BMA
Klaas Enno Stephan
Variational Bayes (VB)

Idea: find an approximate density $q(\theta)$ that is maximally similar to the true posterior $p(\theta \mid y)$.

This is often done by assuming a particular form for $q$ (fixed form VB) and then optimizing its sufficient statistics.
**Variational Bayes**

\[
\ln p(y) = \underbrace{\text{KL}[q||p]}_{\text{divergence}} + \underbrace{F(q,y)}_{\text{neg. free energy}} \geq 0
\]

(unknown) (easy to evaluate for a given \(q\))

\(F(q,y)\) is a functional wrt. the approximate posterior \(q(\theta)\).

Maximizing \(F(q,y)\) is equivalent to:

- minimizing \(\text{KL}[q||p]\)
- tightening \(F(q,y)\) as a lower bound to the log model evidence

When \(F(q,y)\) is maximized, \(q(\theta)\) is our best estimate of the posterior.
Bayesian system identification

**Neural dynamics**
\[ \dot{x} = f(x, u, \theta_x) \]

**Observer function**
\[ y = g(x, \theta_y) + \epsilon \]

**Posterior distribution**
\[ p(y|\theta, m) = \mathcal{N}(g(x, \theta_y), \Sigma(\theta_\sigma)) \]

**Prior distribution**
\[ p(\theta|m) = \mathcal{N}(\mu_\theta, \Sigma_\theta) \]

**Inference on model structure**
\[ p(y|m) = \int p(y|\theta, m)p(\theta|m)d\theta \]

**Inference on parameters**
\[ p(\theta|y, m) = \frac{p(y|\theta, m)p(\theta|m)}{p(y|m)} \]

**Diagram**
- **Design experimental inputs**
- **Define likelihood model**
- **Specify priors**
- **Invert model**
- **Make inferences**
Model estimation machinery

- Posterior parameters: $p(\theta|y,m)$
- Model evidence: $p(y|m)$

**Expectation-Maximization algorithm**

**Iterative procedure:**
1. Compute model response using current set of parameters
2. Compare model response with data
3. Improve parameters, if possible
Generative models & model selection

• any DCM = a particular generative model of how the data (may) have been caused

• generative modelling: comparing competing hypotheses about the mechanisms underlying observed data
  → a priori definition of hypothesis set (model space) is crucial
  → determine the most plausible hypothesis (model), given the data

• model selection ≠ model validation!
  → model validation requires external criteria (external to the measured data)
Multifactorial design: explaining interactions with DCM

Let’s assume that an SPM analysis shows a main effect of stimulus in $X_1$ and a stimulus $\times$ task interaction in $X_2$.

How do we model this using DCM?
Simulated data

Stephan et al. 2007, *J. Biosci.*
Note: GLM vs. DCM

DCM tries to model the same phenomena (i.e. local BOLD responses) as a GLM, just in a different way (via connectivity and its modulation).

No activation detected by a GLM
→ no motivation to include this region in a deterministic DCM.

However, a stochastic DCM could be applied despite the absence of a local activation.

Model comparison: Synesthesia

- "projector" synesthetes experience color externally co-localized with a presented grapheme
- "associators" report an internally evoked association

van Leeuwen et al., J Neurosci 2011
Model comparison: Synesthesia

• “projector” synesthetes experience color externally co-localized with a presented grapheme

• “associators” report an internally evoked association

• across all subjects: no evidence for either model

• but splitting into synesthesia types gives very strong evidence for bottom-up (projectors) and top-down (associators) mechanisms, respectively

van Leeuwen et al., J Neurosci 2011
“All models are wrong, but some are useful.”

George E.P. Box (1919-2013)
Hierarchical strategy for model validation

1. In silico: numerical analysis & simulation studies

2. Humans: cognitive experiments

3. Animals & humans: experimentally controlled system perturbations

4. Patients: clinical utility

For DCM: >15 published validation studies (incl. 6 animal studies):

- Infers site of seizure origin (David et al. 2008)
- Infers primary recipient of vagal nerve stimulation (Reyt et al. 2010)
- Infers synaptic changes as predicted by microdialysis (Moran et al. 2008)
- Infers fear conditioning induced plasticity in amygdala (Moran et al. 2009)
- Tracks anaesthesia levels (Moran et al. 2011)
- Predicts sensory stimulation (Brodersen et al. 2010)
DCM – graphical overview

neuronal dynamics

haemodynamics

state-space model

priors

Bayesian Model Inversion

fMRI data

posterior parameters

model comparison

priors
One slide summary

- Combining the neural and hemodynamic states gives the complete forward model.
- Observation model includes measurement error $e$ and confounds $X$ (e.g. drift).
- Bayesian inversion: parameter estimation variational Bayes or MCMC
- Result 1: A posteriori parameter distributions $p(\theta|y, m)$, characterised by mean $\eta_{\theta|y}$ and covariance $C_{\theta|y}$
- Result 2: Estimate of model evidence $p(y|m)$.

\[ \hat{x} = (A + \sum u_j B^j) x + Cu \]

stimulus function $u$

neural state equation

hidden states $z = \{x, s, f, v, q\}$

state equation $\dot{z} = F(x, u, \theta)$

parameters $\theta^h = \{\kappa, \gamma, \tau, \alpha, \rho\}$
$\theta^e = \{A, B^1...B^e, C\}$
$\theta = \{\theta^h, \theta^e\}$

activity-dependent vasodilatory signal $\dot{s} = z - \kappa s - \gamma(f-1)$

flow-induction (rCBF) $\dot{f} = s$

changes in volume $\dot{v} = f - v^{1+a}$

changes in dHB $\dot{q} = f E(f, \rho)/\rho - v^{1+a}q/v$

modelled BOLD response $y = \lambda(x)$

observation model $y = h(u, \theta) + X\beta + e$
Summary

1. enforces mechanistic thinking: how could the data have been caused?

2. mechanistically explains observed effects (modulations).

3. generate synthetic data (observations) by sampling from the prior – can model explain certain phenomena at all?

4. inference about model structure: formal approach to disambiguating mechanisms → $p(m|y)$

5. inference about parameters → $p(\theta|y)$
Many thanks to Hanneke den Ouden, Andreea Diaconescu, Jean Daunizeau and Klaas Enno Stephan for some of the slides!

Thank you!
Useful references 1

Useful references 2


DCM developments – for your reference

- Nonlinear DCM for fMRI: Could connectivity changes be mediated by another region? (Stephan et al. 2008, Neuroimage)
- Clustering DCM parameters: Classify patients, or even find new sub-categories (Brodersen et al. 2011, Neuroimage)
- Embedding computational models in DCMs: DCM can be used to make inferences on parametric designs like SPM (den Ouden et al. 2010, J Neuroscience)
- Integrating tractography and DCM: Prior variance is a good way to embed other forms of information, test validity (Stephan et al. 2009, Neuroimage)
- Stochastic DCM: Model resting state studies / background fluctuations (Li et al. 2011 Neuroimage, Daunizeau et al. Physica D 2009)
- Resting state DCM: Model second order interactions directly (Friston et al. 2014, Neuroimage)
- DCM for layered BOLD: Model high resolution fMRI to resolve layers (Heinzle et al. 2016, Neuroimage)
Note: The evolution of DCM in SPM

- DCM is not one specific model, but a framework for Bayesian inversion of dynamic system models
- The default implementation in SPM is evolving over time
  - better numerical routines for inversion
  - change in priors to cover new variants (e.g., stochastic DCMs, endogenous DCMs etc.)

To enable replication of your results, you should ideally state which SPM version (incl. release) you are using when publishing papers.
Matlab: [ver,release]=spm('ver');
Attention to motion in the visual system

Paradigm

Stimuli radially moving dots

Pre-Scanning
5 x 30s trials with 5 speed changes
Task - detect change in radial velocity

Scanning (no speed changes)
F A F N F A F N S ....
F - fixation
S - observe static dots + photic
N - observe moving dots + motion
A - attend moving dots + attention

Parameters
- blocks of 10 scans
- 360 scans total
- TR = 3.22 seconds
Attention to motion in the visual system

Paradigm

- fixation only
- observe static dots + photic → V1
- observe moving dots + motion → V5
- task on moving dots + attention → V5 + parietal cortex

Results

Attention - No attention

Büchel & Friston 1997, Cereb. Cortex
Büchel et al. 1998, Brain
Quiz: can this DCM explain the data?
Attention to motion in the visual system

Paradigm

Ingredients for a DCM

Specific hypothesis/question
Model: based on hypothesis
Timeseries: from the SPM
Inputs: from design matrix
Attention to motion in the visual system

**DCM – GUI basic steps**

1 – Extract the time series (from all regions of interest)

2 – Specify the model

3 – Estimate the model

4 – Repeat steps 2 and 3 for all models in model space
Quiz: can this DCM explain the data?
Additional material – not covered in lecture
Stephan et al. 2010, *NeuroImage*
Mean field assumption

Factorize the approximate posterior \( q(\theta) \) into independent partitions:

\[
q(\theta) = \prod_i q_i(\theta_i)
\]

where \( q_i(\theta_i) \) is the approximate posterior for the \( i^{th} \) subset of parameters.

For example, split parameters and hyperparameters:

\[
p(\theta, \lambda \mid y) \approx q(\theta, \lambda) = q(\theta)q(\lambda)
\]
VB in a nutshell (mean-field approximation)

1. Neg. free-energy approx. to model evidence.
   \[ \ln p(y \mid m) = F + KL\left[ q(\theta, \lambda), p(\theta, \lambda \mid y) \right] \]
   \[ F = \langle \ln p(y, \theta, \lambda) \rangle_q - KL\left[ q(\theta, \lambda), p(\theta, \lambda \mid m) \right] \]

2. Mean field approx.
   \[ p(\theta, \lambda \mid y) \approx q(\theta, \lambda) = q(\theta)q(\lambda) \]

3. Maximise neg. free energy wrt. \( q = \) minimise divergence, by maximising variational energies
   \[ q(\theta) \propto \exp(I_\theta) = \exp\left[ \langle \ln p(y, \theta, \lambda) \rangle_{q(\lambda)} \right] \]
   \[ q(\lambda) \propto \exp(I_\lambda) = \exp\left[ \langle \ln p(y, \theta, \lambda) \rangle_{q(\theta)} \right] \]

4. Iterative updating of sufficient statistics of approx. posteriors by gradient ascent.