ELSEVIER

NeuroImage

www.elsevier.com/locate/ynimg
NeuroImage 23 (2004) S264–S274

Modelling functional integration: a comparison of structural equation and dynamic causal models

W.D. Penny,* K.E. Stephan, A. Mechelli, and K.J. Friston

Wellcome Department of Imaging Neuroscience, University College London, London, United Kingdom

Available online 25 September 2004

The brain appears to adhere to two fundamental principles of functional organisation, functional integration and functional specialisation, where the integration within and among specialised areas is mediated by effective connectivity. In this paper, we review two different approaches to modelling effective connectivity from fMRI data, structural equation models (SEMs) and dynamic causal models (DCMs). In common to both approaches are model comparison frameworks in which inferences can be made about effective connectivity per se and about how that connectivity can be changed by perceptual or cognitive set. Underlying the two approaches, however, are two very different generative models. In DCM, a distinction is made between the 'neuronal level' and the 'hemodynamic level'. Experimental inputs cause changes in effective connectivity expressed at the level of neurodynamics, which in turn cause changes in the observed hemodynamics. In SEM, changes in effective connectivity lead directly to changes in the covariance structure of the observed hemodynamics. Because changes in effective connectivity in the brain occur at a neuronal level DCM is the preferred model for fMRI data. This review focuses on the underlying assumptions and limitations of each model and demonstrates their application to data from a study of attention to visual motion.

Keywords: Structural equation; Dynamic causal model; Functional integration

Introduction

Human brain mapping has been used extensively to provide functional maps showing which regions are specialised for specific functions (Frackowiak et al., 2003). A classic example is the study by Zeki et al. (1991) who identified V4 and V5 as specialised for the processing of colour and motion, respectively. More recently, these analyses have been augmented by functional integration studies, which describe how functionally specialised areas interact and how these interactions depend on changes of context.

Early analyses of functional integration used principal component analysis (PCA) to decompose neuroimaging data into a set of modes that are mutually uncorrelated both spatially and temporally. The modes are ordered according to the amount of variance they explain. By comparing the temporal expression of the first few modes with the variation in experimental factors, a distributed functional system associated with various factors can be identified (Friston et al., 1993). A more sophisticated use of PCA occurs in the context of generalised eigenimage analysis (Friston et al., 1997), where the principal component is found which is maximally expressed in one experimental condition or population and minimally expressed in another (e.g., control versus patient groups). If there are more than two experimental factors, this approach can be extended using a canonical variates analysis (CVA) or partial least squares (PLS) (MacIntosh et al., 1996).

More recently, independent component analysis (ICA) has been used to identify modes describing activity in a sparsely distributed network (McKeown et al., 1998). Such PCA/ICA-based methods are called analyses of functional connectivity as they are data-driven transform methods, which make no assumptions about the underlying biology. They are therefore of greatest practical use when it is not clear which regions are involved in a given task.

In contrast, analyses of 'effective connectivity' (see the following sections) are based on statistical models that make anatomically motivated assumptions (e.g., knowledge of structural connectivity) and restrict their inferences to networks comprising a number of preselected regions. Effective connectivity analyses are hypothesis driven rather than data driven and are most applicable when one can specify the relevant functional areas (e.g., from analyses of functional specialisation). The presence of connections, in the model, can be inferred from data obtained by invasive tracing procedures in primates, assuming homology between certain areas in the human and monkey brain. New imaging methodologies such as diffusion tensor imaging also hold the promise of providing information about anatomical connections for the human brain directly (Ramnani et al., 2004).

Detailed discussions of functional versus effective connectivity approaches can be found in chapters 48–53 of (Frackowiak et al., 2003). In this paper, we review the most widely used method for making inferences about functional integration from fMRI, namely, structural equation modelling (SEM). We also review dynamic

* Corresponding author. Fax: +20 7813 1420.
E-mail address: w.penny@fil.ion.ucl.ac.uk (W.D. Penny).
Available online on ScienceDirect (www.sciencedirect.com.)

1053-8119/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2004.07.041

causal modelling (DCM), a new approach that has been designed specifically for the analysis of fMRI time series.

The paper is structured as follows. The sections Structural equation models and Dynamic causal models describe the theoretical foundations of SEM and DCM, and the Attention to visual motion section presents exemplar analyses on fMRI data. We conclude with a discussion of the relative merits of the models in the Discussion section.

Notation

We use uppercase letters to denote matrices and lowercase to denote vectors. I_K denotes the $K \times K$ identity matrix, 1_K is a $1 \times K$ vector of 1's and 0_K is a $1 \times K$ vector of zeros. If X is a matrix, $\mathrm{Tr}(X)$ denotes its trace, $|X|$ its determinant, X_{ij} the i,jth element, X^T the matrix transpose, X^{-1} the matrix inverse, X^{-T} the transpose of the matrix inverse, $\mathrm{vec}(X)$ returns a column vector comprising its columns and $\otimes$ denotes the Kronecker product. The operator $\mathrm{diag}(x)$ returns a diagonal matrix with leading diagonal elements given by the vector x. $\log x$ denotes the natural logarithm. If $p(x) = N(x; \mu, \Sigma)$ then the d-dimensional random vector x is drawn from a multivariate Gaussian distribution with mean μ and covariance Σ. This is given by

$$N(x;\mu,\Sigma) = (2\pi)^{-d/2}|\Sigma|^{-1/2}\exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right) \quad (1)$$

Structural equation models

Structural equation models (SEMs) were developed in the field of econometrics and first applied to imaging data by McIntosh and Gonzalez-Lima (MacIntosh and Gonzalez-Lima, 1991). They comprise a set of regions and a set of directed connections. Importantly, a causal semantics is ascribed to these connections where an arrow from A to B means that A causes B. Causal relationships are thus not inferred from the data but are assumed a priori (Pearl, 1998).

An SEM with particular connection strengths implies a particular set of instantaneous correlations among regions. One can therefore set the connection strengths so as to minimise the discrepancy between the observed and implied correlations and thereby fit a model to data. If, for example, one partitions a given fMRI data set into those scans obtained under two different levels of an experimental factor, then one can attribute differences in the estimated connection strengths to that factor, and so conclude that a pathway has been activated. To date, SEMs have been the most widely used model for connectivity analyses in neuroimaging (Goncalves and Hull, 2003).

Generative model

We consider networks comprising N regions in which the activity at time t is given by the $N \times 1$ vector y_t. If there are T time points and Y is an $N \times T$ data matrix comprising $t = 1 \ldots T$, then the likelihood of the data is given by

$$p(Y|\theta) = \prod_{t=1}^{T} p(y_t|\theta) \quad (2)$$

where θ are the parameters of an SEM. This is the first SEM equation and is important as it embodies the key assumption that network activity is independent from sample to sample. This is certainly valid for PET data but is, at best, questionable for fMRI as samples are known to be temporally autocorrelated (Woolrich et al., 2001). There are however heuristics that allow one to overcome this problem, as we shall see in the following section.

The second SEM equation specifies the generative model at time t

$$p(y_t|\theta) = N(y_t; 0, \Sigma(\theta)) \quad (3)$$

which denotes that the activities are zero mean Gaussian variates with a covariance, $\Sigma(\theta)$, that is, a function of the connectivity matrix θ. The form of this function is specified implicitly by the regression equation that describes how activity in one area is related to activity in other areas via a set of path coefficients, M, as

$$y_t = My_t + e_t \quad (4)$$

where e_t are zero mean Gaussian innovations or errors of covariance R. Typically, R will be a diagonal matrix and we write the error variance in region i as σ_i^2. Regions are connected together via the $N \times N$ path coefficient matrix M where the M_{ij} denotes a connection from region j to region i. The parameters of an SEM, θ, are the unknown elements of M and R. The above equation is an unusual regression equation as the dependent variable appears on both sides of the equality. By subtracting My_t from both sides and multiplying by $(I_N - M)^{-1}$, where I_N is the identity matrix, the equation can be rearranged as follows

$$y_t = (I_N - M)^{-1}e_t \quad (5)$$

This form is particularly useful as it shows us how to generate data from the model. Firstly, we generate the Gaussian variates e_t and then premultiply by $(I_N - M)^{-1}$. This is repeated for each t. This form also allows us to express the covariance of y_t as a function of θ

$$\Sigma(\theta) = (I_N - M)^{-1}R(I_N - M)^{-T}. \quad (6)$$

Estimation

Given a set of parameters θ, we can compute the likelihood of a data set from Eqs. 2, 3 and 6. Given a data set one can therefore find the connectivity matrix that maximises the likelihood using standard optimisation methods such as Pseudo-Newton algorithms or simplex methods [9]. However, optimisation of Eq. 2 would soon run into numerical problems as the probabilities are so small. It is therefore better to maximise the log-likelihood

$$L(\theta) = \log p(Y|\theta) \quad (7)$$

$$= \sum_{t=1}^{T} \log p(y_t|\theta) \quad (8)$$

which, being monotically related to the likelihood, has the same maximum. By plugging in the Gaussian density from Eqs. 1 and 3 we get

$$L(\theta) = -\frac{T}{2}\log|\Sigma(\theta)| - \frac{NT}{2}\log 2\pi - \frac{1}{2}\sum_{t=1}^{T} y_t^T\Sigma(\theta)^{-1}y_t \quad (9)$$

If we define the sample covariance as

$$S = \frac{1}{T}\sum_{t=1}^{T} y_t\, y_t^T \tag{11}$$

then, by noting that the last term is a scalar and that the trace of a scalar is that same scalar value and using the circularity property of the trace operator [i.e., Tr (AB) = Tr (BA)], we can write

$$L(\theta) = -\frac{T}{2}\log|\Sigma(\theta)| - \frac{NT}{2}\log 2\pi - \frac{T}{2}\mathrm{Tr}\left(S\Sigma(\theta)^{-1}\right). \tag{12}$$

If we use unbiased estimates of the sample covariance matrix (which corresponds, for example, to the assumption that S is drawn from a Wishart distribution—see page 134 in Bollen, 1989) then we replace T's in the above equation by $T-1$'s. If we now also drop those terms that are not dependent on the model parameters we get

$$L(\theta) = -\frac{T-1}{2}\left(\log|\Sigma(\theta)| + \mathrm{Tr}\left(S\Sigma(\theta)^{-1}\right)\right). \tag{13}$$

Maximum likelihood estimates can therefore be obtained by maximising the above function.

We close this section with some remarks about identifiability. A model is identifiable if there is a unique parameter vector, θ, that maximises the likelihood. SEMs without loops (e.g., reciprocal connections) are identifiable. They are however biologically uninteresting and it is difficult to establish sufficient conditions for the identifiability of SEMs with loops. Because of this SEM modellers appeal to the concept of 'local identifiability'. This is the condition that in the neighborhood of θ, there are no other parameter vectors with equivalent likelihood. SEMs are locally identifiable if the Hessian (the matrix of second order partial derivatives of the log-likelihood with respect to the parameters) is nonsingular. This provides a practical test for local identifiability.

Inference

One can then use a likelihood ratio (LR) test to assess the merits of one model versus another (for a given specificity, no other test has higher sensitivity). If $p(Y|\theta,m=i)$ and $p(Y|\theta,m=j)$ are the likelihoods of the fitted models $m=i$ and $m=j$, then the likelihood ratio for comparing models i and j is

$$R_{ij} = \frac{p(Y|\theta,m=i)}{p(Y|\theta,m=j)}. \tag{14}$$

If $L(\theta_i)$ and $L(\theta_j)$ are the corresponding log-likelihoods, then the log of the likelihood ratio is

$$\log R_{ij} = L(\theta_i) - L(\theta_j). \tag{15}$$

Under the null hypothesis that the models are identical, and for large T, $-2\log R_{ij}$ is distributed as a chi-squared variable having degrees of freedom equal to the difference in number of parameters between the models (see p. 265 in Bollen, 1989). This provides a mechanism for model comparison. An important caveat is that the models must be nested.

A special case of the above test arises when one wishes to evaluate the goodness of fit of a single model. We will denote this as 'model 0'. This can be achieved by comparing the likelihood of model 0 to the likelihood of the least restrictive (most complex) model one could possibly adopt. This alternative model, denoted 'model 1', obtains simply by setting the model covariance equal to the sample covariance, that is, $\Sigma(\theta) = S$. The likelihood of this extravagant model is

$$L_1 = -\frac{T-1}{2}\left(\log|S| + \mathrm{Tr}\left(SS^{-1}\right)\right) \tag{16}$$

$$= -\frac{T-1}{2}(\log|S| + N). \tag{17}$$

The corresponding (log) likelihood ratio is

$$\log R_{01} = -\frac{T-1}{2}\left(\log|\Sigma(\theta)| + \mathrm{Tr}\left(S\Sigma(\theta)^{-1}\right) - \log|S| - N\right) \tag{18}$$

which in turn has a corresponding chi-squared value

$$\chi^2 = (T-1)F(\theta) \tag{19}$$

where

$$F(\theta) = \log|\Sigma(\theta)| + \mathrm{Tr}\left(S\Sigma(\theta)^{-1}\right) - \log|S| - N. \tag{20}$$

The corresponding degrees of freedom are equal to the degrees of freedom in model 1, k, minus the degrees of freedom in model 0, q. For an N-dimensional covariance matrix, there are $k = N(N+1)/2$ degrees of freedom. For model 0, q equals the total number of connectivity and variance parameters to be estimated. The associated χ^2 test provides one way of assessing if an SEM model is good enough. We reject our model, model 0 (or the null model), if the associated P value is less than, for example, 0.05. In simple terms, we reject our model if its implicit covariance is significantly different from the sample covariance.

We also note that it is possible to obtain maximum likelihood estimates by minimising $F(\theta)$. This is because ignoring terms that are fixed for a given data set, $F(\theta) = -L(\theta)$. The quantity $F(\theta)$ is proportional to the Kullback–Liebler (KL) divergence between the probability density of the samples and the probability density of the model. Thus, maximising the likelihood is equivalent to minimising this KL divergence.

For more general model comparisons, the χ^2 statistic associated with the LR test can be written as

$$\chi^2 = (T-1)(F(\theta_1) - F(\theta_2)). \tag{21}$$

For this reason, the LR test is also known as the chi-square difference test.

A caveat concerning these inferences is that they are based on the assumption that the data points are independent samples. However, as discussed in the previous section, fMRI data are serially correlated. So rather than there being T-independent samples, the 'effective' number of independent samples is somewhat less, say v. This provides a rationale for the heuristic in which the above tests are implemented using v in place of T, and v is calculated based on an autoregressive model of the serial correlation (Friston et al., 1995). The above tests can then be used to make inferences about effective connectivity.

To make inferences about changes in effective connectivity, one can also use the model comparison approach. This is sometimes called the 'stacked model' approach. It involves partitioning or splicing the original data, according to any experimental factor that causes a putative change in connectivity. In this paper, for example,

we will look at the effect of 'attention' on connectivity. This factor has two levels, 'attention' and 'no attention', so two partitions are created. One then constructs a 'null model' in which path coefficients are constrained to be equal between conditions ('no attention' and 'attention') and an 'alternative model' in which coefficients of interest can be different. The null and alternative models are then compared using an LR test. The philosophy behind this approach is identical to that used in analysis of variance (ANOVA) in which two regression models are compared, one having identical coefficients over the levels of a factor and one having different coefficients.

One can also make inferences about changes in effective connectivity via the use of moderator variables, as described in (Buchel and Friston, 1997). This involves creating 'dummy regions' whose activities are specified as follows. If one wishes to test whether experimental factor X changes the connectivity from region A to region B, one creates a dummy region C containing the mean-corrected data in A multiplied by the mean-corrected factor X. This represents the interaction between X and the physiological variate from A. This is formally identical to the explanatory variable in psychophysiological interactions (PPIs) (Friston et al., 1997) and plays a similar role to bilinear effects in DCMs that are discussed below. If, in the fitted SEM model containing A, B and C, the path coefficient from C to B is significantly nonzero, then there is a significant modulation.

One problem with the moderator variable approach, however, is that the data in the dummy regions are highly non-Gaussian (due to the multiplication with a discrete variable). This therefore violates the assumptions of the model. One way around this is to modify the generative model and so to maximise a different objective function (see, for example, chapter 4 in Bollen, 1989). A more fundamental problem, however, is that we may not have enough degrees of freedom to fit SEMs with a sufficiently rich structure (e.g., models containing reciprocal connections). These richer models may be specified in the stacked model approach because the available degrees of freedom is larger by a factor K, where K is the number of partitions. This is because the models must explain K sample covariances. This issue will be revisited in the Attention to visual motion section.

An alternative approach that enables more complex models to be fitted is to set the variance parameters, σ_i^2, to arbitrary values rather than estimating them. In McIntosh and Gonzalez-Lima (1994), for example, it is suggested that they be set to between 35% and 50% of the total variance in that region. Alternatively, one can employ more complex connectivity patterns but constrain sets of connections to be identical. An example of this is the use of reciprocally connected networks where the path coefficients are constrained to be symmetric (Rowe et al., 2002). The disadvantage of these heuristics is that they may result in poorer model fits.

We have described SEM as implemented in the majority of applications to functional brain imaging data. Further details of these applications can be found in (Bullmore et al., 2000). In the wider SEM world, however, SEMs vary in their generative models, estimation procedures and styles of inference. It is possible, for example, to define SEMs with exogenous variables. Further, one school of SEM modelling employs a Bayesian approach where priors are placed over model parameters and the aim of estimation is to find the maximum posterior (rather than maximum likelihood) parameters. This body of work (see, for example, (Scheines et al., 1999)) is of particular interest because the Bayesian framework is also used in DCM.

Dynamic causal models

Whereas SEM was developed in econometrics, dynamic causal modelling (DCM) (Friston et al., 2003) has been specifically designed for the analysis of functional imaging time series. The term 'causal' in DCM arises because the brain is treated as a deterministic dynamical system (see, for example, Section 1.1 in Friston et al. (2003)) in which external inputs cause changes in neuronal activity, which in turn cause changes in the resulting blood oxygen level-dependent (BOLD) signal that is measured with fMRI. The term 'cause' is therefore used quite differently than in SEM (Pearl, 1998).

Generative model

Current DCMs for fMRI comprise a bilinear model for the neurodynamics and an extended balloon model (Friston, 2002; Buxton et al., 1998) for the hemodynamics. The neurodynamics are described by the following multivariate differential equation

$$\dot{z}_t = \left(A + \sum_{j=1}^{J} u_t(j) B^j \right) z_t + C u_t \tag{22}$$

where t indexes continuous time, the dot notation denotes a time derivative, z_t is neuronal activity, $u_t(j)$ is the jth of J inputs at time t and A, B^j and C are connectivity matrices that will be described below. This is known as a bilinear model because the dependent variable, $\dot{z}_t$, is linearly dependent on the product of z_t and u_t. That u_t and z_t combine in a multiplicative fashion endows the model with 'nonlinear' dynamics that can be understood as a nonstationary linear system that changes according to u_t. Importantly, because u_t is known, parameter estimation is tractable. The neuronal activity z_t is an $N \times 1$ vector comprising activity in each of the N regions and the input u_t is a $J \times 1$ vector comprising the scalar inputs $u_t(j)$ where $j = 1 \ldots J$. Exemplar neuronal time series are shown for a simple two-region DCM in Fig. 1.

The effective connectivity in DCM is characterised by a set of 'intrinsic connections', A, that specify which regions are connected and whether these connections are unidirectional or bidirectional. These are analagous to the path coefficients, M, in SEM. Unlike SEM (as used in fMRI to date), however, we also define a set of input connections, C, that specify which inputs are connected to which regions, and a set of modulatory connections, B^j, that specify which intrinsic connections can be changed by which inputs. The overall specification of input, intrinsic and modulatory connectivity comprises our assumptions about model structure. This in turn represents a scientific hypothesis about the structure of the large-scale neuronal network mediating the underlying sensorimotor or cognitive function.

The values in the connectivity matrices can be concatenated into the connectivity vector

$$\theta^c = \begin{bmatrix} \mathrm{vec}\,(A) \\ \mathrm{vec}\,(B) \\ \mathrm{vec}\,(C) \end{bmatrix} \tag{23}$$

where, for example, vec(A) returns a column vector comprising the columns of A. Model structure is defined by specifying which entries in the above matrices are allowed to take on nonzero values, that is, which inputs and regions are connected. A given model, say model m, is then defined by its pattern of connectivity. Note that only connections which are allowed to be nonzero will appear in

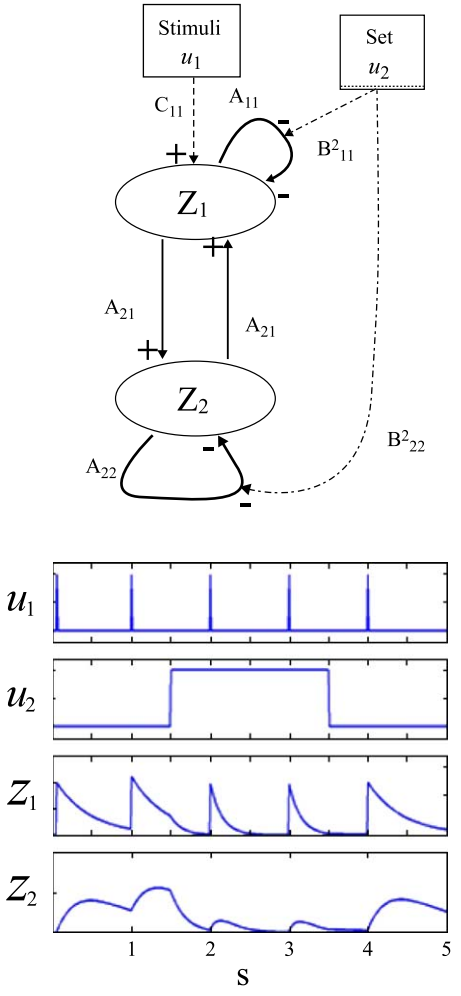

Fig. 1. DCM Neurodynamics. The top panel shows a dynamic causal model comprising $N = 2$ regions and $M = 2$ inputs. The input variable u_1 drives neuronal activity z_1. Informally, neuronal activity in this region then excites neuronal activity z_2, which then reactivates activity in region 1. Formally, these interactions take place instantaneously according to Eq. 22. The time constants are determined by the values of the intrinsic connections A_{11}, A_{12}, A_{21} and A_{22}. Input 2, typically a contextual input such as instructional set, then acts to change the intrinsic dynamics via the modulatory connections B^2_{11} and B^2_{22}. In this example, the effect is to reduce neuronal time constants in each region as can be seen in the neuronal time series in the bottom panel. The y-axis scale is in arbitrary units and the x-axis is in units of seconds.

θ^c. For a network with N_a intrinsic, N_b modulatory and N_c input connections θ^c will have $N_\theta = N_a + N_b + N_c$ entries.

In DCM, neuronal activity gives rise to hemodynamic activity by a dynamic process described by an extended balloon model. This involves a set of hemodynamic state variables, state equations and hemodynamic parameters θ^h (for details, see (Friston et al., 2003)). Exemplar hemodynamic responses are shown in Fig. 2.

We can concatenate all neurodynamic and hemodynamic parameters into the overall p-dimensional parameter vector

$$\theta = \begin{bmatrix} \theta^c \\ \theta^h \end{bmatrix} \tag{24}$$

This vector contains all the parameters of a DCM model that we need to estimate.

For given input u and DCM parameters θ, model predictions, $h(\theta,u)$ can be produced by integrating the state equation as described in (Friston et al., 2003; Friston, 2002). This numerical integration is efficient because most fMRI experiments employ input vectors that are highly sparse by experimental design. For a data set with N_s scans, we can then create an $NN_s \times 1$ vector of model predictions $h(\theta,u)$ covering all time points and all areas (in the order all time points from region 1, region 2, etc.). The observed data y, also formatted as an $NN_s \times 1$ vector, is then modelled as

$$y = h(\theta, u) + X\beta + w \tag{25}$$

where w is an $NN_s \times 1$ vector of Gaussian prediction errors with mean zero and covariance matrix C_e, X contains effects of no interest and β is the associated parameter vector. The matrix X would include, for example, regressors to model scanner-related low-frequency drifts in fMRI time series that are neurobiologically irrelevant. The error covariance is given by $C_e = I_{Ns} \otimes \Lambda$ where $\otimes$ denotes the Kronecker product, Λ is an $N \times N$ diagonal matrix with Λ_{ii} denoting error variance in the ith region.

To generate data from a DCM, one integrates the neurodynamics that are described by Eq. 22 together with the hemodynamics that are described by Eqs. (3) and (4) in Friston et al. (2003). Effects of no interest are then added according to 25. Fig. 1 shows a neuronal time series generated from a simple two-region DCM.

Priors are placed on the A and B^j matrices so as to encourage parameter estimates that result in a stable dynamic system (for a discussion, see Section 2.3.1 in Friston et al. (2003)). For each connection in A and B^j, the prior is

$$p(A_{ik}) = N(A_{ik}; 0, v_a) \tag{26}$$

$$p\left(B^j_{ik}\right) = N\left(B^j_{ik}; 0, v_b\right)$$

where the prior variance v_a is set to ensure stability with high probability (for a discussion of this issue, see Appendix A.3 in Friston et al. (2003)). For each connection in C the prior is

$$p\left(C_{ij}\right) = N\left(C_{ij}; 0, v_c\right). \tag{27}$$

These priors are so-called 'shrinkage-priors' because the posterior estimates shrink towards the prior mean, which is zero.

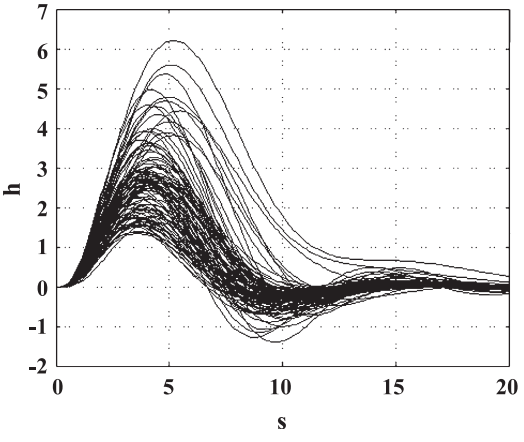

Fig. 2. DCM Hemodynamics. These distributions characterise our expectations about what the hemodynamic responses, h, should look like as a function of time (seconds). We first generated neuronal transients from a single-region DCM according to Eq. 22. Then, for each transient we drew a sample θ^h from the prior over θ^h (see Ref. [11]) and generated a hemodynamic response. We repeated this 100 times to produce the 100 curves shown in the figure.

The size of the prior variance determines the amount of shrinkage. The above information can be concatenated into the overall prior

$$p(\theta^c) = N\left(\theta_p^c, C_p^c\right) \tag{28}$$

where the p subscripts denote priors and

$$\theta_p^c = [0_{N_\theta}]^T$$

$$C_p^c = \text{diag}[v_a 1_{N_a}, v_b 1_{N_b}, v_c 1_{N_c}]. \tag{29}$$

In the above equations, 1_K is a $1 \times K$ vector of 1's and 0_K is a $1 \times K$ vector of zeros. The choice of the prior mean, θ_p^h, and covariance, C_p^h, is usually based on empirical estimates and estimates from physiological studies as described in Friston et al. (2003). Samples of hemodynamics responses from this prior are shown in Fig. 2. Consequently, the overall prior mean and covariance for a DCM are given by

$$\theta_p = \begin{bmatrix} \theta_p^c \\ \theta_p^h \end{bmatrix} \tag{30}$$

$$C_p = \begin{bmatrix} C_p^c & 0 \\ 0 & C_p^h \end{bmatrix}.$$

The prior and likelihood distributions for a given DCM model, say model m, are therefore

$$p(\theta|m) = N\left(\theta; \theta_p, C_p\right) \tag{31}$$

$$p(y|\theta, m) = N(y; h(\theta, u) + X\beta, C_e)$$

These are used in Bayes rule to form the posterior distribution, as described in the Estimation section.

Relation to SEM

If we assume that (i) the neurodynamics are directly observable, that is,. $y_t = z_t$,, and (ii) that the direct inputs are stochastic, that is,. $e_t = Cu_t$ then the generative model for DCM becomes

$$\dot{y}_t = Ay_t + e_t \tag{32}$$

In this context, the effect of modulatory inputs would be accommodated by splitting the data into different partitions, each partition having its own intrinsic connectivity matrix.

Further, if we decompose the intrinsic connectivity matrix into $A = \Theta - I_N$ where Θ is an off-diagonal matrix, then we have unit self-inhibition within regions and arbitrary connections between regions. This gives

$$\dot{y}_t = (\Theta - I_N)y_t + e_t. \tag{33}$$

If we now assume that the dynamics have converged at the point of observation, that is, $\dot{y}_t = 0$, then the generative model reduces to

$$y_t = \Theta y_t + e_t \tag{34}$$

which is identical to SEM (cf. Eq. 4). This could be implemented in DCM by having very strong shrinkage priors (cf. Eq. (26)). Therefore, a second perspective on SEMs is that they correspond to DCMs with stochastic inputs where neuronal states can be directly observed. But unfortunately for SEM, what we observe are hemodynamics not neurodynamics. It is also assumed that these dynamics have reached equilibrium at each point of observation. In other words, the dynamics are assumed to occur over a time scale that is short relative to the fMRI sampling interval. This is also not the case as the time scale of hemodynamics is much longer than this.

Estimation

From Bayes' rule the posterior distribution is equal to the likelihood times the prior divided by the evidence (Gelman et al., 1995)

$$p(\theta|y, m) = \frac{p(y|\theta, m)p(\theta|m)}{p(y|m)}. \tag{35}$$

Taking logs gives

$$\log p(\theta|y, m) = \log p(y|\theta, m) + \log p(\theta|m) - \log p(y|m). \tag{36}$$

The parameters that maximise this posterior probability, the maximum posterior (MP) solution, can then be found using a Gauss–Newton optimisation scheme, whereby parameter estimates are updated in the direction of the gradient of the log-posterior by an amount proportional to its curvature (see, e.g., (Press et al., 1992)). The model parameters are initialised to the mean of the prior density. Because the posterior probability consists of two terms, the likelihood and the prior, the maximum posterior solution is the one which optimally satisfies the two constraints (i) that the model prediction errors are minimised and (ii) that the parameters are close to their prior values.

If the proportion of data points to model parameters is sufficiently large, as is the case with DCM models of fMRI time series, then the posterior is well approximated with a Gaussian. The aim of optimisation is then to estimate the mean and covariance of this density, which can be achieved using an expectation–maximisation (EM) algorithm described in Section 3.1 of Friston (2002). In the E-step, the posterior mean and the posterior covariance are updated using a Gauss–Newton step, and in the M-step the hyperparameters of the noise covariance matrix, C_e, are updated. Both the E and M steps can be expressed in closed form as shown in Friston (2002). These steps are iterated until the posterior distribution

$$p(\theta|y, m) = \mathbf{N}(\theta_{\text{MP}}, \Sigma_{\text{MP}}) \tag{37}$$

is reached, where the subscripts MP denote maximum posterior values. The posterior density can then be used to make inferences about the size of connections. Fig. 8, for example, shows the posterior distribution for a modulatory coefficient.

In statistics, approximation of a posterior density by a Gaussian centred on the maximum posterior solution is known as a Laplace approximation (Kass and Raftery, 1993). The parameters of no interest, β, can also be estimated by forming an augmented parameter vector that includes θ and β and an augmented observation model, as described in Eq. (7) of Friston et al. (2003).

We close this section with some remarks about identifiability. As in SEM (see end of the Estimation section), it is difficult to establish the identifiability of DCMs. We note, however, that EM

optimisation always produces positive definite estimates of the posterior covariance matrix. If we define the Hessian as the matrix of second order partial derivatives of the log-posterior (instead of the log-likelihood, cf. SEM in the Estimation section) then because the Hessian is the inverse of the covariance, EM optimisation guarantees 'local identifiability'. We also note that due to the influence of the prior, the log-posterior is a smoother function than the log-likelihood. It is therefore more likely that these local optima correspond to global optima.

Inference

The structure of a DCM model is defined by specifying which regions are connected to each other, via the intrinsic connectivity matrix, and which inputs can alter which connections, via the modulatory matrix. A given model, say model m, is then defined by this pattern of connectivity. Different models can be compared using the evidence for each model. This can be thought of as a second-level of Bayesian inference. The model evidence is computed from

$$p(y|m) = \int p(y|\theta, m)p(\theta|m)d\theta. \tag{38}$$

Note that the model evidence is simply the normalisation term from the first level of Bayesian inference, given in Eq. 35. In Penny et al. (2004), we show how the evidence can be estimated using Akaike's information criterion (AIC), Bayesian information criterion (BIC) or the Laplace approximation.

Model comparison can then take place using evidence ratios. Given models $m = i$ and $m = j$, the 'Bayes factor' comparing model i to model j is defined as (Kass and Raftery, 1993, 1995),

$$B_{ij} = \frac{p(y|m = i)}{p(y|m = j)} \tag{39}$$

where $p(y|m = j)$ is the evidence for model j found by exponentiating AIC, BIC or Laplace approximations to the log-evidence. When $B_{ij} > 1$, the data favour model i over model j, and when $B_{ij} < 1$ the data favour model j. The Bayes factor is a summary of the evidence provided by the data in favour of one scientific theory, represented by a statistical model, as opposed to another. Just as a culture has developed around the use of P values in classical statistics (e.g., $P < 0.05$), so one has developed around the use of Bayes factors. For example, Bayes factors of 20 or more provide strong evidence in favour of one model over another Penny et al. (2004).

The use of Bayes factors or 'evidence ratios' for DCM is analagous to the use of likelihood ratio tests for SEM (cf. Eq. 14 and Eq. 39). The difference is that the likelihood ratio R_{ij} depends on estimated parameters. This means that R_{ij} is a random variable and so inference must be based on its distribution. With Bayes factors, there is no dependence on parameters (they have been integrated out using Eq. 38), and so they can be interpreted directly, as described above.

Finally we note that, in the context of SEM, a statistical test concerning the 'absolute' fit of a model was derived by comparing it to the most complex SEM one could imagine, that is, one where the model covariance is set to the sample covariance. This suggests a test concerning the 'absolute' fit of a DCM that is based on comparing the evidence for that DCM to the evidence of a DCM having the most complex structure one can imagine, that is, full intrinsic connectivity.

Attention to visual motion

In previous work, we have established that attention modulates connectivity in a distributed system of cortical regions mediating visual motion processing (Buchel and Friston, 1997; Friston and Buchel, 2000). These findings were based on data acquired using the following experimental paradigm. Subjects viewed a computer screen that displayed either a fixation point, stationary dots or dots moving radially outward at a fixed velocity. In some epochs of moving dots, they had to attend to changes in the speed of radial motion (which were actually absent), in other epochs they were instructed to simply watch the dots. For the purposes of our analyses in this paper, we can consider three experimental variables. The 'photic stimulation' variable indicates when dots were on the screen, the 'motion' variable indicates that the dots were moving and the 'attention' variable indicates that the subject was attending to possible velocity changes. These are the three input variables that are shown in Fig. 3. In this paper, we model the activity in three regions V1, V5 and superior parietal cortex (SPC). The original 360-scan time series were extracted from the data set of a single subject using a local eigendecomposition and are shown in Fig. 4.

As an example of the sort of modelling one can do, we look at the effect attention has on effectivity connectivity. Specifically, how attention affects the connection between V1 and V5 (Fig. 4).

SEM

We first apply SEM. The first step in this analysis is to partition the data set into (i) periods in which the subject was attending to moving stimuli and (ii) periods in which stimuli were moving but the subject did not attend to that movement. SEM is therefore only provided data for epochs when the stimulus is in motion. Note that this subset of data still contains motion-related responses (not just steady-state responses) because of the hemodynamic delay. There are 80 fMRI samples in each data set and these compose the 'no attention' and 'attention' conditions.

To make inferences about changes in effective connectivity, we apply the model comparison approach. This involves creating a 'null model' in which path coefficients are fixed between conditions ('no attention' and 'attention') and an 'alternative

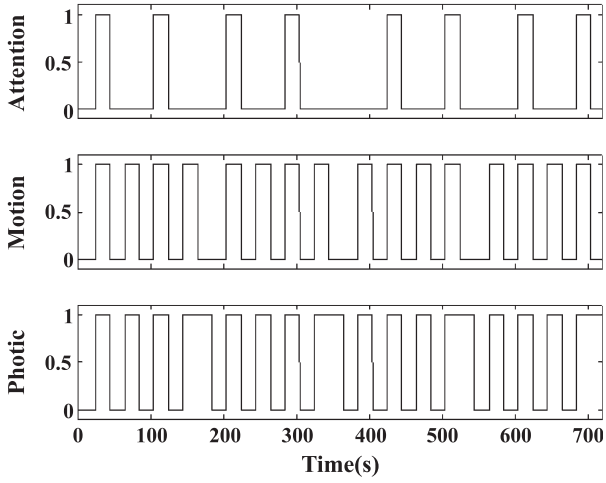

Fig. 3. Attention data: experimental variables. The plots bottom to top show the 'Photic', 'Motion' and 'Attention' inputs, u_i, used in the analysis of the attention to visual motion data.

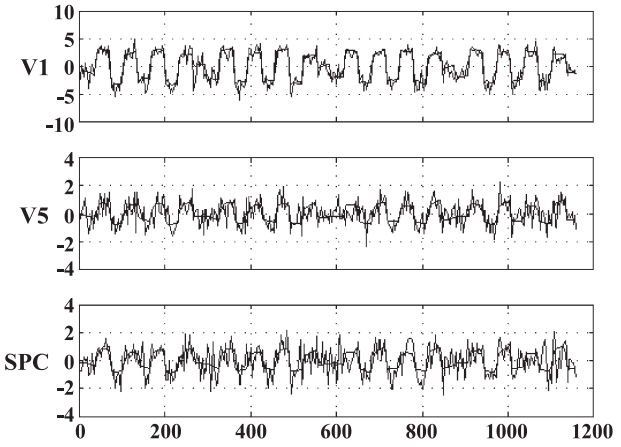

Fig. 4. Attention data: fMRI time series. The plots show fMRI time series (rough solid lines) from regions V1, V5 and SPC and the corresponding estimates from DCM model 2 (smooth solid lines).

model' in which coefficients of interest can vary. In both models, we allow the variance parameters, σ_i^2, to vary between conditions.

In an $N = 3$ region network, there are $k = N(N + 1) / 2 = 6$ degrees of freedom per data set giving a total of $k = 12$ over both conditions. As we are estimating the variance components and allowing them to vary between conditions, this takes six degrees of freedom leaving a maximum of six estimable path coefficients over the two conditions. This provides a limit on the complexity of the SEM one can fit.

With this in mind, we initially constructed an SEM with a purely feedforward structure as shown in Fig. 5. In the null model, the two path coefficients are common between conditions but in the alternative model the path coefficient from V1 to V5, $M_{V1,V5}$, can vary. Testing this model against the null model will allow us to infer whether or not 'attention' changes $M_{V1,V5}$.

In the alternative model, as applied to the two data sets, there are six variance parameters to estimate and three path coefficients, giving $q = 9$. Fig. 5 shows the estimated values of the path coefficients. The overall fit of the alternative model, which captures discrepancies between the sample covariance and implied model covariance matrix (also shown in Fig. 5), was $\chi^2 = 24.6$. This model has $k - q = 3$ degrees of freedom. In contrast, the null model had $\chi^2 = 33.2$ with four degrees of freedom. This leads to the conclusion that attention does indeed significantly change the value of this connection ($\chi^2 = 8.6$, $df = 1$, $P = 0.003$).

If we look at the absolute fit of the alternative model, however, then an LR test tells us that its implied covariance matrix is significantly different from the sample covariance matrix ($\chi^2 = 24.6$, $df = 3$, $P = 2e^{-5}$). This means that it is not a good model. Indeed, one can see from the covariance matrices in Fig. 5 that the covariance between V1 and SPC is not modelled accurately in either the 'attention' or 'no attention' conditions.

We then set up an SEM with reciprocal connectivity between regions as shown in Fig. 6. The assumption of reciprocal connectivity is biologically more realistic. In the null model, the four path coefficients are common between conditions but in the alternative model $M_{V1,V5}$ is allowed to vary. Again, testing this model against the null model allows us to infer whether or not 'attention' changes $M_{V1,V5}$.

In the alternative model, as applied to the two conditions, there are six variance parameters to estimate and five path coefficients, giving $q = 11$. Fig. 6 shows the estimated values of the path coefficients. The overall fit of the alternative model, which captures discrepancies between the sample covariance and implied model covariance matrix (also shown in Fig. 6), was $\chi^2 = 3.9$. This model has $k - q = 1$ degrees of freedom. In contrast, the null model had $\chi^2 = 23.6$ with two degrees of freedom. This leads to the conclusion that attention does indeed significantly change the value of this connection ($\chi^2 = 19.7$, $df = 1$, $P = 9e^{-6}$).

If we look at the absolute fit of the alternative model, then an LR test tells us that its implied covariance matrix is not significantly different from the sample covariance matrix ($\chi^2 = 3.9$, $df = 1$, $P = 0.05$). This means that it is a good model (although, with $P = 0.05$ as a cutoff point, one could argue that its a borderline case). Indeed, one can see from the covariance matrices in Fig. 6 that the covariance between V1 and SPC is now modelled much more accurately than before.

Note that it would not be possible to fit the reciprocal model to the data if we had used the 'moderator variable' approach (see Inference section). This is because, without partitioning the data into two subsets, there would only be $k = 6$ degrees of freedom instead of $k = 12$.

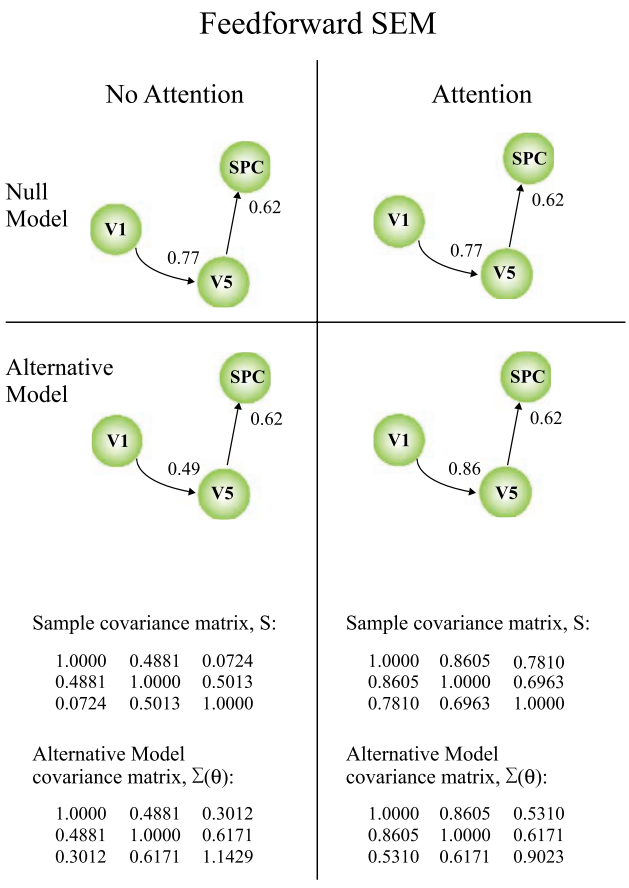

Fig. 5. Feedforward SEM. The figures show null and alternative SEMs fitted to the 'no-attention' and 'attention' periods of the data set. In the alternative model, the V1 to V5 path coefficient was allowed to vary between conditions. The entries in the covariance matrices are ordered V1, V5 and SPC.

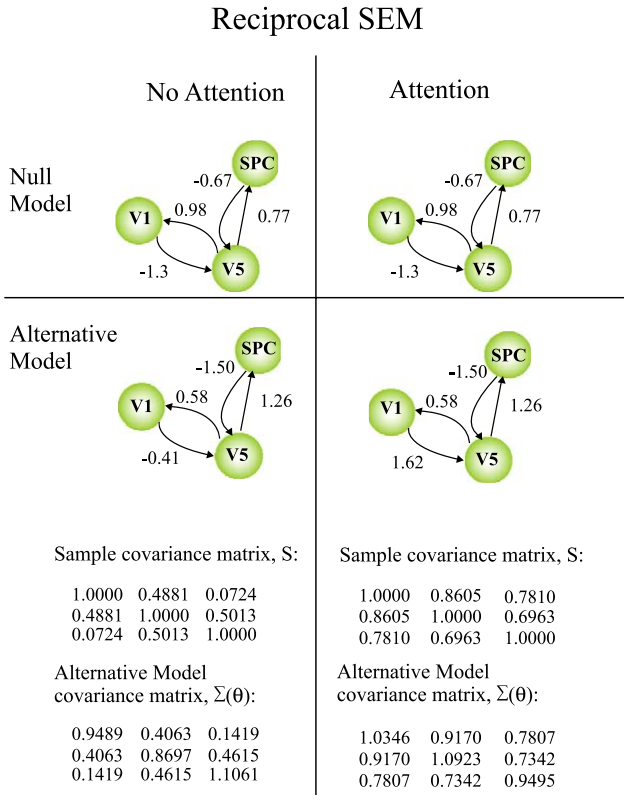

Fig. 6. Reciprocal SEM. The figures show null and alternative SEMs fitted to the 'no-attention' and 'attention' periods of the data set. In the alternative model, the V1 to V5 path coefficient was allowed to vary between conditions. The entries in the covariance matrices are ordered V1, V5 and SPC.

DCM

We then applied DCM to this data set. With DCM, there is no need to partition the time series into selected periods of interest as inferences about changes in connectivity can be made based on the strength of modulatory connections.

Three different DCMs were applied, each differing in their intrinsic connectivity structure; model 1 has a feedforward structure, model 2 a reciprocal structure and model 3 a fully connected structure. The models and their estimated parameters are shown in Fig. 7. All models assume that both motion and attention modulate the connection from V1 to V5.

A Bayes factor, comparing model 2 against model 1, of the order 10^{20} provides decisive evidence in favour of the network with reciprocal connections. That the modulatory connection (due to attention) within this model is significant can be seen by looking at the posterior distribution shown in Fig. 8. This allows one to make statements about our belief that this modulation is larger than some threshold, γ. For example, the probability that this effect is larger than zero is 0.98 and the probability that it is larger than 0.17 is 0.78. This sort of inference can also be made via a model comparison approach in which we compare model 2 to the same model but without the modulatory connection.

A comparison of model 2 against model 3 resulted in no consistent evidence either way. This suggests that the model with reciprocal connections is a sufficiently good model of the data. This is analagous to the LR test for SEM models that assesses the 'absolute' fit of a single model. Both SEM and DCM approaches conclude that models with reciprocal connections are good models per se and are superior to feedforward models. They also both conclude that attention significantly modulates the connectivity from V1 to V5.

Generative models

A final perspective on the different modelling approaches is offered by revisiting the generative models associated with SEM and DCM (see the Generative model section under the Structural equation models and Dynamic causal models sections). Here, we use SEM and DCM models that were fitted to the attention data. Fig. 9a shows a new data set generated from the reciprocal SEM, and Fig. 9b shows a new data set generated from the reciprocal DCM.

Firstly, we note that SEM only provides data for periods in which the stimulus is in motion (hence the absences of data in Fig. 9a). Secondly, the 'time series' it produces are very spiky. This is because, in SEM, samples are assumed to be statistically independent. These time series bear little resemblance to the actual fMRI traces in Fig. 4. We would therefore conclude from this comparison that the generative model underlying SEM is poorly suited to fMRI.

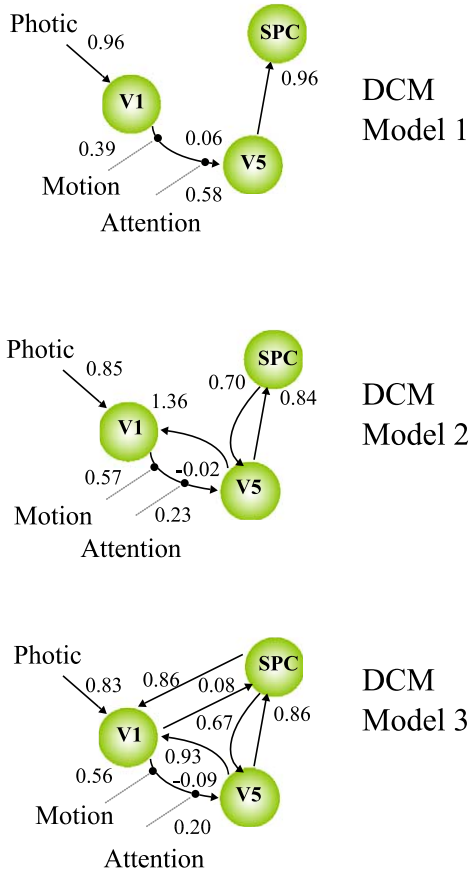

Fig. 7. DCM models. In all models, photic stimulation enters V1 and the motion and attention variables modulate the connection from V1 to V5. Models 1, 2 and 3 differ in their intrinsic connections, model 1 having feedforward structure, model 2 having reciprocal and hierarchically organised intrinsic connectivity and model 3 having full connectivity.

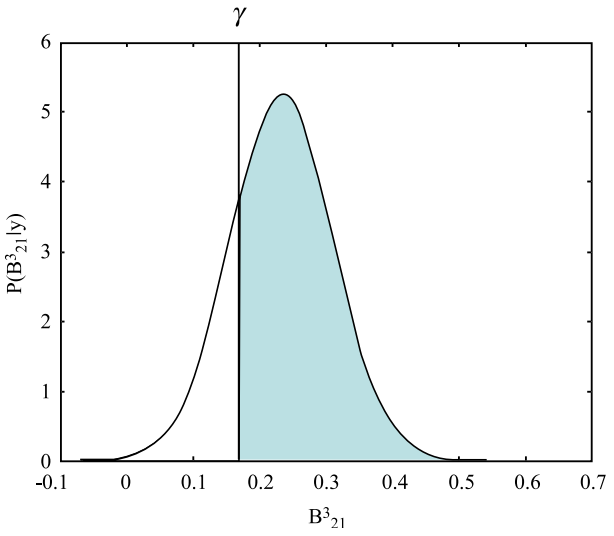

Fig. 8. DCM model: posterior distribution. The plot shows the posterior probability distribution of the parameter B_{21}^{1}. This is the connection from region 1 (V1) to region 2 (V5) that is modulated by attention (the third input). The mean value of this distribution is 0.23. This is also shown in Fig. 7. We can use this distribution to compute our belief that this connection is larger than some threshold γ. If we choose, for example, γ = (log 2) / 4 = 0.17, then this corresponds to computing the probability that this modulatory effect occurs within 4 s. In DCM, faster effects are mediated by stronger connections (see, for example, Eq. 22). For our data, we have $p(B_{21}^{3} > \gamma) = 0.78$.

In contrast, DCM models the entire time series, captures its underlying regularities and produces very similar traces to the actual fMRI data in Fig. 4. One attractive option is to use DCM as a modelling laboratory in which one can investigate the putative effects of experimental manipulations on changes in effective connectivity. This can take place before any data are collected and can provide an aid to experimental design. Furthermore, prior to starting such a study, one can verify that a given model, designed to test a particular hypothesis, is sufficiently sensitive to detect the effects of interest given typical signal to noise ratios (Penny et al., 2004).

Discussion

In this paper, we have compared the use of SEM and DCM for making inferences about changes in effective connectivity from fMRI time series. On our fMRI attention to visual motion data, both SEM and DCM approaches led to the same conclusions (i) that reciprocal models are superior to feedforward models, (ii) that models with reciprocal connections provide a good fit to the data and (iii) that attention significantly modulates the connectivity from V1 to V5.

There are data sets, however, where DCM will be able to make inferences that cannot be made with SEM. Such an example is given in Friston et al. (2003) in which DCM was used to make inferences about changes in connectivity in a three-region auditory system network. Of key interest was whether the repeated presentation of auditory stimuli reduced activity via neuronal saturation in addition to hemodynamic saturation (it did). Such an inference is clearly impossible, even in principle, in SEM as no distinction is made between 'neuronal' and 'hemodynamic' levels.

More generally, as compared to SEM, DCM has the following advantages. Firstly, DCM models interactions at the neuronal rather than the hemodynamic level. As well as being biologically accurate, this is important because neuronal interactions do not necessarily lead to detectable hemodynamic interactions (Gitelman et al., 2003). DCMs are able to work at the neuronal level because they employ a 'forward model' (with hemodynamic parameters) relating neuronal activity to fMRI activity, and this model is inverted during the model fitting process. Secondly, in DCM, one can postulate arbitrarily complex connectivity patterns between regions. This both results in better fitting models and is, again, biologically more realistic. Thirdly, because DCM uses Bayesian model comparison, one can compare nonnested network models (Penny et al., 2004). Finally, DCM uses a sufficiently rich generative model that one can use it as a modelling laboratory in which one can investigate the putative effects of experimental manipulations on changes in effective connectivity. This can take place before any data are collected and provides an aid to experimental design. Because of these advantages, DCM is the preferred method for making inferences about changes in effective connectivity from fMRI data. SEM is, however, appropriate for PET data.

A current limitation of DCM is that model fitting is computationally demanding. As implemented in SPM2 (Friston, 2002), one is limited to modelling networks comprising approximately up to eight regions. Parameter estimation in these models takes of the

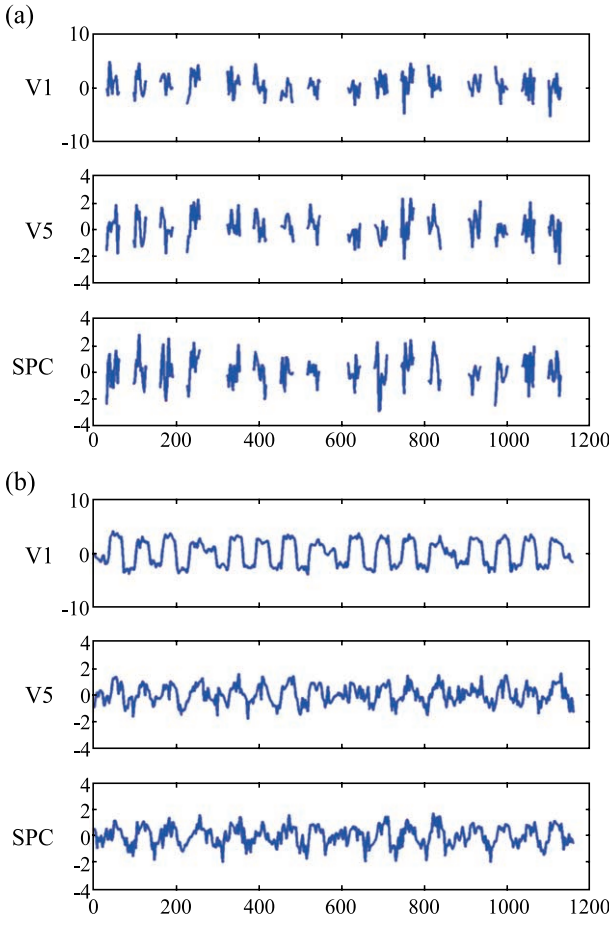

Fig. 9. Data from generative models. The top plot shows data generated from an SEM and the bottom plot data from a DCM. Both of these models were fitted to the attention data set.

order of tens of minutes whereas estimation in comparable SEM networks takes of the order of minutes.

A second current limitation of DCM is that neurodynamics in each region are characterised by a single state variable ('neuronal activity'). This prohibits inferences that can be meaningfully linked to specific neurotransmitter systems as these would require multiple state variables in each region that might, for example, describe activity in excitatory and inhibitory subpopulations. The parameters of such models would best be identified via DCMs that use high temporal resolution data such as EEG. The development of such models may therefore depend on integration of information from fMRI (to find out where activity occurs) and from EEG (to find out when it occurs). This is an exciting area for future research that would significantly strengthen the bridge between modalities in imaging neuroscience and our understanding of the neurobiology underlying cognitive processing.

Acknowledgments

This study was funded by the Wellcome Trust. A. Mechelli is supported by grant MH64445 from the National Institutes of Health (USA).

References

Bollen, K.A., 1989. Structural Equations with Latent Variables. John Wiley, New York.

Buchel, C., Friston, K.J., 1997. Modulation of connectivity in visual pathways by attention: cortical interactions evaluated with structural equation modelling and fMRI. Cerebral Cortex 7, 768–778.

Bullmore, E., Horwitz, B., Honey, G., Brammer, M., Williams, S., Sharma, T., 2000. How good is good enough in path analysis of fMRI data? NeuroImage 11, 289–301.

Buxton, R.B., Wong, E.C., 1998. Dynamics of blood flow and oxygenation changes during brain activation: The Balloon Model. Magnetic Resonance in Medicine 39, 855–864.

Frackowiak, R.S.J., Friston, K.J., Frith, C.D., Dolan, R.J., Mazziotta, J.C. (Eds.), 1997. Human Brain Function Academic Press, USA.

Frackowiak, R.S.J., Friston, K.J., Frith, C.D., Dolan, R., Price, C., Zeki, S., Ashburner, J., Penny, W. (Eds.), 2003. Human Brain Function, (2nd ed.) Academic Press.

Friston, K.J., 2002. Bayesian estimation of dynamical systems: an application to fMRI. NeuroImage 16, 513–530.

Friston, K.J., Buchel, C., 2000. Attentional modulation of effective connectivity from V2 to V5/MT in humans. Proceedings of the National Academy of Sciences of the United States of America 97 (13), 7591–7596.

Friston, K.J., Holmes, A.P., Poline, J.-B., Grasby, P.J., Williams, S.C.R., Frackowiak, R.S.J., Turner, R., 1995. Analysis of fMRI time series revisited. NeuroImage 2, 45–53.

Friston, K.J., Beuchel, C., Fink, G.R., Morris, J., Rolls, E., Dolan, R.J., 1997. Psychophysiological and modulatory interactions in neuroimaging. NeuroImage 6, 218–229.

Friston, K.J., Harrison, L., Penny, W., 2003. Dynamic causal modelling. NeuroImage 19 (4), 1273–1302.

Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 1995. Bayesian Data Analysis. Chapman and Hall.

Gitelman, D.R., Penny, W.D., Ashburner, J., Friston, K.J., 2003. Modeling regional and psychophsyiologic interactions in fMRI: the importance of hemodynamic deconvolution. NeuroImage 19, 200–207.

Goncalves, M.S., Hull, D.A., 2003. Connectivity analysis with structural equation modeling: an example of the effects of voxel selection. NeuroImage 20, 1455–1467.

Kass, R.E., Raftery, A.E., 1993. Bayes factors and model uncertainty. Technical Report 254, University of Washington. http://www.stat.washington.edu/tech.reports/tr254.ps.

Kass, R.E., Raftery, A.E., 1995. Bayes factors. Journal of the American Statistical Association 90, 773–795.

MacIntosh, A.R., Gonzalez-Lima, F., 1991. Structural modeling of functional neural pathways mapped with 2-deoxyglucose: effects of acoustic startle habituation of the auditory system. Brain Research 547, 295–302.

McIntosh, A.R., Gonzalez-Lima, F., 1994. Structural equation modeling and its application to network analysis in functional brain imaging. Human Brain Mapping 2, 2–22.

MacIntosh, A.R., Bookstein, F.L., Haxby, J.V., Grady, C.L., 1996. Spatial pattern analysis of functional brain images using partial least squares. NeuroImage 3, 143–157.

Pearl, J., 1998. Graphs, causality, and structural equation models. Sociological Methods and Research 27, 226–284.

Penny, W.D., Stephan, K.E., Mechelli, A., Friston, K.J., 2004. Comparing dynamics causal models. Neuroimage (Accepted for publication).

Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.V.P., 1992. Numerical recipes in C. Cambridge.

Ramnani, N., Behrens, T.E.J., Penny, W., Matthews, P.M., 2004. New approaches for exploring anatomic and functional connectivity in the human brain. Biological Psychiatry (Accepted for publication).

Rowe, J., Friston, K., Frackowiak, R., Passingham, R., 2002. Attention to action: specific modulation of corticocortical interactions in humans. NeuroImage 17, 988–998.

Scheines, R., Hoijtink, H., Boomsma, A., 1999. Bayesian estimation and testing of structural equation models. Psychometrika 64, 37–52.

Woolrich, M.W., Ripley, B.D., Brady, M., Smith, S.M., 2001. Temporal autocorrelation in univariate linear modelling of fMRI data. NeuroImage 14 (6), 1370–1386.

Zeki, S., Watson, J.D., Lueck, C.J., Friston, K.J., Kennard, C., Frackowiak, R.S., 1991. A direct demonstration of functional specialization in human visual cortex. Journal of Neuroscience 11, 641–649.