

A Stochastic Geometry Model for Functional Magnetic Resonance Images.

Niels Væver Hartvig*
University of Aarhus

Revised version
March 9th 2001

Abstract

In functional magnetic resonance imaging, spatial activation patterns are commonly estimated using a non-parametric smoothing approach. Significant peaks or clusters in the smoothed image are subsequently identified by testing the null hypothesis of lack of activation in every volume element of the scans. A weakness of this approach is the lack of a model for the activation pattern; this makes it difficult to determine the variance of estimates, to test specific neuroscientific hypotheses or to incorporate prior information about the brain area under study in the analysis. These issues may be addressed by formulating explicit spatial models for the activation and using simulation methods for inference. We present one such approach, based on a marked point process prior. Informally, one may think of the points as centres of activation, and the marks as parameters describing the shape and area of the surrounding cluster. We present an MCMC algorithm for making inference in the model, and compare the approach with a traditional non-parametric method, using both simulated and visual stimulation data. Finally we discuss relevant extensions of the model and the inferential framework to account for non-stationary responses and spatio-temporal correlation.

Keywords: Functional magnetic resonance imaging; Stochastic geometry model; Marked point process; Markov chain Monte Carlo; State space model.

1 Introduction

Functional magnetic resonance imaging (fMRI) uses the different magnetic properties of oxy- and deoxyhaemoglobin to visualize localized changes in blood flow, blood volume and blood oxygenation in the brain. These are in turn indicators for local changes in

*Department of Mathematical Sciences, University of Aarhus, Ny Munkegade, DK-8000 Aarhus C, email: vaever@imf.au.dk

neural activity. By exposing a subject to controlled stimuli, which are carefully designed to affect only certain brain functions, it is possible to estimate the anatomical location of neurones involved in the corresponding functions. Brain function may then be mapped to brain anatomy by combining fMRI scans with anatomical scans obtained by conventional MRI.

The technique is quite new; one of the first experiments was reported by Kwong *et al.* (1992), and since then the number of publications in the field has grown extremely fast. Today fMRI is one of the most important modalities for imaging the brain, since it is completely non-invasive, has a reasonable temporal resolution (about 2 sec) and an excellent spatial resolution (about 2 mm). See Lange (1996) or Hartvig (2000) for introductions to the subject.

The data obtained in an fMRI experiment is a time series of three dimensional scans of the brain, as well as covariates describing the presentation of stimuli. In the analysis of the data, the signal of interest is a spatio-temporal process, where the temporal profile is coupled to the stimulation rhythm through the haemodynamic response to neural activation (also known as the BOLD effect). The response lags the neural activation with about 6 sec., and is more smooth than the latter. Empirically, the impulse response has been found to look roughly like a Gamma density, but despite attempts to explain this quantitatively (Buxton *et al.*, 1998), there is still not a fully accepted biological model for the process.

Generally the neurones involved in a specific task are expected to possess spatial structure, yielding a spatially correlated neural activation process. The resulting haemodynamic oxygenation changes contribute further to this correlation, as they diffuse in the venous side of the capillary system, spreading over several millimetres. Modelling the spatial structure of this haemodynamic process is a difficult task: Firstly the overall pattern will of course depend on the type of stimulation, and it is difficult to impose structure on this in a general setting. Secondly the complex geometry of the cortical surface makes it difficult to define relevant neighbourhoods in the space of volume elements (or *voxels*) of the scanned brain.

Instead, a common approach is to estimate the activation magnitude separately in each voxel by a one dimensional time series model, see for instance Worsley and Friston (1995), Lange and Zeger (1997), Bullmore *et al.* (1996) or Genovese (2000). The spatial structure of the data is included in a second step, when the image (or volume) of marginal estimates is convolved with a smoothing kernel, to obtain a non-parametric spatial estimate. Subsequently, significant peaks or clusters in the image are identified, by testing the null hypothesis of lack of activation at each voxel, and the final estimate may consist of voxels that are significantly higher than what would be expected by chance. Here the significance level is corrected for the large number of hypotheses tested, using results for Gaussian random fields (Worsley, 1995).

The fundamental problem in this approach is the lack of a model for the activation, i.e. there is no model for the distribution of the statistics under the alternative hypothesis that a voxel is active, and no assumptions are made about the distribution of shape and size of activated regions. Without an explicit spatial model, concepts such as uncertainty of the estimated pattern or the testing of high-level hypotheses are very difficult to study.

In this paper we propose a spatial model, by which some of these problems may be addressed. The model is motivated by two fundamental assumptions in the fMRI literature, which are based partly on the spatial structure on a neuronal level, and partly on the haemodynamic origin of the signal: 1) The activated areas have a spatial extent of several millimetres and 2) the activation pattern is “smooth”. Using these, we will model the activation surface as a collection of Gaussian functions, which to some extent represent individual centres in the brain. This is formulated as a stochastic geometry model based on marked point process prior (Baddeley and van Lieshout, 1993), where the points stand for the locations and the marks describe the shape and height of the centres. The inference in the model is based on simulation techniques, by which we can estimate the posterior mean of functions of interest, such as the mean activation pattern.

One advantage, compared to the typical analysis outlined above, is a more precise estimate of the spatial pattern. This may be particularly relevant for short time series, in experiments with many different types of stimuli or in situations where signal estimation is more important than just signal detection. The latter is the case for instance in pre-surgical planning or when fMRI is combined with other imaging modalities. A further motivation is the extended inferential scope, which allows us to assess the uncertainty of estimates in a Bayesian framework, or to quantify the belief in more specific hypotheses by estimating posterior probabilities. Finally the haemodynamic response function may be modelled in a semi-parametric way, which allows for non-stationarities and non-linearities. With the latter approach explicit knowledge of the stimulation paradigm is not required, and we can hence estimate activation which is not time-locked to the stimulation rhythm.

The paper is organized as follows: We first present a typical set of fMRI data and its preprocessing in Section 2. In Section 3 we formulate the basic model for the spatial activation pattern and combine this with a simple model for the temporal response to obtain a spatio-temporal model. The temporal pattern is assumed to be known and described by a convolution model. The posterior inference is done by an MCMC algorithm, which is described in Section 4. In Section 5 we apply the model to simulated data, which is used for estimating prior parameters, and to visual stimulation data. Finally we discuss relevant extensions in Section 6, to account for correlated noise or non-stationary responses, and give a conclusion in Section 7.

2 fMRI data and its preprocessing

For illustration, we will consider data acquired in a well studied experimental design, namely a visual stimulation presented periodically in blocks of 20 seconds. The stimulus was a light, flashed with 7 Hz in front of the right eye of the subject. 90 so-called Echo-Planar Imaging scans were acquired during a 3 minute period, with an inter-scan time (or repetition time) of 2 seconds. The stimulation was arranged in blocks of 20 seconds off, 20 seconds on, 20 seconds off etc., with 4 complete on-off cycles during the session. Each volume of scans consists of 5 slices of thickness 5 mm, each comprised of 128 by 128 voxels of size 1.875×1.875 mm. We discarded the first 5 scans, due to initial instabilities in the magnetization level of the tissue.

Figure 1 is a graphical illustration of the data. The left panel displays a section of a slice from one of the scans, oriented in an oblique axial-coronal direction with the posterior part of the brain in the top of the image. The right panels display two voxel time series; one is located in the visual cortex, which is known to process visual impressions, the other is located in an area where activation is not expected. The fluctuation in the former series is evident, as is the haemodynamic delay and dispersement.

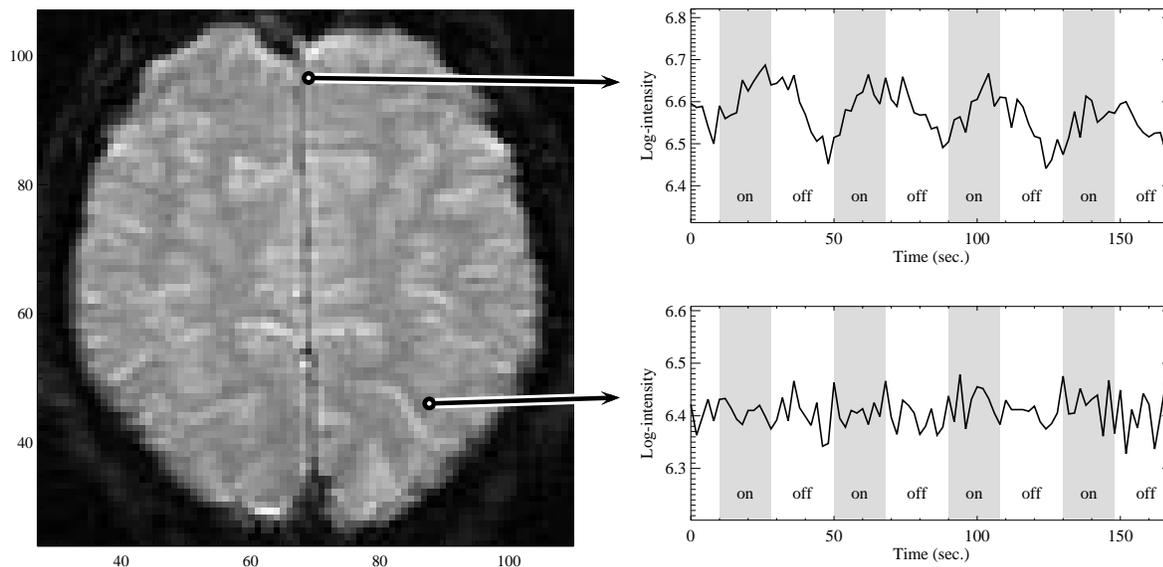


Figure 1: Graphical illustration of visual stimulation fMRI data. Left: A section of an MR scan of a slice of the brain. Right: Time series of respectively an active (top) and non-active (bottom) voxel. The stimulation epochs are indicated by grey bars.

The scans are almost always preprocessed before the statistical analysis to reduce artifacts caused by the scanner or by subject movement, or to map individual brains to a standard atlas. To correct for movement artifacts, we have used a simple procedure, where each image is aligned to a reference image by minimizing the squared difference between the two images over all translations and rotations. Often different types of trends and low-frequency fluctuations may be observed in the voxel time series. These may be caused by scanner instability, by physiological processes or by aliased cardiac and respiratory pulsations. We chose a very simple correction for this, by subtracting a fitted linear trend term in each voxel time series. More general trend and fluctuation models, such as a cosine basis proposed by Holmes *et al.* (1997), may be applied. As will be clear later, however, our focus is partly to model general temporal response patterns, and hence we are cautious not to remove any fluctuations related to the haemodynamic response. A linear term is a good compromise in this context.

Finally we log-transformed the data to stabilize variances. Furthermore, this transformation is motivated by the fact that the units of the MR scanner are arbitrary and commonly variation is quantified as percent of baseline intensity. The units on the log-scale are directly interpretable in accordance with this, namely as relative increments of

the original series.

Let $Y = \{Y_{it}, i \in V, t = 1, \dots, m\}$ denote the preprocessed fMRI time series. Here V is the set of voxels covering brain tissue, $V \subseteq S$, where S represents a two dimensional slice or a three dimensional volume of the brain, and m is the number of scans. We will let π_t denote the stimulation function, where $\pi_t = 1$ indicates stimulation and $\pi_t = 0$ no stimulation at time t .

3 The model

Our general model has the form

$$Y_{it} = (A_i(X) + \eta_i)\varphi_t + \varepsilon_{it},$$

where $\eta = \{\eta_i, i \in V\}$ and $\varepsilon = \{\varepsilon_{it}, i \in V, t = 1, \dots, m\}$ are Gaussian processes. Here $A(X) = \{A_i(X), i \in V\}$ is the magnitude of activation, which is parametrized by a marked point process X , and $\varphi = \{\varphi_t, t = 1, \dots, m\}$ is the temporal variation caused by the BOLD effect. We will describe in detail how the spatial and temporal patterns are modelled in the following.

3.1 A model for the spatial activation pattern

Consider first the case where data only represent a two dimensional slice of the brain, that is $V \subseteq S \subseteq \mathbb{R}^2$. We will describe the spatial activation pattern by a marked point process $X = \{X_1, X_2, \dots, X_n\}$, where $X_k = (\mu_k, a_k, d_k, r_k, \theta_k)$. A point X_k may to some extent be considered as a centre of activation with location $\mu_k \in S$, and where the four marks $(a_k, d_k, r_k, \theta_k)$ describe the magnitude and shape of the centre. The activation pattern $\{A_i(X)\}_{i \in V}$ is assumed to have a specific geometry, namely a sum of Gaussian functions, $A_i(X) = \sum_{k=1}^n h(i; X_k)$, where

$$h(i; X_k) = a_k \exp \left\{ -\frac{\pi \log 2}{d_k} \left(\frac{j_1^2}{r_k/(1-r_k)} + \frac{j_2^2}{(1-r_k)/r_k} \right) \right\}, \quad (1)$$

here $(j_1, j_2) = R(-\theta_k)(i - \mu_k)$ and $R(\theta)$ is a rotation with angle θ . This representation is motivated by the common assumptions of smoothness and spatial extent of the activation, and the heuristic idea is that a general smooth activation surface with few localized peaks, may be well approximated by a collection of Gaussian functions. The interpretation of the parameters is that $a_k \in \mathbb{R}_+$ is the height of the Gaussian bell at the centre μ_k , $d_k \in \mathbb{R}_+$ is the area of the contour ellipse at half height, $r_k \in (0, 1)$ is a measure of the eccentricity of the ellipse, more precisely the ratio of the first principal axis and the sum of the two axes, and $\theta_k \in [-\pi/4, \pi/4]$ is the orientation of the ellipse.

We have specific prior knowledge on the parameters of the model. The magnitude is typically about 2%-5% of the baseline intensity, we expect the activation clusters to cover at least a few voxels, and we may often have a strong prior idea of where they will occur, based on previous experiments on the same subject or on general knowledge of the brain

function under study. Furthermore a typical point configuration should contain only a moderate number of points. This may be included in the model in a Bayesian framework, and we will therefore consider the process X as a realization of a random variable with a prior distribution.

Each centre X_k is a point in $\mathcal{X} = S \times M$, where $M = [0, C_a] \times [0, C_d] \times (0, 1) \times [-\pi/4, \pi/4]$. Here C_a and C_d are natural bounds for the height and area, respectively. Let \mathcal{X} be equipped with the Borel σ -field $\mathcal{S} \times \mathcal{M}$ and the Lebesgue measure $\lambda_2 \times \lambda_4$, and let Ω denote the exponential space over \mathcal{X} , that is the set of finite sets $\{x_1, \dots, x_n\}$ where $x_i \in \mathcal{X}$ for all i . The process X is then a point process in Ω or, equivalently, a marked point process with point space S and mark space M . We will assume that the prior distribution has density

$$p(x) \propto \prod_{k=1}^n (\beta(\mu_k) p(a_k) p(d_k) p(r_k)), \quad x \in \Omega, \quad (2)$$

with respect to the unit rate Poisson process on Ω , where $n = n(x)$ is the number of points in x . Here $\beta(\cdot)$ is an intensity function, which may give preference to specific cortical areas, or may be constant if there is no prior knowledge of where the activation is likely to occur, and $p(\cdot)$ is a generic notation for prior densities of the three mark parameters a_k , d_k and r_k .

3.1.1 Priors for the marks

The priors for a and d should be as uniform as possible, yet penalizing values close to zero. The inverse Gamma distribution is a suitable choice in this context, with its light tail near zero and quite heavy tail for large values. Hence we will assume that $a^{-1} \sim \Gamma(2, \beta_a)$ and $d^{-1} \sim \Gamma(2, \beta_d)$ with the restrictions that $a \in (0, C_a]$ and $d \in (0, C_d]$. The density of d is

$$p(d) = \exp(\beta_d/C_d) (\beta_d/C_d + 1)^{-1} \beta_d^2 d^{-3} \exp(-\beta_d/d), \quad d \in (0, C_d].$$

The upper-bounds C_a and C_d are natural bounds for the magnitude and size of activation clusters. The prior mean of d is $\beta_d/(1 + \beta_d/C_d)$, or approximately β_d when C_d is large.

As for the axis ratio r we wish to discourage very eccentric ellipses. This can be obtained by a Beta-prior, $r \sim \text{Beta}(\beta_r, \beta_r)$.

3.1.2 The intensity function

The intensity function provides a flexible tool for incorporating substantial prior information on the position of the activation. One possibility is to use anatomical covariates obtained from a high-resolution scan of the brain, acquired simultaneously with the functional scans. A relevant anatomical constraint is to restrict activation to the gray matter sheet of the cortical surface (Kiebel *et al.*, 2000). A simple approach to addressing this, is to segment the high-resolution scan, and let the intensity function favour points located in gray matter. Using a soft constraint like this allows for some degree of variation in the position, which is advantageous given the uncertainty in the classification of the high-resolution scan and in the alignment of the images.

Another interesting possibility is to use functional prior information, for instance obtained from a previous experiment on the same subject. This fits well with the Bayesian paradigm of sequential updating of information, and may improve efficiency in the analysis by effectively restricting focus to relevant areas, yet allowing for variations from one experiment to the other. Since the activation pattern may be quite variable even in replicated experiments, also in this case a soft constraint in a Bayesian framework is appropriate.

3.2 A model for the temporal pattern

The simplest model for the temporal pattern φ is a fixed regression model. The response is to a good approximation time-invariant and additive (Boynton *et al.*, 1996), which leads to a convolution model, where an impulse response function is convolved with the stimulation function to obtain φ . Based on empirical studies, Friston *et al.* (1995) suggested to use a Gaussian density with mean 6 sec. and variance 9 sec.² as impulse response, to model the delay and dispersion of the haemodynamic response. We will adopt this choice here, and thus let

$$\varphi_t = \sum_i \pi_{t-i} \frac{T}{\sqrt{2\pi}3} \exp\left(-\frac{(iT-6)^2}{18}\right), \quad (3)$$

where T is the repetition time. This simplicity of this model makes it an attractive starting point, as it allows us to focus on the spatial pattern in the inference, which simplifies the simulation procedure significantly.

The basic assumption made here, is that the spatio-temporal activation profile is separable, i.e. that the response function is the same in all voxels. We will discuss later, how the model may be extended to relax this assumption, and to account for a non-stationary response that changes over time.

3.3 Combining the spatial and temporal models

Given the centres X and the haemodynamic response function φ , the model for the intensity Y is,

$$Y_{it} = (A_i(X) + \eta_i)\varphi_t + \varepsilon_{it}, \quad (4)$$

where $\eta = \{\eta_i, i \in V\} \sim N(0, \tau^2 I_{|V|})$, $\varepsilon = \{\varepsilon_{it}, i \in V, t = 1, \dots, m\} \sim N(0, \sigma^2 I_{|V|} \otimes I_m)$ and η and ε are independent. The likelihood function is given by

$$p(Y|x) = (2\pi\sigma^2)^{-\frac{(m-1)|V|}{2}} \exp\left\{-\frac{1}{2\sigma^2} \sum_{i \in V} \sum_{t=1}^m (Y_{it} - \tilde{Y}_i \varphi_t)^2\right\} \\ \times (2\pi(\sigma^2 + \tau^2 \text{ss}_\varphi))^{-\frac{|V|}{2}} \exp\left\{-\frac{1}{2(\sigma^2/\text{ss}_\varphi + \tau^2)} \sum_{i \in V} (\tilde{Y}_i - A_i(x))^2\right\}. \quad (5)$$

Here \tilde{Y}_i is the coefficient of the projection of $\{Y_{it}, t = 1, \dots, m\}$ on the vector space $L = \text{span}\{\varphi\}$,

$$\tilde{Y}_i = \sum_{t=1}^m Y_{it}\varphi_t/\text{ss}_\varphi, \quad \text{ss}_\varphi = \sum_{t=1}^m \varphi_t^2. \quad (6)$$

Notice that the likelihood function factorizes into two terms, involving only the projection of Y onto L and onto the orthogonal complement to L , respectively, with X only entering in the former. Hence we find that the regression image $\{\tilde{Y}_i, i \in V\}$, which is known as a Statistical Parametric Map (SPM) in the brain imaging literature (Friston *et al.*, 1994), is sufficient for X . As mentioned earlier, usually the SPM is smoothed with a Gaussian kernel to obtain a non-parametric estimate of the activation, and the present setup may thus be viewed as an alternative analysis of the SPM, based on a parametric model.

The purpose of the random effect term η_i is to regularize the estimate of X . To see why this is necessary, consider the log posterior distribution of X , which up to an additive constant is given by

$$\log p(x|Y) = -\frac{1}{2(\sigma^2/\text{ss}_\varphi + \tau^2)} \sum_{i \in V} \left(\tilde{Y}_i - A_i(x) \right)^2 + \log p(x).$$

Suppose for a moment that $\tau = 0$, corresponding to omitting the random effect η_i above. By inserting sufficiently many small bells, we can obtain a configuration where $A_i(x) = \tilde{Y}_i$ when the latter is positive, and $A_i(x) = 0$ elsewhere. This will minimize the sum of squares above. Even if the prior density of such a pathological point configuration is very small, it will be the maximum a posteriori estimate in the limit as m , and hence ss_φ , tends to infinity, since the sum of squares will dominate in the limit. By assuming a fixed positive value for τ^2 this undesirable property of the posterior distribution is removed. Intuitively τ^2 is a measure of how well we expect the actual activation surface to be described by a reasonable collection of Gaussian functions, while the purpose of the prior for X is to quantify what we mean by a reasonable collection.

We will use simple estimates for the variance parameters. An unbiased and consistent estimator for σ^2 is given by

$$\hat{\sigma}^2 = \frac{1}{(m-1)|V|} \sum_{i \in V} \sum_{t=1}^m \left(Y_{it} - \tilde{Y}_i \varphi_t \right)^2 \sim \sigma^2 \chi^2(f)/f, \quad f = (m-1)|V|. \quad (7)$$

As for τ^2 , we will estimate $\sigma^2/\text{ss}_\varphi + \tau^2$ by considering the regression coefficients \tilde{Y}_i . These are mutually independent and distributed as

$$\tilde{Y}_i \sim N(A_i(x), \sigma^2/\text{ss}_\varphi + \tau^2), \quad i \in V.$$

Letting ∂i denote the 9-voxel neighbourhood of i , we will let

$$\bar{Y}_i = \frac{1}{9} \sum_{j \in \partial i} \tilde{Y}_j \sim N(\bar{A}_i(x), \frac{1}{9}(\sigma^2/\text{ss}_\varphi + \tau^2))$$

for $i \in V^\circ$, where $V^\circ = \{i \in V \mid \partial i \subseteq V\}$. By assuming that the activation surface $A_i(x)$ can be approximated by a plane locally around i , we have that $A_i(x) = \bar{A}_i(x)$ and hence that

$$\frac{9}{8|V^\circ|} \sum_{i \in V^\circ} \left(\tilde{Y}_i - \bar{Y}_i \right)^2 \quad (8)$$

is an unbiased and consistent estimator for $\sigma^2/\text{ss}_\varphi + \tau^2$. When the approximation is not exact, we will get a slight positive bias in the estimate for τ^2 .

4 Simulating from the posterior distribution

In order to explore the posterior distribution of the activation centres given the data, we have designed a Metropolis-Hastings algorithm based on the Geyer and Møller (1994) algorithm for general finite point processes. This algorithm is a special case of the reversible-jump algorithm of Green (1995), where the Jacobian term is always one. Let x be the current point configuration. We will then propose to 1) insert a new point, 2) remove an existing point or 3) change an existing point, with probabilities p_1 , p_2 and p_3 respectively, where $p_1 + p_2 + p_3 = 1$. By ‘‘change an existing point’’ we mean that one of the coordinates of the point is changed, either the position or one of the marks.

Let $q_m(x' \mid x)$ denote the proposal density of a new configuration x' based on the current configuration x with move type $m = 1, 2, 3$. The probability of accepting the move is then respectively

$$\begin{aligned} \alpha_1(x, x') &= \min \left\{ \frac{p(x' \mid Y)q_2(x \mid x')p_2}{p(x \mid Y)q_1(x' \mid x)p_1}, 1 \right\}, \\ \alpha_2(x, x') &= \min \left\{ \frac{p(x' \mid Y)q_1(x \mid x')p_1}{p(x \mid Y)q_2(x' \mid x)p_2}, 1 \right\}, \\ \alpha_3(x, x') &= \min \left\{ \frac{p(x' \mid Y)q_3(x \mid x')}{p(x \mid Y)q_3(x' \mid x)}, 1 \right\}. \end{aligned}$$

If the move is rejected, the Markov chain stays in x . The proposal distributions are described in detail in the following.

4.1 Insertion of a point

With probability p_1 we propose to add a new point $\xi = (\mu, a, d, r, \theta)$ to the existing point configuration $x = \{x_1, \dots, x_n\}$. In order to obtain a reasonable acceptance rate for this move, we wish to perform a Gibbs-like update and sample the parameters from a density proportional to the conditional intensity $p(x \cup \xi \mid Y)/p(x \mid Y)$. However this is a distribution on the six dimensional space of points and marks and it is not possible to simulate directly from it. Instead, we will propose the parameters (μ, a, d, r, θ) sequentially, such that the proposal $q_1(x \cup \xi \mid x)$ is a combination of the terms

$$q_1(x \cup \xi \mid x) = q(\mu \mid x)q(a \mid \mu, x)q(d \mid \mu, a, x)q(r \mid \mu, a, d, x)q(\theta \mid \mu, a, d, r, x), \quad (9)$$

where we use the generic symbol $q(\cdot|\cdot)$ for a proposal density. We will choose the proposal of a single parameter, a say, such that it resembles the conditional intensity of a point $(\mu, a, d_0, r_0, \theta_0)$ given the current configuration x , where (d_0, r_0, θ_0) are fixed typical values for the remaining parameters and μ is the proposed position of the point. In our applications we have chosen $a_0 = 0.01$, $d_0 = 50 \text{ mm}^2$ (corresponding to about 14 voxels in our data), $r_0 = 0.5$ and $\theta_0 = 0$. Generally we simulate from discretized approximations to the conditional intensities, the details are given below.

Using (5) we find that when ignoring the priors, the conditional intensity of a new point ξ given x is

$$\frac{p(Y|x \cup \xi)}{p(Y|x)} = \exp \left\{ -\frac{1}{2(\sigma^2/\text{ss}_\varphi + \tau^2)} \left(\sum_{i \in V} h(i; \xi)^2 - 2 \sum_{i \in V} h(i; \xi)(\tilde{Y}_i - A_i(x)) \right) \right\}. \quad (10)$$

By approximating the discrete sum by an integral, we find

$$\begin{aligned} \sum_{i \in V} h(i; \xi)^2 &\simeq \iint a^2 \exp \left\{ -\frac{2\pi \log 2}{d} \left(\frac{x_\theta^2}{r/(1-r)} + \frac{y_\theta^2}{(1-r)/r} \right) \right\} dx dy / (v_x v_y) \\ &= \iint a^2 \exp \left\{ -\frac{2\pi \log 2}{d} (x^2 + y^2) \right\} dx dy / (v_x v_y) \\ &= a^2 d / (2 \log 2 v_x v_y) = a^2 \tilde{d} / (2 \log 2), \end{aligned} \quad (11)$$

where v_x and v_y are the length of the voxel sides in mm and $\tilde{d} = d/(v_x v_y)$ is the area measured in voxels. Above (x_θ, y_θ) represents a translation and rotation of (x, y) , and the second equality follows since this transformation together with the coordinate scaling has Jacobian one.

When proposing the position μ we will fix the remaining parameters at $(a_0, d_0, r_0, \theta_0)$ and approximate the intensity in (10) with a voxel-wise constant density;

$$q(\mu|x) \propto \exp \left\{ \frac{1}{(\sigma^2/\text{ss}_\varphi + \tau^2)} \sum_{i \in V} h(i; \mu, a_0, d_0, r_0, \theta_0)(\tilde{Y}_i - A_i(x)) \right\} \quad \text{for } \mu \in V.$$

The log-proposal is thus proportional to the convolution of the residual image with the typical activation centre, which may be calculated efficiently in the Fourier domain, see Press *et al.* (1992). Considering (10) as a function of the height a , the proposal density is

$$\begin{aligned} q(a|\mu, x) \\ \propto \exp \left\{ -\frac{1}{2(\sigma^2/\text{ss}_\varphi + \tau^2)} \left(\frac{a^2 d_0}{2 \log 2 v_x v_y} - 2a \sum_{i \in V} h(i; \mu, 1, d_0, r_0, \theta_0)(\tilde{Y}_i - A_i(x)) \right) \right\}, \end{aligned} \quad (12)$$

which is a Gaussian distribution,

$$a|\mu, x \sim N \left(\frac{\sum_{i \in V} h(i; \mu, 1, d_0, r_0, \theta_0)(\tilde{Y}_i - A_i(x))}{\tilde{d}_0/(2 \log 2)}, \frac{\sigma^2/\text{ss}_\varphi + \tau^2}{\tilde{d}_0/(2 \log 2)} \right),$$

restricted to the interval $(0, C_a]$. As for the three remaining parameters (d, r, θ) we will approximate the conditional intensity with a piecewise log-linear intensity, and sample from the corresponding distribution. When proposing d we will select a grid $(\delta_0, \dots, \delta_m)$ such that $\delta_0 = 0$, $\delta_m = C_d$ and let

$$q(d|\mu, a, x) \propto \exp \left\{ p_{i-1} + \frac{p_i - p_{i-1}}{\delta_i - \delta_{i-1}}(d - \delta_{i-1}) \right\} \quad \text{for } d \in (\delta_{i-1}, \delta_i],$$

where

$$p_i = -\frac{1}{2(\sigma^2/\text{ss}_\varphi + \tau^2)} \left(\delta_i \frac{a^2}{2 \log 2 v_x v_y} - 2 \sum_{i \in V} h(i; \mu, a, \delta_i, r_0, \theta_0)(\tilde{Y}_i - A_i(x)) \right) - 3 \log \delta_i - \beta_d/\delta_i, \quad (13)$$

for $i = 1, \dots, m-1$, $p_0 = p_1$ and $p_m = p_{m-1}$. Above the last two terms stem from the prior for d . The expressions for $q(r|\mu, a, d, x)$ and $q(\theta|\mu, a, d, r, x)$ are derived similarly.

4.2 Removal of a point

With probability p_2 we propose to remove a point. If the current configuration x is empty we do nothing, otherwise we select the candidate from the points in x with equal probability $1/n(x)$.

4.3 Moving a point

With probability p_3 we propose to change a parameter of a randomly selected point. We choose one of the parameters μ, a, d, r or θ with equal probability and a new value is proposed by considering the conditional distribution given the other parameters.

Suppose for instance that a point $\xi = (\mu, a, d, r, \theta) \in x$ has been selected and we wish to propose a new position μ' for ξ . Corresponding to the insertion of a new point above, we will then propose the position by simulating from a distribution which has voxel-wise constant density

$$q(\mu'|x) \propto \exp \left\{ \frac{1}{(\sigma^2/\text{ss}_\varphi + \tau^2)} \sum_{i \in V} h(i; \mu', a, d, r, \theta)(\tilde{Y}_i - A_i(x \setminus \xi)) \right\}, \quad \mu' \in V.$$

For the parameters r, d and θ we consider a neighbourhood of the current value, and approximate the conditional density as in (13) above. In our application, we have chosen a neighbourhood of 100 mm^2 for d , 0.3 for r and 0.35 for θ .

Finally, the height a is simulated from a normal distribution as when proposing a new point,

$$a|x \sim N \left(\frac{\sum_{i \in V} h(i; \mu, 1, d, r, \theta)(\tilde{Y}_i - A_i(x \setminus \xi))}{\tilde{d}/(2 \log 2)}, \frac{\sigma^2/\text{ss}_\varphi + \tau^2}{\tilde{d}/(2 \log 2)} \right).$$

5 Applications

5.1 A simulation study

We simulated data from the model (4) with a known activation pattern A , displayed in Figure 2. The pattern mimics a “true” activation image, in the sense that it has coherent regions of activation of both small and moderate size, and by the fact that it is more complex than any single realization of the stochastic geometry model.

We performed an informal sensitivity analysis, where different parameters of the prior were studied. We used three different values of a constant intensity $\beta(\mu) = \beta$ (0.01, 0.1, 1.0) and two different values of respectively β_a (0.02, 0.05) and β_d (50 mm², 200 mm²) (assuming a voxel-size of 3.52 mm² as for the real data.) For all runs, we set $\beta_r = 5$. For each parameter combination, we obtained 600000 iterations of the MCMC algorithm as described in the previous section, where at each step one of the points were updated, or a change in the number of points was proposed. By diagnostic plots, the chains were judged to be stationary after a burn-in of 100000 iterations, and we subsampled every 100'th iteration from this point to obtain 5000 samples.

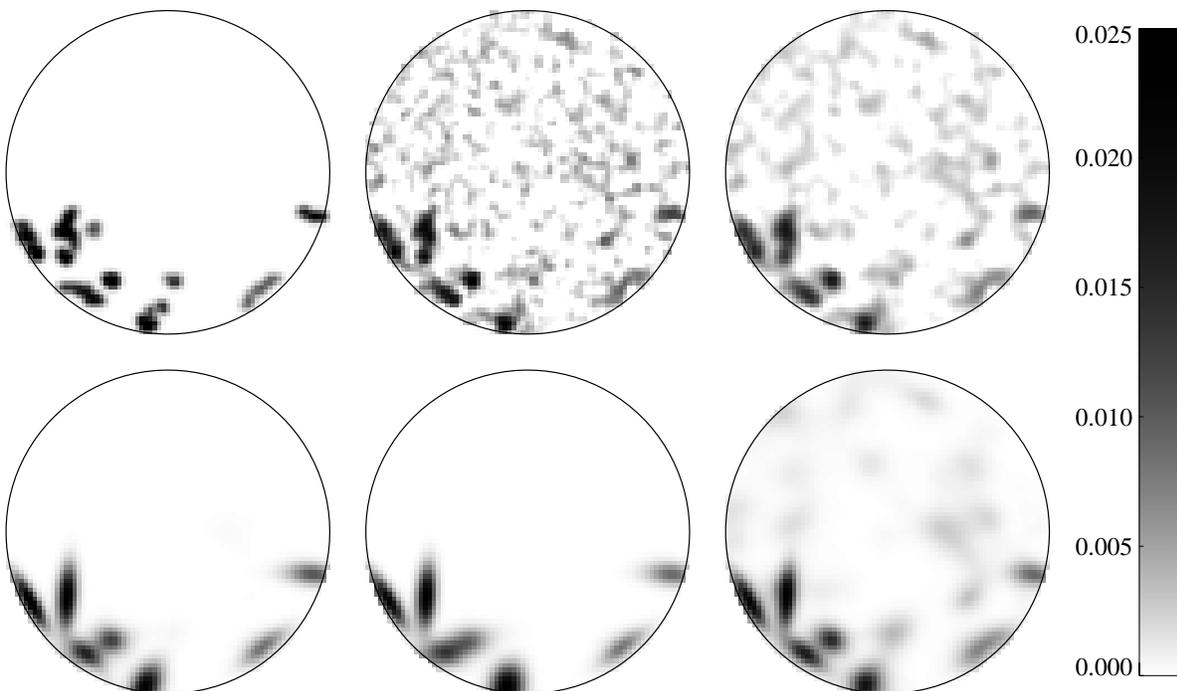


Figure 2: Top row: True activation pattern (left) and non-parametric estimates with kernel-width 2 voxels (middle) and 3 voxels (right). Bottom row: Estimates of posterior mean activation from the model. Left: $\beta = 0.1$, $\beta_d = 200$ mm², $\beta_a = 0.05$, middle: $\beta = 0.01$, $\beta_d = 200$ mm², $\beta_a = 0.05$, right: $\beta = 1.0$, $\beta_d = 50$ mm², $\beta_a = 0.02$. For display purposes, the intensities in the images in the top row are clipped to the range of the colour scale. The actual ranges are respectively [0,0.045], [-0.009,0.025] and [-0.007,0.018].

Figure 2 shows usual non-parametric estimates, obtained by smoothing the regression image $\{\tilde{Y}_i\}$ with a Gaussian kernel of full-width-at-half-maximum (FWHM) 2 and 3 voxels respectively. These are typical choices of the kernel-width in the neuroimaging literature; often a width of 3 voxels is used to ensure that the estimate is sufficiently smooth to approximate a continuous random field. The figure displays three posterior mean activation estimates as well, where the effect of the prior parameters is evident: When restricting the number of points by reducing β (middle panel), the two regions in the lower left part of the image, is merged to one. On the other hand, when the insertion of new bells is encouraged, by increasing β and reducing β_d and β_a (right panel), the activation pattern is sensitive to noise, and is more similar to the non-parametric estimates. The parameter setting of the left-most panel is the best compromise between robustness and sensitivity in this case, as measured by the posterior mean L^2 -distance between the model activation pattern and the true pattern (Table 1). As a typical summary statistic of interest, we also consider the posterior mean and standard deviation of the integrated activation, i.e. the integral of the activation surface. The true value is 3.49, which is within one standard deviation from the mean for the first parameter combination.

Table 1: Goodness-of-fit (GOF) of the model with different parameter values. The mean L^2 -distance between the model activation and the true pattern is used as a goodness-of-fit measure. Standard errors due to simulation are given in parentheses. The last column contains the posterior mean and standard deviation of the integrated activation, i.e. the total mass under the activation surface. The true value is 3.49. The two last lines display corresponding values for the non-parametric estimates.

β	β_d	β_a	GOF* (st.err.)	Int. act. (st.dev.)
0.1	200	0.05	129.2 (0.23)	3.72 (0.34)
1.0	200	0.05	130.2 (0.21)	4.04 (0.36)
0.01	200	0.05	135.9 (0.28)	3.65 (0.27)
1.0	50	0.05	130.3 (0.21)	3.96 (0.35)
0.1	50	0.05	131.4 (0.21)	3.67 (0.30)
1.0	50	0.02	138.5 (0.22)	5.63 (0.43)
1.0	200	0.02	141.5 (0.25)	5.92 (0.48)
NP, FWHM 2			174.2	3.02
NP, FWHM 3			143.0	2.99

*Scaled by 10^3

As it is often the case in Bayesian image analysis, we have no rigorous method for selecting the parameters of the prior. A simple approach is to use simulation studies like this, to determine sensible combinations. Furthermore, the parameters β_a and β_d are directly related to the size and magnitude of activation clusters, hence sensible values of these may be determined from previous experience.

Clearly the model-based approach is very different from the non-parametric procedure, as the inferential tool in the former is an entire distribution rather than only a point

estimate. We may, however, still perform an isolated comparison of estimates obtained by the two procedures, both visually in Figure 2 and by the summary values in Table 1. In both cases the model-based estimate seems to be much more precise than the FWHM 2 estimate, and slightly better than the FWHM 3 estimate. An argument against this comparison, however, is the fact that the non-parametric images are often thresholded, and only the set of supra-threshold voxels is used as an activation estimate. The plot of the number of supra-threshold voxels as a function of threshold level in Figure 3 allows for a comparison of the estimates from this point of view. Here there is only a slight difference between the FWHM 2 estimate and the model-based one. The FWHM 3 estimate, however, tends to oversmooth the activation pattern much more than the two others.

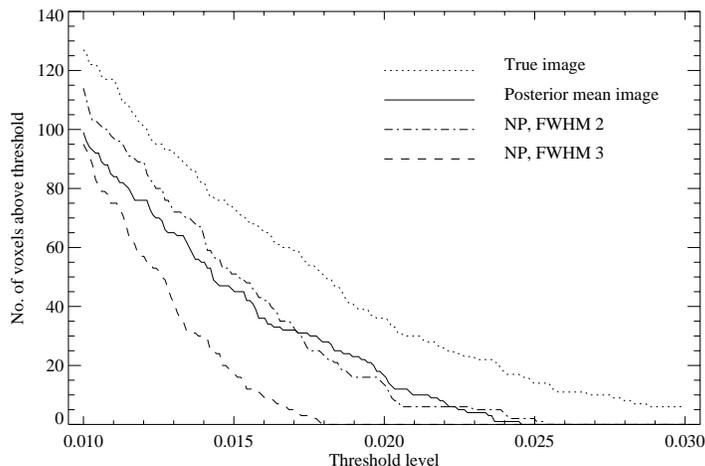


Figure 3: The number of supra-threshold voxels as function of threshold level for the true image, the posterior mean estimate and for the non-parametric estimates with FWHM 2 and 3 voxels respectively.

5.2 An analysis of visual stimulation data

We selected one of the five slices in the visual stimulation data described in Section 2, and analyzed it by the stochastic geometry model. The variances were estimated to $\hat{\sigma} = 0.0294$ and $\hat{\tau} = 0.00421$, and based on the simulation study, we set $\beta(\mu) = 0.1$, $\beta_d = 200 \text{ mm}^2$, $\beta_a = 0.05$ and $\beta_r = 5$.

Since problems with mixing was more prominent with this data, we chose a more elaborate rule for updating points in the MCMC algorithm. At each iteration, where a point-update was proposed, all the points were considered after turn, in a random ordering, and all parameters of each point were updated. One iteration of this kind thus corresponds to a collection of about 100 simple single-parameter updates used in the previous section, depending on the number of points. We found that a systematic update of all points in this fashion improved the acceptance rates for insertion and deletion of

points, and thus the mixing of the algorithm. After a burn-in run, we generated 75000 iterations from this modified algorithm and subsampled every 10'th sample. The plots in Figure 4 show diagnostic variables of the simulated point process, namely the number of points and the log-posterior and log-prior densities. Though the autocorrelation is reasonably low, the mixing of the algorithm may be improved, for instance by simulated tempering (Geyer and Thompson, 1995). The acceptance probabilities for the different move types are listed in Table 2.

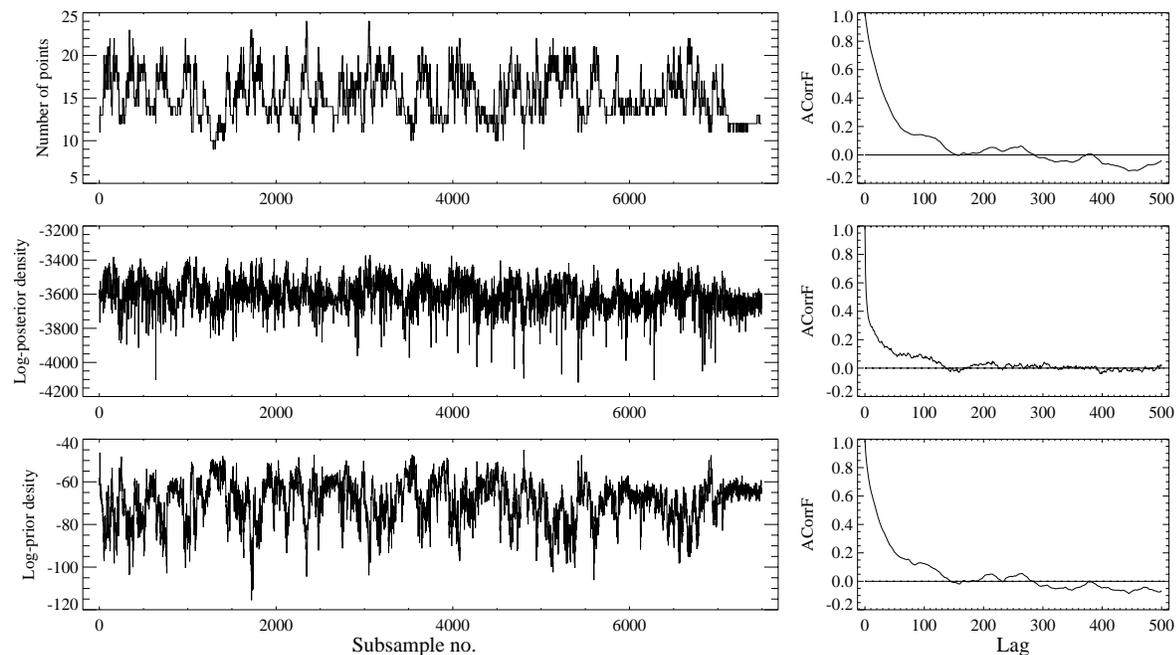


Figure 4: Time series of summary statistics obtained by subsampling every 10th iteration of the modified MCMC algorithm. Left: Simulated values of the number of points, log-posterior and log-prior densities. Right: Empirical auto-correlation functions for the time series.

The images in Figure 5 show the posterior mean activation level, calculated voxel-wise in the present model with uncorrelated noise and in a model with a separable spatio-temporal correlation function to be described later. The area with high activation intensity in the back of the brain corresponds well with the visual cortex, which is known to process visual impressions. Displayed for comparison is also the non-parametric estimate with kernel-width 3 voxels (FWHM). The problem with oversmoothing seems to be present also in this example, though the true scene is not known for these data. On the other hand, the posterior mean images possess some very eccentric regions, which are most likely artifacts. The simplistic noise model is one likely reason for this. This is both intuitively clear, since it is well known that the noise may be complex in fMRI data, and can also be seen by comparing the improvement in the estimate in the model with correlated noise compared to the independent noise model. Furthermore we did not observe this effect in the simulated data, where the noise was independent. We thus

Table 2: Acceptance probabilities for the different move types in the MCMC algorithm. The correlated noise model will be described in Section 6.2.

Move type	Acceptance (%)	Acceptance (%)
	Independent noise	Correlated noise
Insert point	15.7	13.9
Delete point	15.6	14.1
Update position	10.6	11.5
Update height	40.8	45.2
Update area	30.6	34.3
Update angle	64.7	66.6
Update ratio	55.9	59.4

suggest that more elaborated noise modelling is a key factor in addressing this, and will discuss an extension to correlated noise in Section 6.2. Another approach is to penalize eccentric ellipses more in the prior distribution. This seems sensible, since relatively few voxels are affected by a thin ellipse, and thus the likelihood function contains little information on these. One problem, however, is the fact that the imaged slice is a section through the convoluted cortical surface. Eccentric ellipses may thus arise naturally when a circular area on the cortical surface is transected orthogonally by the image-plane. A natural, but also much more ambitious extension, is thus to extract the geometry of the cortical surface from high-resolution anatomical scans, and formulate the model directly on the two dimensional surface. Recent works along this line are Andrade *et al.* (2000) and Kiebel *et al.* (2000).

A more fundamental difference between the model-based and non-parametric approach, is the possibility of attaching estimates of uncertainty to the images. As an example of this, displayed is also a conservative activation estimate, obtained by subtracting two times the voxel-wise standard deviation from the mean image. This may be considered as a thresholded version of the original mean image, but where the threshold is voxel-dependent, to reflect the voxel-wise uncertainty. This is only one possible way of visualizing the posterior variance of activation pattern, and the flexibility of the MCMC approach may of course be exploited in a range of other ways. One possibility is to estimate summary characteristics of the activation pattern with associated standard errors, which may be used to quantify how well data support specific neuroscientific hypothesis.

6 Extensions

So far we have established a basic framework for analyzing fMRI data by a simple spatio-temporal model. However, both the noise and signal may possess more complex structures, which this model does not account for, necessitating several refinements and extensions. We will briefly discuss relevant extensions in the following and illustrate how the inferential framework may be modified suitably.

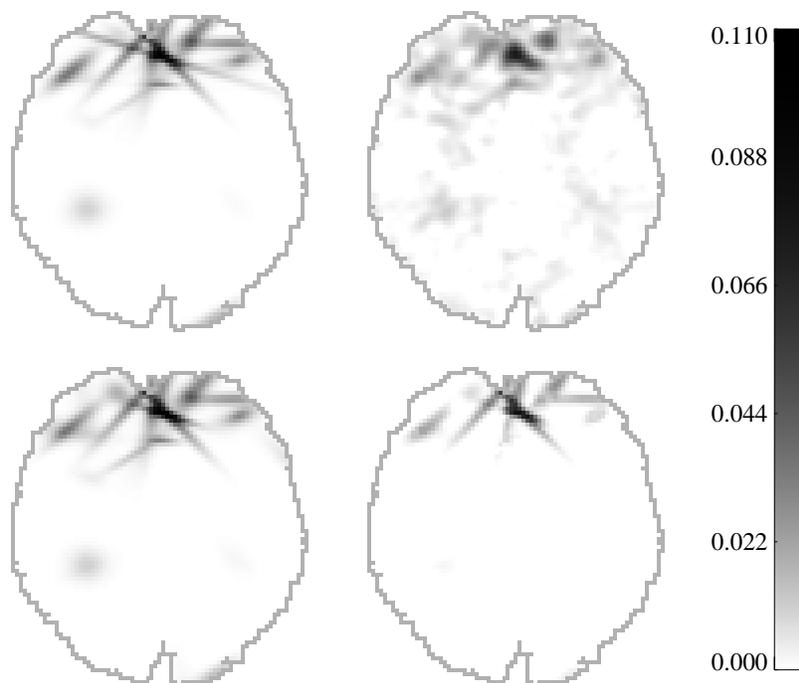


Figure 5: Activation estimates for the visual stimulation data. Top: Mean posterior activation in the stochastic geometry model with uncorrelated noise (left) and the non-parametric estimate with kernel-width 3 voxels (right). Bottom: Mean posterior activation in the model with a separable spatio-temporal covariance function described in Section 6.2 (left), and a conservative estimate in this model, obtained by subtracting twice the standard deviation from the mean activation level in each voxel (right).

6.1 Non-stationary responses

We have assumed that the temporal pattern φ is the same for all voxels, to obtain a simple spatio-temporal model. It is well known, that this is only an approximation, and more general approaches are studied for instance by Lange and Zeger (1997) and Genovese (2000), who explicitly account for differences in delay from one voxel to another. In practice, the approximation will be relatively good for blocked paradigms, where the stimulus is presented for longer periods of time, but problematic for so-called event-related designs, where the presentation changes rapidly. A natural extension to improve this, is to linearly combine the simple response function with its derivatives with respect to different parameters (Friston *et al.*, 1998). These may, in a Taylor-like fashion, account for small voxel-wise differences in the delay and dispersion.

A fundamental question, which is more challenging to address, is whether the response is stationary over time, or if it changes with general alertness and learning as suggested by Gaschler-Markefski *et al.* (1997). To model non-stationary responses we may consider

the state-space model

$$\varphi_t = \lambda_t + \nu_t, \quad \nu_t - \nu_{t-1} \sim N(0, \kappa^2), \quad t = 1, 2, \dots, m, \quad (14)$$

where $\{\nu_t - \nu_{t-1}\}$ are independent and $\nu_0 = 0$. The mean λ_t is the simple model (3) which reflects the overall temporal structure, but φ is allowed to deviate from this via the random walk structure of the noise terms ν_t . The variance κ^2 governs the smoothness of the residual process $\varphi_t - \lambda_t$.

By combining this with the spatial prior (2), we can make inference on (X, φ) through the joint posterior distribution $P(X, \varphi|Y)$. For computational reasons we will in fact consider the posterior of (X, φ, η) , where $\eta = \{\eta_i, i \in V\}$ are the random intercepts in the model (4). This is given by

$$p(X, \varphi, \eta|Y) \propto P(Y|X, \varphi, \eta)P(X)P(\varphi)P(\eta),$$

where the likelihood term is obtained by conditioning on η in (4). A Markov chain with the posterior as invariant distribution may be generated by a variable-at-a-time Metropolis-Hastings algorithm, where iteratively one parameter is updated given the two others. When updating X the proposals are as described earlier, though with the modification that we replace $A_i(x)$ with $A_i(x) + \eta_i$ and set $\tau^2 = 0$ in the formulas in Section 4 to condition on η . A similar modification applies to the likelihood function in (5), when calculating the acceptance ratio. When updating η or φ we can simulate directly from the conditional distributions, as it can easily be verified that

$$\eta_i|Y, X, \varphi \sim N\left(\frac{\tau^2}{\tau^2 + \sigma^2/ss_\varphi}(\tilde{Y}_i - A_i(X)), \tau^2\left(1 - \frac{\tau^2}{\tau^2 + \sigma^2/ss_\varphi}\right)\right), \quad (15)$$

with all η_i 's conditionally independent. The simulation of φ may be carried out recursively by simulating $(\varphi_t|\varphi_{t+1}, \dots, \varphi_m, X, \eta, Y)$ for $t = m, m-1, \dots, 1$. These are all normal distributions, and the moments may be calculated efficiently with the Kalman smoother (West and Harrison, 1989).

We estimated the temporal response using the visual stimulation data of Section 2, though preprocessed in a slightly different way, as we removed some low-frequency trends with large magnitude to stabilize the algorithm. The plots in Figure 6 illustrate a simple least squares estimate and the posterior mean of φ . The simple estimate is obtained by assuming that the spatial activation pattern is fixed and given by the mean image in the top left panel in Figure 5. The posterior mean is based on 75000 simulations of the Markov chain described above, where the update rules for the points are as in Section 5.2. Both plots indicate that the actual response may in fact not be described by a fixed time-invariant model as the last peak is higher than the three first, and the dip below baseline is more prominent after the first and third cycle than after the second.

A consequence of modelling φ in this way, is that the stimulation function is only partly included in the model. Though it may seem inefficient to ignore a relevant covariate like this, in some experiments the *actual* stimulation is not directly controllable by the experimenter and hence precise information on this is not available. This may be the case in mental experiments, where it is not possible to end the stimulation at an exact

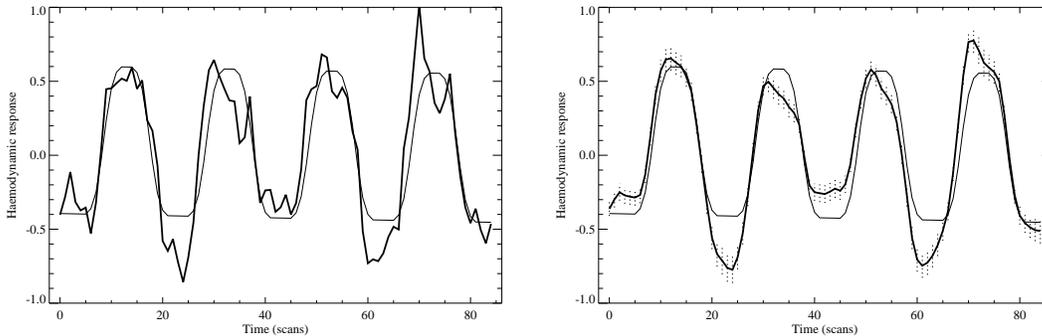


Figure 6: Left: Least squares estimate of the haemodynamic response, assuming a known spatial activation pattern given by the top left panel in Figure 5. Right: Monte Carlo estimate of the posterior mean of the haemodynamic response function based on 7500 subsamples of 75000 iterations of the MCMC algorithm (see text.) Overlaid are pointwise 95%-credibility regions based on the posterior variance. The initial stationary model (3) is overlaid as a thin line in both plots.

time point. Furthermore with this formulation, we may detect subtle activation patterns, which depends on the paradigm in more complex ways. An example of the latter is the XOR signal of Lange *et al.* (1999).

As a further extension a collection of different response functions could be modelled by a multidimensional state space model for φ . By assigning different functions to different groups of centres, one may account for regional differences in the response. At least for moderate dimensions of φ the recursive simulation routine would still be very efficient.

6.2 Correlated noise

As discussed earlier the initial model with uncorrelated noise is too simple in practice. The noise sources in fMRI data are both of physiological and physical origin. The pixel values are constructed by inverse Fourier transforms of measurements of currents in a coil over a short time period. Hence there is no physical separation of the pixels, which could justify independence. The temporal correlation arises from physiological sources, but also intrinsically in the MR scanner.

The main problem with modelling a general covariance function is the fact that the likelihood function must be easily calculated, in order that the MCMC algorithm is reasonably fast. A tractable starting point is to assume a separable covariance function. We have previously studied the empirical correlation of the visual stimulation data (Hartvig, 1999), and found evidence that the temporal covariance varied slightly with spatial location, but that a separable model was a reasonable approximation to the true covariance function. Letting $\varepsilon = \{\varepsilon_{it}, i \in V, t = 1, \dots, m\}$ be the noise terms in (4) regarded as a $|V| \times m$ matrix, we will thus consider the model

$$\varepsilon \sim N_{|V| \times m}(0, \sigma^2 \Lambda \otimes \Gamma).$$

where \otimes denotes the Kronecker product and where Λ and Γ are the spatial and temporal correlation matrices of dimension $|V| \times |V|$ and $m \times m$, respectively.

Let L and M denote the lower-triangular Cholesky square roots of Λ and Γ . Consider the data Y as a $|V| \times m$ matrix and let $Y^\circ = Y(M^{-1})'$, and $\varphi^\circ = M^{-1}\varphi$. Then the conditional likelihood function, where we condition on η , is given by

$$p(Y|x, \varphi, \eta) = (2\pi\sigma^2)^{-\frac{m|V|}{2}} |M|^{-|V|} |L|^{-m} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{t=1}^m \|L^{-1}(Y_{*t}^\circ - \tilde{Y}^\circ \varphi_t^\circ)\|^2 \right\} \\ \times \exp \left\{ -\frac{1}{2\sigma^2/ss_\varphi^\circ} \|L^{-1}(\tilde{Y}^\circ - A - \eta)\|^2 \right\}, \quad (16)$$

where

$$\tilde{Y}_i^\circ = \sum_{t=1}^m Y_{it}^\circ \varphi_t^\circ / ss_\varphi^\circ \quad \text{and} \quad ss_\varphi^\circ = \sum_{t=1}^m \varphi_t^{\circ 2}$$

are defined equivalently to (6), and $Y_{*t}^\circ = \{Y_{it}^\circ\}_{i \in V}$. Notice that equations of the form $v = L^{-1}w$ may be solved easily, due to the lower-triangularity of L . Suppose good estimates are available for Γ and Λ , such that the uncertainty of these can be ignored. By simply inserting the expression above in the Metropolis-Hasting ratio instead of (5), the MCMC algorithm will converge to the posterior distribution in the correlated noise model.

The problem is thus reduced to obtaining estimates of Γ and Λ , or rather the corresponding Cholesky decompositions of these. Due to the large number of voxels, it is generally very hard to decompose Λ , which preclude the use of Gaussian models parametrized in terms of elements of the covariance matrix (Cressie, 1991). As a pragmatic alternative, in Hartvig (1999) we proposed a stationary moving average-type model, parametrized directly by L . For a given ordering of the voxel indices (corresponding to the ordering of the elements of the matrix Λ), we let $L_{ij} = l_{i-j}$, where $l_k = 0$ unless $k \in D$ and D is a set of neighbours in the positive direction. We chose the lexicographic ordering of voxel indices, and though the covariance function will to some extent depend on this choice, the computational advantages of parametrizing the model in terms of L are considerable: By construction, the covariance function will always be positive definite, which simplifies parameter estimation, and both the covariance and inverse covariance matrices may easily be calculated, despite the large dimensionality.

We fitted a stationary model with six parameters to the spatial covariance and an AR(1)-model to the temporal covariance function. Using the visual stimulation data, we performed 75000 iterations of the MCMC algorithm as described in Section 5.2. The mean posterior activation image is displayed in Figure 5, and as discussed earlier, it is slightly improved compared to the original estimate.

6.3 Negative haemodynamic responses

We have restricted the activation level a_k to be positive, but in fact a decrease in the intensity during stimulation is sometimes observed. This may either be due to the fact

that an area is more active during the “rest”-condition than the stimulation condition, or it may be attributed to special haemodynamic effects. In order to extend the model to account for this, the activation may be represented as $A^+ - A^-$, where A^+ and A^- are positive surfaces describing positive and negative activation respectively. For identifiability reasons we have to incorporate an interaction term in the prior that separates the two surfaces; if not, overlapping positive and negative sites will be highly correlated, and the interpretation of the activation surface becomes very difficult.

We propose simply to use an interaction prior, which makes the two surfaces conditionally independent given the data. For the independent noise model, the prior is of the form

$$p(X^+, X^-) \propto f(X^+)f(X^-) \exp(-\alpha \sum_{i \in V} A_i(X^+)A_i(X^-)),$$

where $f(\cdot)$ is the density in (2) and where X^+ and X^- are two point processes determining the positive and negative surfaces respectively. When $\alpha > 0$, the last term penalizes configuration where $A_i(X^+)$ and $A_i(X^-)$ are both large for some $i \in V$. Let $\alpha^{-1} = \sigma^2/ss_\varphi + \tau^2$, the variance of \tilde{Y}_i , then it is easy to see that X^+ and X^- are independent given the data Y . Hence we can make inference about X^+ and X^- in their respective marginal distributions, and afterwards combine estimates using the independence of the two point processes. Clearly other types of interactions may be considered, but the present is appealing because of the advantages of marginalizing the inference, namely a reduction of the dimensionality of the point processes and improved properties of the simulation algorithm.

6.4 Interaction between centres

In Bayesian object recognition it is well known, that unless an interaction term is included in the prior, the estimate may tend to contain clusters of almost identical objects (Baddeley and van Lieshout, 1993). An extension of the prior in this respect may have the following form

$$p(x) \propto \prod_{k=1}^n \beta(\mu_k) \prod_{k < j} \phi(x_k, x_j) \prod_{k=1}^n (p(a_k)p(d_k)p(r_k)), \quad x \in \Omega,$$

where ϕ is an interaction function, which prevents centres from clustering. It is natural to consider a function with a hard-core property, which prohibits pairs of centres with distances close to zero. This is achieved with the model of Ogata and Tanemura (1984) where $\phi(x_k, x_j) = 1 - \exp\{-\delta(x_k, x_j)/\rho\}^p$, $p \geq 2$, with respect to a distance $\delta(\cdot, \cdot)$ on \mathcal{X} . Here $\rho > 0$ is an interaction radius. A hard-core Strauss model is obtained by setting $p = \infty$, while finite values of p yield an interaction function which increases continuously from 0 to 1 with the distance between two points. A natural definition of the distance δ is to let two centres be close, if they are close in space *and* have similar size and shape. One way of assessing this is by the J-divergence (Kullback, 1959) of the corresponding Gaussian functions. This does not satisfy the triangle equality, and is thus strictly speaking not a distance (in the sense of a metric), but for the present purpose the symmetry-, uniqueness- and positivity-properties are sufficient.

6.5 Generalization to three dimensions

The spatial model can straightforwardly be generalized to a three dimensional setting where $S \subseteq \mathbb{R}^3$. In this case a centre is given by $x = (\mu, a, d, r_1, r_2, \theta_1, \theta_2)$, and the contribution to the activation volume is

$$h(i; x) = a \exp \left\{ -\log 2 \left(\frac{4\pi}{3d} \right)^{2/3} \left(\frac{j_1^2}{(r_1^2/r_2 r_3)^{2/3}} + \frac{j_2^2}{(r_2^2/r_1 r_3)^{2/3}} + \frac{j_3^2}{(r_3^2/r_1 r_2)^{2/3}} \right) \right\}.$$

Here $r_3 = 1 - r_1 - r_2$, $r_k > 0$ for $k = 1, 2, 3$ and

$$(j_1, j_2, j_3) = \begin{pmatrix} \cos \theta_1 \cos \theta_2 & -\sin \theta_1 & -\cos \theta_1 \sin \theta_2 \\ \sin \theta_1 \cos \theta_2 & \cos \theta_1 & -\sin \theta_1 \sin \theta_2 \\ \sin \theta_2 & 0 & \cos \theta_2 \end{pmatrix} (i - \mu).$$

With this parametrization d is the volume of the contour ellipsoid at height $a/2$, and r_k is the ratio of the k th main axis and the sum of the three axes. The angles θ_1 and θ_2 are the rotations in the xy -plane and xz -plane respectively, which are restricted to the interval $[-\pi/4, \pi/4]$. The natural extension of the priors is to assume that $(r_1, r_2) \sim D_2(\beta_r, \beta_r)$ where D_2 is the two dimensional Dirichlet distribution.

7 Discussion

We have proposed a spatio-temporal model for fMRI data which explicitly accounts for the fact that signal changes are locally coherent in both space and time. This assumption is often implicitly included in the analysis, when spatial and temporal filtering are applied, but rarely formulated explicitly in a model. The relation (5) shows that in the simplest setting we are effectively fitting Gaussian functions of different sizes and orientations to a regression image, and assessing the significance of these. The random field theory has counterparts to this, namely the search for local maxima in both scale and space (Siegmund and Worsley, 1995), and in the space of ellipses with different orientation and shape (Shafie *et al.*, 1998). The method is, however, fundamentally different from the random field approach. The latter provides a framework for signal detection, by testing multiple null hypotheses with correction for the large number of tests performed. As was pointed out by Keith Worsley in the discussion to Lange and Zeger (1997), what is really an estimation problem is thus answered by hypothesis testing, with corresponding conceptual and mathematical problems. With the proposed method the focus is shifted towards estimation of the activation pattern by standard Bayesian methods.

Since the amount of data in fMRI experiments may be enormous, there is a compromise between model complexity and the computational burden of the analysis. In an attempt to formulate a relatively simple model, we have made specific assumptions on the spatial pattern, and clearly these may not be fully satisfied by the true activation. Though the random intercept surface η will account for minor deviations from the point process representation, the model should be applied to a range of other fMRI data for further validation.

As mentioned earlier, Kiebel *et al.* (2000) and Andrade *et al.* (2000) have recently studied global models, where the geometry of the cortical surface is used to model the haemodynamic effects. These were not formulated in a parametric framework, and assessing the uncertainty of estimated activation patterns were not considered. The geometry of the cortical surface is a very relevant covariate to be included in our setup also, for instance by formulating the model on the two dimensional surface. A major challenge with this extension, however, is the increased computational burden, which is already at the limit of what is acceptable for practical purposes.

Alternatively local spatial models have been proposed by Descombes *et al.* (1998), Salli *et al.* (2001) and Hartvig and Jensen (2000). These are Markov random field-type models, where MAP estimates are obtained iteratively or in closed form, whence MCMC is not required. The computational burden is thus much reduced, but so is the inferential scope, since only a point estimate is obtained. In the present setup, the significance of hypotheses of interest within single subjects may be quantified, or estimates and standard errors of relevant features of the activation in different experiments may be obtained, for comparing different groups of subjects.

Modelling the temporal response in a non-parametric setting with few assumptions seems relevant, given the uncertainty about the haemodynamic effects in different stimulation types. Also the fact that the modelled response depends only partly on the specified paradigm is an advantage when analyzing data where the actual paradigm is difficult to determine. The approach has some similarity with non-parametric multivariate methods, such as principal component analysis (PCA), where a representative time course and the corresponding spatial pattern is estimated directly from the data. In our setup, the time-course is also estimated from the data, but unlike in PCA, the assumptions of spatial smoothness and coherency is simultaneously taken into account.

With noticeable exceptions (Genovese, 2000; Frank *et al.*, 1998), Bayesian analyses of fMRI data are rare. We are of the opinion that a Bayesian approach to this data makes sense for several reasons. Firstly there is substantial prior information on the activation pattern, which should of course be used in the analysis. This may either be general knowledge of the functional organization of the brain, or results from earlier experiments on the same subject. The ease by which data can be acquired even allows us to design experiments according to this, by performing pilot studies to generate detailed prior information before the actual experiment. Secondly often large inter- and intra-subject variation is observed, which makes it more natural to consider the activation pattern as a realization of stochastic variable than as a fixed unknown parameter. A similar interpretation is made in the currently applied random effect analyses of Holmes and Friston (1998).

Acknowledgements

This work was supported by MaPhySto, Centre for Mathematical Physics and Stochastics, funded by a grant from the Danish National Research Foundation. The author is grateful for inspiring and constructive discussions with Jens Ledet Jensen and Rasmus

Waagepetersen. Thanks to Hans Stødkilde-Jørgensen from the MR-ResearchCentre at Skejby Sygehus for kindly providing the data, and to an anonymous referee for valuable suggestions for improving the paper.

References

- Andrade, A., Kherif, F., Mangin, J.F., Worsley, K.J., Simon, O., Dehaene, S., Le Bihan, D. and Poline, J.B. (2000) Detection of fMRI activation using cortical surface mapping. Research report, Service Hospitalier Frédéric Joliot, CEA, Orsay. Submitted.
- Baddeley, A.J. and van Lieshout, M.N.M. (1993) Stochastic geometry models in high-level vision. In K.V. Mardia and G.K. Kanji (eds.), *Statistics and Images*, vol. 1, chap. 11, pp. 231–256, Appl. Statist.
- Boynton, G.M., Engel, S.A., Glover, G.H. and Heeger, D.J. (1996) Linear systems analysis of functional magnetic resonance imaging in human V1. *J. Neurosci.*, **16**, 4207–4221.
- Bullmore, E., Brammer, M., Williams, S.C., Rabe-Hesketh, S., Janot, N., David, A., Mellers, J., Howard, R. and Sham, P. (1996) Statistical methods of estimation and inference for functional MR image analysis. *Magn. Reson. Med.*, **35**, 261–277.
- Buxton, R.B., Wong, E.C. and Frank, L.R. (1998) Dynamics of blood flow and oxygenation changes during brain activation: The balloon model. *Magn. Reson. Med.*, **39**, 855–864.
- Cressie, N.A.C. (1991) *Statistics for spatial data*. Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. New York: John Wiley & Sons, Inc.
- Descombes, X., Kruggel, F. and von Cramon, D.Y. (1998) fMRI signal restoration using a spatio-temporal Markov random field preserving transitions. *NeuroImage*, **8**, 340–349.
- Frank, L.R., Buxton, R.B. and Wong, E.C. (1998) Probabilistic analysis of functional magnetic resonance imaging data. *Magn. Reson. Med.*, **39**, 132–148.
- Friston, K.J., Jezzard, P. and Turner, R. (1994) The analysis of functional MRI time-series. *Human Brain Mapping*, **1**, 153–171.
- Friston, K.J., Holmes, A.P., Poline, J.B., Grasby, P.J., Williams, S.C.R., Frackowiak, R.S.J. and Turner, R. (1995) Analysis of fMRI time-series revisited. *NeuroImage*, **2**, 45–53.
- Friston, K.J., Fletcher, P., Josephs, O., Holmes, A., Rugg, M.D. and Turner, R. (1998) Event-related fMRI: Characterizing differential responses. *NeuroImage*, **7**, 30–40.

- Gaschler-Markefski, B., Baumgart, F., Tempelmann, C., Schindler, F., Stiller, D., Heinze, H.J. and Scheich, H. (1997) Statistical methods in functional magnetic resonance imaging with respect to nonstationary time-series auditory cortex activity. *Magn. Reson. Med.*, **38**, 811–820.
- Genovese, C.R. (2000) A Bayesian time-course model for functional magnetic resonance imaging data. *J. Amer. Statist. Assoc.*, **95**, 691–719. With discussion and a reply by the author.
- Geyer, C.J. and Møller, J. (1994) Simulation procedures and likelihood inference for spatial point processes. *Scand. J. Stat.*, **21**, 359–373.
- Geyer, C.J. and Thompson, E.A. (1995) Annealing Markov chain Monte Carlo with applications to ancestral inference. *J. Amer. Statist. Assoc.*, **90**, 909–920.
- Green, P.J. (1995) Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*, **82**, 711–732.
- Hartvig, N.V. (1999) A stochastic geometry model for fMRI data. Research report 410, Department of Theoretical Statistics, University of Aarhus.
- Hartvig, N.V. (2000) *Parametric Modelling of Functional Magnetic Resonance Imaging Data*. Ph.D. thesis, Department of Theoretical Statistics, University of Aarhus.
- Hartvig, N.V. and Jensen, J.L. (2000) Spatial mixture modelling of fMRI data. *Human Brain Mapping*, **11**, 233–248.
- Holmes, A.P. and Friston, K.J. (1998) Generalisability, random effects and population inference. *NeuroImage*, **7**, S754.
- Holmes, A.P., Josephs, O., Büchel, C. and Friston, K.J. (1997) Statistical modelling of low-frequency confounds in fMRI. *NeuroImage*, **5**, S480.
- Kiebel, S.J., Goebel, R. and Friston, K.J. (2000) Anatomically informed basis functions. *NeuroImage*, **11**, 656–667.
- Kullback, S. (1959) *Information Theory and Statistics*. John Wiley & Sons, Inc.
- Kwong, K.K., Belliveau, J.W., Chesler, D.A., Goldberg, I.E. *et al.* (1992) Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proc. Natl. Acad. Sci. USA*, **89**, 5675–5679.
- Lange, N. (1996) Tutorial in biostatistics. Statistical approaches to human brain mapping by functional magnetic resonance imaging. *Statistics in Medicine*, **15**, 389–428.
- Lange, N. and Zeger, S.L. (1997) Non-linear Fourier time series analysis for human brain mapping by functional magnetic resonance imaging. *Appl. Statist.*, **46**, 1–29. With discussion and a reply by the authors.

- Lange, N., Strother, S.C., Anderson, J.R., Nielsen, F.A., Holmes, A.P., Kolenda, T., Savoy, R. and Hansen, L.K. (1999) Plurality and resemblance in fMRI data analysis. *NeuroImage*, **10**, 282–303.
- Ogata, Y. and Tanemura, M. (1984) Likelihood analysis of spatial point patterns. *J. R. Statist. Soc. B*, **46**, 496–518.
- Press, W.H., Teukolsky, S.A., Vetterling, W.T. and Flannery, B.P. (1992) *Numerical Recipes in C*. Cambridge University Press, second edn.
- Salli, E., Korvenoja, A., Visa, A., Katila, T. and Aronen, H.J. (2001) Reproducibility of fMRI: Effect of the use of contextual information. *NeuroImage*, **13**, 459–471.
- Shafie, K., Worsley, K.J., Wolforth, M. and Evans, A.C. (1998) Rotation space: Detecting functional activation by searching over rotated and scaled filters. *NeuroImage*, **7**, S755.
- Siegmund, D.O. and Worsley, K.J. (1995) Testing for a signal with unknown location and scale in a stationary Gaussian random field. *Ann. Stat.*, **23**, 608–639.
- West, M. and Harrison, J. (1989) *Bayesian Forecasting and Dynamic Models*. Springer-Verlag, New York.
- Worsley, K.J. (1995) Estimating the number of peaks in a random field using the Hadwiger characteristic of excursion sets, with applications to medical images. *The Annals of Statistics*, **23**, 640–669.
- Worsley, K.J. and Friston, K.J. (1995) Analysis of fMRI time-series revisited — again. *NeuroImage*, **2**, 173–181.