Appl. Statist. (2007)
56, *Part* 3, *pp.* 327–345

Adaptive Gaussian Markov random fields with applications in human brain mapping

A. Brezger, L. Fahrmeir and A. Hennerfeind

Ludwig-Maximilians-Universität München, Germany

[Received December 2005. Revised February 2007]

Summary. Functional magnetic resonance imaging has become a standard technology in human brain mapping. Analyses of the massive spatiotemporal functional magnetic resonance imaging data sets often focus on parametric or non-parametric modelling of the temporal component, whereas spatial smoothing is based on Gaussian kernels or random fields. A weakness of Gaussian spatial smoothing is underestimation of activation peaks or blurring of high curvature transitions between activated and non-activated regions of the brain. To improve spatial adaptivity, we introduce a class of inhomogeneous Markov random fields with stochastic interaction weights in a space-varying coefficient model. For given weights, the random field is conditionally Gaussian, but marginally it is non-Gaussian. Fully Bayesian inference, including estimation of weights and variance parameters, can be carried out through efficient Markov chain Monte Carlo simulation. Although motivated by the analysis of functional magnetic resonance imaging data, the methodological development is general and can also be used for spatial smoothing and regression analysis of areal data on irregular lattices. An application to stylized artificial data and to real functional magnetic resonance imaging data from a visual stimulation experiment demonstrates the performance of our approach in comparison with Gaussian and robustified non-Gaussian Markov random-field models.

Keywords: Adaptive weights; Human brain mapping; Inhomogeneous Markov random fields; Markov chain Monte Carlo methods; Space-varying coefficient models; Spatiotemporal modelling

1. Introduction

The methodological development of this paper is motivated by functional magnetic resonance imaging (FMRI), a current standard technology in human brain mapping and neuroscience. One of the major aims of this technique is the non-invasive, individual localization of functional brain areas, i.e. the detection of regions that are activated by sensory, motor and cognitive functions. Main objectives of FMRI studies are to improve knowledge about the complex structure of neuronal circuits to guide neurosurgical interventions in brain tumours and to reveal disease-specific alterations in neuronal processing. In FMRI experiments, a subject is exposed to controlled external stimuli. A local increase of neural activity is indicated by a local increase of blood oxygenation in activated areas, and this blood oxygenation level dependent effect can be visualized by FMRI. In classical experiments the stimulus is a boxcar paradigm, i.e. a sequence of *off* and *on* periods. The scanner records images of several slices of the brain. Each slice is about 5 mm thick and consists of 64×59 pixels. Slices usually have a distance of several millimetres, and their images are obtained sequentially in time. Therefore, slices are often analysed separately. For each pixel of a slice, an FMRI experiment with a boxcar stimulus generates an

Address for correspondence: A. Hennerfeind, Institut für Statistik, Ludwig-Maximilians-Universität München, Ludwigstrasse 33, 80539 München, Germany.
E-mail: andreah@stat.uni-muenchen.de

 0035–9254/07/56327

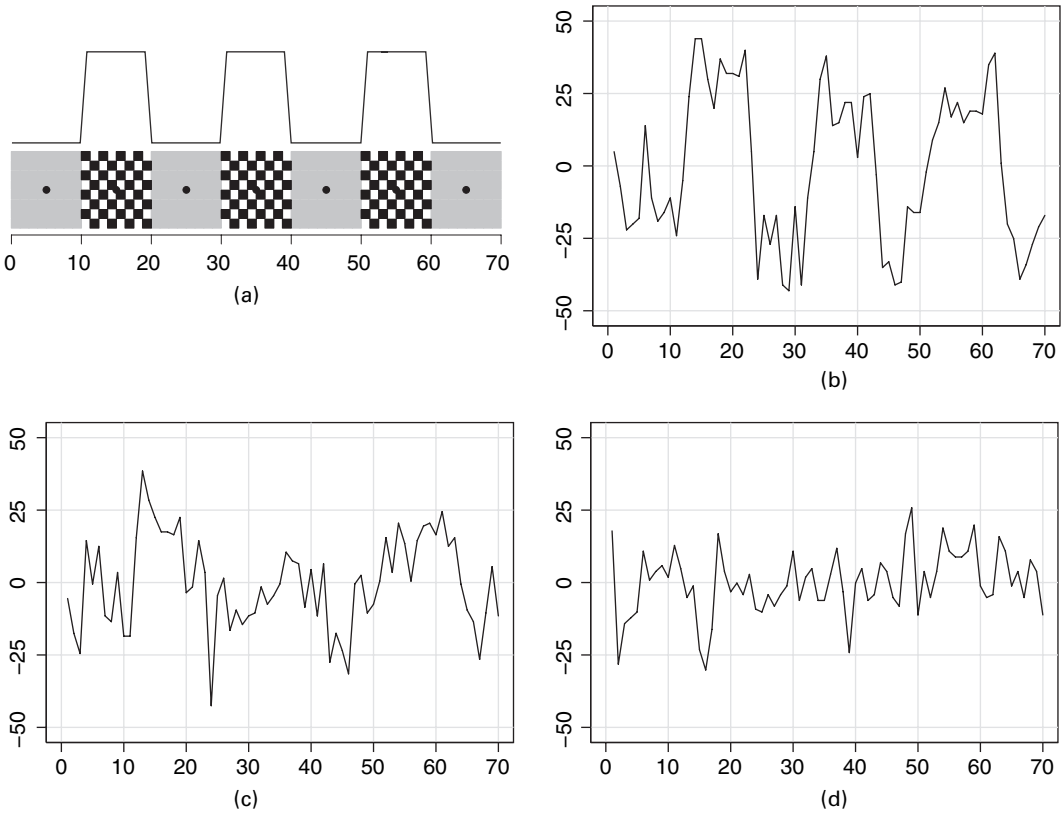

Fig. 1. Visual FMRI: (a) an 8-Hz flickering rectangular checker-board (on) is presented to the subject, alternating every 10 time units (1 unit = 3 s) with a uniformly dark background and a fixation point (off); an experiment consists of four off periods and three on periods; addtionally shown are centred representative MR signal time courses (in arbitrary units) from (b) strongly, (c) weakly and (d) non-activated pixels

MR signal time series, with an increase during the on periods compared with the control or rest condition off. Our application in Section 5 analyses data for one slice from a visual experiment. Fig. 1 shows the boxcar stimulus and three MR time series of length 70, observed at three pixels, which are selected from the centre of the activated region (Fig. 1(b)), near to its boundary (Fig. 1(c)) and from a non-activated region (Fig. 1(d)). Obviously, the activation effect of the stimulus on the MR signal is high in activated areas and is not present in non-activated areas.

As outlined in more detail in Section 2, the standard approach of statistical parametric mapping (see Friston *et al.* (1995)) for assessing brain activity employs separate parametric time series regression models at each pixel, with the MR signal as response and a transformed version of the stimulus as the regressor of primary interest. The value of the corresponding regression coefficient is considered as the 'intensity' or 'amplitude' of activation at the pixel. Spatial correlation between pixels is taken into account by stationary Gaussian random fields in a post-processing step. To allow for a more flexible modelling of time trends and time-varying activation effects, Genovese (2000) and Gössl *et al.* (2000) suggested semiparametric time series regression based on regression splines and state space models respectively. Again, however, these semiparametric models are fitted separately for each pixel of a slice. To take into account spatial correlation between pixels and to combine the pixelwise regressions in a joint model, Gössl *et al.* (2001) and Fahrmeir and Gössl (2002) suggested Bayesian space-varying coefficient models based on

Gaussian Markov random fields (MRFs). A potential weakness of Gaussian random-field priors is underestimation of peaks and smoothing over edges, discontinuities or unsmooth parts of underlying functions. To overcome these problems, a conceptually different approach based on spatial Bayesian variable selection has been developed in Smith *et al.* (2003), but without a data-driven method for selection of the degree of spatial interaction parameters. Smith and Fahrmeir (2007) have shown how this issue can be solved by numerical approximation.

In this paper, we follow previous ideas and combine separate pixelwise regressions through a Bayesian space-varying coefficient model. To enhance spatial adaptivity for the activation effects array, we introduce in Section 3 a new class of inhomogeneous or compound MRF priors where the interaction weights, determining the degree of spatial correlations between adjacent pixels, are allowed to vary stochastically as continuous, non-negative random variables. Conditional on these weights, the prior is an intrinsic Gaussian MRF, but marginally it is a non-Gaussian MRF with edge preserving properties. All model parameters, including the adaptive interaction weights, can be estimated in a fully Bayesian setting by using Markov chain Manto Carlo (MCMC) techniques. As a key feature we show how to compute the normalizing constant of the intrinsic Gaussian MRF, depending on the weights, in correct and computationally efficient manner. This allows us to combine efficient sampling from Gaussian MRFs as described in Rue (2001) and Rue and Held (2005) with Metropolis–Hastings (MH) steps for updating interaction weights (Section 4).

Although this methodological development was motivated through challenges in FMRI data analysis, it is general and can be useful in other settings such as spatial smoothing and regression analysis of geographic data or boundary analysis, which is also referred to as areal wombling (Lu *et al.*, 2005). In Section 5, we apply our approach to a stylized synthetic spatiotemporal data set and to FMRI data from a visual stimulation experiment, and we compare its performance with results that were obtained with commonly used (unweighted) Gaussian MRFs as well as with spatially more adaptive Laplace and Geman–Reynolds priors. We conclude that our spatially adaptive modelling approach performs well and should be a promising alternative in other applications also.

2. Space-varying coefficient models in functional magnetic resonance imaging

We start this section with a brief description of the standard approach for analysing FMRI data, where FMRI time series $\{y_{it}, t = 1, \ldots, T\}$ at each pixel or voxel i are fitted separately by a time series regression model. Figs 1(b), 1(c) and 1(d) show FMRI time series for three selected pixels together with an external visual on–off stimulus $\{x_t, t = 1, \ldots, T\}$. In its basic form, such a regression model for pixel i is defined through the measurement model

$$y_{it} = \mathbf{u}_t'\boldsymbol{\alpha}_i + z_{it}\beta_i + \varepsilon_{it}, \qquad \varepsilon_{it} \sim N(0, \sigma_i^2) \tag{1}$$

for $t = 1, \ldots, T$. The first term $\mathbf{u}_t'\boldsymbol{\alpha}_i$ models a smooth base-line trend $f_i(t) = \mathbf{u}_t'\boldsymbol{\alpha}_i$, where the vector $\mathbf{u}_t$ consists of a few simple basis functions, e.g. a moderate number of B-splines or some terms of a Fourier expansion, evaluated at $t = 1, \ldots, T$. The second term, which is sometimes called the activation profile, is the product of the 'activation effect' β_i at pixel i and the covariate z_{it}, which is a smoothed and delayed version of the original on–off stimulus $x_t, t = 1, \ldots, T$, that is visualized in Fig. 1(a).

The transformation from x_t to z_{it} takes into account that the cerebral blood flow, the source of the MR signal, increases only approximately 6–8 s after the onset of the stimulus, and that flow responses do not occur suddenly, but more continuously and delayed. We shall use the most common transformation, which is obtained by a delayed convolution with a so-called

haemodynamic response function $h(s;\theta)$, i.e.

$$z_{it} = \sum_{s=0}^{t-d_i} h(s;\theta_i) x_{t-d_i-s}. \tag{2}$$

Usually, Poisson ($\mathrm{Po}(\lambda_i)$) or gamma ($\mathrm{Ga}(\lambda_i, u_i)$) densities are chosen for h. The parameters $\theta_i = \lambda_i$ or $\boldsymbol{\theta}_i = (\lambda_i, u_i)'$ and the time lag d_i generally depend on pixel i and are estimated in a pilot step; see Gössl *et al.* (2001).

The idea is that the (estimated) activation effect $\hat{\beta}_i$ that is generated by the (transformed) stimulus should be large in strongly activated pixels (Fig. 1(b)), medium for weakly activated pixels (Fig. 1(c)) and close to zero ('non-significant') for non-activated pixels (Fig. 1(d)). The estimates $\{\hat{\beta}_i, i = 1, \ldots, I\}$ for all pixels in a slice can be represented by an 'activation map' or, after standardization, by a 't-map' $\{t_i = \hat{\beta}_i / \mathrm{se}(\hat{\beta}_i), i = 1, \ldots, I\}$. In the standard statistical parametric mapping approach (see Friston *et al.* (1995)), spatial correlation that is induced by neighbouring pixels can be taken into account by applying Gaussian random-field theory to the t-map, leading to an adjusted t-map. An adaptive smoothing procedure for pixelwise activation maps with good edge preserving properties has been suggested in recent work by Tabelow *et al.* (2005).

The basic model (1) has been refined by relaxing assumptions on the time series structure, e.g. by introducing autoregressive error terms, or using regression splines with adaptive knot selection (Genovese, 2000) and state space models (Gössl *et al.*, 2000), for non-parametric or semiparametric estimation of base-line trends $f_i(t)$ and even time-varying activation effects. Still, however, FMRI data are fitted separately at each pixel i on the basis of (seemingly) unrelated regression or time series models in these approaches.

In this paper we look at the measurement model (1) as a *joint model* for all pixels $i = 1, \ldots, I$. Then the coefficients $\{\boldsymbol{\alpha}_i, i = 1, \ldots, I\}$ and $\{\beta_i, i = 1, \ldots, I\}$ are spatially varying over the grid of pixels in a certain slice. To allow for activation effects also to vary smoothly with time during the FMRI experiment, we slightly extend model (1) to

$$\begin{aligned} y_{it} &= f_i(t) + b_i(t) z_{it} + \varepsilon_{it} \\ &= \mathbf{u}_t' \boldsymbol{\alpha}_i + z_{it} \mathbf{v}_t' \boldsymbol{\beta}_i + \varepsilon_{it}, \qquad \varepsilon_{it} \sim N(0, \sigma_i^2), \end{aligned} \tag{3}$$

for $t = 1, \ldots, T$ and all pixels $i = 1, \ldots, I$. In addition to the time-varying base-line trend $f_i(t) = \mathbf{u}_t' \boldsymbol{\alpha}_i$, model (3) also admits a time-varying activation effect $b_i(t) = \mathbf{v}_t' \boldsymbol{\beta}_i$ at pixel i. Similarly to $\mathbf{u}_t$, the components of $\mathbf{v}_t$ are simple basis functions such as Fourier terms or B-splines, evaluated at $t = 1, \ldots, T$. Model (3) is a joint spatiotemporal regression model for FMRI data, with *space-varying activation coefficients* $\{\boldsymbol{\beta}_i, i = 1, \ldots, I\}$. In Gössl *et al.* (2001) a hierarchy of spatial and spatiotemporal Bayesian models based on homogeneous Gaussian MRF priors for space-varying coefficients has been presented and applied to FMRI data. A drawback of Gaussian priors is that they tend to oversmooth peaks and to blur edges or areas of high curvature between activated and non-activated areas. To improve spatial adaptivity, we suggest MRF priors with stochastically varying interaction weights for adjacent pixels. Conditional on the weights, these are intrinsic Gaussian MRFs, but marginally they are non-Gaussian MRF priors; see the following section.

3. Adaptive Gaussian Markov random-field priors

In our application to FMRI data, we shall assume independent MRF priors for the components β_{ik} of the space-varying coefficient vectors $\boldsymbol{\beta}_i = (\beta_{i1}, \ldots, \beta_{il})$, $i = 1, \ldots, I$, in model (3). Therefore, we focus on a scalar component β_i, $i = 1, \ldots, I$, as in the simpler model (1) with a time

constant activation effect. Possible extensions to multivariate MRFs will be shortly discussed at the end of the section.

In what follows, we consider adaptive versions of intrinsic Gaussian MRFs for the random field $\boldsymbol{\beta} = \{\beta_i, i = 1, \ldots, I\}$. Its full conditional distributions have the form

$$\beta_i | \beta_{i \neq j}, \tau, \mathbf{w} \sim N\left(\sum_{i \sim j} \frac{w_{ij}\beta_j}{w_{i+}}, \frac{\tau^2}{w_{i+}} \right), \qquad i = 1, \ldots, I, \tag{4}$$

where $w_{i+} = \Sigma_j w_{ij}$, and the summation extends over neighbours $j \sim i$ of i on a regular or irregular lattice. In our application to FMRI, the lattice of pixels is regular, with the four next pixels as neighbours. Our methodological development, however, also covers irregular lattices and other neighbourhood structures. The weights $w_{ij} > 0$ are symmetric ($w_{ij} = w_{ji}$) and measure 'similarity' of neighbours (i, j), $i \neq j$, or strength of interactions between β_i and β_j, $i \neq j$. The most common choice for the weights on a regular lattice is $w_{ij} = 1$, if (i, j) are neighbours (and $w_{ij} = 0$ otherwise). Valid full conditional distributions (4) are obtained for many other specifications. In epidemiological or geographic applications, for example, the weights might be chosen to be inversely proportional to the distance between centroids of neighbouring regions; see Rue and Held (2005). In any case, the weights are specified in advance or computed in a preprocessing step. The variance (or inverse precision) τ^2 acts as a global smoothing parameter, which is usually estimated from the data together with the MRF $\boldsymbol{\beta}$.

The joint distribution for $\boldsymbol{\beta}$ that is derived from the full conditional distribution is a partially improper Gaussian distribution given by

$$p(\boldsymbol{\beta} | \tau^2, \mathbf{w}) \propto \left(\prod_{i=2}^{I} \lambda_i \right)^{1/2} \exp\left(-\frac{1}{2\tau^2} \boldsymbol{\beta}' \mathbf{K} \boldsymbol{\beta} \right). \tag{5}$$

The precision or penalty matrix has elements

$$k_{ij} = \begin{cases} w_{i+}, & i = j, \\ -w_{ij}, & i \sim j, \\ 0 & \text{otherwise}, \end{cases}$$

and is singular, so expression (5) is a singular multivariate Gaussian density. Its normalizing constant (up to a factor not depending on the weights) is given by the square root of the product of the non-zero eigenvalues of the penalty matrix. If the weights are assumed to be given or known, this normalizing constant is irrelevant in Bayesian computations and is mostly omitted in the proportionality (5). Is it essential, however, for correct Bayesian posterior inference involving the adaptive Gaussian MFRs that are considered in what follows and used in our application to FMRI data.

Adaptive modifications of the intrinsic Gaussian MRF model (5) can be obtained by allowing the interaction weights w_{ij} to vary stochastically in a further stage of the hierarchical model. Two main concepts may be distinguished. Both are inspired by the literature as well as our own work in the one-dimensional framework of state space modelling for time series and spatially adaptive non-parametric regression as outlined in what follows. Let

$$y(t) = f(t) + \varepsilon(t), \qquad t = 1, \ldots, n,$$

be the classical time series or non-parametric regression model, where $f(t)$ is an unknown time trend or regression curve to be estimated from the data, and the errors $\varepsilon(t) \sim N(0, \sigma^2)$ are independent and identically distributed (IID). A simple state space approach, which can also

be interpreted as a Bayesian approach to non-parametric regression, assumes a random-walk prior of first or second order for the sequence $\{f(t), t = 1, \ldots, n\}$, i.e.

$$f(t) = f(t-1) + u(t)$$

or

$$f(t) = 2f(t-1) - f(t-2) + u(t).$$

The classical assumption for the random-walk errors $u(t)$ is that they are independent and $N(0, \tau^2)$ distributed. The variance τ^2 acts as a global smoothing parameter that is estimated jointly with the trend function from the data. The random-walk priors are one-dimensional Gaussion MRFs with first- or second-order neighbourhood structure; see Rue and Held (2005), chapter 3.

The first spatially adaptive approach for fitting time trends with jumps or abrupt changes in level and trend was developed by Carter and Kohn (1996) by assuming (conditionally) independent random errors

$$u_t | w_t, \tau^2 \sim N\left(0, \frac{\tau^2}{w_t}\right)$$

with IID gamma priors $w_t \sim \text{Ga}(\nu/2, \nu/2)$. Thus, the marginal distribution of the errors is a (heavy-tailed) Student distribution with ν degrees of freedom. The case $\nu = 1$ or a Cauchy distribution is of special interest as a robust prior.

In the second, conceptually different, approach Lang *et al.* (2002) introduced spatially correlated smoothness parameters, by assuming random-walk models of first or second order for log-variances $v_t = \log(w_t)$, or, in other words, Gaussion MRF priors for v_t in a second hierarchy. They compared spatial adaptivity of the IID gamma prior approach and their random-walk prior approach in simulation studies based on stylized jump functions or highly oscillating functions (the 'Doppler' curve, which is popular in the wavelet shrinkage literature). The conclusions are as follows. For jump functions, the IID gamma priors slightly outperform the random-walk prior approach in terms of mean-square error, and they are computationally much less demanding with respect to computing times and implementation. For Doppler curves with low spatial variability both approaches still have comparable performance, whereas varying variances are preferable for Doppler curves with high spatial variability. Analogous findings have been obtained by Lang and Brezger (2004) and Baladandayuthapani *et al.* (2005) in the related context of spatially adaptive Bayesian penalized regression splines.

Conceptionally, both approaches can be extended to define adaptive MRFs in higher dimensions. Methodological and computational demands are considerably increasing, however, even with the simpler IID gamma priors. The main reason why inference becomes more difficult is the normalizing constant in expression (5), involving the non-zero eigenvalues of the penalty matrix $\mathbf{K} = \mathbf{K}(\mathbf{w})$, and therefore depending on the (high dimensional) vector $\mathbf{w}$ of weights in a complicated non-linear way. Before presenting both approaches in more detail, we state a lemma which helps to circumvent repeated expensive calculation of eigenvalues in MH sampling steps and should be useful on its own in related settings.

Lemma 1. Consider two vectors $\mathbf{w}$ and $\mathbf{w}^*$ of weights w_{ij} and w_{ij}^* respectively, and denote corresponding penalty matrices and their eigenvalues evaluated at $\mathbf{w}$ and $\mathbf{w}^*$ by $\mathbf{K}$ and $\mathbf{K}^*$, and λ_i and λ_i^* respectively. Then

$$\prod_{i=2}^{I} \lambda_i^* \Big/ \prod_{i=2}^{I} \lambda_i = |\mathbf{K}_{11}^*|/|\mathbf{K}_{11}|, \tag{6}$$

where $|\mathbf{K}_{11}^*|$ and $|\mathbf{K}_{11}|$ are the determinants of the submatrices of $\mathbf{K}^*$ and $\mathbf{K}$ after deleting the last row and the last column.

The advantage arising from lemma 1 is that, instead of an expensive computation of eigenvalues of order $O(n^3)$, the ratio can be obtained by the computationally much more efficient Cholesky decomposition of band matrices, which is of order $O(n)$.

The following proof of lemma 1 shows that equation (6) remains true, if we delete any row i and column i, $i = 1, \ldots, I$. We follow Brezger (2005), where equation (6) has been shown for regular lattices. The spectral decomposition of $\mathbf{K}$ gives

$$\begin{aligned}\mathbf{K} &= \mathbf{\Gamma}\mathbf{\Lambda}\mathbf{\Gamma}' \\ &= (\mathbf{\Gamma}_1\mathbf{\Gamma}_2)\begin{pmatrix}\mathbf{\Lambda}_{11} & 0\\ 0 & 0\end{pmatrix}\begin{pmatrix}\mathbf{\Gamma}_1'\\ \mathbf{\Gamma}_2'\end{pmatrix} \\ &= \mathbf{\Gamma}_1\mathbf{\Lambda}_{11}\mathbf{\Gamma}_1'\end{aligned}$$

where $\mathbf{K}$ and $\mathbf{\Gamma}$ are partitioned as

$$\mathbf{K} = \begin{pmatrix}\mathbf{K}_{11} & \mathbf{K}_{12}\\ \mathbf{K}_{21} & k_{22}\end{pmatrix},$$

$$\mathbf{\Gamma} = (\mathbf{\Gamma}_1\mathbf{\Gamma}_2) = \begin{pmatrix}\mathbf{\Gamma}_{11} & \mathbf{\Gamma}_{12}\\ \mathbf{\Gamma}_{21} & \gamma_{22}\end{pmatrix},$$

k_{22} and γ_{22} are scalars and $\mathbf{\Gamma}_2$ is a column vector. The matrix $\mathbf{\Gamma}_1$ contains the eigenvectors of $\mathbf{K}$ corresponding to the non-zero eigenvalues of $\mathbf{K}$ and $|\mathbf{\Lambda}_{11}| = \prod_{i=1}^{I-1} \lambda_i$ is the product of the non-zero eigenvalues of $\mathbf{K}$. Furthermore, $\mathbf{\Gamma}$ is orthonormal and $\mathbf{\Gamma}'\mathbf{\Gamma} = \mathbf{\Gamma}\mathbf{\Gamma}' = \mathbf{I}$, i.e. $\mathbf{\Gamma}^{-1} = \mathbf{\Gamma}'$.

For a proof of equation (6) the following statements are required.

(a) $|\mathbf{K}_{11}| = |\mathbf{\Gamma}_{11}|^2|\mathbf{\Lambda}_{11}|$, and $|\mathbf{K}_{11}| > 0$. This follows from $\mathbf{K}_{11} = \mathbf{\Gamma}_{11}\mathbf{\Lambda}_{11}\mathbf{\Gamma}_{11}'$, and $|\mathbf{\Gamma}_{11}| > 0$ and $|\mathbf{\Lambda}_{11}| > 0$.

(b) $|\mathbf{\Gamma}_{11}|^2 = \gamma_{22}^2$. To prove this, since $|\mathbf{\Gamma}_{11}| > 0$, it holds that

$$|\mathbf{\Gamma}| = |\mathbf{\Gamma}_{11}|(\gamma_{22} - \mathbf{\Gamma}_{21}\mathbf{\Gamma}_{11}^{-1}\mathbf{\Gamma}_{12}) = \frac{1}{\gamma_{22}}|\mathbf{\Gamma}_{11}|. \tag{7}$$

The second equality can be derived from the fact that, if $\mathbf{A}$ is a non-singular quadratic matrix and

$$\mathbf{A} = \begin{pmatrix}\mathbf{A}_{11} & \mathbf{A}_{21}\\ \mathbf{A}_{12} & \mathbf{A}_{22}\end{pmatrix},$$

and

$$\mathbf{A}^{-1} = \mathbf{B} = \begin{pmatrix}\mathbf{B}_{11} & \mathbf{B}_{21}\\ \mathbf{B}_{12} & \mathbf{B}_{22}\end{pmatrix},$$

then

$$\mathbf{A}_{22}^{-1} = \mathbf{B}_{22} - \mathbf{B}_{21}\mathbf{B}_{11}^{-1}\mathbf{B}_{12},$$

where

$$\mathbf{B} = \mathbf{\Gamma}$$

and

$$\mathbf{A} = \mathbf{\Gamma}^{-1} = \mathbf{\Gamma}' = \begin{pmatrix} \mathbf{\Gamma}'_{11} & \mathbf{\Gamma}'_{21} \\ \mathbf{\Gamma}'_{12} & \gamma_{22} \end{pmatrix}.$$

From equation (7) we obtain that $|\mathbf{\Gamma}_{11}|^2 = \gamma_{22}^2$, since the orthonormality of $\mathbf{\Gamma}$ implies that $|\mathbf{\Gamma}|^2 = 1$.

(c) $\gamma_{22} = 1/\sqrt{I}$. To prove this we note that the row sums of $\mathbf{K}$ equal 0, i.e. $\mathbf{K}(1, \ldots, 1)' = \mathbf{\Gamma}_1 \mathbf{\Lambda}_{11} \mathbf{\Gamma}'_1 (1, \ldots, 1)' = 0$. Since the columns of $\mathbf{\Gamma}_1 \mathbf{\Lambda}_{11}$ are linearly independent, it follows that $\mathbf{\Gamma}'_1 (1, \ldots, 1)' = 0$. This means that the elements of each column of $\mathbf{\Gamma}_1$ sum to 0 and therefore any vector with constant elements is orthogonal to all columns of $\mathbf{\Gamma}_1$. Since $\mathbf{\Gamma}_2$ must be orthogonal to all columns of $\mathbf{\Gamma}_1$, i.e. $\mathbf{\Gamma}'_1 \mathbf{\Gamma}_2 = 0$, and additionally fulfil $\mathbf{\Gamma}'_2 \mathbf{\Gamma}_2 = 1$, it follows that $\mathbf{\Gamma}_2 = (1/\sqrt{I}, \ldots, 1/\sqrt{I})'$ and thus $\gamma_{22} = 1/\sqrt{I}$.

Now, combining properties (a)–(c) gives

$$|\mathbf{\Lambda}_{11}| = I |\mathbf{K}_{11}|,$$

and statement (6) follows immediately.

Adaptive Gaussian MRFs can now be defined by analogy with the one-dimensional situation that was discussed before. In the simpler formulation, we specify the weights to be IID random variables following a gamma hyperprior

$$p(w_{ij}) \propto w_{ij}^{\nu/2-1} \exp\left(-w_{ij} \frac{\nu}{2}\right); \tag{8}$$

in brief $w_{ij} \sim \text{Ga}(\nu/2, \nu/2)$, with small degrees of freedom, in particular $\nu = 1$. The resulting marginal or compound MRF prior $p(\beta|\tau^2)$ is non-Gaussian and does not admit a simple analytic form. This approach allows varying strength of interactions between neighbouring sites $i \sim j$, but without spatial prior structure. In our FMRI application it gives 'additional flexibility when pixel i is near the border' of an activation area, where some neighbours $j \sim i$ have similar activation effects whereas others may be only weakly or not activated.

The second, more complex, formulation is to assume that the (log-) weights w_{ij} vary spatially as well. Following the ideas that were discussed for state space time series models, we could assume that log-weights $v_{ij} = \log(w_{ij})$ follow another Gaussian MRF model

$$p(\mathbf{v}|\phi) \propto \exp\left(-\frac{1}{2\phi} \mathbf{v}' \mathbf{L} \mathbf{v}\right) \tag{9}$$

in a second hierarchy, but now on the graph of links (edges) between pixels or regions (nodes) of the original graph. Knorr-Held (2003) pointed out some potential problems with such a formulation, due to the increased complexity: the edge neighbourhood structure should encourage the model to favour continuous boundaries, but there are several alternative ways to achieve this. For example, edge segments on a regular lattice may be defined as adjacent if they are connected to one another, and for irregular lattices we could consider links as adjacent if they share common regions. In any case, the graph of edges or links may become rather large. Because of limited information in the data care needs to be taken when specifying the MRF prior for links; for example, it may be necessary to use a proper MRF prior to guarantee propriety of posteriors. Furthermore, efficient MH algorithms must be developed for updating the weights. We outline possibilities for MH algorithms in Section 4.

Similar problems arise in Bayesian *areal wombling* via adjacency modelling; see for example

Ma *et al.* (2005) and Lu *et al.* (2005). Here, weights w_{ij} are modelled as 0–1-variables, with $w_{ij}=1$ if an edge between i and j is present and $w_{ij}=0$ if not, i.e.

$$w_{ij}|p_{ij} \sim \text{Bernoulli}(p_{ij}), \tag{10}$$

where $\log\{p_{ij}/(1-p_{ij})\}$ may follow a Gaussian MRF in a second hierarchy. Lemma 1 should also be useful in this framework.

For our application to FMRI data, we prefer the simpler IID gamma prior formulation. From previous analyses based on non-adaptive Gaussian MRFs it is known that activation surfaces are closer to two-dimensional 'block functions' than to highly oscillating functions. On the basis of our experience in the one-dimensional case that was outlined at the beginning of the section, and in view of the open potential problems with an MRF prior for the weights, we therefore decided to rely on the simpler model.

We end the section with some remarks on possible alternatives and extensions. Intrinsic Gaussian MRFs are a special case of pairwise interaction MRFs

$$p(\boldsymbol{\beta}|\tau) \propto \exp\left\{ -\sum_{i\sim j} \Phi\left(\frac{\beta_i-\beta_j}{\tau}\right) \right\}, \tag{11}$$

where τ is a scale or variance parameter and Φ is symmetric with $\Phi(u)=\Phi(-u)$. The Laplace prior with $\Phi(u)=|u|$ is considered to have improved edge preserving properties as well as the Geman–Reynolds prior

$$\Phi(u) = -\lambda/(1+|u|^p), \tag{12}$$

with $p=2$ or $p=1$, and λ as a tuning parameter, which was introduced by Geman and Reynolds (1992). A practical problem with this prior is the appropriate choice of hyperparameters by data-driven methods. The normalizing constant is analytically intractable, making inclusion in a fully Bayesian MCMC algorithm difficult.

In our application to FMRI, we compare results that were obtained with adaptive Gaussian MRFs with results with Laplace and Geman–Reynolds priors in the simpler setting of time constant effects β_i in model (1). We avoided precomputing normalizing constants for the latter priors but selected the tuning parameters p, τ and λ carefully, on the basis of simulation results as well as on the data at hand.

In our application to FMRI, we stick to the assumption that coefficients $\beta_{i1}, \ldots, \beta_{ik}$ are independent *a priori*. This is quite a natural assumption because our 'covariates' in the space-varying models are (orthogonal) basis functions from a Fourier expansion. This situation corresponds to Bayesian shrinkage methods, such as ridge regression, in the situation of one ($I=1$) regression model with a ridge penalty on regression coefficients.

In other applications it may be reasonable to assume that regression coefficients are correlated *a priori*. Then adaptive versions of multivariate Gaussian MRFs as suggested by Carlin and Banerjee (2003), Jin *et al.* (2005) and, in the context of space-varying coefficient models, Assunção (2003) appear to be a promising approach. A closer look at adaptive multivariate Gaussian MRF priors with stochastically varying weights reveals drastically increasing, challenging computational problems in constructing efficient MH algorithms.

4. Posterior inference for space-varying coefficient models

Bayesian specification of two-dimensional space-varying coefficient models of the form (3) is completed by specifying priors for the remaining parameters. In principle (adaptive) Gaussian MRFs could also be chosen for the base-line parameters $\boldsymbol{\alpha}_i = (\alpha_{i1}, \ldots, \alpha_{im})$. However, because

the focus is on activation effects $\boldsymbol{\beta}_i$, we only assign separate, independent diffuse priors $p(\boldsymbol{\alpha}_i) \propto$ constant or highly dispersed normal priors. For the observation variances σ_i^2, we follow the common simple choice of weakly informative, independent inverse gamma priors $\sigma_i^2 \sim \text{IG}(a, b)$. As for the weights w_{ij}, we might consider a spatial Gaussian MRF prior for the field of log-variances $\theta_i = \log(\sigma_i^2)$. Again, modelling and computational problems are drastically increased, because MH algorithms with block updating will be needed instead of simply drawing from inverse gamma full conditionals. Therefore, we choose the simpler formulation with IID inverse gamma variances. Finally, we also assume weakly informative inverse gamma priors IG(c, d) for the variance parameter τ_k^2 of the (adaptive) Gaussian MRFs for the activations effects $\boldsymbol{\beta}_{(k)} = (\beta_{1k}, \ldots, \beta_{Ik})'$.

Gathering parameters in vectors $\boldsymbol{\alpha} = (\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_I)$, $\boldsymbol{\beta} = (\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_I)$, where $\boldsymbol{\beta}_i = \{\beta_{ik}, k = 1, \ldots, l\}$, $\boldsymbol{\sigma}^2 = (\sigma_1^2, \ldots, \sigma_I^2)$, $\boldsymbol{\tau}^2 = (\tau_1^2, \ldots, \tau_l^2)$ and $\mathbf{w} = (w_{ij}, i \sim j)$, and observations in $\mathbf{y} = (y_{it}, i = 1, \ldots, I, t = 1, \ldots, T)$, fully Bayesian inference is based on the posterior

$$p(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\sigma}^2, \boldsymbol{\tau}^2, \mathbf{w}|\mathbf{y}) \propto L(\mathbf{y}|\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\sigma}^2)\ p(\boldsymbol{\alpha})\ p(\boldsymbol{\beta}|\boldsymbol{\tau}^2, \mathbf{w})\ p(\boldsymbol{\sigma}^2)\ p(\boldsymbol{\tau}^2)\ p(\mathbf{w}).$$

The likelihood $L(\mathbf{y}|\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\sigma}^2)$ is determined by the observation model and the other factors by the priors, together with conditional independence assumptions.

Inference is performed by MCMC simulation through repeated drawings from univariate or multivariate full conditionals. The general strategy is as follows.

Step 1: draw the blocks $\boldsymbol{\alpha}_{(k)} = \{\alpha_{1k}, \ldots, \alpha_{Ik}\}$ from the Gaussian full conditionals, $k = 1, \ldots, m$.
Step 2: draw the blocks $\boldsymbol{\beta}_{(k)} = \{\beta_{1k}, \ldots, \beta_{Ik}\}$ from the (multivariate) Gaussian full conditionals, given current iterates for the weights $\mathbf{w}$, $k = 1, \ldots, l$.
Step 3: draw the weights w_{ij} via MH steps that are described below.
Step 4: draw the variance parameters $\boldsymbol{\sigma}^2$ and the hyperparameters $\boldsymbol{\tau}^2$ from their corresponding inverse gamma full conditionals.

Given the current iterates for the weights in the MCMC algorithm of Section 3, steps 1, 2 and 4 are in complete analogy with estimation in spatial models with latent Gaussian MFRs; see for example Lang and Brezger (2004) and Rue and Held (2005). Thereby we make efficient use of sparse matrix operations for block updating in step 2. In what follows we focus on details of the MH step 3.

For the IID gamma prior formulation (8), the full conditional for the weight $w_{ij}^{(k)}$ is

$$\begin{aligned} p(w_{ij}^{(k)}|\cdot) &\propto p(\boldsymbol{\beta}_k|\cdot)\ p(w_{ij}^{(k)}) \\ &\propto \left(\prod_{i=2}^{I} \lambda_i\right)^{1/2} (w_{ij}^{(k)})^{\nu/2-1} \exp\left[-w_{ij}^{(k)} \left\{\frac{\nu}{2} + \frac{(\beta_{ik} - \beta_{jk})^2}{2\tau_k^2}\right\}\right]. \end{aligned} \tag{13}$$

Here λ_i, $i = 2, \ldots, I$, are the non-zero eigenvalues of the penalty matrix $\mathbf{K}_k(\mathbf{w}^{(k)})$ corresponding to the joint prior $p(\boldsymbol{\beta}_k|\mathbf{w}^{(k)}, \tau_k^2)$ derived from the conditional prior (5). We explicitly denote the penalty matrix by $\mathbf{K}_k(\mathbf{w}^{(k)})$ to emphasize its dependence on the weights $\mathbf{w}$. Note that expression (13) is a gamma Ga(a_k', b_k') density with parameters

$$\begin{aligned} a_k' &= \frac{\nu}{2}, \\ b_k' &= \frac{\nu}{2} + \frac{(\beta_{ik} - \beta_{jk})^2}{2\tau_k^2}, \end{aligned} \tag{14}$$

multiplied by $(\Pi_{i=2}^{I} \lambda_i)^{1/2}$. To sample from this distribution we employ an MH step and use a

gamma distribution with the parameters in expression (14) as proposal density. Therefore the acceptance probability reduces to

$$\alpha = \min\left\{1, \left(\prod_{i=2}^{I}\lambda_i^* \Big/ \prod_{i=2}^{I}\lambda_i\right)^{1/2}\right\}$$
$$= \min(1, |\mathbf{K}_{11}^*|^{1/2}/|\mathbf{K}_{11}|^{1/2}).$$

The λ_i^* denote the non-zero eigenvalues of the penalty matrix $\mathbf{K}_k(w_{ij}^{*(k)})$ resulting from a proposed weight $w_{ij}^{*(k)}$, and the second equality follows from lemma 1 in Section 3. Thus acceptance probabilities can be computed by using efficient Cholesky decomposition of order $O(n)$. Additionally, we exploit the fact that it is sufficient to start the Cholesky decomposition in the row corresponding to the position where a proposed new weight is located. Block updating of several weights in one step speeds up computation considerably, since the ratio (6) must be evaluated only once per block. Since the proposal densities $p(w_{ij}^{(k)})$ are independent no further difficulties are imposed by sampling from a block of weights. In our application to human brain mapping in Section 5 joint updating of six weights still yields an acceptance rate of greater than 50%.

For the formulation with spatially correlated weights, we assume a Gaussian MRF prior (9) with a (proper or improper) penalty matrix $\mathbf{L}$ derived from the specification of the neighbourhood structure for edges or links between adjacent pixels or regions. The full conditional $p(\mathbf{v}|\cdot)$ for $\mathbf{v}$ given the data and the remaining parameters is determined by

$$p(\mathbf{v}|\cdot) \propto p(\boldsymbol{\beta}|\tau^2, \mathbf{v})\ p(\mathbf{v}|\phi)$$
$$\propto \left(\prod_{i=1}^{I}\lambda_i\right)^{1/2} \exp\left\{-\frac{1}{2\tau^2}\boldsymbol{\beta}'\mathbf{K}(\mathbf{v})\boldsymbol{\beta}\right\} \exp\left(-\frac{1}{2\phi^2}\mathbf{v}'\mathbf{L}\mathbf{v}\right),$$

where $\mathbf{K}(\mathbf{v})$ is obtained from the original penalty matrix $\mathbf{K}$ through the reparameterization $w_{ij} = \exp(v_{ij})$. Owing to the non-linearities in $\mathbf{K}(\mathbf{v})$, sampling from this full conditional must be based on some approximation along similar lines to those in Rue (2001) and Knorr-Held and Rue (2002). For example, using a second-order Taylor series expansion $\exp(v_{ij}) \approx 1 + v_{ij} + v_{ij}^2/2$ leads to a Gaussian proposal density

$$p_{\text{prop}}(\mathbf{v}) \propto \exp\{-\tfrac{1}{2}(\mathbf{v}-\mathbf{m})'\mathbf{M}(\mathbf{v}-\mathbf{m})\}$$

with appropriately defined $\mathbf{m}$ and $\mathbf{M}$. However, $\mathbf{M}$ and $\mathbf{m}$ will depend on τ and ϕ and therefore must be recomputed in each iteration of the sampling steps. Moreover, the band structure for $\mathbf{M}$ becomes lost, making Cholesky decompositions for drawing from the proposal density less efficient. Instead of this block move MH step, single-move MH steps or MH steps for smaller blocks based on conditional prior proposals (Knorr-Held, 1999) may be constructed. In any case, lemma 1 in Section 3 will again be useful to avoid direct computation of the eigenvalues.

5. Application

We illustrate our approach based on adaptive Gaussian MRF (or adaptive Gauss) priors with $\nu = 1$ by application to an FMRI data set from a visual stimulation experiment as described in Section 1. Visual paradigms are known to elicit great activation amplitudes in the visual cortical areas, which are sharply separated from other functional areas.

However, before returning to these real data we present some results from a pretest on the basis of a synthetic spatiotemporal data set that was generated according to a stylized structure

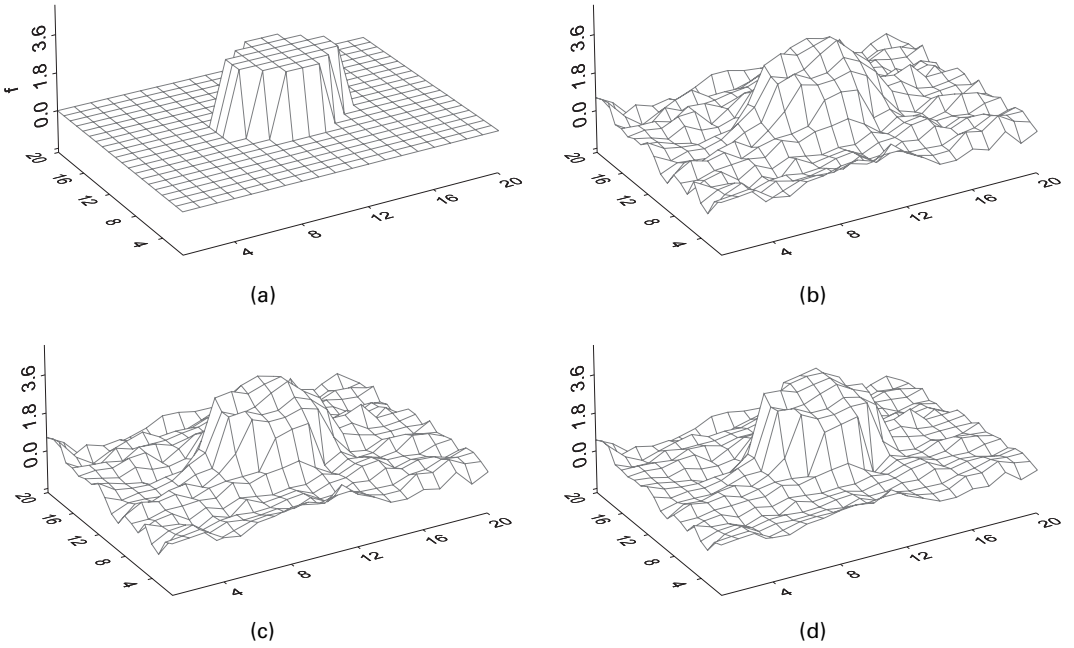

Fig. 2. (a) True surface $f(i)$ together with (b) Gauss, (c) Laplace and (d) adaptive Gauss surfaces $\hat{f}(i)$

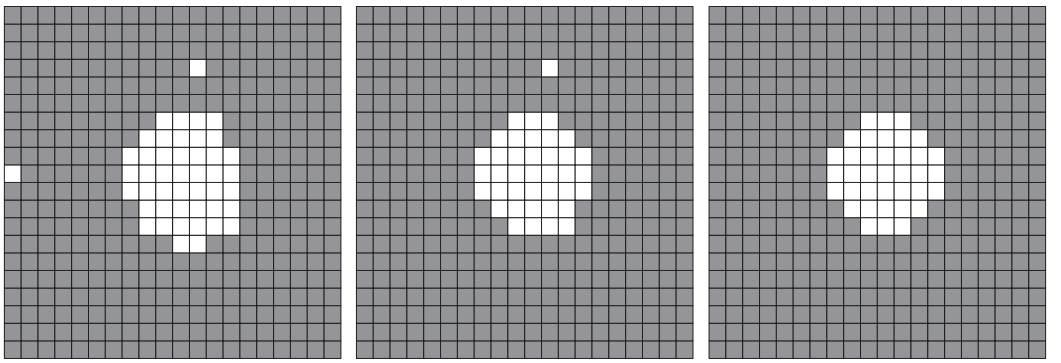

Fig. 3. Posterior probabilities of the spatial effects, with white areas indicating that at least 95% of the sample estimates were positive: (a) Gauss; (b) Laplace; (c) adaptive Gauss

resembling FMRI data. The surface $f(\cdot)$ in Fig. 2(a) stylizes an activation area by a cylinder on a 20×20 grid. Its height is the activation effect. Outside the cylinder, activation is zero. For each pixel i, data were generated by

$$y_{it} = f(i)z_{it} + \varepsilon_{it}, \qquad \varepsilon_{it} \sim N(0, \sigma_i^2), \quad t = 1, \ldots, 70, \quad i = 1, \ldots, 400,$$

where z_{it} is a (not transformed) 0–1 stimulus as displayed in Fig. 1(a). The pixel-specific variances $\sigma_i^2 = 25 + 2\,N(0, 1)$ are generated in advance and are sufficiently large to achieve a realistically low signal-to-noise ratio. The aim of this pretest was to compare the performance of spatial smoothers with different priors. Figs 2(b)–2(d) show the posterior mean estimates $\hat{f}(i)$ of the surfaces by using a Gauss, Laplace and adaptive Gauss prior. Fig. 3 displays appropriate pos-

terior probabilities, with white pixels indicating that at least 95% of the sample estimates were positive. Remaining grey pixels are considered to be 'non-significant'. Compared with the Gauss prior, the Laplace and the adaptive Gauss prior have improved edge preserving properties. This is also confirmed by a comparison of the mean-squared error. Whereas the Gauss prior yields a mean-squared error of 0.19, the Laplace prior results in a value of 0.14 and the adaptive Gauss prior yields an even smaller mean-squared error of 0.09.

Note that we do not show estimates resulting from Geman–Reynolds priors here since a fair comparison would include a data-driven choice of the hyperparameters λ and δ. Although the cylinder may be recovered almost perfectly with a suitable choice of the hyperparameters, other choices result in undersmoothed and oversmoothed surfaces. However, retrieving $f(i)$ almost perfectly is also possible with the adaptive Gauss prior in case the smoothing parameter τ^2 is not estimated simultaneously from the data but set to an adequate, prefixed, value. This does not apply to the Gauss prior, since it does not have edge preserving properties.

Turning to the FMRI data set, we apply the parametric observation model (1) with time constant activation effect β_i in the first step, where the transformed stimulus z_{it} was determined through a pilot estimate; see Gössl (2001) page 33, for details. The base-line trend was modelled parametrically by

$$f_i(t) = \mathbf{u}_i'\boldsymbol{\alpha}_i = \alpha_{i0} + \alpha_{i1}t + \alpha_{i2}\sin(\pi t/16) + \alpha_{i3}\cos(\pi t/25) + \alpha_{i4}\cos(\pi t/40),$$
$$i = 1, \ldots, 2949, \quad t = 1, \ldots, 70.$$

In the second step we apply observation model (3) with a time-varying activation profile $b_i(t) = z_{it}\mathbf{v}_t'\boldsymbol{\beta}_i$. We replace β_i by

$$\mathbf{v}_t'\boldsymbol{\beta}_i = \beta_{i0} + \beta_{i1}t + \beta_{i2}\cos(\pi t/25) + \beta_{i3}\cos(\pi t/40).$$

All frequencies were selected through stepwise selection procedures.

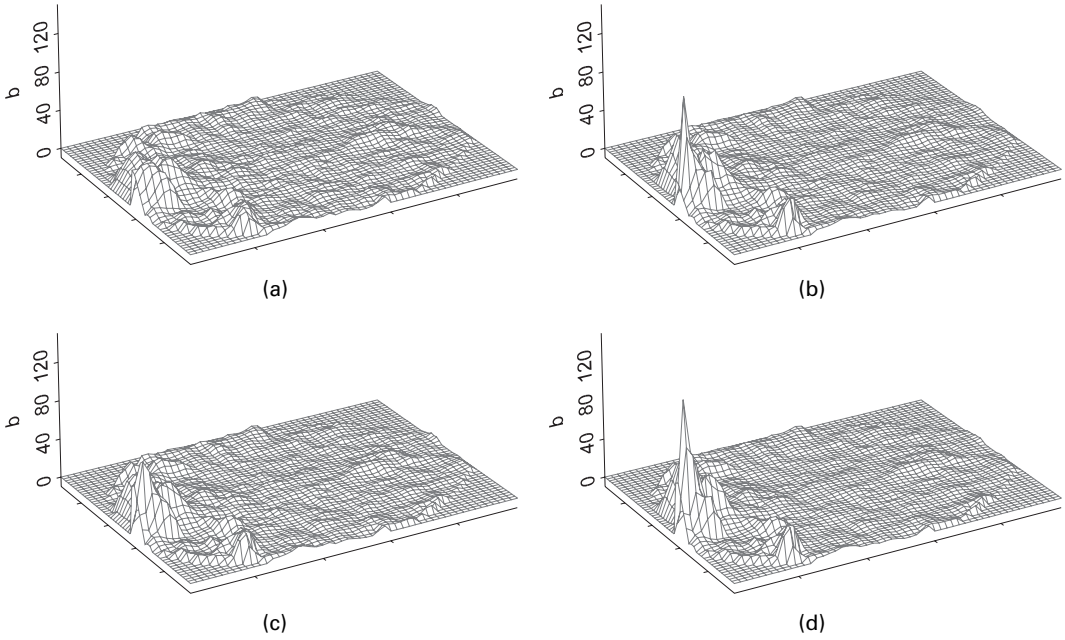

Fig. 4. Estimated surfaces $\hat{\beta}_i$ for the time constant model (note that the origin of the x–y-co-ordinates is at the z-axis): (a) Gauss; (b) adaptive Gauss; (c) Laplace; (d) Geman–Reynolds

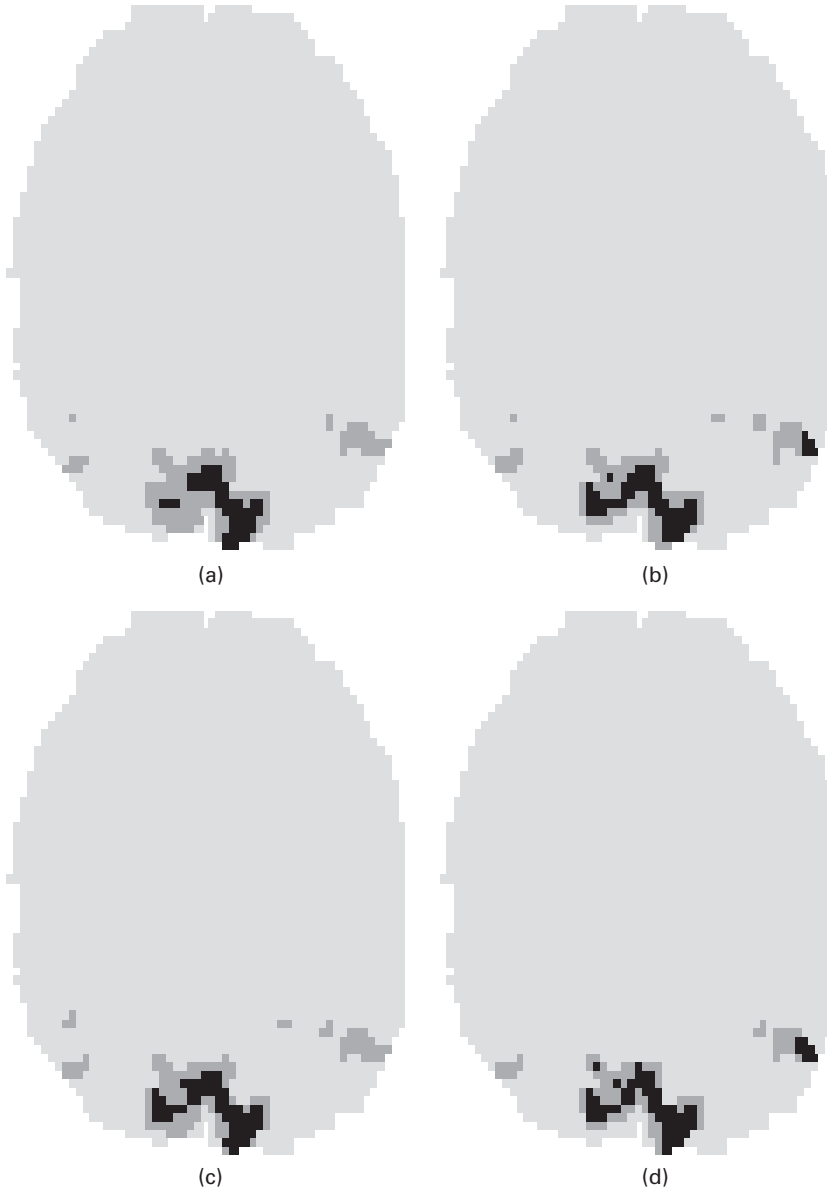

Fig. 5. Map of activated pixels (□, non- or weakly activated pixels ($\hat{\beta}_i \leqslant 15$); ▩, activated pixels with $15 < \hat{\beta}_i \leqslant 30$; ■, strongly activated pixels with $\hat{\beta}_i > 30$): (a) Gauss; (b) adaptive Gauss; (c) Laplace; (d) Geman–Reynolds

For the time constant model we estimated the activation surface $\{\hat{\beta}_i\}$ by using the Gauss, adaptive Gauss, Laplace and a Geman–Reynolds prior. Inverse gamma priors with $a = b = 0.001$ were chosen for the error variances σ_i^2 in all four cases, as well as for the variance parameter τ^2 in the case of the Gauss and adaptive Gauss priors. Other choices for those priors, such as uniform priors on the standard deviations σ_i and τ, yielded analogous results. On the basis of simulation results for a grid of values of the tuning parameters p, τ and λ of the Geman–Reynolds prior as well as on visual inspection for the data at hand we set them to

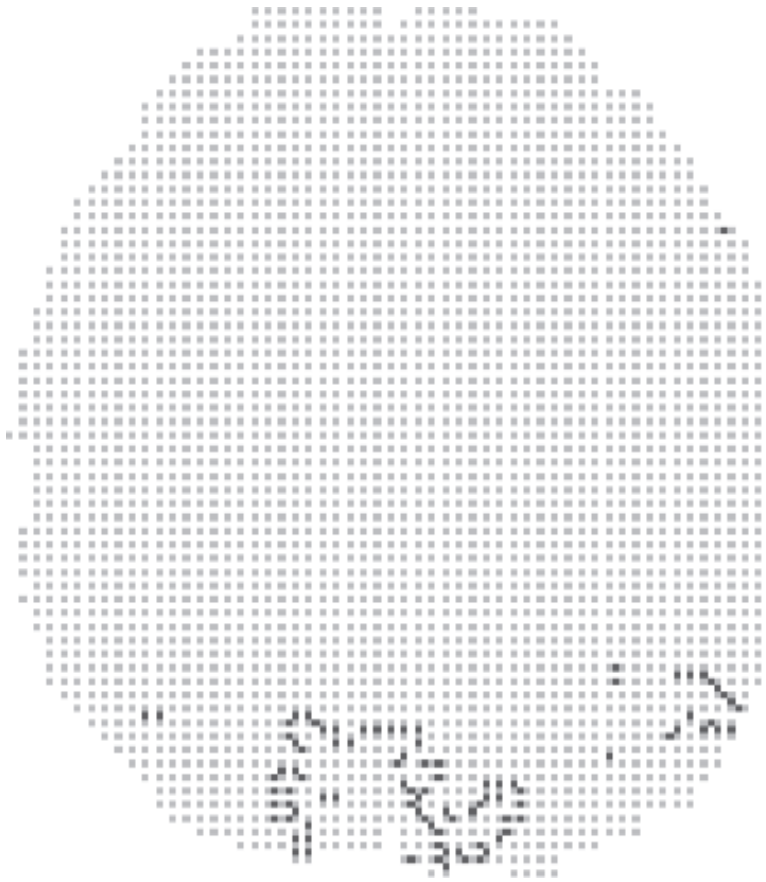

Fig. 6. Boundary map resulting from the adaptive Gauss prior: □, pixels i; ■, weights w_{ij} with $\hat{w}_{ij} < 0.1$ (whereat this threshold is chosen by visual judgment)

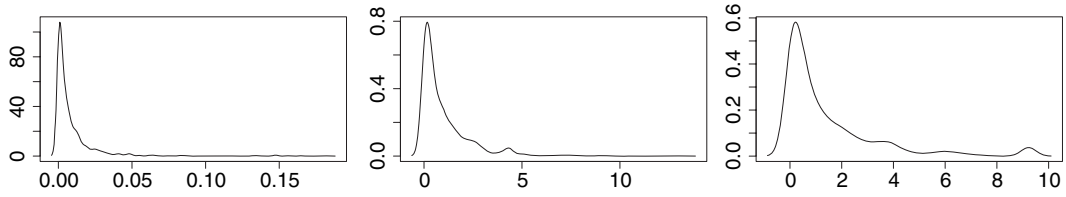

Fig. 7. Kernel density estimates of the posterior distributions of three weights w_{ij} with (a) small ($w_{ij} = 0.009$), (b) medium ($w_{ij} = 1.00$) and (c) large ($w_{ij} = 1.44$) posterior means

$p = 2$, $\tau = 5$ and $\lambda = 3$. Concerning convergence and mixing of the MCMC samples, 12000 iterations with a burn-in of 2000 and a thinning of 10 proved to be sufficient for all our models.

Fig. 4 shows posterior mean estimates $\{\hat{\beta}_i, i = 1, \ldots, I\}$ for these four models. Obviously, the Gauss prior oversmooths the sharp peaks and ridges as well as steep slopes in the area of the central visual cortex (on the left-hand side of the activation surface), whereas it undersmooths in non-activated areas, resulting in a comparably rough estimated surface. The result for the Laplace prior shows better local adaptivity, but this prior still oversmooths the central activation area. Both the adaptive Gauss and the Geman–Reynolds prior exhibit the features

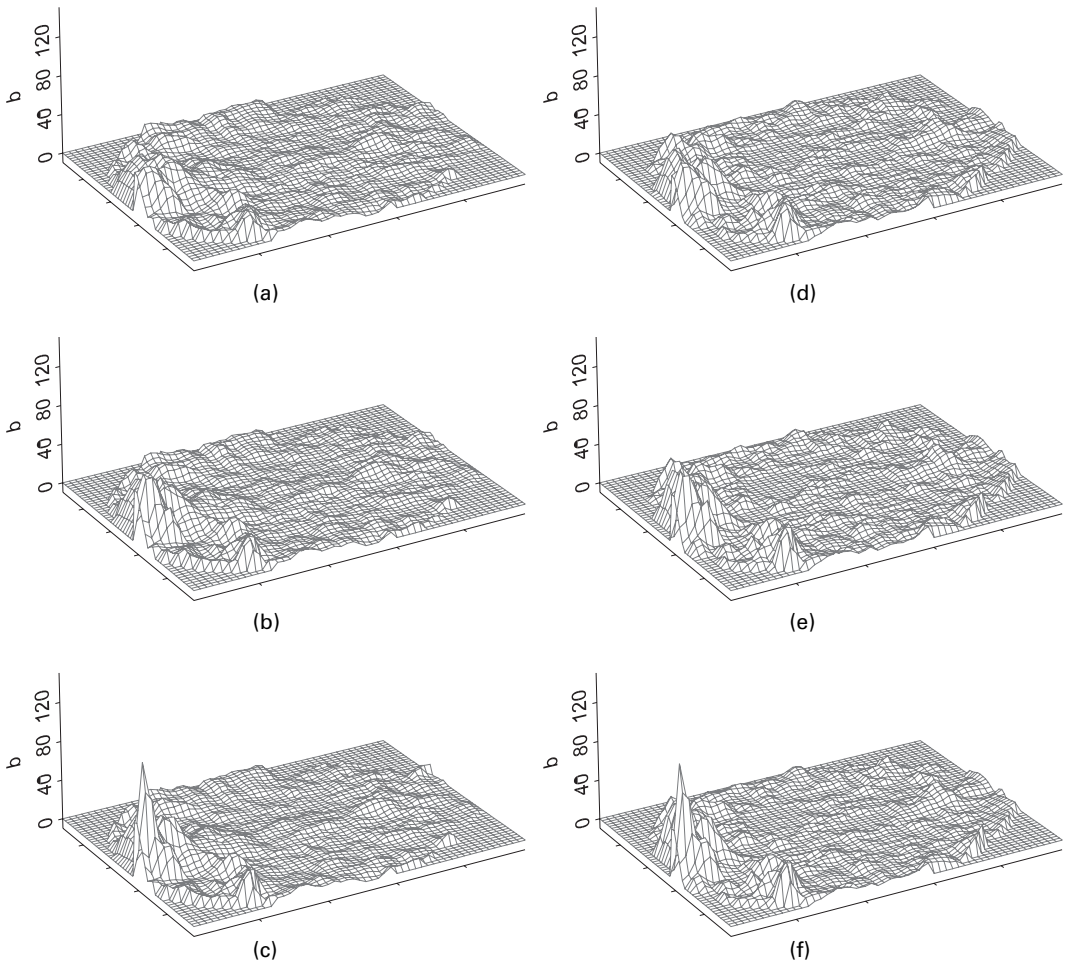

Fig. 8. Estimated activation effects $\{\mathbf{v}_t'\hat{\beta}_i, i = 1, \ldots, I\}$ at (a)–(c) $t = 18$ and (d)–(f) $t = 58$ for (a), (d) Gauss, (b), (e) Laplace and (c), (f) adaptive Gauss priors (note that the origin of the x–y-co-ordinates is at the z-axis)

desired: non-activated areas are smooth and close to zero, but activated areas, in particular the distinct physiologically known peak in the central visual cortex, remain. Fig. 5 displays a summary of the activated areas distinguishing two grades of activation, whereat thresholds are chosen by visual judgment. The difference between the distinct models can be seen most clearly when looking at the left-hand half of the middle activation centre. Compared with the Gauss prior the robust and adaptive priors yield a smaller number of activated pixels but a greater number of strongly activated pixels. Again this provides evidence that the Gauss prior over-smooths slopes and peaks. The advantage of the adaptive Gauss prior is that the construction of boundary maps (on the basis of the 5769 estimated weights $\hat{w}_{ij}$) as displayed in Fig. 6 is straightforward. Indicating dissimilarity of pixels i and j, small values of the estimated weights suggest boundaries. Estimated weights $\hat{w}_{ij}$ range from 0.0004 to 1.8 with a mean of 1.01. Kernel density estimates of the posterior distributions of three selected weights with small, medium and large posterior mean are displayed in Fig. 7. This direct access to boundary maps indicates that our approach may also be useful for areal wombling, as an alternative to the Bernoulli weights in distribution (10). A further advantage of the adaptive Gauss prior is that hyperparameters

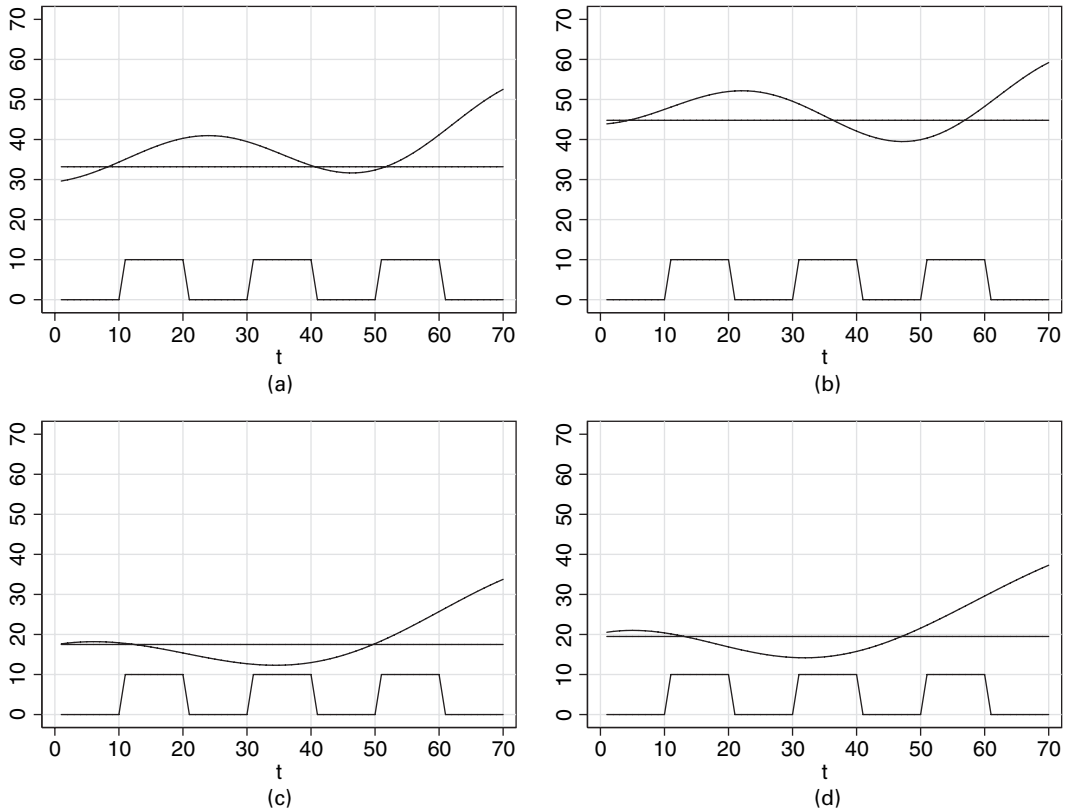

Fig. 9. Estimated time-varying activation effects $\{\mathbf{v}_t'\hat{\beta}_i, t = 1, \ldots, 70\}$ at (a), (b) pixel $i = 98$ and (c), (d) pixel $i = 327$ for (a), (c) Gauss and (b), (d) adaptive Gauss priors (additional curves indicate corresponding estimated time constant activation effects)

can be estimated easily from the data. This is much more difficult for the Geman–Reynolds prior with its complicated normalizing constant.

This disadvantage becomes even more crucial when applying the time-varying model (3), because we would have to specify hyperparameters for each of the components of the parameter vector β_i. Therefore we carry out comparative analyses for Gauss, Laplace and adaptive Gauss priors only. With the same choice of inverse gamma priors as in the time constant case, we can estimate all space-varying parameters with the technique that was outlined in Section 4. Fig. 8 shows the estimated activation surfaces $\{\mathbf{v}_t'\hat{\beta}_i, i = 1, \ldots, I\}$ for two different points of time which refer to the first and third activation period. Comparing the smoothing qualities of the different priors, it can again be seen that the adaptive Gauss prior has distinctly better local adaptivity properties. Fig. 9 displays posterior means of time-varying activation effects at two selected characteristic pixels. Figs 9(a) and 9(b) show estimated effects at an active pixel ($i = 98$; co-ordinates (35,4)) resulting from the Gauss and adaptive Gauss model respectively. Although the estimated activation is consistently greater with the adaptive Gauss model, the shape of the estimated curves is highly comparable. Figs 9(c) and 9(d) show estimated effects at pixel $i = 327$ (co-ordinates (12,11)). This pixel lies within the smaller left-hand activation centre (see Fig. 5) and seems to be activated within the third activation period mainly. Note that the Laplace prior results in curves of the same shape lying between the curves that are displayed in Fig. 9 (and are not shown).

6. Conclusion

Gaussian MRFs with adaptive interaction weights as developed in this paper seem to be a promising framework. They should also be of use in other spatial problems where local adaptivity and edge preservation are required. Interaction weights, estimated jointly with spatial effects, can provide valuable information about the local strength of interaction between neighbouring pixels or regions. This may be used for detecting areas of high local curvature or for boundary analysis (wombling), as in our application. This additional feature must be balanced against increasing computation times, compared with global Gaussian MRF priors, or simple robust priors such as the Laplace prior. Additional computational effort is moderate for problems with a moderate number of spatial units, but it can increase considerably (from hours to days) in applications as considered in the paper.

Acknowledgements

We thank Dorothee Auer and Christoff Gössl (Max Planck Institute of Psychiatry) for providing the data and for many discussions, motivating and guiding this work, and the German National Science Foundation for financial support through grants from the Sonderforschungsbereich 386 'Statistical analysis of discrete structures'. Our thanks go to two referees for their helpful and supportive comments and suggestions to improve the first version substantially.

References

Assunção, R. (2003) Space varying coefficient models for small areal data. *Environmetrics*, **14**, 453–473.
Baladandayuthapani, V., Mallick, B. K. and Carroll, R. J. (2005) Spatially adaptive Bayesian penalized regression splines. *J. Computnl Graph. Statist.*, **14**, 378–394.
Brezger, A. (2005) Bayesian P-splines in structured additive regression models. *Dissertation*. University of Munich, Munich.
Carlin, B. P. and Banerjee, S. (2003) Hierarchical multivariate CAR models for spatio-temporally correlated survival data. In *Bayesian Statistics 7* (eds J. M. Bernardo, A. P. Dawid, J. O. Berger, M. West, D. Heckerman, M. J. Bayarri and A. F. M. Smith), pp. 45–64. Oxford: Oxford University Press.
Carter, C. and Kohn, R. (1996) Markov chain Monte Carlo for conditionally Gaussian state space models. *Biometrika*, **83**, 589–601.
Fahrmeir, L. and Gössl, C. (2002) Semiparametric Bayesian models for human brain mapping. *Statist. Modllng*, **2**, 235–250.
Friston, K. J., Holmes, A. P., Poline, J.-B., Grasby, P., Williams, S. C. R., Frackowiak, R. S. J. and Turner, R. (1995) Analysis of fMRI time-series revisited. *NeuroImage*, **2**, 45–53.
Geman, D. and Reynolds, G. (1992) Constrained restoration and the recovery of discontinuities. *IEEE Trans. Pattn Anal. Mach. Intell.*, **14**, 367–383.
Genovese, C. R. (2000) A Bayesian time-course model for functional magnetic resonance imaging data (with discussion). *J. Am. Statist. Ass.*, **95**, 691–719.
Gössl, C. (2001) Bayesian models in functional magnetic resonance imaging: approaches for human brain mapping. *Dissertation*. University of Munich, Munich.
Gössl, C., Auer, D. P. and Fahrmeir, L. (2000) Dynamic models in fMRI. *Magn. Resnce Med.*, **43**, 72–81.
Gössl, C., Auer, D. P. and Fahrmeir, L. (2001) Bayesian spatio-temporal inference in functional magnetic resonance imaging. *Biometrics*, **57**, 554–562.
Jin, X., Carlin, B. P. and Banerjee, S. (2005) Generalized hierarchical multivariate CAR models for areal data. *Biometrics*, **61**, 950–961.
Knorr-Held, L. (1999) Conditional prior proposals in dynamic models. *Scand. J. Statist.*, **26**, 129–144.
Knorr-Held, L. (2003) Some remarks on Gaussian Markov random field models for disease mapping. In *Highly Structured Stochastic Systems* (eds P. Green, N. Hjort and S. Richardson). Oxford: Oxford University Press.
Knorr-Held, L. and Rue, H. (2002) On block updating in Markov random field models for disease mapping. *Scand. J. Statist.*, **29**, 597–614.
Lang, S. and Brezger, A. (2004) Bayesian P-splines. *J. Computnl Graph. Statist.*, **13**, 183–212.
Lang, S., Fronk, E.-M. and Fahrmeir, L. (2002) Function estimation with locally adaptive dynamic models. *Computnl Statist.*, **17**, 479–500.
Lu, H., Reilly, C. S., Banerjee, S. and Carlin, B. P. (2005) Bayesian areal wombling via adjacency modeling. *Research Report 2005-015*. Division of Biostatistics, University of Minnesota, Minneapolis.

Ma, H., Virnig, B. and Carlin, B. P. (2005) Spatial methods in geographic administrative data analysis. *Research Report 2005-016*. Division of Biostatistics, University of Minnesota, Minneapolis.
Rue, H. (2001) Fast sampling of Gaussian Markov random fields. *J. R. Statist. Soc.* B, **63**, 325–338.
Rue, H. and Held, L. (2005) *Gaussian Markov Random Fields; Theory and Applications*. Boca Raton: Chapman and Hall–CRC.
Smith, M. and Fahrmeir, L. (2007) Spatial Bayesian variable selection with application to functional magnetic resonance imaging. *J. Am. Statist. Ass.*, to be published.
Smith, M., Pütz, B., Auer, D. and Fahrmeir, L. (2003) Assessing brain activity through spatial Bayesian variable selection. *NeuroImage*, **20**, 812–815.
Tabelow, K., Polzehl, J. and Spokoiny, V. (2005) Analysing fMRI experiments with structural adaptive smoothing procedures. *Preprint 1079*. Weierstrass Institute for Applied Analysis and Stochastics, Berlin.