ELSEVIER

NeuroImage

www.elsevier.com/locate/ynimg
NeuroImage 30 (2006) 1255 – 1272

Dynamic causal modeling of evoked responses in EEG and MEG☆

Olivier David,[1] Stefan J. Kiebel,* Lee M. Harrison, Jérémie Mattout, James M. Kilner, and Karl J. Friston

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

Received 9 May 2005; revised 27 July 2005; accepted 11 October 2005
Available online 9 February 2006

Neuronally plausible, generative or forward models are essential for understanding how event-related fields (ERFs) and potentials (ERPs) are generated. In this paper, we present a new approach to modeling event-related responses measured with EEG or MEG. This approach uses a biologically informed model to make inferences about the underlying neuronal networks generating responses. The approach can be regarded as a neurobiologically constrained source reconstruction scheme, in which the parameters of the reconstruction have an explicit neuronal interpretation. Specifically, these parameters encode, among other things, the coupling among sources and how that coupling depends upon stimulus attributes or experimental context. The basic idea is to supplement conventional electromagnetic forward models, of how sources are expressed in measurement space, with a model of how source activity is generated by neuronal dynamics. A single inversion of this extended forward model enables inference about both the spatial deployment of sources and the underlying neuronal architecture generating them. Critically, this inference covers long-range connections among well-defined neuronal subpopulations.

In a previous paper, we simulated ERPs using a hierarchical neural-mass model that embodied bottom-up, top-down and lateral connections among remote regions. In this paper, we describe a Bayesian procedure to estimate the parameters of this model using empirical data. We demonstrate this procedure by characterizing the role of changes in cortico-cortical coupling, in the genesis of ERPs. In the first experiment, ERPs recorded during the perception of faces and houses were modeled as distinct cortical sources in the ventral visual pathway. Category-selectivity, as indexed by the face-selective N170, could be explained by category-specific differences in forward connections from sensory to higher areas in the ventral stream. We were able to quantify and make inferences about these effects using conditional estimates of connectivity. This allowed us to identify where, in the processing stream, category-selectivity emerged.

In the second experiment, we used an auditory oddball paradigm to show that the mismatch negativity can be explained by changes in connectivity. Specifically, using Bayesian model selection, we assessed changes in backward connections, above and beyond changes in forward connections. In accord with theoretical predictions, there was strong evidence for learning-related changes in both forward and backward coupling. These examples show that category- or context-specific coupling among cortical regions can be assessed explicitly, within a mechanistic, biologically motivated inference framework.

Keywords: Electroencephalography; Magnetoencephalography; Neural networks; Nonlinear dynamics; Causal modeling; Bayesian inference

Introduction

Event-related fields (ERFs) and potentials (ERPs) have been used for decades as putative magneto- and electrophysiological correlates of perceptual and cognitive operations. However, the exact neurobiological mechanisms underlying their generation are largely unknown. Previous studies have shown that ERP-like responses can be reproduced by brief perturbations of model cortical networks (David et al., 2005; Jansen and Rit, 1995; Jirsa, 2004; Rennie et al., 2002). The goal of this paper was to demonstrate that biologically plausible dynamic causal models (DCMs) can explain empirical ERP phenomena. In particular, we show that changes in connectivity, among distinct cortical sources, are sufficient to explain stimulus- or set-specific ERP differences. Adopting explicit neuronal models, as an explanation of observed data, may afford a better understanding of the processes underlying event-related responses in magnetoencephalography (MEG) and electroencephalography (EEG).

Abbreviations: DCM, dynamic causal Model(ing); EEG, electroencephalography; ERF, event-related field; ERP, event-related potential; MEG, magnetoencephalography; MMN, Mismatch negativity.

☆ Software note: The analyses presented in this paper will be available as a toolbox, distributed with the next release [SPM5] of the SPM software (http://www.fil.ion.ucl.ac.uk/spm).

* Corresponding author. Fax: +44 207 813 1420.
E-mail address: skiebel@fil.ion.ucl.ac.uk (S.J. Kiebel).
[1] Present address: INSERM U594 Neuroimagerie Fonctionnelle et Métabolique, Université Joseph Fourier, CHU-Pavillon B-BP 217, 38043 Grenoble Cedex 09, France.

Available online on ScienceDirect (www.sciencedirect.com).

1053-8119/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2005.10.045

Functional vs. effective connectivity

The aim of dynamic causal modeling (Friston et al., 2003) is to make inferences about the coupling among brain regions or sources and how that coupling is influenced by experimental factors. DCM uses the notion of *effective connectivity*, defined as the influence one neuronal system exerts over another. DCM represents a fundamental departure from existing approaches to connectivity because it employs an explicit generative model of measured brain responses that embraces their nonlinear causal architecture. The alternative to causal modeling is to simply establish statistical dependencies between activity in one brain region and another. This is referred to as *functional connectivity*. Functional connectivity is useful because it rests on an operational definition and eschews any arguments about how dependencies are caused. Most approaches in the EEG and MEG literature address functional connectivity, with a focus on dependencies that are expressed at a particular frequency of oscillations (i.e., coherence). See Schnitzler and Gross (2005) for a nice review. Recent advances have looked at nonlinear or generalized synchronization in the context of chaotic oscillators (e.g., Rosenblum et al., 2002) and stimulus-locked responses of coupled oscillators (see Tass, 2004). These characterizations often refer to phase-synchronization as a useful measure of nonlinear dependency. Another exciting development is the reformulation of coherence in terms of autoregressive models. A compelling example is reported in Brovelli et al. (2004) who were able show that "synchronized beta oscillations bind multiple sensorimotor areas into a large-scale network during motor maintenance behavior and carry Granger causal influences from primary somatosensory and inferior posterior parietal cortices to motor cortex." Similar developments have been seen in functional neuroimaging with fMRI (e.g., Harrison et al., 2003; Roebroeck et al., 2005).

These approaches generally entail a two-stage procedure. First an electromagnetic forward model is inverted to estimate the activity of sources in the brain. Then, a *post hoc* analysis is used to establish statistical dependencies (i.e., functional connectivity) using coherence, phase-synchronization, Granger influences or related analyses such as (linear) directed transfer functions and (nonlinear) generalized synchrony. DCM takes a very different approach and uses a forward model that explicitly includes long-range connections among neuronal subpopulations underlying measured sources. A single Bayesian inversion allows one to infer on parameters of the model (i.e., effective connectivity) that mediate functional connectivity. This is like performing a biological informed source reconstruction with the added constraint that the activity in one source has to be caused by activity in other, in a biologically plausible fashion. This approach is much closer in sprit to the work of Robinson et al. (2004) who show that "model-based electroencephalographic (EEG) methods can quantify neurophysiologic parameters that underlie EEG generation in ways that are complementary to and consistent with standard physiologic techniques." DCM also speaks to the interest in neuronal modeling of ERPs in specific systems. See for example Melcher and Kiang (1996), who evaluate a detailed cellular model of brainstem auditory evoked potentials (BAEP) and conclude "it should now be possible to relate activity in specific cell populations to psychophysical performance since the BAEP can be recorded in behaving humans and animals." See also Dau (2003). Although the models presented in this paper are more generic than those invoked to explain the BAEP, they share the same ambition of understanding the mechanisms of response generation and move away from phenomenological or descriptive quantitative EEG measures.

Dynamic causal modeling

The central idea behind DCM is to treat the brain as a deterministic nonlinear dynamical system that is subject to inputs, and produces outputs. Effective connectivity is parameterized in terms of coupling among unobserved brain states, i.e., neuronal activity in different regions. Coupling is estimated by perturbing the system and measuring the response. This is in contradistinction to established methods for estimating effective connectivity from neurophysiological time series, which include structural equation modeling and models based on multivariate autoregressive processes (Harrison et al., 2003; Buchel and Friston, 1997; Mcintosh and Gonzalez-Lima, 1994). In these models, there is no designed perturbation and the inputs are treated as unknown and stochastic. Although the principal aim of DCM is to explain responses in terms of context-dependent coupling, it can also be viewed as a biologically informed inverse solution to the source reconstruction problem. This is because estimating the parameters of a DCM rests on estimating the hidden states of the modeled system. In ERP studies, these states correspond to the activity of the sources that comprise the model. In addition to biophysical and coupling parameters, the DCM parameters cover the spatial expression of sources at the sensor level. This means that inverting the DCM entails a simultaneous reconstruction of the source configuration and their dynamics.

Implicit in the use of neural-mass models is the assumption that the data can be explained by random fluctuations around population dynamics that are approximated with a point mass (i.e., the mean or expected state of a population). This is usually interpreted in relation to the dynamics of an ensemble of neurons that constitute sources of signal. However, in the context of modeling ERPs and ERFs, there is also an ensemble of trials that are averaged to form the data. The mean-field-like assumptions that motivate neural mass models can be extended to cover ensembles of trials. This sidesteps questions about the trial-to-trial genesis of ERPs. However, we have previously addressed these questions using the same neural-mass model used in this paper (David et al., 2005), by dissociating "the components of event-related potentials (ERPs) or event-related fields (ERFs) that can be explained by a linear superposition of trial-specific responses and those engendered nonlinearly (e.g., by phase-resetting)." See David et al. (2005) for further details.

Because DCMs are not restricted to linear or instantaneous systems, they generally depend on a large number of free parameters. However, because it is biologically grounded, parameter estimation is constrained. A natural way to embody these constraints is within a Bayesian framework. Consequently, DCMs are estimated using Bayesian inversion and inferences about particular connections are made using their posterior or conditional density. DCM has been previously validated with functional magnetic resonance imaging (fMRI) time series (Friston et al., 2003; Riera et al., 2004). fMRI responses depend on hemodynamic processes that effectively low-pass filter neuronal dynamics. However, with ERPs, this is not the case and there is sufficient information, in the temporal structure of evoked responses, to enable precise conditional identification of quite complicated DCMs. In this study, we use a