NeuroImage 49 (2010) 442–456

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

# Bayesian spatiotemporal model of fMRI data

Alicia Quirós [a,*], Raquel Montes Diez [a], Dani Gamerman [b]

[a] Departamento de Estadística e Investigación Operativa, Universidad Rey Juan Carlos, Madrid, Spain
[b] Instituto de Matemática, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil

ARTICLE INFO

*Article history:*
Received 9 September 2008
Revised 26 June 2009
Accepted 9 July 2009
Available online 29 July 2009

*Keywords:*
Activation inference
Bayesian analysis
fMRI
GMRF
MCMC
Spatiotemporal model

ABSTRACT

This research describes a new Bayesian spatiotemporal model to analyse block-design BOLD fMRI studies. In the temporal dimension, we parameterise the hemodynamic response function's (HRF) shape with a potential increase of signal and a subsequent exponential decay. In the spatial dimension, we use Gaussian Markov random fields (GMRF) priors on activation characteristics parameters (location and magnitude) that embody our prior knowledge that evoked responses are spatially contiguous and locally homogeneous. The result is a spatiotemporal model with a small number of parameters, all of them interpretable. Simulations from the model are performed in order to ascertain the performance of the sampling scheme and the ability of the posterior to estimate model parameters, as well as to check the model sensitivity to signal to noise ratio. Results are shown on synthetic data and on real data from a block-design fMRI experiment.



## Introduction

This research focuses on determining which parts of the brain show activation in response to a stimulus in BOLD fMRI data. By observing the relation between a stimulus paradigm (in an experiment) and the hemodynamic response based on BOLD effect (Ogawa et al., 1990), fMRI provides a measure of brain activation. In particular, we consider a block-design fMRI study (Lazar 2008), in which scans are acquired under two different conditions, alternating periods when the stimulus is on with those when the stimulus is off.

Inference about brain activity in fMRI data is commonly addressed through the General Linear Model (GLM) analysis, introduced by Friston et al. (1995), in which a linear dependency of the BOLD signal and the hemodynamic response function (HRF) is assumed.

Generally, the stimulus pattern is fit simply as a box-shaped wave, slightly delayed, in order to account for the lapse of time between the stimulus onset and the arrival of the blood to the activated area. This box-shaped wave is convolved with a HRF template for which several kernels have been considered, including Poisson (Friston et al., 1994), Gaussian (Friston et al., 1995) and gamma (Lange and Zeger 1997; Boynton et al., 1996). The convolution approach is attractive for its simplicity. However, it imposes restrictions to the model, e.g. it forces antisymmetry and monotonicity on each half cycle, as mentioned by Crellin, Hastie and Johnstone in the published discussions of Lange and Zeger (1997).

The result of this analysis is a statistical parametric map (SPM), i.e. a $T$, $F$ or $Z$ estimate for each voxel. The next step is to threshold this SPM (leading to a multiple comparisons problem), in order to decide, at a given level of significance, which parts of the brain are activated. The $p$-value used to make inference represents the probability of obtaining a result at least as extreme as the observed data given the truth of the null hypothesis of no activation and it is sometimes misinterpreted as the likelihood of activity presence. Furthermore, in this approach it is not possible to infer that no activation occurs (Friston et al., 2002; de Pasquale et al., 2008).

The spatial aspect of the hemodynamic response is usually taken into account by spatially smoothing the data with a fixed Gaussian kernel as a pre-processing step. This helps to improve the signal to noise ratio. However, smoothing the data can result in drawbacks: too much smoothing will blur activations, while too little will leave unnecessary noise in the data (Flandin and Penny, 2007). As univariate methods do not take into account the spatial structure on data, several approaches have been proposed in the literature to model spatial dependencies (Bowman et al., 2008; Flandin and Penny, 2007; Harrison et al., 2007; Penny et al., 2005; Hartvig and Jensen, 2000; Gossl et al., 1999).

Multivariate methods provide an alternative to individual analysis of voxels, reducing the whole spatiotemporal data set into certain multivariate components with similar temporal characteristics, see for example the reference book on multivariate methods, Mardia et al. (1979). Multivariate techniques applied to the analysis of fMRI data include principal component analysis (Friston, 1994; Sjöstrand et al., 2006), independent component analysis (Beckmann and Smith, 2004; Esposito et al., 2003; McKeown et al., 2003; Calhoun et al., 2001; Porill et al., 1999) and cluster analysis (Goutte et al., 1999). Interpretation of

* Corresponding author. Edificio Dept. II, Despacho 049, Universidad Rey Juan Carlos, C/Tulipan, s/n, 28933 Mostoles (Madrid), Spain. Fax: +34 91 488 7626.
*E-mail addresses:* alicia.quiros@urjc.es (A. Quirós), raquel.montes@urjc.es (R.M. Diez), dani@im.ufrj.br (D. Gamerman).

1053-8119/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2009.07.047

results must be treated with care when using these techniques, as there is a high risk of data overfitting.

The fMRI data analysis involves the spatiotemporal relationship between a stimulus or cognitive task and the cerebral response measured with fMRI. In spite of the obvious spatiotemporal nature of data, there are few spatiotemporal models, see for example, Bowman (2007), Penny et al. (2005), Katanoda et al. (2002) and Woolrich et al. (2004a). Although all these models are based on convolution, the four of them present a different modelisation of the spatiotemporal correlation structure between voxels. Bowman (2007) incorporates a functionally defined distance metric into a parametric structure for spatial correlations within a ROI (with the difficulty of choosing the ROI and the functional distance between voxels in it) and includes temporal correlations between scans. Penny et al. (2005) propose a fully Bayesian model with spatial priors defined over regression coefficients of a GLM, using Gaussian Markov random fields (GMRF), and the errors are modelled as an autoregressive process. Katanoda et al. (2002) propose a spatiotemporal regression model for each voxel that involves the time series of the neighbouring voxels together with its own. Woolrich et al. (2004a) present a fully Bayesian approach, incorporating spatiotemporal noise modelling. A general problem of spatiotemporal models is the large number of parameters and, as a consequence, the great computational burden.

The Bayesian paradigm provides an appropriate framework for making inference using complex models and to overcome the multiple comparisons problem. It also constitutes a natural but rigorous theory for combining prior and experimental information. Most Bayesian approaches to the modelling of fMRI data use GMRF as prior distributions, in order to account for the spatial structure present in the data, e.g. Gossl et al. (2001) use GMRF to spatially regularise regression coefficients and Woolrich et al. (2004b) to spatially regularise AR coefficients. Moreover, several choices of the precision matrix of the GMRF prior on regression coefficients of a GLM-AR model have been considered in the Bayesian literature, these include uninformative priors (Penny et al., 2003), global-shrinkage priors (Friston and Penny, 2003) and Laplacian priors (Penny et al., 2005) among others. An interesting comparison of these priors can also be found in Penny et al. (2005).

The package SPM is a set of tools created by Frackowiak et al. (2004), and described in Friston et al. (2006), for the implementation of statistical analysis of parametric maps, and multivariate techniques, based on GLM. Although no recognised standard has yet been set, SPM has become a standard as the most complete package available.

In his PhD. thesis, Kornak (2000) proposed a two-stage model to analyse fMRI data. In the first stage, he summarises the temporal information present in the data at individual voxels and forms voxel maps for each parameter, using least squares estimation. He incorporates these maps in the second stage using GMRF as prior distributions for the parameters in a Bayesian spatial model. A fully Bayesian extension of Kornak's work can be seen in Quirós et al. (2006).

In this paper, we propose a new Bayesian spatiotemporal model to determine active areas into the brain by merging the two stages of the model first proposed by Kornak (2000) and extended by Quirós et al. (2006). In the temporal dimension, we parameterise the HRF shape with a potential increase of signal and a subsequent exponential decay. In addition, the delay of the HRF is not fixed in advance but modelled as an unknown parameter. In the spatial dimension, we use GMRF priors on activation characteristics parameters (location and magnitude) that embody our prior knowledge that evoked responses are spatially contiguous and locally homogeneous. In this way, smoothing is included as a part of the model and it is not left to a pre-processing step. Despite being spatiotemporal, the proposed model has a small number of parameters and all of them are interpretable.

The paper is arranged as follows: in the following section, we state the model and prior distributions of the parameters in the model. We examine the results obtained by applying the model to simulations, synthetic data and to real data in the results section. This is followed by the discussion and conclusions. The posterior distribution and the details regarding MCMC methods used to sample from the posterior distribution are included in two appendices to the document.

## Materials and methods

### *The model*

In this work we consider a block-design fMRI study, where stimulation blocks (in which the stimulus is on during several seconds) are alternated with resting blocks (in which the stimulus is off). Two consecutive blocks (one with stimulus and one without) are called a cycle. Let $C$ be the number of cycles in the experiment, $T$ the number of images in each cycle and $N\times M$ the dimension of each image. Notice that this is a 2D approach (analysing axial slices) and that half of the images in a cycle ($T/2$) are taken when the stimulus is on and half, when the stimulus is off. So that the whole data of the experiment may be given by

$$y = \left\{ y_{s,\tau,c} : s = 1,\ldots,N\times M; \tau = 1,\ldots,T; c = 1,\ldots,C \right\}, \tag{1}$$

where $y_{s,\tau,c}$ is the value of voxel $s$, in the image number $\tau$ of the cycle $c$ and where each cycle $c$ is assumed to be identically distributed.

Our proposal to analyse fMRI data is the following model for $c=1,\ldots,C$:

$$y_{\tau,c} = k + a \odot h_\tau(d) + \varepsilon 1, \quad \varepsilon \sim \mathcal{N}\left(0,\sigma^2\right) \tag{2}$$

where $\odot$ means element by element product and $\mathbf{1}$ is a $N\times M$-matrix of ones.

Field $k=\{k_s : s=1,\ldots, N\times M\}$ represents the baseline level for each voxel $s$.

Field $a=\{a_s : s=1,\ldots, N\times M\}$ defines the spatial variation of brain activity by classifying voxels as active or non-active and by providing the magnitude of the response in active voxels. This is achieved by defining

$$a = z \odot x, \tag{3}$$

where, field $z=\{z_s : s=1,\ldots, N\times M\}$ is a binary random field, defining the presence (1) or absence (0) of activity (in practice, $z$-field is defined by thresholding a continuous field, $w$, i.e., $z_s = I_{\{w_s > 0\}}$, and it provides a summary of location of activity) and where field $x=\{x_s : s=1,\ldots, N\times M\}$ is a continuous random field that models the spatially varying response level of activated regions.

For active voxels, we parameterise $h_\tau(d)$ for $\tau=1,\ldots,T$, with the shape of a Poisson probability density function with mean $d$, that is, a potential increase of signal and a subsequent exponential decay,

$$h_\tau(d_s) = \frac{e^{-d_s} d_s^{\tau-1}}{(\tau-1)!} \tag{4}$$

for each voxel $s$. Consequently, parameter $d$ can be interpreted as the delay of the response with respect to the stimulus onset. In conclusion, when a voxel is active ($z_s=1$), the data are modelled as the sum of the baseline $k_s$ and the expression $x_s \frac{e^{-d_s} d_s^{\tau-1}}{(\tau-1)!}$, which parameterises the hemodynamic response for that voxel. Otherwise, if the voxel $s$ is non-active ($z_s=0$), the data are only modelled by $k_s$.

### *Prior distributions*

Under the Bayesian framework, it is necessary to specify prior distributions for the image of interest, in this case, the $x$-field and the generator field of $z$, $w$. GMRF have been widely used for this purpose (for instance, see Winkler (2003)) as they are able to incorporate prior