Voxelwise fMRI Node Simulations
Specification File

Sam Harrison   November 19, 2013

Contents

1 Modular Workflow

We define a 'node' as a set of co-activating regions with a common neural[1] time course. There will be a number of these nodes activating and interacting over the course of any fMRI scan. The fundamental node equation to infer is

$$D^{(sr)} = h\left(P^{(s)}A_n^{(sr)}\right) + \varepsilon^{(sr)} \ . \tag{1}$$

We wish to test the efficacy of methods that try to recover these nodes from resting state scans, where in general none of the four terms on the right of equation 1 are known.

The obvious way to simulate data to test these methods is:

1. Generate a set of spatial maps (section 2)

2. Generate a set of neural time courses (section 3)

3. Convolve the overall neural captivity with an HRF (section 4)

4. Add noise and confounds (section 5)

1.1 Two stage approach

While the fundamental approach implied by equation 1 contains the general principles, it may be useful to explicitly consider special cases. For example, consider the model implied by the TFM approach. In this case, a high dimensional node parcellation is acquired (usually by sICA, but not necessarily), at which point the second step combines these nodes into *modes*[2]. The equation for this case is

$$D^{(sr)} = h\left(\left(P^{(s)}M^{(s)}\right)A_n^{(sr)}\right) + \varepsilon^{(sr)} \ . \tag{2}$$

In this case, $P^{(s)}$, represents a hard, atlas-based parcellation of the brain, with subject variability arising from registration errors. This set of nodes is then combined into spatially overlapping modes by the action of $M^{(s)}$.

2 Spatial Maps

The maps should have the following properties, which should be easy to vary:

- Sparsity

- Spatial coherence

- Correlation (and higher order statistics)

---

[1]Neural is taken to mean the activity before HRF convolution.

[2]The distinction between modes and nodes is usually drawn along the lines of modes spatially overlapping and containing anticorrelations.

- Subject variability

- Registration errors

There also needs to be a way to test atlas based methods. Could either map all of the above into a brain space, or perhaps generate an atlas in a purely abstract sense.
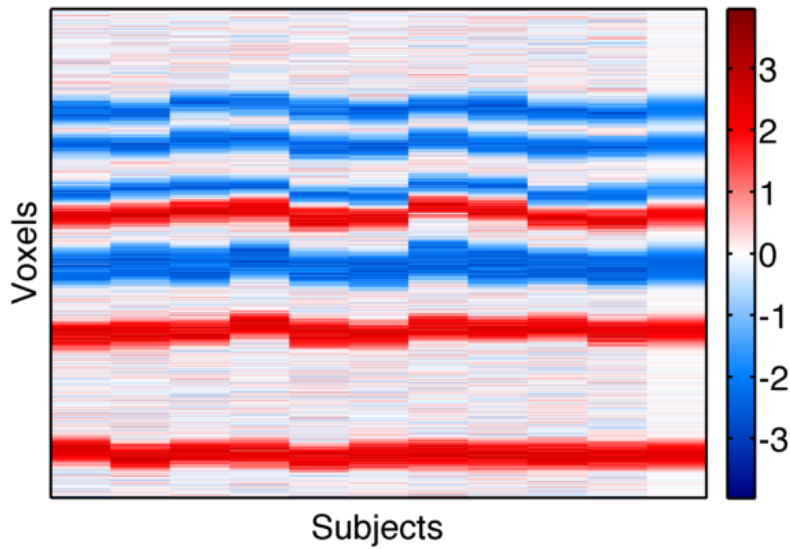

Figure 1: Example of subject variability for one spatial map. The observed maps, including registration errors, are shown for 10 subjects. The mean map, over subjects, is shown on the far right.

2.1 Two stage node and mode generation

As described in section 1.1, it may be advantageous to split the spatial maps into a node / mode formulation. In that case, the properties of the nodes should be:

- Binary weights

- Spatial coherence

- Non-overlapping

- Registration errors

Then, the modes need to introduce the properties that are missing compared to the list in section 2. Therefore, the modes should introduce:

- Correlation between modes (and higher order statistics)

- Anti-correlation within modes

- Subject variability

One of the advantages of this formulation is that atlas based methods can be introduced relatively painlessly—the node maps, without the subject misalignments, can form the atlas.

3   Neural Time Courses

These represent the neural activity, so should be simulated at a higher sampling rate than TR. They should also be simulated for long enough that there are no transient HRF effects (e.g. from zero-padding the convolution). Some of the required properties are:

- Non-Gaussianity

- Correlation

- Non-stationarity

- Coloured spectra

- Presence / absence in subjects—c.f. left / right lateralised language

4   HRF Convolution

This operates on the voxelwise neural activity, and the following should be available, computation permitting:

- Linear / nonlinear

- Subject / spatial variability

- Spatio-temporal structure?

- High frequency BOLD?

5   Noise & Confounds

This is simply added to the 'clean' BOLD signal generated in the previous steps. The simplest model would be spatially and temporally white noise, but some of the following might be useful to include:

- Temporal / spatial structure

4

- Non-Gaussianities

- Voxelwise variance changes

- Interaction with BOLD signal e.g. noise power modulated by signal power

- Physiological confounds—aliased cardiac, very low frequency drifts etc

- Global signal

- Artefactual nodes e.g. motion, multiband etc

6   Pre-processing

Want to simulate the standard pre-processing pipeline, so de-meaning, variance normalisation, high-pass filtering etc. Could also investigate less common procedures e.g. global signal regression.

Could also investigate different concatenation approaches for methods that assume that.

7   Scoring

In the case of a nonlinear HRF there is not an obvious ground truth for methods that infer a linear decomposition. The best case is probably the set of subject maps (including registration errors). The ground truth time courses should probably be just taken as the spatial maps regressed out of the clean BOLD signal.

A complementary scoring mechanism might be the accuracy in separating the BOLD signal from the noise (which is say what the PCA step before an ICA decomposition aims to do). Again, it's hard to formulate what a good comparison between an atlas method and say ICA would be—this possibly relies on the ability to separate the signal from the noise.

Maybe a sensible question to ask with these various methods is: do all the voxels within the spatial map contain the same BOLD time course, whatever that may be?