ELSEVIER

NeuroImage

www.elsevier.com/locate/ynimg
NeuroImage 30 (2006) 1273 – 1284

Dynamic causal modelling of evoked responses in EEG/MEG with lead field parameterization

Stefan J. Kiebel,* Olivier David, and Karl J. Friston

Wellcome Department of Imaging Neuroscience, Functional Imaging Laboratory, 12 Queen Square, London WC1N 3BG, UK

Received 31 May 2005; revised 18 October 2005; accepted 20 December 2005
Available online 21 February 2006

Dynamical causal modeling (DCM) of evoked responses is a new approach to making inferences about connectivity changes in hierarchical networks measured with electro- and magnetoencephalography (EEG and MEG). In a previous paper, we illustrated this concept using a lead field that was specified with infinite prior precision. With this prior, the spatial expression of each source area, in the sensors, is fixed. In this paper, we show that using lead field parameters with finite precision enables the data to inform the network's spatial configuration and its expression at the sensors. This means that lead field and coupling parameters can be estimated simultaneously. Alternatively, one can also view DCM for evoked responses as a source reconstruction approach with temporal, physiologically informed constraints. We will illustrate this idea using, for each area, a 4-shell equivalent current dipole (ECD) model with three location and three orientation parameters. Using synthetic and real data, we show that this approach furnishes accurate and robust conditional estimates of coupling among sources and their orientations.

Keywords: Electroencephalography; Magnetoencephalography; Generative model; Hierarchical networks; Nonlinear dynamics

Introduction

In David et al. (2006), we described dynamic causal modeling (DCM) for event-related fields (ERFs) and potentials (ERPs). This new approach is grounded on a neuronally plausible, generative model that can be used to estimate and make inferences about category- or context-specific coupling among cortical regions. Context-specific coupling changes as a function of condition (i.e., experimental context such as "new" vs. "old" in memory paradigms) or stimulus-bound attributes (i.e., "house" vs. "face"). These changes can reconfigure neuronal interactions and produce different evoked responses for each category or context. The coupling parameters embody bottom-up, top-down, and lateral connections among remote cortical regions. Parameters are estimated with a Bayesian procedure using empirical data (ERPs/ERFs). With Bayesian model selection, one can use model evidences to compare competing models and identify the model that best explains the data.

In David et al. (2006), we constructed the spatial forward model using distributed dipole modeling on the grey matter surface. This procedure has the advantage of using the precise anatomical structure of the head. The subject's anatomy was derived from the high-resolution structural magnetic resonance imaging (sMRI). Critically, each area's lead field was predetermined so that each area had a fixed spatial expression in the sensors. Although this approach provides spatially precise expressions in the sensors, the true spatial configuration of an area may be different from our model and lead to biased conditional estimates of other [e.g., coupling] parameters. For example, the spatial model can be wrong because its parameters like location, orientation, or extent are specified inaccurately.

Alternatively, each lead field or its underlying spatial parameters can be regarded as a parameter of the model. In a Bayesian context, the above procedure is equivalent to using zero prior variance (i.e., infinite precision which expresses our belief that the specified lead field mediated the sensor data). If this belief is not supported by the data, the optimization algorithm will, at worst, fail to provide a good solution and compensate for the misspecified spatial model by biasing conditional estimates of other parameters like coupling. A way to avoid this is to decrease our strong belief in a specific lead field and use finite precision priors on the lead field parameters. There are several ways to parameterize the lead field. Although we could employ a surface-based forward model, we use equivalent current dipoles (ECDs). This has distinct advantages over other models. First, ECDs' spatial expression is analytic, i.e., the forward model computation is fast (Mosher et al., 1999). Secondly, the model is based on electrode positions only and does not need information from a structural MRI. Thirdly, many authors reported ECD location and orientation for specific ERP/ERF experiments in the peer-reviewed literature (e.g., Valeriani et al., 2001): Within our

* Corresponding author. Fax: +44 207 813 1420
E-mail address: skiebel@fil.ion.ucl.ac.uk (S.J. Kiebel).
Available online on ScienceDirect (www.sciencedirect.com).

1053-8119/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2005.12.055

approach, these locations and orientations could be employed as prior expectations on ECD parameters. Finally, ECDs are a natural way to specify nodes in the probabilistic graphs that DCMs represent.

One can also view DCM for evoked responses as a source reconstruction approach with temporal, physiologically informed constraints imposed by our assumption that a hierarchical network of discrete areas generated the data. The reconstructed source activities over time fall out naturally as the system's states. Typically, most current source reconstruction approaches for EEG/MEG data are based exclusively on constraints given by the spatial forward model (Darvas et al., 2004). However, recently models have been proposed which use (spatio-) temporal constraints to invert the model (Darvas et al., 2001; Galka et al., 2004). These spatiotemporal approaches are closer to DCM but use generic constraints derived from temporal smoothness considerations and autoregressive modeling.

This paper is structured as follows. In the Theory section, we will describe briefly the temporal generative model for ERP/ERFs (for a detailed description, see David et al., 2005). This is followed by a description of the spatial forward model, its parameterization and typical prior distributions we adopt for ERP data. In the Results section, we illustrate the operational details of the procedures on two ERP datasets. In the first ERP experiment, we repeat the analysis of an auditory oddball dataset (David et al., 2006) to show that the mismatch negativity can be explained by changes in connectivity to and from the primary auditory cortex. This analysis shows that biologically meaningful results can be obtained in terms of the parameters governing the neuronal architectures generating ERPs. In the second experiment, we establish face validity in terms of the spatial parameters; we analyze sensory-evoked potentials (SEPs) elicited by unilateral median nerve stimulation and measured with EEG. With this model, we can explain the observed SEP to 200 ms. We find strong connectivity among areas during the course of the SEP. The estimated orientations of these sources conform almost exactly to classical estimates in the literature. Furthermore, we observe short transmission delays among sources within the contralateral hemisphere (~6 ms) but long delays (~50 ms) between homologous sources in both hemispheres. Finally, using synthetic data, we show that finite precision priors on lead field parameters result in models with greater evidence and more accurate and robust conditional estimates, in relation to models with infinitely precise priors.

Theory

Intuitively, the DCM scheme regards an experiment as a designed perturbation of neuronal dynamics that are promulgated and distributed throughout a system of coupled anatomical sources to produce region-specific responses. This system is modeled using a dynamic input–state–output system with multiple inputs and outputs. Responses are evoked by deterministic inputs that correspond to experimental manipulations (i.e., presentation of stimuli). Experimental factors (i.e., stimulus attributes or context) can also change the parameters or causal architecture of the system producing these responses. The state variables cover both the neuronal activities and other neurophysiological or biophysical variables needed to form the outputs. Outputs are those components of neuronal responses that can be detected by MEG/EEG sensors. In our model, these components are depolarizations of a 'neural mass' of pyramidal cells.

DCM starts with a reasonably realistic neuronal model of interacting cortical regions. This model is then supplemented with a spatial forward model of how neuronal activity is transformed into measured responses, here, MEG/EEG scalp-averaged responses. This enables the parameters of the neuronal model (i.e., effective connectivity) to be estimated from observed data. For MEG/EEG data, this spatial model is a forward model of electromagnetic measurements that accounts for volume conduction effects (Mosher et al., 1999).

Hierarchical MEG/EEG neural mass model

We have developed a hierarchical cortical model to study the influence of forward, backward, and lateral connections on ERFs/ERPs (David et al., 2004). This model is used here as a DCM and embodies directed extrinsic connections among a number of sources, each based on the Jansen and Rit (1995) model, using the connectivity rules described in Felleman and Van Essen (1991). These rules, which rest on a tri-partitioning of the cortical sheet into supra-, infra-granular layers and granular layer 4, have been derived from experimental studies of monkey visual cortex. Under these simplifying assumptions, directed connections can be classified as (i) bottom-up or forward connections that originate in agranular layers and terminate in layer 4; (ii) top-down or backward connections that connect agranular layers; (iii) lateral connections that originate in agranular layers and target all layers. These long-range or extrinsic cortico-cortical connections are excitatory and comprise the axonal processes of pyramidal cells. For simplicity, we do not consider thalamic connections but model thalamic output as a function operating on the input (see below).

The Jansen and Rit (1995) model emulates the MEG/EEG activity of a cortical source using three neuronal subpopulations. A population of excitatory pyramidal (output) cells receives inputs from inhibitory and excitatory populations of interneurons, via intrinsic connections (intrinsic connections are confined to the cortical sheet). Within this model, excitatory interneurons can be regarded as spiny stellate cells found predominantly in layer 4 and in receipt of forward connections. Excitatory pyramidal cells and inhibitory interneurons occupy agranular layers and receive backward and lateral inputs. Using these connection rules, it is straightforward to construct any hierarchical cortico-cortical network model of cortical sources.

The ensuing DCM is specified in terms of its state equations and an observer or output equation

$$\dot{x} = f(x, u, \theta)$$

$$h = g(x, \theta) \tag{1}$$

where x are the neuronal states of cortical areas, u are exogenous inputs, and h is the output of the system. θ are quantities that parameterize the state and observer equations (see also below under 'Prior assumptions'). The state equations are ordinary second-order differential equations and are derived from the behavior of the three neuronal subpopulations which operate as linear damped oscillators. The integration of the differential equations pertaining to each subpopulation can be expressed as a convolution (David and Friston, 2003). This convolution trans-

forms the average density of its presynaptic inputs into an average postsynaptic membrane potential. The convolution kernel is given by

$$p(t)_e = \begin{cases} \frac{H_e}{\tau_e} t\exp(-t/\tau_e) & t \geq 0 \\ 0 & t < 0 \end{cases} \quad (2)$$

where subscript "e" stands for "excitatory". Similarly subscript "i" is used for inhibitory synapses. H controls the maximum post-synaptic potential, and τ represents a lumped rate constant. An operator S transforms the potential of each subpopulation into firing rate, which is the input to other subpopulations. This operator is assumed to be an instantaneous sigmoid nonlinearity

$$S(x) = \frac{1}{1+\exp(-rx)} - \frac{1}{2} \quad (3)$$

where $r = 0.56$ determines its form. Interactions, among the subpopulations, depend on internal coupling constants $\gamma_{1,2,3,4}$, which control the strength of intrinsic connections and reflect the total number of synapses expressed by each subpopulation. The integration of this model to form predicted response rests on formulating these two operators (Eqs. (2) and (3)) in terms of a set of differential equations as described in David et al. (2004). A DCM, at the neuronal level, obtains by coupling areas with extrinsic forward, backward and lateral connections as described above.

These equations, for all areas, can be integrated using the matrix exponential of the systems Jacobian as described in the appendices of David et al. (2005). Critically, the integration scheme allows for conduction delays on the connections, which are free parameters of the model. The output of area i is the depolarization of pyramidal cells, over all time bins (David et al., 2006).

Event-related input and event-related response-specific effects

To model event-related responses, the network receives inputs via input connections. These connections are exactly the same as forward connections and deliver inputs u to the spiny stellate cells in layer 4. In the present context, inputs u model afferent activity relayed by subcortical structures and are modeled with two components: The first is a gamma density function (truncated to peri-stimulus time). This models an event-related burst of input that is delayed with respect to stimulus onset and dispersed by subcortical synapses and axonal conduction. Being a density function, this component integrates to unity over peri-stimulus time. The second component is a discrete cosine set modeling systematic fluctuations in input, as a function of peri-stimulus time. In our implementation, peri-stimulus time is treated as a state variable, allowing the input to be computed explicitly during integration. Critically, the event-related input is exactly the same for all ERPs.

The effects of experimental factors are mediated through event-related response (ERR)-specific changes in connection strengths. This models experimental effects in terms of differences in forward, backward, or lateral connections that confer a selectivity on each source, in terms of its response to others. The experimental or ERR-specific effects are modeled by coupling gains. By convention, we set the gain of the first ERP to unity, so that subsequent ERR-specific effects are relative to the first.

Spatial forward model

The dendritic signal of the pyramidal subpopulation of the ith source $x_0^{(i)}$ is detected remotely on the scalp surface in MEG/EEG. The relationship between scalp data h and pyramidal activity is linear and instantaneous

$$h = g(x,\theta) = L(\theta^L)Kx_0 \quad (4)$$

where L is a lead field matrix (i.e., spatial forward model), which accounts for passive conduction of the electromagnetic field (Mosher et al., 1999). The diagonal matrix $K = \mathrm{diag}(\theta^K)$ models the contribution of relative density of synapses proximate and distal to the cell body on current flow induced by pyramidal cell depolarization. The contribution matrix K has positive or negative weights to allow for the average dipole orientation to be parallel or anti-parallel with the assumed orientation.

The key contribution of this work is to make the lead field a function of some parameters $L(\theta^L)$. Here, we assume that the spatial expression of each area is caused by one equivalent current dipole (ECD). The head model for the dipoles is based on four concentric spheres, each with homogeneous and isotropic conductivity. The four spheres approximate the brain, skull, cerebrospinal fluid (CSF), and scalp. The parameters of the model are the radii and conductivities for each layer. Here, we use as radii 71, 72, 79, and 85 mm, with conductivities 0.33, 1.0, 0.0042, and 0.33 S/m respectively. The potential at the sensors requires an evaluation of an infinite series which can be approximated using fast algorithms (Mosher et al., 1999; Zhang, 1995). The lead field of each ECD is a function of three location and three orientation or moment parameters $\theta^L = (\theta^{\mathrm{pos}}, \theta^{\mathrm{mom}})$. For the ECD forward model, we used a Matlab (Mathworks) routine that is freely available as part of the FieldTrip package (http://www2.ru.nl/fcdonders/fieldtrip/, see also (Oostenveld, 2003)) under the GNU general public license.

The dipole parameters are naturally visualized in brain 3D space. In the Results section below, we display dipole locations and their orientations as arrows on a structural MRI template.

Dimension reduction

For computational reasons, it is expedient to reduce the dimensionality of the sensor data while retaining the maximum amount of information. This is assured by projecting the data onto a subspace defined by its principal eigenvectors E

$$\begin{aligned} y &\leftarrow Ey \\ L &\leftarrow EL \\ \varepsilon &\leftarrow E\varepsilon \end{aligned} \quad (5)$$

where ε is the observation error (see next subsection). The eigenvectors are computed using principal component analysis or singular value decomposition (SVD). Because this projection is orthonormal, the independence of the projected errors is preserved, and the form of the error covariance components assumed by the observation model remains unchanged. In this paper, we reduce the sensor data to three or four modes, which usually contain the interesting ERR components.

Observation equations

In summary, our DCM comprises a state equation that is based on neurobiological heuristics and an observer equation based on an

electromagnetic forward model. By integrating the state equation and passing the ensuing states through the observer equation, we generate a predicted measurement. This corresponds to a generalized convolution of the inputs to generate an output $h(\theta)$ (Eq. (4)). This generalized convolution furnishes an observation model for the vectorized data[1] y and the associated likelihood

$$y = \mathrm{vec}\left(h(\theta) + X\theta^X + \varepsilon\right)$$

$$p(y\,|\,\theta,\lambda) = N\left(\mathrm{vec}\left(h(\theta) + X\theta^X\right),\ \mathrm{diag}(\lambda)\otimes V\right). \tag{6}$$

Measurement noise ε is assumed to be zero mean Gaussian and independent over channels, i.e., $\mathrm{Cov}(\mathrm{vec}(\varepsilon)) = \mathrm{diag}(\lambda)\otimes V$, where λ is an unknown vector of channel-specific variances. V represents the error temporal autocorrelation matrix, which we assume is the identity matrix. This is tenable because we downsample the data to about 8 ms. Low-frequency noise or drift components are modeled by X, which is a block diagonal matrix with a low-order discrete cosine set for each ERP and channel. The order of this set can be determined by Bayesian model selection (see below).

This model is fitted to data by tuning the free parameters θ to minimize the discrepancy between predicted and observed MEG/EEG time series under model complexity constraints (more formally, the parameters minimize the Variational Free Energy—see below). These parameters specify the constants in the state and observation equations above. In addition to minimizing the prediction error, the parameters are constrained by a prior specification of the range they are likely to lie in (Friston et al., 2003). These constraints, which take the form of a prior density $p(\theta)$, are combined with the likelihood $p(y\,|\,\theta)$, to form a posterior density $p(\theta\,|\,y) \propto p(y\,|\,\theta)p(\theta)$ according to Bayes' rule. It is this posterior or conditional density we want to estimate. Gaussian assumptions about the errors in Eq. (6) enable us to compute the likelihood from the prediction error. The only outstanding quantities we require are the priors, which are described next.

Prior expectations

The connectivity architecture is constant over peri-stimulus time and defines the dynamical behavior of the DCM. We have to specify prior assumptions about these constant parameters to estimate their posterior distributions. Priors have a dramatic impact on the landscape of the objective function to be extremized: precise prior distributions ensure that the objective function has a global minimum that can be attained robustly. Under Gaussian assumptions, the prior distribution $p(\theta_i)$ of the ith parameter is defined by its mean and variance. The mean corresponds to the prior expectation. The variance reflects the amount of prior information about the parameter. A tight distribution (small variance) corresponds to precise prior knowledge.

The parameters of the state equation can be divided into six subsets: (i) extrinsic connection parameters, which specify the coupling strengths among areas, and (ii) intrinsic connection parameters, which reflect our knowledge about canonical micro-circuitry within an area, (iii) conduction delays, (iv) synaptic parameters controlling the kinetics within an area, and (v) input parameters, which control the subcortical delay and dispersion of event-related responses, (vi) spatial parameters which determine the expression of the observable network state in the sensors. Critically, all the constants, apart from the spatial parameters, are positive. To ensure positivity, we estimate the log of these constants under Gaussian priors, using the same prior distributions as in David et al. (2006). For the spatial parameters θ^{pos}, θ^{mom}, and θ^k, refer to Table 1.

The expectation of the location prior is usually given in millimeter in some standard brain space. In this paper, we use, if not otherwise specified, tight spherical priors of $v_x^{\mathrm{pos}} = v_y^{\mathrm{pos}} = v_z^{\mathrm{pos}} = 8$ and 0 for the covariances. Although not employed in this paper, one can define location priors that have nonspherical distributions. This can be useful, if one expresses the location prior in terms of a 3D ellipsoid[2], not necessarily aligned with the imaging axes. Similarly, for the prior on the moment, one can use priors that point in one principal direction, while the other two directions have only small variances. In the present paper, we exclusively employ uninformative moment priors, i.e., $v_x^{\mathrm{mom}} = v_y^{\mathrm{mom}} = v_z^{\mathrm{mom}} = 8$. In this paper, we use only spherical prior distributions. Note that the moment parameters are formulated as projections onto the three axes of stereotactic space. This parameterizes not only the orientation but also the magnitude of the dipole and therefore induces some redundancy with respect to the contribution parameters θ_i^K for each source. We have chosen to leave this redundancy in the parameterization because it allows for fixed dipole orientations without necessarily fixing the magnitude. Because of this redundancy, we can choose some quite arbitrary prior variance for θ_i^K, which will further add to the prior uncertainty imposed by the already uninformative prior of θ_i^{mom}.

Estimation, inference, and model comparison

For a given DCM, say model m, parameter estimation corresponds to approximating the moments of the posterior distribution given by Bayes rule

$$p(\theta\,|\,y,m) = \frac{p(y\,|\,\theta,m)p(\theta,m)}{p(y\,|\,m)}. \tag{7}$$

The estimation procedure employed in DCM is described in Friston (2002). The posterior moments (conditional mean η and covariance Σ) are updated iteratively using variational Bayes under a fixed-form Laplace (i.e., Gaussian) approximation to the conditional density $q(\theta) = N(\eta,\Sigma)$. This can be regarded as an Expectation-Maximization (EM) algorithm that employs a local linear approximation of Eq. (6) about the current conditional expectation. The **E**-step conforms to a Fisher-scoring scheme (Press et al., 1992) that performs a descent on the variational free energy $F(q,\lambda,m)$ with respect to the conditional moments. In the **M**-Step, the error variances λ are updated in exactly the same way. The estimation scheme can be summarized as follows:

Repeat until convergence

$$\mathbf{E}-\mathrm{Step}\ q \leftarrow \min_q F(q,\lambda,m)$$

$$\mathbf{M}-\mathrm{Step}\ \lambda \leftarrow \min_\lambda F(q,\lambda,m) = \max_\lambda L(\lambda,m)$$

$$\begin{aligned} F(q,\lambda,m) &= \langle \ln q(\theta) - \ln p(y\,|\,\theta,\lambda,m) - \ln p(\theta\,|\,m)\rangle_q \\ &= D(q\,||\,p(\theta|y,\lambda,m)) - L(\lambda,m) \\ L(\lambda,m) &= \ln p(y|\lambda,m). \end{aligned} \tag{8}$$

[1] Concatenated column vectors of data from each channel.

[2] Note that an ellipsoid can also have one or two axes lengths close to zero to define a one or two dimension manifold that contains the source.

Note that the free energy is simply a function of the log-likelihood and the log-prior for a particular DCM and $q(\theta)$. The expression $\langle\cdot\rangle_q$ denotes the expectation under the density q. $q(\theta)$ is the approximation to the posterior density $p(\theta|y, \lambda, m)$ we require. The **E**-step updates the moments of $q(\theta)$ (these are the variational parameters η and Σ) by minimizing the variational free energy. The free energy is the Kullback–Leibler divergence (denoted by $D(\cdot||\cdot)$), between the real and approximate conditional density minus the log-likelihood. This means that the conditional moments or variational parameters maximize the log-likelihood $L(\lambda, m)$ while minimizing the discrepancy between the true and approximate conditional density. Because the divergence does not depend on the covariance parameters, minimizing the free energy in the **M**-step is equivalent to finding the maximum likelihood estimates of the covariance parameters. This scheme is identical to that employed by DCM for fMRI, the details of which can be found in Friston et al. (2002), (2003).

Bayesian inference proceeds using the conditional or posterior density estimated by the EM algorithm. Usually this involves specifying a parameter or compound of parameters as a contrast $c^T\eta$. Inferences about this contrast are made using its conditional covariance $c^T\Sigma c$. For example, one can compute the probability that any contrast is greater than zero or some meaningful threshold, given the data. This inference is conditioned on the particular model specified. In other words, given the data and model, inference is based on the probability that a particular contrast is bigger than a specified threshold. In some situations, one may want to compare different models. This entails Bayesian model comparison.

Different models are compared using their evidence (Penny et al., 2004). The model evidence is

$$p(y|m) = \int p(y|\theta, m)p(\theta|m)d\theta. \tag{9}$$

Note that the model evidence is simply the normalization term in Eq. (7). The evidence can be decomposed into two components: an accuracy term, which quantifies the data fit, and a complexity term, which penalizes models with a large number of parameters. Therefore, the evidence embodies the two conflicting requirements of a good model, that it explains the data and is as simple as possible. In the following, we approximate the model evidence for model m, under the Laplace approximation, by

$$\ln p(y|m) \approx \ln p(y|\lambda, m). \tag{10}$$

This is simply the maximum value of the objective function attained by EM (see the **M**-Step in Eq. (8)). The most likely model is the one with the largest log-evidence. This enables Bayesian model selection. Model comparison rests on the likelihood ratio of the evidence for two models. This ratio is the Bayes factor B_{ij}. For models i and j

$$\ln B_{ij} = \ln p(y|m = i) - \ln p(y|m = j). \tag{11}$$

Conventionally, strong evidence in favor of one model requires the difference in log-evidence to be three or more (cf., Table 1 in Penny et al., 2004).

Summary

A DCM is specified through its priors. These are used to specify (i) how regions are interconnected, (ii) which regions receive subcortical inputs, (iii) which cortico-cortical connections change with the levels of experimental factors, and (iv) how the observable states express themselves spatially in the sensors. Usually, the most interesting questions pertain to changes in cortico-cortical coupling that explain differences in ERPs. Posterior distributions are estimated using an Expectation-Maximization algorithm that operates on the conditional distributions. After model estimation, we use the model evidence to compare alternative models. Typically, alternative models would be chosen to have different connectivity, different combinations of sources or different dipole location priors. Inference about a specific model is made using the posterior distribution of contrasts. We will illustrate the operational details in the following section.

Table 1
Prior densities for lead field parameters of the ith area

$$\theta_i^{\text{pos}} \sim N\left(\begin{pmatrix} x^{\text{pos}} \\ y^{\text{pos}} \\ z^{\text{pos}} \end{pmatrix}, \begin{pmatrix} v_x^{\text{pos}} & v_{xy}^{\text{pos}} & v_{xz}^{\text{pos}} \\ v_{xy}^{\text{pos}} & v_y^{\text{pos}} & v_{yz}^{\text{pos}} \\ v_{xz}^{\text{pos}} & v_{yz}^{\text{pos}} & v_z^{\text{pos}} \end{pmatrix}\right)$$

$$\theta_i^{\text{mom}} \sim N\left(\begin{pmatrix} x^{\text{mom}} \\ y^{\text{mom}} \\ z^{\text{mom}} \end{pmatrix}, \begin{pmatrix} v_x^{\text{mom}} & v_{xy}^{\text{mom}} & v_{xz}^{\text{mom}} \\ v_{xy}^{\text{mom}} & v_y^{\text{mom}} & v_{yz}^{\text{mom}} \\ v_{xz}^{\text{mom}} & v_{yz}^{\text{mom}} & v_z^{\text{mom}} \end{pmatrix}\right)$$

$$\theta_i^K \sim N(1,1)$$

Results

In this section, we illustrate the use of DCM using two real ERP datasets. Furthermore, we use two synthetic ERP datasets to show that DCM with a parameterized lead field can furnish more accurate and robust coupling estimates and models with greater evidence. This rest on using synthetic data where one knows the true model and true parameters. We use the first real dataset to address the face validity of the neuronal (i.e., coupling) parameter estimates and the second to look at the spatial parameters.

The first ERP dataset was acquired using an oddball paradigm (David et al., 2006). The data show a mismatch negativity (MMN) and P300 component in response to rare stimuli, relative to frequent (Debener et al., 2002; Linden et al., 1999). In this example, we attribute changes in coupling to plasticity underlying the learning of frequent or standard stimuli.

The second data are somatosensory-evoked potentials (SEPs) following unilateral median nerve stimulation. For these data, we show that one can model the first 150 ms in peri-stimulus time with a 3-area network including contralateral primary somatosensory cortex (SI) and bilateral secondary somatosensory cortices (SII). We use this model to cross-validate the estimated ECD orientations with the existing literature on SEPs.

An advantage of generative models is that it is easy to generate synthetic data. This can be achieved by integrating the system using parameters estimated from real data (the SEP data). Here, we use the posterior means of the parameters and add noise generated using the estimated error covariance. This procedure provides realistic-looking synthetic data. Note that because we use subspace

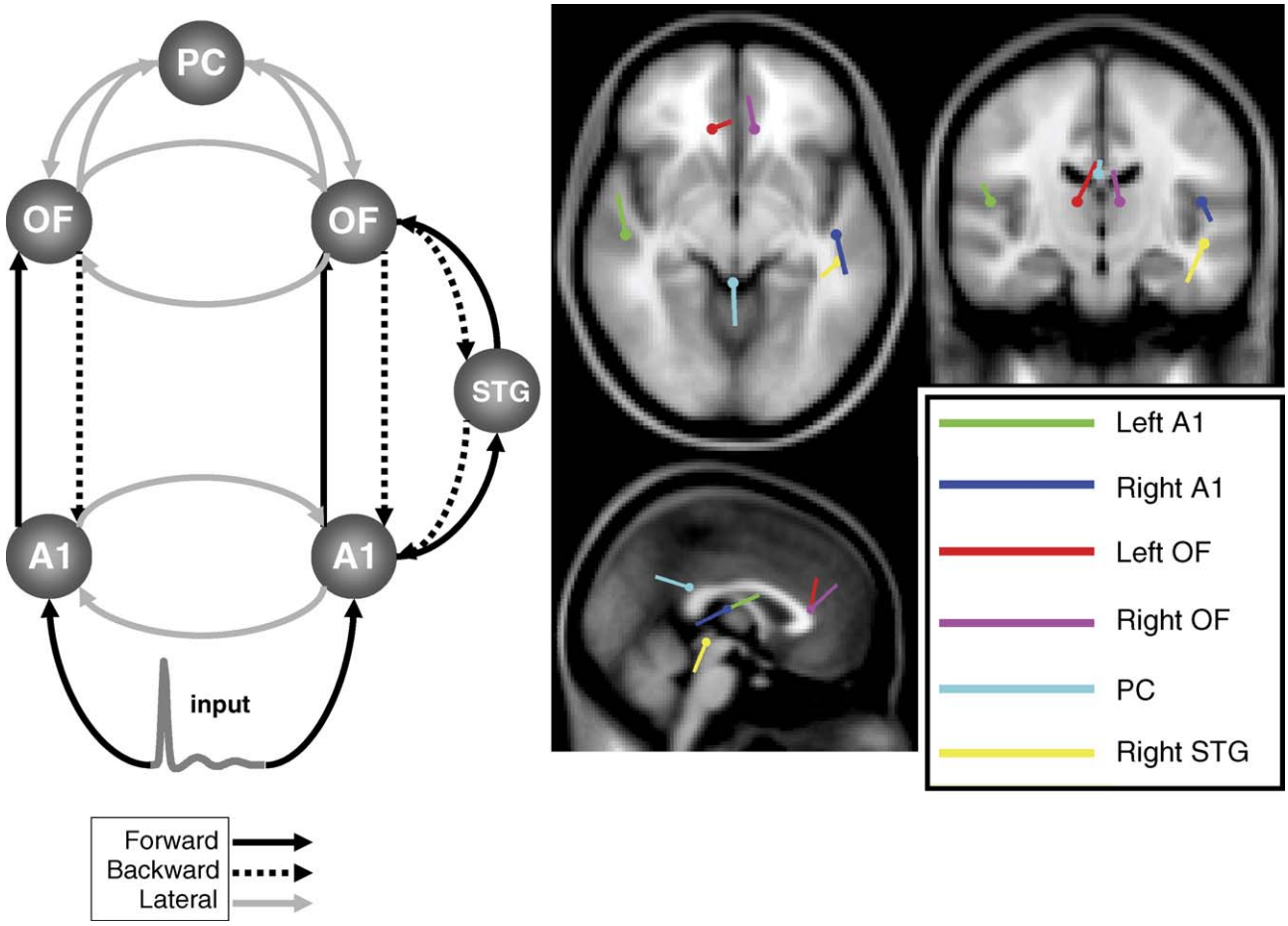

Fig. 1. DCM specification for the auditory oddball paradigm. Left: graph depicting the sources and connections of the DCM: A1: primary auditory cortex, OF: orbitofrontal cortex, PC: posterior cingulate cortex, STG: superior temporal gyrus. A bilateral extrinsic input acts on primary auditory cortex which project to orbitofrontal regions. In the right hemisphere, an indirect pathway was specified, via a relay in the superior temporal gyrus. At the highest level in the hierarchy, orbitofrontal and left posterior cingulate cortices were assumed to be laterally and reciprocally connected. Right: dipole locations and orientations (conditional means).

projection (Dimension reduction section), the synthetic data also exists in this low-dimensional space[3].

In the first set of simulations, we illustrate that a DCM with (uninformative) zero-mean priors on dipole orientations can identify the true dipole moments with high precision. In the second simulation, we repeat the simulation but specify priors whose means deviate from the true parameters. We find that even with these ill-informed priors, DCM can recover the true spatial configuration of modeled sources. In contrast, a model that assumes the (inaccurate) lead field, leads to sub-optimal results. These findings indicate that a lead field does not have to be fully specified before fitting the full spatiotemporal data but can be estimated simultaneously with the neuronal and connectivity parameters.

ERP data

Oddball data—mismatch negativity

This single-subject dataset was acquired using 128 EEG electrodes and 2048-Hz sampling. Auditory stimuli, 1000- or 2000-Hz tones with 5-ms rise and fall times and 80-ms duration, were presented binaurally for 15 min, every 2 s in a pseudo-random sequence. 2000-Hz tones (oddballs) occurred 20% of the time (120 trials) and 1000-Hz tones (standards) 80% of the time (480 trials). The subject was instructed to keep a mental record of the number of 2000-Hz tones. Before averaging single trials between −100 to 500 ms in peri-stimulus time, data were referenced to mean activity, downsampled to 125 Hz, and band-pass filtered between 0.5 and 25 Hz. Trials showing ocular artefacts (~30%), and 11 bad channels were removed from further analysis.

The mismatch negativity was observed around 140 ms at frontal electrodes. Other late components, also characteristic of rare events, were seen in most frontal electrodes, centered on 250 ms to 350 ms post-stimulus. As reported classically, early components (i.e., the N100) were almost identical for rare and frequent stimuli.

We modeled the data between 8 and 496 ms in peri-stimulus time. We followed (David et al., 2006) and constructed the following DCM (see Fig. 1, left): an extrinsic (thalamic) input entered bilateral primary auditory cortex (**A1**) which was connected to ipsilateral orbitofrontal cortex (**OF**). In the right hemisphere, an indirect forward pathway was specified from **A1** to **OF** through the superior temporal gyrus (**STG**). All these

[3] For visualization, the data can be back-projected to the ERP measurement space.

connections were reciprocal. At the highest level in the hierarchy, **OF** and left posterior cingulate cortex (**PC**) was laterally and reciprocally connected.

We found that this model is potentially over-parameterized when using all six parameters per dipole (see Discussion). To add more constraints and reduce the number of parameters, we assigned all dipole locations to their anatomically designated area (Fig. 1) with a zero-variance prior. Similarly, we assumed the orientations of the primary auditory cortices to be known. These were derived with an auxiliary analysis using just two ECDs (bilateral auditory cortices) modeling the first 140 ms of the data. The assumption underlying this procedure is that the temporal component around 100 ms (N100) is nearly exclusively generated by primary auditory cortex. This approach is often used in classical dipole fitting (Valeriani et al., 2001). Alternatively, we could have specified priors based on ECD fits of the N100 component reported in the literature. The prior distribution of the orientations of the four remaining dipoles (12 parameters) was noninformative with zero mean and variance of 2 mm.

After projecting the data to its first three spatial eigenvectors, we created three DCMs that differed in terms of which connections could show putative learning-related changes. The three models allowed changes in forward **F**, backward **B** and forward and backward **FB**. The log-evidences for a Bayesian model comparison (Penny et al., 2004) were −852.67 (**F**), −898.96 (**B**), −846.10 (**FB**), i.e., there is strong evidence for the **FB** model. In Fig. 2 (left), we show the conditional estimates and posterior confidences for this model. They reveal a profound increase, for rare events, in several connections. We can be over 95% confident that these connections increased. In Fig. 1 (right), we show the conditional means of the dipole locations and orientations overlaid on an MRI template. The numbers alongside each connection are the estimated coupling-gain during oddball processing.

In summary, this analysis suggests that a sufficient explanation for mismatch responses is an increase in connectivity, in particular to and from primary auditory cortex. This could represent a failure to suppress prediction error induced by unexpected or oddball stimuli, relative to predictable or learned stimuli, which can be predicted more efficiently.

Median nerve stimulation—sensory-evoked potentials

We now present a DCM of data generated by a neuronal network that has been well characterized in terms of its spatial deployment: The sensory-evoked potential in response to a median nerve stimulus. We model only the first 150 ms of the response because we assume that, at early peri-stimulus times, the response can be modeled by fewer areas than in later peri-stimulus times, when higher areas may come into play. We model just the response

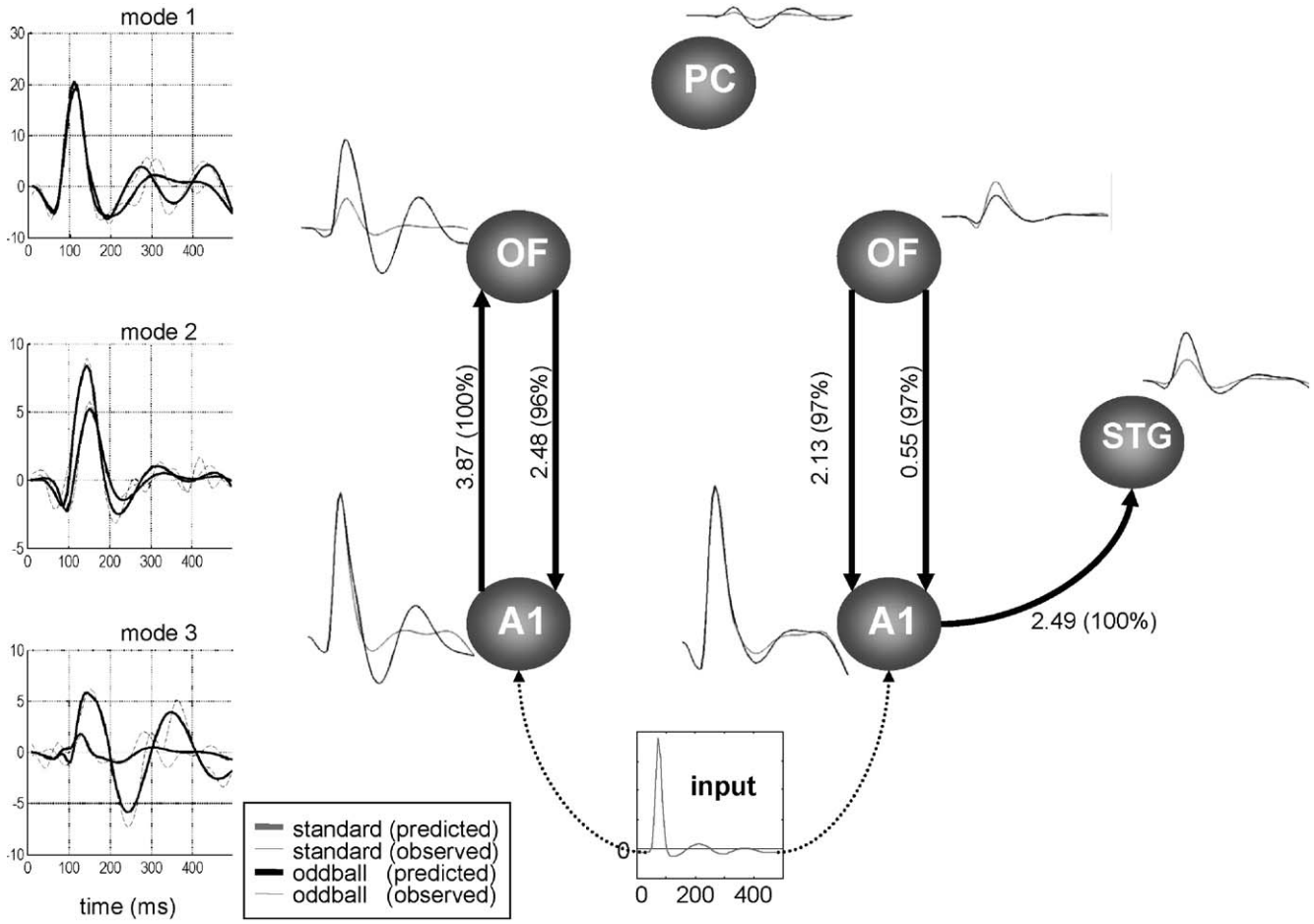

Fig. 2. Oddball data and corresponding source activity. Left: temporal expression of the first three modes at the sensor level. These capture 82% of the variance of the data. Right: reconstructed responses for each source and changes in coupling. We show the gain in coupling when rare relative to frequent events are presented. The percent conditional confidence that this ratio is greater than zero is shown in brackets. Only changes with 95% confidence or more are reported.

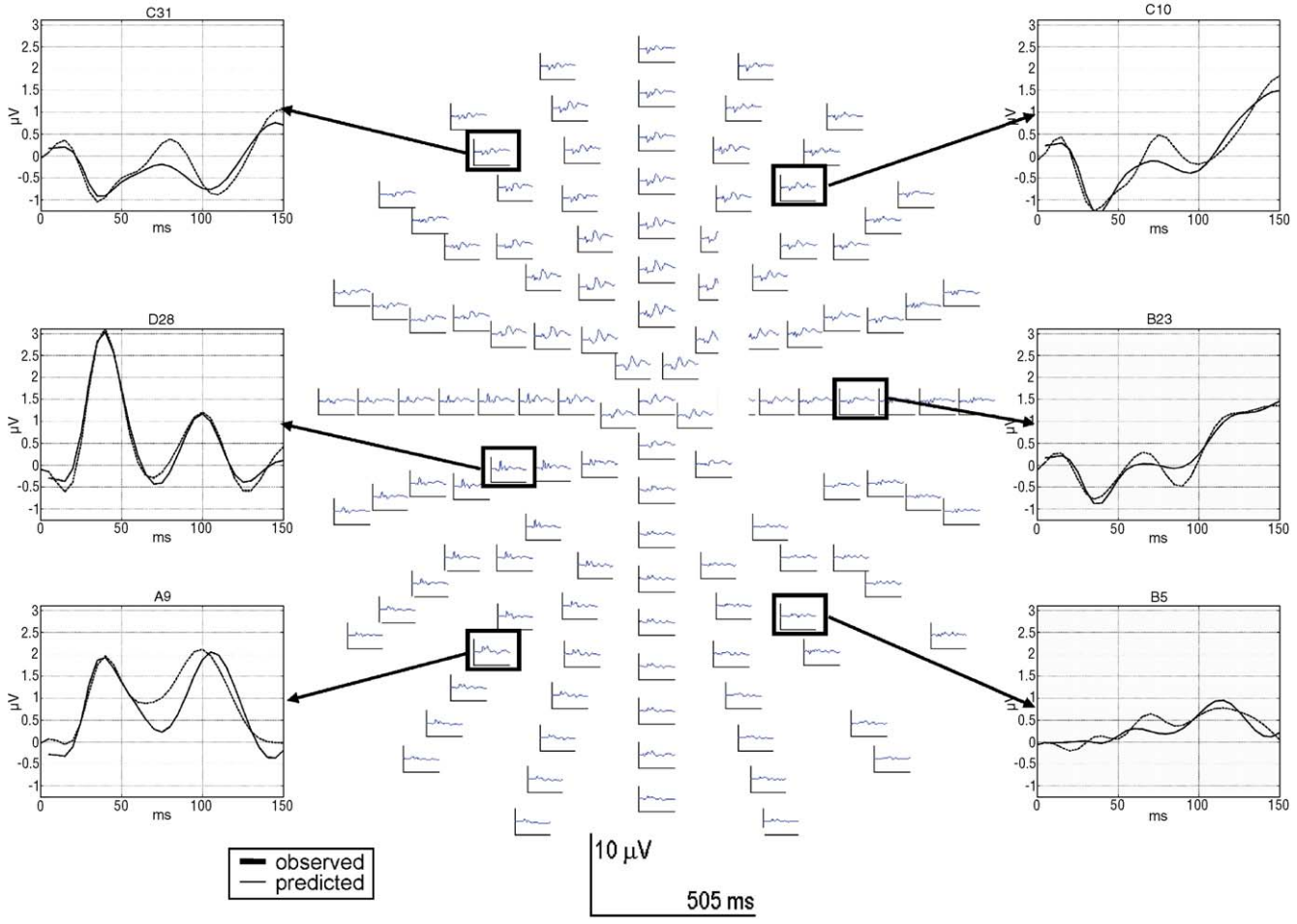

Fig. 3. Plot of SEP data and DCM fit in sensor space for representative channels. Middle: plots of the ERP channel data from −100 to 400 ms in peri-stimulus time. Five bad channels are excluded. Left and right: data and DCM fit for six channels, between 5 and 150 ms in peri-stimulus time.

to a median nerve stimulus. In the present context, this is interesting for cross-validation purposes because there is substantial literature on ECD modeling of SEPs using either EEG or MEG (Lin and Forss, 2002; Mauguiere et al., 1997; Valeriani et al., 2001).

The data were preprocessed using the current version of Statistical Parametric Mapping (SPM5b). Before averaging, the data were epoched between −100 to 150 ms, downsampled to 200 Hz, filtered between 0.5 and 35 Hz and thresholded at 100 μV (removing ca. 30% of all single trials). Five bad channels were removed from analysis. The ERP is shown in Fig. 3 (middle). With DCM, we modeled peri-stimulus time 5 to 150 ms of one condition, right median nerve stimulation.

We assumed three areas generated the SEP: contralateral SI (cSI), and bilateral SII (cSII and iSII). Exogenous input is received by cSI after passing through subcortical structures. We model forward and backward connectivity between contralateral SI and SII with lateral, reciprocal connections between the SII cortices. The structure of this neuronal network or graph is shown in Fig. 5 (right).

For analysis, we projected the data to a three-dimensional subspace spanned by the principal eigenvariates (see Fig. 4 for the corresponding eigen- or spatial modes). The components capture 95.04% of the data's variance. For all three dipoles, we used strong priors on location (prior variance of 8 mm^2) based on findings of Mauguiere et al. (1997). In MNI space, we specified the following

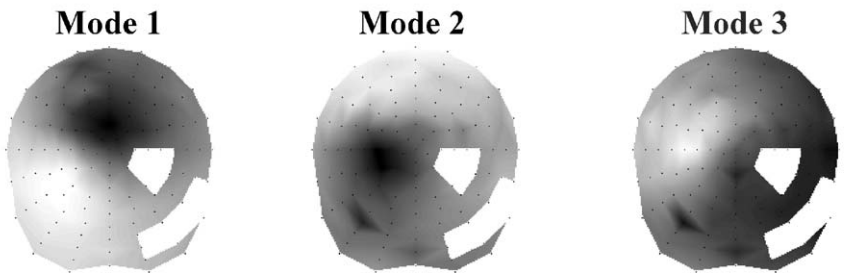

Fig. 4. First three spatial modes of SEP following median nerve stimulation. The 'holes' are caused by missing data due to bad channels.

coordinates: [0, 50, 45] for cSI, [20, 40, 10] for cSII, and [20, −40, 10] for iSII.

In Fig. 5, we show the near-perfect fit to the projected data (left) and the reconstructed source activities (right). The early response around 40 ms is modeled nearly exclusively by contralateral SI. The activity from cSI feeds forward to cSII, which contributes to the component around 80 ms, with smaller contributions from the other two areas. The component peaking at 115 ms is generated mainly by both SII areas. The connectivity estimates reveal a strong interaction between areas: the forward connection strength from cSI to cSII is 27.68 (with a conditional confidence of 100% in the coupling being greater than zero). The backward connection from cSII to cSI is 2.67 (100%). The lateral connection from cSII to iSII is 3.57 (99%). The reverse connection from iSII to cSII is apparently not engaged, with a conditional expectation of 0.95 (53%).

In terms of the spatial configuration of these sources, the conditional mean of two delays on extrinsic connections deviated considerably from their prior mean of 16 ms. The intra-hemispheric connection from cSI to cSII had a propagation delay (conditional mean) of 6.53 ms, and the lateral inter-hemispheric connection from cSII to iSII of 50.44 ms. The precisions of these estimates were high as indicated by the conditional variances of 0.008 (cSI→cSII) and 0.029 (cSII→iSII). These delays lie in a physiologically plausible range and point to the possibility of estimating propagation or conduction delays from ERP/ERF data (see Discussion).

The conditional means of the dipole orientations are shown in Fig. 5. It is pleasing to note that all three dipole orientations are very close to the ones reported in Mauguiere et al. (1997), see their Fig. 2. This result speaks to the face validity of the current DCM approach.

Note that, in our data, there is little evidence of an N20 component, which is usually seen in median nerve stimulation experiments. The reason is that the N20 is a high-frequency response component, and our specific preprocessing removes very high-frequency components. We used this processing for reasons of computational expediency because we were interested mainly in low-frequency components. However, one can see some evidence of the N20 (Fig. 3), but it is not fitted well. The likely reason is that the subspace used to reduce the dimensionality of the data did not span these early, high-frequency spaces. We will assess models of the early-latency components like the N20 in future work looking at the effects of preprocessing and dimension reduction.

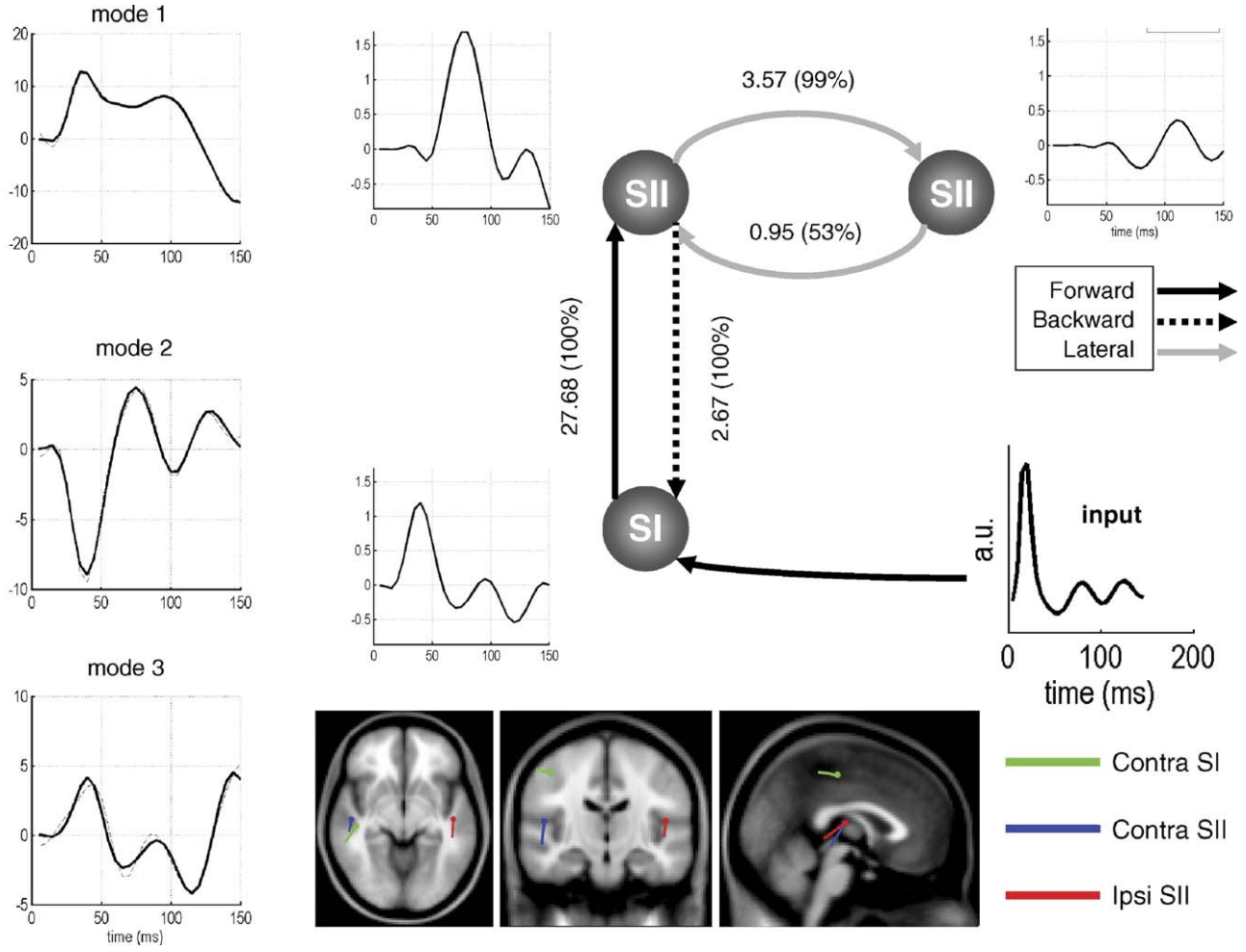

Fig. 5. Data and DCM for SEP data. Left: the time-courses of the first three modes of the data (see Fig. 4 for their spatial expression) and their fit using DCM. Right: DCM using three areas. The input (unilateral left median nerve stimulation) feeds into contralateral SI (cSI). cSI projects to cSII and receive backward connections from cSII. cSII projects laterally, reciprocally to ipsilateral SII. The numbers correspond to connection strength (Hz) and the (confidence that these parameters are greater than zero).

Simulations

True lead field

In this simulation, we added noise to the systems response generated using the conditional means of the above SEP DCM. We specified two different models. The first used the true lead field as known and fixed, i.e., all dipole parameter distributions have zero prior variance with expectations equal to the true parameters (see Fig. 5 for true locations and orientations). The second model has, for each dipole, informed, tight priors on the location and uninformed, broad priors on the moment parameters. For the locations, we used the true parameters as prior mean with prior variances of 8. For the orientation priors, we used means of zero with variances of 8. For both models, the prior for the contribution matrix K was set to its prior expectation of the identity matrix. The purpose of this simulation pair was to show that one can still obtain precise conditional estimate of neuronal parameters without knowing the spatial orientation of the sources.

We modeled the synthetic data in its space of the first three modes of the original data. As expected, the 'known lead field' model fits the data extremely well (mean percent error of <1%, plot not shown). All parameter estimates, including the coupling and intra-area parameters, are close to the true values. This means that all parameters can be recovered from the data when the spatial forward model is accurate and known. The log-evidence for this model was −273.96.

For the second model with unspecified orientations, the connectivity estimates are close to their true values. The estimated lead field is also very close to the true lead field. The log-evidence for this model was −292.53. The lower model evidence reflects the use of more parameters. The dipole orientation estimates are most interesting because we used a broad zero mean prior on these. The error (converted to angles around axes in degrees) ranged between 0.1 and 13.8° (see Table 2). These deviations are small and do not change the overall shape of the lead field. In summary, one can estimate not only coupling parameters from the data but also, simultaneously, important lead field parameters like the moment of equivalent current dipoles.

False lead field

In these simulations, we generate synthetic data as above. However, this time, the prior expectation of the orientation parameters deviates from the true parameters (a relative rotation of 60° of all dipoles around the x-axis). The first model uses zero variance priors as before, and a prior mean that now embodies false assumptions about the lead field. The second model has the same priors except that we acknowledge prior uncertainty about the orientations by using prior variances of 8. The purpose of this pair of simulations was to show that failing to properly encode prior uncertainty can lead to a low model evidence and low-confidence estimates of coupling parameters that are biased by conditional dependencies among the parameters.

Table 2
First simulation: deviations of dipole orientations of contralateral SI/SII and ipsilateral SII from true orientations (angles around axes in degrees)

	cSI	cSII	iSII
Angle x	4.4	3.1	10.8
Angle y	13.8	1.0	5.2
Angle z	4.1	0.1	3.3

Table 3
Second simulation: deviations of dipole orientations of contralateral SI/SII and ipsilateral SII from true orientations (angles around axes in degrees)

	cSI	cSII	iSII
Angle x	1.3	1.8	0.9
Angle y	3.3	9.7	6.6
Angle z	3.5	2.0	3.2

The first model, as expected, cannot retrieve the true coupling parameters from the data because the spatial model is not the true lead field. For example, the conditional mean of the forward connectivity from cSI→cSII is 8.99 (true: 24.34) and of backward cSII→cSI is 1.20 (true: 3.57). The log-evidence for this model is −573.31. This is much lower than for the final model, which accurately estimates the connectivity parameters and the lead field parameters (see Table 3), with a log-evidence of −249.83.

Discussion

In this paper, we have presented dynamical causal modeling (DCM) for event-related potentials and fields using equivalent current dipole models. We have shown that this Bayesian approach can be used to estimate parameters for a generative ERP model. Importantly, one can estimate, simultaneously, source activity, extrinsic connectivity, its modulation by context, and spatial lead field parameters from the data. An alternative view of DCM for ERP/ERF is to consider it a source reconstruction algorithm with biologically grounded temporal constraints. We have used simulated and real ERP data to show the usefulness and validity of the approach. Although we have not applied DCM to ERF data in the present paper, we note that the model can be adapted to ERFs by adjusting the electromagnetic component of the forward model.

DCM embodies several advantages over existing approaches. To start with, the generative model describes the full spatiotemporal data after projection to a low-dimensional subspace. Importantly, a single parameter estimation encompasses all the model parameters. This is in contrast to classical ECD fitting approaches, where dipoles are sequentially fitted to the data interactively using user-selected periods and/or channels of the data (Valeriani et al., 2001). Classical approaches must proceed in this way because there is usually too much spatial and temporal dependency among the sources to identify the parameters precisely. With our approach, we place temporal constraints on the model that are consistent with the way signals are generated biophysically. As we have shown above, these allow the simultaneous fitting of multiple dipoles to the data.

Furthermore, we can specify priors for all parameters. This informs the model about our belief concerning the parameters. With respect to lead field parameters, we used zero mean priors for the dipole moments. These are uninformative priors, i.e., we use the data to estimate orientations. For location parameters, we used priors with a high precision, i.e., we have a strong belief about where areas should be located. We return to this issue below.

One output of the DCM is the conditional density of the parameters. This can be used to express certainty about the parameter estimates. For example, if we find that the posterior variance of a dipole's moment is much smaller than its prior variance, we can be certain about its orientation. As illustrated above, we use Bayesian confidence intervals (also called credible

intervals) to express this certainty. The computation of these confidence intervals can be seen as a by-product of the Expectation-Maximization algorithm, which uses the Jacobian of the model parameters (i.e., how changes in the parameters are expressed in measurement space). Other methods that do not use this first-order approximation typically use Monte-Carlo and parametric bootstrap methods for the computation of dipole confidence intervals (Braun et al., 1997; Fuchs et al., 2004).

As a Bayesian technique, DCM computes the model evidence. As we have shown above, we can use model evidences of competing models to assess which is the most likely given some data. Model comparisons are important because they can be used to answer questions about how the data were generated. For example, in the oddball data, we found strong evidence for a network with extensive stimulus-dependent forward and backward connectivity, as opposed to a network with changes in feedforward connections only.

We used the equivalent current dipole (ECD) model because it is analytic, fast to compute and a quasi-standard when source-reconstructing ERP or ERF data. However, the ECD model is just one candidate for spatial forward models. Given some parameterization of the lead field, one can use any spatial model in the observation equation (Eq. (4)). A further example would be some linear distributed approach (Baillet and Garnero, 1997; Phillips et al., 2002), where a 'patch' of dipoles, confined to the cortical surface, would act as the spatial expression of one area. Possible parameters include the extent of the patch and location on the surface. With DCM, one could use different forward models for different areas in a single model (hybrid models). For example, one could employ the ECD model for early responses while using a distributed forward model for higher areas.

An advantage of the ECD model is that there is no need for structural information from magnetic resonance imaging (MRI). In practice, this means that a full DCM analysis can proceed automatically in less than an hour. In contrast, linear distributed methods often rely on surface tessellation of the individual's MRI, which can be time consuming, even if automated. The drawback of the ECD model is a potentially less accurate localization and a failure to model distributed, nondipole-like responses. However, note that dipole models for MEG are seen, with respect to localization error, as an adequate alternative to realistic boundary element methods (Darvas et al., 2004; Leahy et al., 1998). Furthermore, with DCM, exact localization is not necessarily the primary goal. Our experience suggests that precise Bayesian inversion only requires that each ECD is located roughly in some designated anatomical region. Also, as found by other authors, dipole parameters can identified precisely if neighboring areas have different orientations (e.g., Forss et al., 1996).

An important observation is that, after projection to a few spatial modes, the orientation of the dipoles matters more than their location for modeling the data. By this, we mean that the conditional precision for orientation is much higher than for the location parameters. Intuitively, the orientation determines most of the topology of a dipole's spatial expression in sensor space. Therefore, with the first few modes of the data, orientation can be estimated with high precision, whereas we cannot determine location from the reduced data. However, this is not an issue with typical DCM studies because the location of each dipole is implicit in the hypothesis (i.e., graphical model) the DCM represents. Locations can be derived from the literature (EEG/MEG, fMRI, PET). Note that a critical advantage of the present approach is that one can use model comparison to select the best model among different, plausible networks. In the present paper, we used tight priors on location. On basis of Bayesian model comparison using models with and without tight location priors, we recommend fixing dipole locations, i.e., to treat them as known with zero prior variance. For the orientations, we suggest uninformative priors (see previous section). In cases where one wishes to use further spatial constraints, one can use informative priors on orientation. Such priors could be derived from the literature, in particular for early- and medium-latency responses (<200 ms), for which the inter-subject variance of dipole parameters seems to be low (e.g., Forss et al., 1996), see their Fig. 4. In summary, our intuition based on Bayesian model selection and inversion is that the data contain relatively little information about the location of sources but are very sensitive to their orientation. This means that questions that are framed in terms of sources with known [roughly] location are more likely to be answered with conditional certainty.

The number of SVD components chosen for dimension reduction is user-dependent. One reason to perform a subspace projection is to make the DCM approach computationally efficient or rather, computationally feasible with high-density EEG or MEG measurements (128 to 300 channels). However, another reason to remove modes is that they cannot be modeled. For example, modes can contain artefacts or activity from nonmodeled higher areas. With the data described here, we found that our model is usually good at explaining up to the first 3–4 SVD components, especially for medium-latency peri-stimulus times. Typically, these components represent around 80–95% of the data's variance and are a sensible representation of the evoked response. One way of explaining more of the data is to add more areas. For instance, with the SEP data, we could have added areas located in the parietal cortex or frontal areas (Mauguiere et al., 1997) to make a more accommodating model, especially at later peri-stimulus times. However, such an approach does not necessarily lead to higher model evidence because of the increased model complexity. Another way of potentially improving the model is to use generative models of typical artefacts, e.g., muscular or ocular artefacts. However, such models are difficult to formulate because of their inherent physical complexity and variable expression over subjects and acquisition settings. In this situation, a good approach is to rely on a generic blind deconvolution algorithm and separate the data into components of interest and no interest. In this paper, we have used singular value decomposition. One can also consider independent component analysis (ICA) which provides a more constrained decomposition of the data (Makeig et al., 1997; Tang et al., 2005).

We have shown that it is possible to use DCM to estimate propagation delays between cortical areas. For these delays, we chose a prior mean of 16 ms. For the SEP data, the delay between cSI→cSII was estimated as 6 ms, and the delay from cSII→iSII as 50 ms. These estimates seem plausible given the trans-callosal connection between the SII cortices. Delay estimation is used routinely for diagnostic purposes. For example, the latency of the N20 component of the SEP measured on the scalp is used to make clinical inferences about the conduction delay from the periphery to the somatosensory cortex. At the sensor level, delay estimation of this sort is difficult at later peri-stimulus times, for which the responses of multiple areas overlap in time and space. DCM can provide estimates of inter-area delays because it is informed about this spatiotemporal dispersion.

Conclusion

DCM is useful for estimating connectivity in a hierarchical network based on evoked responses measured with EEG and MEG data. The parameterization of the lead field using equivalent current dipoles results in an accurate and robust estimation of both connectivity and dipole parameters. One can view DCM for evoked responses as a source reconstruction approach with temporal, physiologically informed constraints.

Acknowledgments

This work was supported by the Wellcome Trust. We thank Akaysha Tang and Felix Blankenburg for their valuable discussions and Robert Oostenveld for providing us with his Matlab implementation of ECD forward models.

References

Baillet, S., Garnero, L., 1997. A Bayesian approach to introducing anatomo-functional priors in the EEG/MEG inverse problem. IEEE Trans Biomed. Eng. 44, 374–385.

Braun, C., Kaiser, S., Kincses, W.E., Elbert, T., 1997. Confidence interval of single dipole locations based on EEG data. Brain Topogr. 10, 31–39.

Darvas, F., Schmitt, U., Louis, A.K., Fuchs, M., Knoll, G., Buchner, H., 2001. Spatio-temporal current density reconstruction (stCDR) from EEG/MEG-data. Brain Topogr. 13, 195–207.

Darvas, F., Pantazis, D., Kucukaltun-Yildirim, E., Leahy, R.M., 2004. Mapping human brain function with MEG and EEG: methods and validation. NeuroImage 23 (Suppl. 1), S289–S299.

David, O., Friston, K.J., 2003. A neural mass model for MEG/EEG: coupling and neuronal dynamics. NeuroImage 20, 1743–1755.

David, O., Cosmelli, D., Friston, K.J., 2004. Evaluation of different measures of functional connectivity using a neural mass model. NeuroImage 21, 659–673.

David, O., Harrison, L., Friston, K.J., 2005. Modelling event-related responses in the brain. NeuroImage 25, 756–770.

David, O., Kiebel, S.J., Harrison, L.M., Mattout, J., Kilner, J.M., and Friston, K.J., 2006. Dynamic Causal Modelling of Evoked Responses in EEG and MEG. Neuroimage 30; 1255–1272.

Debener, S., Kranczioch, C., Herrmann, C.S., Engel, A.K., 2002. Auditory novelty oddball allows reliable distinction of top-down and bottom-up processes of attention. Int. J. Psychophysiol. 46, 77–84.

Felleman, D.J., Van Essen, D.C., 1991. Distributed hierarchical processing in the primate cerebral cortex. Cereb. Cortex 1, 1–47.

Forss, N., Merlet, I., Vanni, S., Hamalainen, M., Mauguiere, F., Hari, R., 1996. Activation of human mesial cortex during somatosensory target detection task. Brain Res. 734, 229–235.

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

Friston, K.J., Penny, W., Phillips, C., Kiebel, S., Hinton, G., Ashburner, J., 2002. Classical and Bayesian inference in neuroimaging: theory. NeuroImage 16, 465–483.

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

Fuchs, M., Wagner, M., Kastner, J., 2004. Confidence limits of dipole source reconstruction results. Clin. Neurophysiol. 115, 1442–1451.

Galka, A., Yamashita, O., Ozaki, T., Biscay, R., Valdes-Sosa, P., 2004. A solution to the dynamical inverse problem of EEG generation using spatiotemporal Kalman filtering. NeuroImage 23, 435–453.

Jansen, B.H., Rit, V.G., 1995. Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns. Biol. Cybern. 73, 357–366.

Leahy, R.M., Mosher, J.C., Spencer, M.E., Huang, M.X., Lewine, J.D., 1998. A study of dipole localization accuracy for MEG and EEG using a human skull phantom. Electroencephalogr. Clin. Neurophysiol. 107, 159–173.

Lin, Y.Y., Forss, N., 2002. Functional characterization of human second somatosensory cortex by magnetoencephalography. Behav. Brain Res. 135, 141–145.

Linden, D.E., Prvulovic, D., Formisano, E., Vollinger, M., Zanella, F.E., Goebel, R., Dierks, T., 1999. The functional neuroanatomy of target detection: an fMRI study of visual and auditory oddball tasks. Cereb. Cortex 9, 815–823.

Makeig, S., Jung, T.P., Bell, A.J., Ghahremani, D., Sejnowski, T.J., 1997. Blind separation of auditory event-related brain responses into independent components. Proc. Natl. Acad. Sci. U. S. A. 94, 10979–10984.

Mauguiere, F., Merlet, I., Forss, N., Vanni, S., Jousmaki, V., Adeleine, P., Hari, R., 1997. Activation of a distributed somatosensory cortical network in the human brain. A dipole modelling study of magnetic fields evoked by median nerve stimulation: Part I. Location and activation timing of SEF sources. Electroencephalogr. Clin. Neurophysiol. 104, 281–289.

Mosher, J.C., Leahy, R.M., Lewis, P.S., 1999. EEG and MEG: forward solutions for inverse methods. IEEE Trans. Biomed. Eng. 46, 245–259.

Penny, W.D., Stephan, K.E., Mechelli, A., Friston, K.J., 2004. Comparing dynamic causal models. NeuroImage 22, 1157–1172.

Phillips, C., Rugg, M.D., Friston, K.J., 2002. Anatomically informed basis functions for EEG source localization: combining functional and anatomical constraints. NeuroImage 16, 678–695.

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

Oostenveldm R. (2003) Improving EEG Source Analysis using Prior Knowledge. Ref Type: Thesis/Dissertation.

Tang, A.C., Sutherland, M.T., McKinney, C.J., 2005. Validation of SOBI components from high-density EEG. NeuroImage 25, 539–553.

Valeriani, M., Le Pera, D., Tonali, P., 2001. Characterizing somatosensory evoked potential sources with dipole models: advantages and limitations. Muscle Nerve 24, 325–339.

Zhang, Z., 1995. A fast method to compute surface potentials generated by dipoles within multilayer anisotropic spheres. Phys. Med. Biol. 40, 335–349.