

Statistical Modelling

<http://smj.sagepub.com/>

Semiparametric Bayesian models for human brain mapping

L Fahrmeir and C Gössl

Statistical Modelling 2002 2: 235

DOI: 10.1191/1471082x02st040oa

The online version of this article can be found at:

<http://smj.sagepub.com/content/2/3/235>

Published by:



<http://www.sagepublications.com>

On behalf of:



Statistical Modelling Society

Statistical Modeling Society

Additional services and information for *Statistical Modelling* can be found at:

Email Alerts: <http://smj.sagepub.com/cgi/alerts>

Subscriptions: <http://smj.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

Citations: <http://smj.sagepub.com/content/2/3/235.refs.html>

>> [Version of Record](#) - Oct 1, 2002

[What is This?](#)

Semiparametric Bayesian models for human brain mapping

L Fahrmeir¹ and C Gössl²

¹Department of Statistics, Ludwig-Maximilians-University Munich, Munich, Germany

²Max-Planck-Institute of Psychiatry, Munich, Germany

Abstract: Functional magnetic resonance imaging (fMRI) has led to enormous progress in human brain mapping. Adequate analysis of the massive spatiotemporal data sets generated by this imaging technique, combining parametric and non-parametric components, imposes challenging problems in statistical modelling. Complex hierarchical Bayesian models in combination with computer-intensive Markov chain Monte Carlo inference are promising tools.

The purpose of this paper is twofold. First, it provides a review of general semiparametric Bayesian models for the analysis of fMRI data. Most approaches focus on important but separate temporal or spatial aspects of the overall problem, or they proceed by stepwise procedures. Therefore, as a second aim, we suggest a complete spatiotemporal model for analysing fMRI data within a unified semiparametric Bayesian framework. An application to data from a visual stimulation experiment illustrates our approach and demonstrates its computational feasibility.

Key words: functional magnetic resonance imaging, human brain mapping, MCMC, semiparametric models, spatiotemporal models

Data and software link available from: <http://stat.uibk.ac.at/SMIJ>

Received October 2001; revised March 2002; accepted August 2002

1 Introduction

The development of modern non-invasive imaging techniques has led to rapid advances in human brain mapping. Functional magnetic resonance imaging (fMRI) aims at the localization of functional brain areas in a living human brain, that is, the detection of areas or regions that are responsible for the processing of certain stimuli. Adequate statistical modeling and analysis of the massive spatiotemporal data sets generated with this imaging technique incorporates challenging tasks for the statistical community. Modern Bayesian approaches, using computer-intensive and realistically complex models, offer promising tools for inference.

This paper has two major aims. First, it provides a review of current general semiparametric Bayesian models for human brain mapping using fMRI data. A characteristic feature of many of these approaches is that they are designed only to

Address for correspondence: L Fahrmeir, Department of Statistics, Ludwig-Maximilians-University Munich, Ludwigstrasse 33, 80539 Munich, Germany. E-mail: fahrmeir@stat.uni-muenchen.de

solve separate parts of the complete problem. Hierarchical Bayesian modelling, however, has the potential to link such separate pieces to a joint model. Therefore, as our second goal we suggest a comprehensive semiparametric spatiotemporal model, based on previously developed submodels, and show its computational feasibility by an application to data from a visual fMRI experiment.

In fMRI experiments it is assumed that the presentation of an external stimulus causes a local increase in blood oxygenation in activated areas. This increase can be visualized with specially sensitized MR sequences and is called the BOLD (blood oxygenation level dependent) effect. Thus, acquiring a complete time series of MR images, activated areas can be identified by a systematic MR signal increase during stimulation in comparison to a rest condition. In most cases the stimulus is presented in a so-called boxcar design of alternating activation and rest periods. For a visual stimulation, Figure 1 shows such a boxcar stimulation paradigm together with representative MR signal time series of different regions of the brain that are strongly, weakly, or not activated.

Conventional standard analyses of fMRI data comprise correlation (Bandettini *et al.*, 1993) or parametric regression models (Friston *et al.*, 1995). Let $i, i = 1, \dots, I$, denote the pixel index. In a parametric linear regression set-up, the time series $\{y_{it}, t = 1, \dots, T\}$ of measured MR signals at pixel i is assumed to be the sum

$$y_{it} = w_t' a_i + z_t b_i + \epsilon_{it}, \quad \epsilon_{it} \sim N(0, \sigma_i^2), \quad i = 1, \dots, I, \quad t = 1, \dots, T \quad (1.1)$$

of a baseline trend $\{w_t' a_i, t = 1, \dots, T\}$, an activation profile $\{z_t b_i, t = 1, \dots, T\}$, and a measurement error $\{\epsilon_{it}, t = 1, \dots, T\}$. The vector w_t consists of the evaluations of a few known simple basis functions such as polynomials or trigonometric functions. The transformed stimulus $\{z_t, t = 1, \dots, T\}$ is linked to the original (boxcar) stimulus

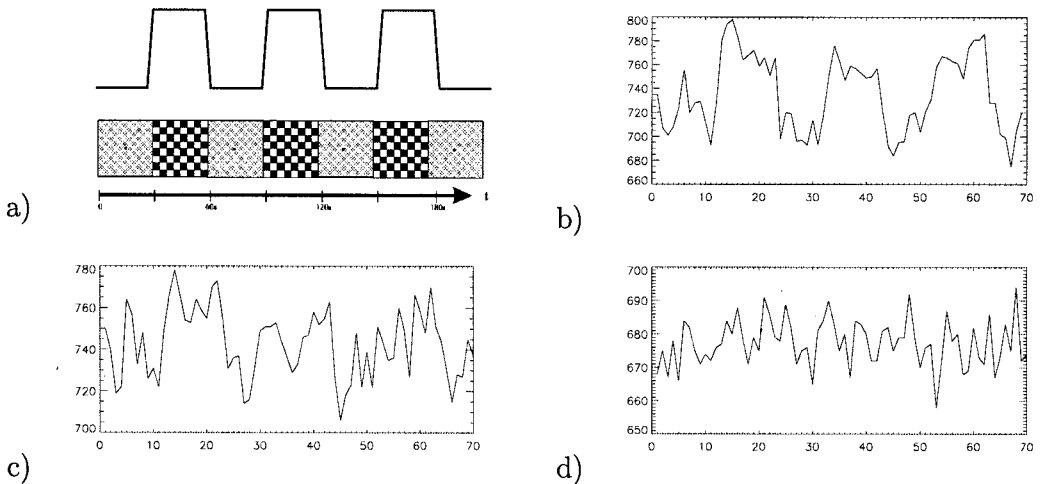


Figure 1 Visual fMRI: (a) An 8Hz flickering rectangular checkerboard (ON) is presented to the subject, alternating every 30 s with a uniformly dark background and a fixation point (OFF); an experiment consists of four OFFs and three ONs; Representative MR signal time courses from strongly (b), weakly (c), and non-activated (d) pixels

$\{x_t, t = 1, \dots, T\}$ of Figure 1a through a so-called hemodynamic response function (HRF). The HRF is used to model the neurovascular coupling, that is, the relation between neuronal activation and the changes in blood oxygenation. These occur only shifted by some time lag d and proceed rather continuously, often described by a convolution with a Poisson or Gamma density b :

$$z_t = \sum_{s=0}^{t-d} b(s, \theta) x_{t-d-s}. \quad (1.2)$$

The parameters d and θ of the HRF are mostly estimated prior to the above regression analysis, as global estimates for the whole brain (Friston *et al.*, 1994) or for each pixel separately (Cohen, 1997; Rajapakse *et al.*, 1998; Gössl *et al.*, 2000). In the latter case, the transformed stimulus depends on pixel i through θ_i , that is, one has $z_{it} = z_t(\theta_i)$. Further, instead of the convolution model direct approaches have also been proposed (Figure 2), which describe increase and decrease of the hemodynamic response directly by a combination of diverse functions or splines (Aguirre *et al.*, 1998; Crellin *et al.*, 1998; Kruggel *et al.*, 2000; Genovese, 2000; Gössl *et al.*, 2001b).

After having estimated the stimulus effects $b = (b_1, \dots, b_T)'$ in the regression model, these are tested to be zero for each pixel separately. Pixels with a significant stimulus effect are assumed to be activated and are displayed in a so-called activation map, which is shown in Figure 3 for a visual stimulation. Furthermore, in many cases spatial postprocessing steps are applied to the field of estimated effects b . They range from simple cluster algorithms (Forman *et al.*, 1995), to reduce the number of false positive activations, to Gaussian random field considerations (Poline *et al.*, 1997), to account for multiple test problems arising from the huge number of tested pixels.

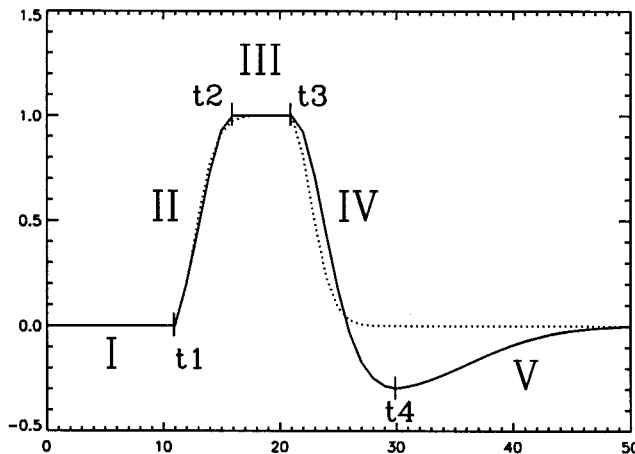


Figure 2 Direct hemodynamic response model, divided into five basis epochs. For comparison, the dotted line displays a classical Poisson convolution model

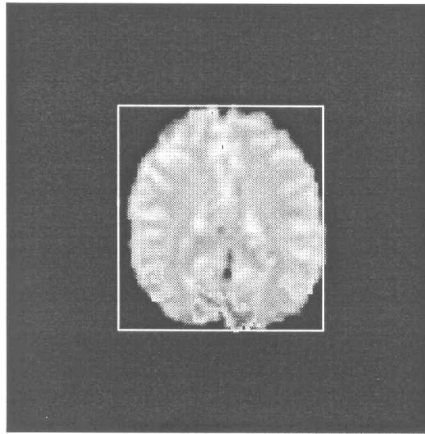


Figure 3 Activation map for a simple linear regression model

Beside other distribution-free, non-parametric, Fourier-based or multivariate techniques, in recent years several variations and extensions of the general regression approach have been proposed, based on generalized additive models. Bullmore *et al.* (1996) model the reference function as a sum of trigonometric basis functions and derive a different test statistic, for which they calculate an empirical distribution by permutation. Genovese (2000) and Gössl *et al.* (2000) describe a pixelwise semiparametric approach to model the baseline more flexibly. Whereas Genovese uses regression splines for this purpose, Gössl *et al.* apply a state space model. In the latter, as an important feature, the activation effect is also modelled dynamically, allowing temporally varying activation effects to be described. To control the multiple test problem, Everitt and Bullmore (1999) and Hartvig and Jensen (2000) describe the calculated map of test statistics by a spatial mixture model of activated and non-activated pixels. Hartvig and Jensen (2000) thereby also account for spatial dependencies between pixels. Spatio-temporal models have been proposed by Gössl *et al.* (2001a), formulated in the form of different priors in a hierarchical Bayesian approach.

A common characteristic of the mentioned models is that statistical modelling and analysis is divided into several separate steps. The main issues are: detrending and smoothing in preprocessing steps prior to the evaluation; determination of the HRF; estimation of the activation profile; and spatial considerations concerning the activation maps. We show how this subdivision can be avoided in a hierarchical Bayesian framework.

The paper is organized as follows. Section 2 first reviews semiparametric regression-based Bayesian model pieces, which are defined pixelwise and neglect spatial aspects, but seem promising for being incorporated into a joint fMRI model (Section 2.1). In Section 2.2, we describe some approaches that deal with spatial aspects, but that are less flexible with respect to the temporal component than the pixelwise models of Section 2.1. In Section 2.3, we introduce a hierarchical Bayesian approach that incorporates flexible spatiotemporal modelling and estimation of the hemodynamic

response function in a comprehensive fMRI model. Section 3 illustrates this method with an application to fMRI data from a visual experiment. The concluding section discusses some perspectives for future research, such as computational issues and analysis of diffusion tensor imaging (DTI) for tracking fiber connections between functional areas.

2 Semiparametric Bayesian models for fMRI experiments

In the following, we first introduce semiparametric Bayesian regression models for fMRI signals from a pixelwise point of view. The main ingredients are flexible forms for the baseline trend and the activation profile, adapting smoothness priors or basis functions approaches to fMRI data. Section 2.2 introduces some models that deal with spatial aspects, but also considers more restrictive forms for pixelwise modelling of baseline and activation. In Section 2.3 we propose a comprehensive spatiotemporal model, incorporating space–time interactions and simultaneous estimation of the hemodynamic response functions. This model will be applied to a visual experiment in Section 3.

2.1 Semiparametric regression models for pixels

All regression-based approaches for the evaluation of fMRI data assume pixelwise additive models of the general form

$$y_{it} = a_{it} + f_{it} + \varepsilon_{it}, \quad i = 1, \dots, I, \quad (2.1)$$

where $\{a_{it}, t = 1, \dots, T\}$ is the baseline trend, $\{f_{it}, t = 1, \dots, T\}$ is the activation profile and $\{\varepsilon_{it}, t = 1, \dots, T\}$ is measurement noise at pixel i . Various models can result from specific assumptions on these three components. A slight extension of the pixelwise parametric linear regression model (1.1) is obtained by setting

$$a_{it} = w'_t a_i, \quad f_{it} = v'_t b_i z_t(\theta_i), \quad (2.2)$$

where $z_t(\theta_i)$ is defined as in equation (1.2) and θ_i is assumed to be known from a pilot estimate. The linear combination $v'_t b_i$ models a time-varying activation effect $b_{it} = v'_t b_i$ in complete analogy to the trend component $a_{it} = w'_t a_i$. For the special choice $v_t \equiv 1$, expression (2.2) reduces to (1.1)

More flexible forms for the unknown functions a_{it} and b_{it} , $t = 1, \dots, T$, are obtained by modelling and estimating them non-parametrically. Bayesian approaches for non- and semiparametric regression and time series analysis are based on appropriate smoothness priors or on basis function representations with adaptive knot selection. Gössl *et al.* (2000) propose a varying coefficient model

$$y_{it} = a_{it} + b_{it} z_t(\theta_i) + \varepsilon_{it}, \quad \varepsilon_{it} \sim N(0, \sigma_i^2), \quad (2.3)$$

where smoothness of the trend $\{a_{it}\}$ and the time-varying activation effect $\{b_{it}\}$ is achieved through pixelwise random walk priors for the unknown parameters a_{it} and b_{it} . Expressing this model in state space form, the baseline trend and a time-varying estimation effect can be estimated by Kalman filtering and smoothing separately at each of the pixels. As an alternative, Bayesian smoothing splines (Hastie and Tibshirani, 2000) or P-splines, developed by Lang and Brezger (2002) as a Bayesian version of Eilers and Marx (1996), might be used.

Random walk priors can be seen as a discretized version of smoothing spline priors. For an equidistant grid as in the analysis of fMRI signals, estimates from both methods are practically indistinguishable. We prefer random walk priors, because they are computationally less demanding and are more naturally combined with Markov random field priors to obtain full spatiotemporal models (see Gössl *et al.* (2001b) and the next subsection).

Bayesian P-splines assume representations of the form

$$a_{it} = \sum_{j=1}^m \tilde{a}_{ij} B_j(t),$$

$$b_{it} = \sum_{j=1}^m \tilde{b}_{ij} B_j(t),$$

with a moderate number $m = r + l$ B-spline basis functions of degree l and r inner knots. Again, smoothness is achieved through random walk priors for the vectors $a_i = (\tilde{a}_{i1}, \dots, \tilde{a}_{im})$ and $b_i = (\tilde{b}_{i1}, \dots, \tilde{b}_{im})$, but m can be considerably smaller than T . This more parsimonious parameterization might be useful for computational reasons.

The hemodynamic response function in (1.2) and the transformed stimulus $z_t(\theta_i)$ are mostly determined in a pilot step. Genovese (2000) includes specification of the hemodynamic response as part of his pixelwise semiparametric regression models. A related approach is suggested in Gössl *et al.* (2001b) in combination with the linear parametric observation model (1.1). We will integrate this HRF model as part of our complete spatiotemporal model in Section 2.3.

A final remark concerns the Gaussian error assumption $\varepsilon_{it} \sim N(0, \sigma_i^2)$. In our experience, Gaussian observation models showed good performance in all applications. However, robustified observation models could be incorporated by (discrete or continuous) mixtures of normals for the error distribution. Also temporal autocorrelations of the error terms have already been described (see, for example, Bullmore *et al.*, 1996; Friston *et al.*, 1995). There is evidence, however, that much of this may be modelled by using a flexible trend model with white noise superimposed. Purdon and Weiskoff (1998) used a similar approach with an AR(1) model for the trend, and Zarahn *et al.* (1997) found that this type of model fitted well to empirical periodograms of noise data. Hartvig (2000, chapter 2) applies flexible nonlinear state space models to decompose the trend into a smooth component and white noise errors.

2.2 Spatial models

This section discusses some models that deal with spatial correlation between neighbouring pixels. Although formulated for special models or problems, they have the potential for being combined with flexible pixelwise models described in the previous section.

Simple parametric spatial models can be defined by assuming a Markov random field prior for the activation effects b_i , $i = 1, \dots, I$, in the linear regression model (1.1). A common MRF prior defined in terms of Gaussian difference priors is (see, for example, Besag *et al.*, 1991):

$$p(b|\lambda) \propto \exp \left\{ -\frac{1}{2} \lambda \sum_{i \sim j} (b_i - b_j)^2 \right\}, \quad (2.4)$$

where $i \sim j$ denotes that i and j are neighbouring pixels. The precision parameter λ controls the amount of smoothness and can be estimated jointly with other parameters by assigning a gamma prior. A certain concern with Gaussian priors (2.4) is over-smoothing in areas of high curvature in the activation surface $\{b_i, i = 1, \dots, I\}$.

Fahrmeir *et al.* (2002) replace Gaussian pairwise difference priors (2.4) by robust versions and investigate their potential use to avoid blurring of edges, peaks, and areas of high curvature in the activation surface. The general form of robust priors is given by

$$p(b|\tau, w) \propto \exp \left\{ -\sum_{i \sim j} w_{ij} \Phi[\tau(b_i - b_j)] \right\},$$

where τ is a scale parameter, Φ is symmetric with $\Phi(u) = \Phi(-u)$, and w_{ij} are weights for neighbours $i \sim j$. Usually, the weights are assumed to be known, for example, by setting them equal to one or by measuring the distance of pixels in some way. In Fahrmeir *et al.* (2002), the weights are also admitted to be random variables obeying a hyperprior. Then, the marginal prior $p(b|\tau)$ is a mixture of pairwise interaction priors. Such priors give additional flexibility when pixel i is near the border of an activated area. There some neighbors $j \in \partial i$ might have similar activation effects, whereas others may be only weakly or not activated. More specifically, the w_{ij} 's are allowed to be i.i.d. variables following a Gamma hyperprior $w_{ij} \sim GA(v/2, v/2)$. For $\Phi(u) = u^2/2$, the resulting mixture distribution is a Student prior with v degrees of freedom, which is a Cauchy prior for $v = 1$. For $\Phi(u) = |u|$, the mixture is a weighted Laplace prior. Experience with simulated and real data applications shows that the robustified priors improve estimation of the activation surface. However, if the main interest is in testing, that is, in the activation maps, there is no substantial gain in using robust instead of Gaussian priors. As an alternative to the i.i.d. priors for the weights, MRF priors as suggested by Aykroyd (1998) could be considered. However, it is unclear whether the additional computational effort is justified and whether estimation of local features is significantly improved.

Hartvig (2000) considers a spatiotemporal semiparametric observation model

$$y_{it} = b_i(\Phi_t + v_t) + \varepsilon_{it},$$

for detrended signals y_{it} . The activation surface $\{b_i, i = 1, \dots, I\}$ is decomposed into the sum $b_i = a_i + \eta_i$ of a smooth part a_i and a random intercept η_i . The smooth part is modelled through a spatial Bayesian basis function approach. The basis functions are bells with centres obeying a spatial point process prior. The hemodynamic response function is decomposed into a known fixed part Φ_t , obtained from a pilot estimate, and a random walk v_t for deviations from Φ_t . Estimation is carried out via MCMC. The model is related to the additive spatiotemporal model in Gössl *et al.* (2001a), where a MRF smoothness prior is assumed for the activation surface $\{b_i\}$.

Hartvig and Jensen (2000) propose a spatial mixture model for the statistical parametric map (SPM) of pixelwise activation estimates t_i , obtained from a parametric regression model as in (1.1). For each pixel i , a latent binary activation indicator A_i is introduced. The marginal distribution of t_i is then specified by a mixture distribution $\pi f(t_i|A_i = 1) + (1 - \pi)f(t_i|A_i = 0)$, where π is the fraction of activated pixels, or by more complex mixture models. Spatial correlations are modelled through adequate priors for the activation profiles. Conceptually, it seems possible to include this approach into the spatial part of a spatiotemporal model, but implementation will probably cause challenging problems.

2.3 A complete spatiotemporal model based on smoothness priors

There are several ways of combining temporal, spatial, and hemodynamic response models as the components of a joint, complete spatiotemporal model for fMRI data. In view of the challenging and massive amount of data, computational feasibility is still an important issue, and some compromise with model complexity is needed. In the following, as in Gössl *et al.* (2001a), temporal and spatial components are modelled through Gaussian random field smoothness priors. Additionally, a hemodynamic response model is incorporated as a further ingredient of the complete model. As illustrated in Section 3, this leads to computationally feasible posterior inference with Metropolis–Hastings techniques.

At least in current implementations, basis function and partition approaches require too much computation time for high-dimensional spatiotemporal models.

We first focus on temporal and spatial smoothness priors for the trend $\{a_{it}\}$ and the activation effect $\{b_{it}\}$ in the semiparametric model (2.3). The hemodynamic response model is described at the end of this subsection.

Since priors for $\{a_{it}\}$ and $\{b_{it}\}$ are essentially identical, we restrict presentation to the activation effect. Let $b_i, i = 1, \dots, I$ denote the vector of all parameters characterizing $\{b_{it}\}$. In a parametric temporal model $b_{it} = \varphi'_t b_i$, the vector b_i is typically low-dimensional or, for a time-constant activation effect $b_{it} = b_i$, scalar. For semiparametric models (2.3) with random walk priors, the entire high-dimensional vector $(b_{i1}, \dots, b_{iT})'$ itself is considered as the unknown b_i . To define spatiotemporal priors, we further

gather the entire set of parameters in $b = (b'_1, \dots, b'_I)'$. Then all priors for b have the same general conditional Gaussian form

$$p(b|K) \propto \exp\left(-\frac{1}{2}b'Kb\right), \tag{2.5}$$

where K is a penalty or precision matrix of known structure, containing unknown hyperparameters. Models of special interest result by specifying particular structures for K .

For block diagonal $K = \text{diag}(K_1, \dots, K_I)$, we obtain pixelwise models with

$$p(b|K) \propto \prod_i \exp\left(-\frac{1}{2}b'_i K_i b_i\right).$$

For example, semiparametric models (2.3) with second order random walk priors are given by

$$K_i = \lambda_i Q^t, \quad Q^t = \begin{pmatrix} 1 & -2 & 1 & & & & \\ -2 & 5 & -4 & 1 & & & \\ 1 & -4 & 6 & -4 & 1 & & \\ & \ddots & \ddots & \ddots & \ddots & \ddots & \\ & & 1 & -4 & 6 & -4 & 1 \\ & & & 1 & -4 & 5 & -2 \\ & & & & 1 & -2 & 1 \end{pmatrix} \tag{2.6}$$

The precision parameter λ_i controls the amount of smoothness. Usually, we assign λ_i a highly non-informative gamma hyperprior $\lambda_i \sim GA(a, b)$. Also the classical parametric pixelwise models (1.1) and spatial models with pairwise difference priors can be written as special cases of the general prior (2.4).

Semiparametric spatiotemporal models can be specified by appropriate combination of precision matrices for temporal and spatial components. In Gössl *et al.* (2001a), an additive and a non-additive model have been proposed. Here, we focus on the latter one, which incorporates space–time interactions. In componentwise formulation, the prior is

$$p(b|\lambda) \propto \exp\left[-\frac{1}{2} \sum_i \left\{ \lambda_i^t \sum_t (\Delta^2 b_{it})^2 + \lambda_i^s \sum_t \sum_{j \in \partial i} (\Delta^2 b_{it} - \Delta^2 b_{jt})^2 \right\}\right], \tag{2.7}$$

where $\Delta^2 b_{it} = b_{it} - 2b_{i,t-1} + b_{i,t-2}$ are second differences. The first term corresponds to (pixelwise) second order random walk priors for the temporal effect, while the second term penalizes differences between neighbouring local trends b_{it} and b_{jt} between pixels i and adjacent pixel j , with $j \in \partial i$ of j . For $\lambda_i^s \equiv 0$, equation (2.7) reduces to a pixelwise model. Assuming for simplicity equal temporal and spatial precision parameters

$\lambda_i \equiv \lambda_i^t \equiv \lambda_i^s$ as in Gössl *et al.* (2001a), the prior (2.7) can be written in the general form (2.4) with

$$K = (Q^s \otimes Q^t) + (\Lambda \otimes Q^t) = (Q^s + \Lambda) \otimes Q^t,$$

$$Q_{ij}^s = \begin{cases} \sum_{k \in \partial_i} (\lambda_i + \lambda_k) & i = j \\ -(\lambda_i + \lambda_j) & i \sim j \\ 0 & \text{else} \end{cases}$$

and $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_I)$.

As in Genovese (2000), we will further include specification of the HRF as part of the observation model. This will allow us to embed estimation of the activation profile within a complete spatiotemporal modelling set-up, in contrast to Gössl *et al.* (2001a), where the hemodynamic response function is estimated in a pilot step. For boxcar paradigms as considered in our application in Section 3, the hemodynamic response function $z_{it}(\theta_i)$ is composed of a sequence of identical hemodynamic response models $w(u; \theta_i)$, linked together on the t -axis, and defined by

$$w(u; \theta_i) = \begin{cases} 0.0 & \text{for } u < t_1 & \text{(I)} \\ \left\{ \exp \left[-\left(\frac{u - t_2}{\theta_1} \right)^2 \right] - c_1 \right\} / (1 - c_1) & \text{for } u \in [t_1, t_2[& \text{(II)} \\ 1.0 & \text{for } u \in [t_2, t_3[& \text{(III)} \\ (1 + \theta_5) \left\{ \exp \left[-\left(\frac{u - t_3}{\theta_2} \right)^2 \right] - c_2 \right\} / (1 - c_2) - \theta_5 & \text{for } u \in [t_3, t_4[& \text{(IV)} \\ -\theta_5 \exp \left[-\left(\frac{u - t_4}{4.0} \right)^2 \right] & \text{for } u \geq t_4 & \text{(V)}, \end{cases}$$

where the above time points and normalizing constants are related to the model parameters θ_1 increase, θ_2 decrease, θ_3 lag, θ_4 plateau, and θ_5 undershoot, according to $t_1 = \theta_3$, $t_2 = \theta_1 + \theta_3$, $t_3 = \theta_1 + \theta_3 + \theta_4$, $t_4 = \theta_1 + \theta_2 + \theta_3 + \theta_4$, $c_1 = c_2 = \exp(-1)$. Figure 2 displays the shape of this hemodynamic response model. It is used by Gössl *et al.* (2001b) in combination with a linear parametric observation model (1.1). A generalization from the simple boxcar paradigm to a more complex sequence of tasks as in Genovese (2000) is conceptually straightforward.

For the observation error variances σ_i^2 and the parameters θ_i of the hemodynamic response model, we assume pixelwise priors

$$\sigma_i^2 \sim IG(\gamma_a, \gamma_b), \quad \theta_{ij} \sim U[0, I_j], \quad j = 1, \dots, 4, \quad \theta_{i5} \sim N(0, \xi^2),$$

where the support $[0, I_j]$ is predefined, for example, $I_j = 15$ s.

Together with usual conditional independence assumptions, the priors and the observation models complete the Bayesian model specification.

3 Applications

To illustrate the spatiotemporal approach of Section 2.3 we use the data set also analysed in Gössl *et al.* (2001a), consisting of an fMRI time series from a visual stimulation experiment. In order to elicit a strong activation in visual cortical areas a standard 8 Hz flickering checkerboard stimulus was used. On a 1.5 T system (Echospeed, GE Medical Systems, Milwaukee), a set of 70 T2* BOLD images was acquired, consisting of seven slices parallel to the intercommissural line, covering the occipital lobes. Each image comprises 128×128 pixels (voxel size: $2.9 \times 2.9 \times 5$ mm). The stimulation paradigm was divided in four rest and three activation periods, each epoch 10 images (30 s) long. Figure 1a shows the study design. To correct for subject's motion, an image registration was performed prior to the analysis.

For the estimation of the above models we use a pixelwise Metropolis–Hastings (MH) algorithm to generate the samples. The parameters a , b , and σ^2 are sampled in Gibbs steps, for the HRF parameters an MH step is introduced. The full conditionals are easy to calculate. All MCMC algorithms consisted of 25 000 iterations with the first 5000 being discarded as burn-in and every 20th iteration included in the final sample. For convergence diagnostics a selection of randomly chosen parameter chains was used. For all samples autocorrelations were less than 0.2 and almost independent of starting values. Parameters of the Gamma hyperpriors were set to $\gamma_a = 1$ and $\gamma_b = 10$, respectively. First order neighbourhoods were employed in the spatial applications. Apart from the undershoot parameter, posteriors for all parameters are approximately normal. To reduce working memory load only posterior means and variances for parameters were updated in each iteration, and the posterior mean was used as point estimate. For the undershoot the mode is applied, because a highly skewed posterior is expected. Activation maps are calculated by testing pixelwise whether the stimulus effect is significantly different from zero or not. For a confidence level α , this is equivalent to checking whether the α quantile covers zero or not, that is, checking $P(b_{it} > 0 | Y) \geq 1 - \alpha$, using approximate normality of the posterior. These maps are the direct Bayesian analogue of the classic activation maps. The maps in our example were thresholded at different levels, to improve qualitative comparison between approaches.

In Figure 3 an activation map ($P(b_i > 0 | Y) = 1 - 10^{-6}$) for a conventional parametric pixelwise approach is displayed. Main activation is located in the primary visual cortex in the occipital lobe, also called V1. Additionally, small activated areas in the extrastriate cortex or secondary visual areas can be found in the posterior parietal areas on both sides of the brain. Rostral to (in front of) V1, a few pixels in the precuneus also show a significant reaction to the stimulus. Primary visual areas are the first cortical relay station for general processing of all visual stimuli. The secondary or associate regions co-activated in our experiment are supposed to handle specific information about, for example, colour and shape of objects. Episodic memory, a part of the declarative memory, is assumed to be located in the precuneus.

The results of the complete spatiotemporal model (2.7) including the HRF estimation are shown in Figure 4 ($P(b_{it} > 0 | Y) \geq 1 - 10^{-5}$). Each of the respective three maps in Figure 4a represents one time point in the stimulation cycle. Time is indicated by the asterisk on the boxcar at the bottom of the figure. We confine ourselves to these three

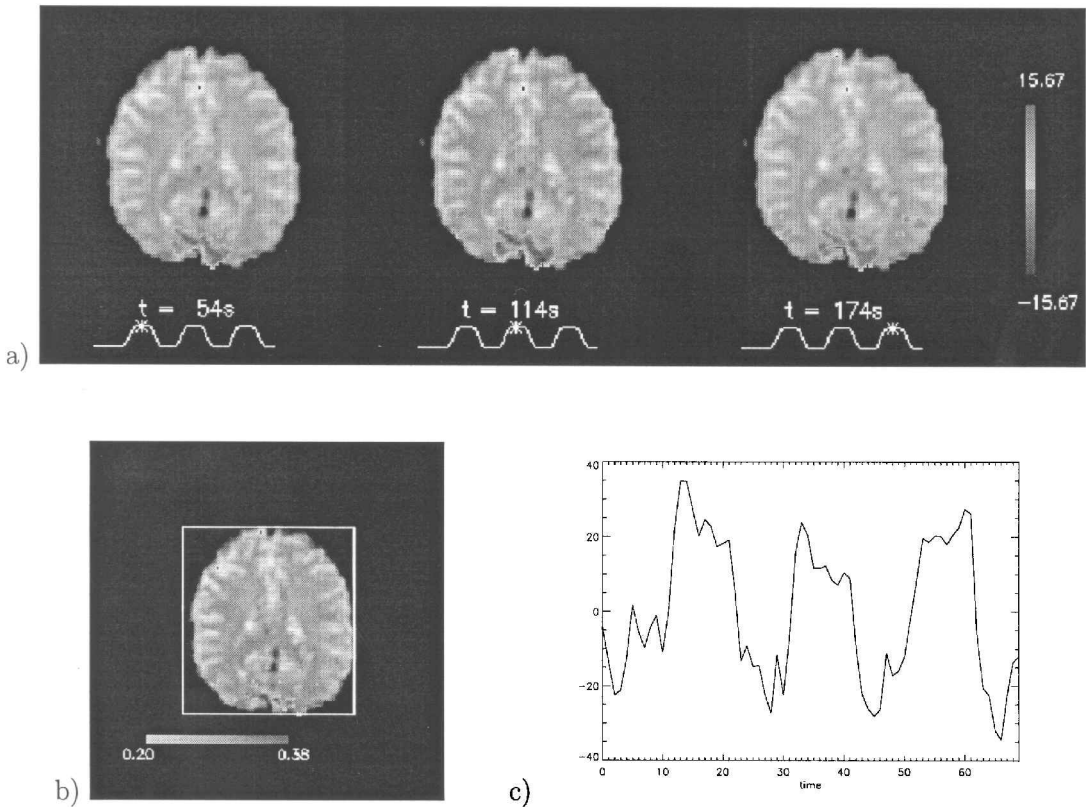


Figure 4 Activation map (a), undershoot map (b), average time course of activated pixels (c) for the spatio-temporal model, including the estimation of the hemodynamic response function

maps because activation varies only slowly in time and most interesting differences can be found between stimulation periods in this activation paradigm. The undershoot map in Figure 4b displays the undershoot parameter θ_5 thresholded at a level of 0.25. Finally, in Figure 4c the average time course of all activated pixels is plotted. Compared to the pixelwise approach, activations are estimated very consistently, especially in the primary visual cortex. In contrast, temporally resolved results reveal a considerable temporal variation of the activation in the secondary visual areas. Small fluctuations in the precuneus can be observed, too. Further, simultaneous estimation of the HRF parameters yields undershoot values greater than 0.25 located in the primary visual cortex, in good agreement with standard functional knowledge. The signal drop beyond the baseline after each activation cycle is described by the undershoot parameter θ_5 . This hemodynamic characteristic is well known to occur, especially in visual stimulations.

When analysing fMRI data of simple sensory paradigms, the main goal is usually the robust individual localization of the primary and secondary sensory cortices. The information provided by parametric methods, pixelwise or preferably spatial, is in general sufficient for this purpose. The activated areas depicted in Figures 3 and 4

mainly coincide. However, some patients (in particular older, or mentally or neurologically impaired) may have difficulties in maintaining focal brain activation during the course of the experiment (3–10 minutes). Temporal variations in alertness, attention, or vigilance can produce severe and unpredictable interferences within an fMRI experiment. At least complicating the analysis, such fluctuations moreover may also provide additional information about the processing of certain stimuli. Semiparametric approaches are thus well suited for such circumstances. These models can capture activation variations during the experiment that simple parametric models would neglect. Furthermore, as transiently activated areas can be assessed, greater sensitivity and precision can be achieved in, for example, the impaired patients. The same holds for the simultaneous modelling of hemodynamic response properties. Standard convolution methods are working robustly. However, additional insight is gained by a more flexible simultaneous modelling of the HRF, as could be shown, for example, for the stimulation paradigm dependence of the post-stimulus undershoot in visual and acoustic stimulations (Gössl *et al.*, 2001b). This approach can easily be further extended to other hemodynamic properties in question, as first applications to pharmacological MR studies proved.

In conclusion, semiparametric spatial methods are well suited for the analysis of functional MRI data. Model choice for a particular data situation has to be based on the specific objectives of each study. If the main goal is the detection of activated areas averaged over time, then simpler spatial models with time-constant activation parameters that perform sufficiently well. If dynamically changing activation effects are to be expected or of interest, then additional information can be gained by using more complex semiparametric spatiotemporal approaches, also allowing for conclusions about hemodynamic response properties.

4 Perspectives

In Section 2, we have already indicated some alternative possibilities for spatiotemporal modelling of fMRI data. To avoid oversmoothing of peaks and areas of high curvature between activated and non-activated regions, the spatial part of Gaussian priors could be replaced by robust priors as outlined in Section 2.3. The pixelwise Bayesian regression spline approach of Genovese (2000) can be extended to a spatiotemporal model by introducing a spatial prior for the B-spline coefficients, given the number and location of knots. The computational effort, however, seems enormous. A further possibility is to transfer the spatial mixture model of Hartvig and Jensen (2000) to define a prior for the activation effect b_i in a regression model by a similar spatial mixture distribution. Generally, however, it is unclear whether additional efforts in implementation and computing time are in balance with additional gain in precision and new insights.

A new perspective and challenging task in human brain mapping is the analysis of diffusion tensor imaging (DTI) data for tracking neuronal fiber pathways in the living human brain, connecting functional brain regions. Knowledge of these links is essential for a full understanding of the human brain and its neuronal circuits. The basic idea in

DTI is that within the brain water diffusion is preferentially directed along the white matter fiber tracts, and that this spatially anisotropic diffusion can be characterized by a 3×3 diffusion tensor (a symmetric, non-negative definite matrix). Diffusion tensor magnetic resonance imaging provides noisy observations of the six unknown elements of the diffusion tensor at each pixel. Recently, several proposals have been made on how to use this information contained in noisy DTI data for reconstructing fiber pathways. A basic idea is that at each pixel, the largest eigenvalue of the diffusion tensor has a similar direction as a fiber trajectory going through this pixel. Preferably, smoothed or regularized versions of the observed diffusion tensors or their largest eigenvalues are used for reconstructing fiber tracts. Currently, regularization of DTI maps is carried out in a preprocessing step, followed by tracking algorithms based on the regularized DTI maps. In Basser *et al.* (2000) the tracking is performed using differential equations, Poupon *et al.* (2000) use a complex Bayesian Markov random field prior for this purpose, resulting in considerable computational and numerical burdens. Gössl *et al.* (2002) propose a tracking algorithm based on linear state space modelling. Applying an inherent smoothness prior, this algorithm allows for a sensitive and accurate tracking of major white matter fiber bundles, as could be shown, for example, for the pyramidal tract. However, the potential of DTI for understanding the neuronal circuits, for the planning of neurosurgical interventions, or for the detection of white matter changes in psychiatric diseases is enormous. Nevertheless, the statistical analysis of diffusion tensor imaging data has just begun.

Acknowledgements

This work was supported by a grant from the German National Science Foundation (DFG), Sonderforschungsbereich 386. We thank an associate editor and a referee for valuable comments, substantially improving the presentation.

References

- Aguirre GK, Zarahn E, D'Esposito M (1998) The variability of human BOLD hemodynamic response. *NeuroImage*, **8**, 360–69.
- Aykroyd (1998) Bayesian estimation for homogeneous and inhomogeneous Gaussian random fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **20**, 533–39.
- Bandettini PA, Jesmanowicz A, Wong EC, Hyde JS (1993) Processing strategies for time-course data sets in functional MRI of the human brain. *Magnetic Resonance in Medicine*, **30**, 161–73.
- Basser PJ, Pajevic S, Pierpaoli C, Duda J, Aldroubi A (2000) In vivo fiber tractography using DT-MRI data. *Magnetic Resonance in Medicine*, **44**, 625–32.
- Besag J, York J, Mollie A (1991) Bayesian image restoration with two applications in spatial statistics (with discussion). *Annals of the Institute of Statistical Mathematics*, **43**(1), 1–59.
- Bullmore E, Brammer M, Williams SCR, Rabe-Hesketh S, Janot N, David A, Mellers J, Howard R, Sham P (1996) Statistical methods of estimation and inference for functional MR image analysis. *Magnetic Resonance in Medicine*, **35**, 261–77.
- Cohen MS (1997) Parametric analysis of fMRI data using linear systems methods. *NeuroImage*, **6**, 93–103.
- Crellin N, Hastie T, Johnstone I (1998) Statistical methods for image sequences. Technical

- report, Department of Statistics, Stanford University.
- Eilers P, Marx B (1996) Flexible smoothing using B-splines and penalized likelihood (with comments and rejoinder). *Statistical Science*, **11**, 89–121.
- Everitt BS, Bullmore ET (1999). Mixture model mapping of the brain activation in functional magnetic resonance images. *Human Brain Mapping*, **7**(1), 1–14.
- Fahrmeir L, Gössl C, Hennerfeind A (2002) Robust spatial smoothing in functional MRI. In: *Exploratory Data Analysis in Empirical Research*, Schwaiger M and Opitz O (eds). Springer Verlag, 50–7.
- Forman SD, Cohen JD, Fitzgerald M, Eddy WF, Mintun MA, Noll DC (1995) Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. *Magnetic Resonance in Medicine*, **33**, 636–47.
- Friston KJ, Holmes AP, Poline J-B, Grasby P, Williams SCR, Frackowiak RSJ, Turner R (1995) Analysis of fMRI time-series revisited. *NeuroImage*, **2**, 45–53.
- Friston KJ, Holmes AP, Worsley KJ, Poline J-B, Williams SCR, Frackowiak, RSJ (1994) Analysis of functional MRI time-series. *Human Brain Mapping*, **1**, 153–71.
- Genovese CR (2000) A Bayesian time-course model for functional magnetic resonance imaging data (with discussion). *Journal of the American Statistical Association*, **95**, 691–719.
- Gössl C, Auer DP, Fahrmeir L (2000) Dynamic models in fMRI. *Magnetic Resonance in Medicine*, **43**, 72–81.
- Gössl C, Auer DP, Fahrmeir L (2001a) Bayesian spatio-temporal inference in functional magnetic resonance imaging. *Biometrics*, **57**(2).
- Gössl C, Fahrmeir L, Auer DP (2001b) Bayesian modeling of the hemodynamic response function in BOLD fMRI. *NeuroImage*, **14**, 140–48.
- Gössl C, Fahrmeir L, Pütz B, Auer L, Auer DP (2002) Fiber tracking from DTI using linear state space models: Detectability of the pyramidal tract. *NeuroImage*, **16**, 378–88.
- Hartvig NV (2000) Parametric modelling of functional magnetic resonance imaging data. Ph.D. Thesis, University of Aarhus.
- Hartvig NV, Jensen JL (2000) Spatial mixture modeling of fMRI data. *Human Brain Mapping*, **11**, 233–48.
- Hastie T, Tibshirani R (2000) Bayesian backfitting. *Statistical Science*, **15**, 193–223.
- Kruggel F, Zysset S, von Cramon DY (2000) Nonlinear regression of functional MRI data: An item recognition task study. *NeuroImage*, **12**, 173–83.
- Lang S, Brezger A (2002) Bayesian P-splines. *Journal of Computational and Graphical Statistics*, to appear.
- Poline JB, Worsley KJ, Evans AC, Friston KJ (1997) Combining spatial extent and peak intensity to test for activations in functional imaging. *NeuroImage*, **5**, 83–96.
- Poupon C, Clark CA, Frouin V, Regis J, Bloch I, LeBihan D, Mangin J-F (2000) Regularization of diffusion-based direction maps for the tracking of brain white matter fascicles. *NeuroImage*, **12**, 184–95.
- Purdon PL, Weiskoff RM (1998) Effect of temporal autocorrelation due to physiological noise and stimulus paradigm on voxel-level false-positive rates in fMRI. *Human Brain Mapping*, **6**, 239–49.
- Rajapakse JC, Kruggel F, Maisog JM, von Cramon DY (1998) Modeling hemodynamic response for analysis of functional MRI time-series. *Human Brain Mapping*, **6**, 283–300.
- Zarahn E, Aguirre GK, D'Esposito M (1997) Empirical analyses of BOLD fMRI Statistics. *NeuroImage*, **5**, 179–97.