doi: 10.1111/j.1467-9469.2006.00554.x
 Vol 34: 587–614, 2007

A Spatio-Temporal Model for Functional Magnetic Resonance Imaging Data – with a View to Resting State Networks

EVA B. VEDEL JENSEN

Department of Mathematical Sciences, The T.N. Thiele Centre, University of Aarhus

THORDIS L. THORARINSDOTTIR

Department of Mathematical Sciences, The T.N. Thiele Centre, and The MR Research Centre, Skejby University Hospital, University of Aarhus

ABSTRACT. Functional magnetic resonance imaging (fMRI) is a technique for studying the active human brain. During the fMRI experiment, a sequence of MR images is obtained, where the brain is represented as a set of voxels. The data obtained are a realization of a complex spatio-temporal process with many sources of variation, both biological and technical. We present a spatio-temporal point process model approach for fMRI data where the temporal and spatial activation are modelled simultaneously. It is possible to analyse other characteristics of the data than just the locations of active brain regions, such as the interaction between the active regions. We discuss both classical statistical inference and Bayesian inference in the model. We analyse simulated data without repeated stimuli both for location of the activated regions and for interactions between the activated regions. An example of analysis of fMRI data, using this approach, is presented.

Key words: Bayesian analysis, functional magnetic resonance imaging, image analysis, interaction, point processes, resting state networks, spatio-temporal modelling

1. Introduction

Functional magnetic resonance imaging (fMRI) is a non-invasive imaging technique that has been available for about 10 years. Cognitive psychologists and neuroscientists have shown an enormous interest in fMRI because it is believed that fMRI can reveal the human brain in action. There is a comprehensive literature on the topic, mainly in *Human Brain Mapping*, *Magnetic Resonance in Medicine* and *NeuroImage*, reporting various empirical findings and new methods of analysis.

During a typical fMRI experiment, the subject is asked to perform specific behavioural tasks (like finger-tapping or calculations) or the subject is exposed to passive stimulus (like flashing light). The experiment is carefully designed with periods of rest (off periods) between periods of stimuli (on periods). The brain is scanned during the experiment and represented as a set of voxels. At each voxel a time series is recorded, showing the local brain activity during the experiment. An informative introduction for statisticians to the design of fMRI experiments can be found in the paper by Genovese (2000).

The analysis of fMRI data is usually aimed at localizing the activated or deactivated parts of the brain during the experiment. The initial analysis is often performed voxel-wise, using the time series available at each voxel. The variation in the local signal intensity is analysed using a temporal model, involving the known design of the experiment and the haemodynamic response function. Using this technique, local activation estimates based on level changes during on and off periods are assessed. Spatial modelling of fMRI data is usually carried out after the image of voxel-wise activation estimates (e.g. an image

of p-values for activation tested by t-tests) is obtained. The most common approach is to use Gaussian random field theory for this part of the modelling (see Friston *et al.*, 1995; Cao & Worsley, 1999). The approach is not without problems as the threshold value will depend on the search volume. This type of procedure, involving generalized linear models, has been implemented in the Statistical Parametric Mapping (SPM) software package. The package has been developed by members and collaborators of the Wellcome Department of Imaging Neuroscience (UCL, UK).

In Genovese (2000), a fully Bayesian analysis of fMRI data is discussed (see also Friston, 2002; Friston *et al.*, 2002a, b). The model still only involves one voxel at a time but is very heavy computationally. In the comments to Genovese (2002) (see Worsley, 2000), it is suggested to try to spatially link the voxel-wise models. In recent times, independent component analysis (ICA) has become quite popular (see Stone, 2002; McKeown *et al.*, 2003; Beckmann & Smith, 2005). See also the early critical comments in Friston (1998). Techniques for detecting functional clusters have been described in Tonini *et al.* (1998).

Especially amongst psychologists there has been some criticism of the localization paradigm. They argue that psychological processes are probably not realized as static constellations. Also, it is believed that the repeated stimulus experiments are artificial. In Greicius *et al.* (2003), the functional connectivity in the resting brain is studied. In particular, the hypothesis of a default mode network is examined. Regions of interest, being deactivated during a cognitive task, are found to be interacting during periods of rest without particular stimulus. This finding is obtained, using an unconventional, but natural type of analysis. The average time series from one region is used as an explanatory variable in the analysis of the time variation in other regions of the brain. Here, it is of interest to try to develop models that can justify this type of data analysis.

In a way, these developments are a consequence of the fact that fMRI is a more mature field now. Instead of seeking the locations of active brain regions, the focus is on the interaction between the active regions. This change of paradigm has consequences for the choice of appropriate method of analysis. Instead of looking for changes in level it seems to be more promising to study the covariation between the time series.

A first attempt to provide a modelling framework for experiments without repeated stimuli is outlined in the present paper. Such an experiment will be called a non-stimulus experiment. A simple simulated example of the experimental situation we have in mind is shown in Fig. 1. Here, the MR signal intensity is observed in a two-dimensional slice of the 'brain' in the time interval (arbitrary units) [0,100]. In Fig. 1, the development of the activity over time is shown, from $t=5$ (upper left) to $t=95$ (lower right) in jumps of 10 time units. Note that three regions of the brain simultaneously light up. The crucial point is that the random time points of activation is unknown to the experimenter. The aim of the analysis of the experiment is to find the areas of the brain that are simultaneously activated.

Our modelling approach is based on spatio-temporal point processes. Purely spatial processes have earlier been used in Taskinen (2001) and Hartvig (2002). Here, the spatial activation is modelled by Gaussian bell functions centred around the points from a spatial point process. As an example, the model studied in Hartvig (2002) is in its simplest form as follows

$$Z_{tx} = \sum_j h(x - x_j)\varphi_t + \sigma\varepsilon_{tx},$$

where Z_{tx} is the observed MR signal intensity at time t and voxel x, $\{x_j\}$ is a point process defined on the brain, h is a Gaussian density function with mean 0 and independent components, φ_t is a regression variable, containing information about the repeated stimulus experiment, and $\varepsilon_{tx} \sim N(0,1)$ represents the noise. For a non-stimulus experiment it seems obvious

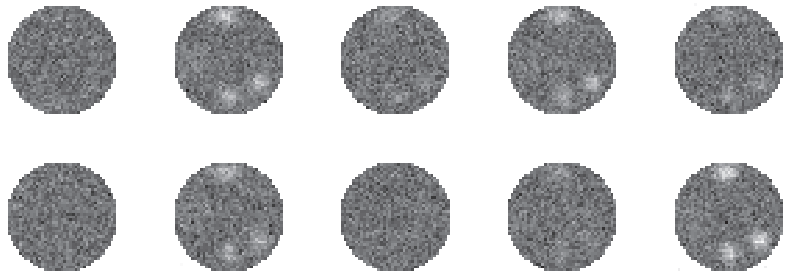

Fig. 1. Development of the activity over time. From left to right and top to bottom: the activity at time $t = 5, 15, \ldots, 95$.

to replace φ_t with a stationary stochastic process $\{F_t\}$. One possibility is to consider stimuli at random time points such that

$$F_t = \sum_i g(t - t_i),$$

where $\{t_i\}$ is a Poisson point process on the real line and g is a haemodynamic response function.

The general model to be described in the present paper is specified, using marked point process theory. The classical repeated stimulus experiments can also be dealt with, using this modelling approach, but this is not our primary objective. Various methods of analysing the model will be discussed, with increasing degree of computational complexity. Inference based on mean values, variances and covariances is relatively easy from a computationally point of view while likelihood or Bayesian methods are more demanding.

In section 2, the suggested spatio-temporal model is described. Models for the temporal and spatial parts of the activation profile are discussed in section 3 while section 4 describes the underlying spatio-temporal point process. First- and second-order properties of Z_{tx} are expressed in terms of corresponding properties of the underlying spatio-temporal point process in section 5 while specific point process models are discussed in section 6. Section 7 describes statistical inference based on mean value and covariance relations as well as Bayesian analysis. A simulation study is presented in section 8 while an analysis of real data can be found in section 9. Future work and perspectives are outlined in section 10. A summary of the main features of the new approach may be found in section 11.

2. The spatio-temporal model

Our general model has the form

$$Z_{tx} = \mu_x + \sum_i f_{tx}(t_i, x_i; m_i) + \sigma_x \varepsilon_{tx}, \tag{1}$$

where μ_x is the baseline signal at voxel x and $\Psi = \{[t_i, x_i; m_i]\}$ is a marked spatio-temporal point process on $\mathbb{R} \times \mathcal{X}$ with marks in $\mathcal{M} \subseteq \mathbb{R}^d$. The observation period of the fMRI experiment is $[0, T]$. The set $\mathcal{X}$ is a bounded subset of $\mathbb{R}^2$ or $\mathbb{R}^3$, representing a two-dimensional slice or a three-dimensional volume of the brain. Furthermore, ε_{tx} is the error term with $\mathbb{E}\varepsilon_{tx} = 0$ and $\mathbb{V}\varepsilon_{tx} = 1$. It is assumed that $\{\epsilon_{tx}\}$ are mutually independent. Various models for correlated noise are discussed in section 10.

According to (1), the activation profile is described by the marked point process Ψ. Each marked point $[t_i, x_i; m_i]$ may be considered as a centre of activation at location $x_i \in \mathcal{X}$. The centre is activated at time t_i and its duration and extension are described by the mark $m_i \in \mathcal{M}$.

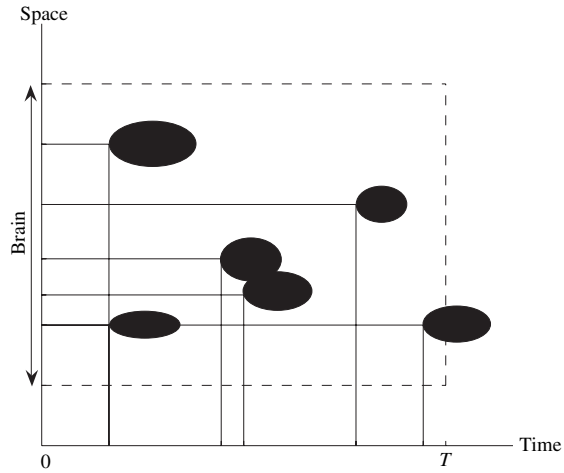

Fig. 2. Illustration of the spatio-temporal point process model. Each ellipse illustrates the set of $(t, x) \in [0, T] \times \mathcal{X}$, affected by the activation in the leftmost point (t_i, x_i) of the ellipse. The mark m_i determines the shape and size of the ellipse. In the illustration, an example of simultaneous activation in two different places of the brain is seen, as well as activation of the same place of the brain at different time points.

In what follows, we let $m_i = (m_i^1, m_i^2) \in \mathcal{M}_1 \times \mathcal{M}_2$, $\mathcal{M}_i \subseteq \mathbb{R}^{d_i}$, say, $i = 1, 2$, where m_i^1 describes the duration and m_i^2 the spatial extension of the ith activation. If two regions $\mathcal{X}_0$ and $\mathcal{X}_1$ of the brain interact, it is expected that activations occur simultaneously in $\mathcal{X}_0$ and $\mathcal{X}_1$. Specific point process models with such long-distance dependencies will be described in section 6. An illustration of the basic set-up may be found in Fig. 2.

In the analysis of fMRI data, spatial smoothing is often performed to reduce the noise of the data. The model (1) is closed under linear smoothing. Thus, suppose the data are smoothed by replacing Z_{tx} with $\tilde{Z}_{tx} = \sum_{z \in \mathcal{X}_x} \omega_{z-x} Z_{tz}$, where $\mathcal{X}_x$ is a neighbourhood around x. We suppose that $\mathcal{X}_x = \mathcal{X}_0 + x$. Furthermore, ω_y, $y \in \mathcal{X}_0$, satisfy $\omega_y \geq 0$ and $\sum_{y \in \mathcal{X}_0} \omega_y = 1$. If $\{Z_{tx}\}$ follow (1), then

$$\tilde{Z}_{tx} = \tilde{\mu}_x + \sum_i \tilde{f}_{tx}(t_i, x_i; m_i) + \tilde{\sigma}_x \tilde{\varepsilon}_{tx},$$

where

$$\tilde{\mu}_x = \sum_{u \in \mathcal{X}_0} \omega_u \mu_{x+u}, \ \tilde{f}_{tx} = \sum_{u \in \mathcal{X}_0} \omega_u f_{t, x+u}, \ \tilde{\sigma}_x^2 = \sum_{u \in \mathcal{X}_0} \omega_u^2 \sigma_{x+u}^2.$$

Our model is therefore closed under smoothing, except for the fact that smoothing introduces correlated errors.

3. Activation profile

Most current fMRI studies rely on the blood oxygenation level-dependent effect (Ogawa *et al.*, 1992) to detect changes in the MR signal intensity. Neural activity initiates a localized inflow of oxygenated blood to the active area, a *haemodynamic response*. This response is detectable in the MR signal due to different magnetic properties of oxygenated and de-oxygenated blood. The biological processes behind the haemodynamic response are not known in detail, but the general structure of the temporal behaviour has been described and reproduced in many studies. The haemodynamic response lags the neuronal activation with several seconds; it

increases slowly to a peak value at about 4–7 seconds after a neuronal impulse, and then returns to baseline again a few seconds after the neuronal impulse ceases. Often a late undershoot is reported as well, in the sense that when the signal drops after the peak value, it drops below baseline for a period before it returns to the baseline value.

Several different methods for modelling the haemodynamic response function (HRF) have been introduced. Perhaps, the most precise models are input-state-output models such as the Balloon model (see Buxton *et al.*, 2004 and references therein for more details). These models are computationally very complex. The simpler models described below are considered to give a fairly good approximation to empirical studies of the HRF (see Friston *et al.*, 1995; Glover, 1999). In these models, g is of the following form

$$g(u;m^1)=\int_0^l \kappa(u-v)\,\mathrm{d}v, \tag{2}$$

where l is the temporal duration of the activation. The mark m^1 includes l and possibly other parameters describing the function κ. As discussed above, $\kappa(t)\approx 0$ for $t\leq 0$, κ increases in the interval from 0 to about 4–7 seconds and then decreases to 0, possibly with a drop below 0 before returning to the value 0.

In the spatio-temporal point process model, each marked point $[t_i, x_i; m_i]$ represents an activation centred around x_i and starting at time t_i, with duration and extension specified by m_i^1 and m_i^2 respectively. Using the above-established results, all voxels x around a voxel x_i activated at time t_i will contribute to the MR signal intensity with a haemodynamic response proportional to the one observed in x_i. We will here assume that the constant of proportionality depends on x and x_i only via $x-x_i$. The resulting model for the activation profile becomes

$$f_{tx}(u,y;m)=g(t-u;m^1)h(x-y;m^2). \tag{3}$$

In the fMRI literature, g is called the HRF and h is the spatial activation function (SAF). For a recent use of (3) in repeated stimulus experiments, see the seminal paper of Hartvig (2002).

The modelling of the HRF and SAF is discussed below.

3.1. Temporal activation

3.1.1. HRF as an integral of Gaussian densities

Based on empirical studies, Friston *et al.* (1995) modelled the delay and dispersion of the haemodynamic response by a Gaussian density with mean 6 s and variance 9 s^2 as impulse response. In our formulation, this gives

$$\kappa(t)=\frac{1}{\sqrt{2\pi 3}}\exp\left(-\frac{(t-6)^2}{18}\right), \tag{4}$$

cf. Fig. 3.

This model assumes that the temporal activation pattern is the same for all activations during the experiment, which is a rather strong assumption. It is not complicated to make (4) slightly more general, by allowing the mean and the variance of the Gaussian density to vary for each activation. That information would then be included in the mark m^1. The response function would however still not be able to account for a haemodynamic response with a late undershoot. A natural extension to improve this is to linearly combine (4) with its derivatives with respect to different parameters as in Friston *et al.* (1998).

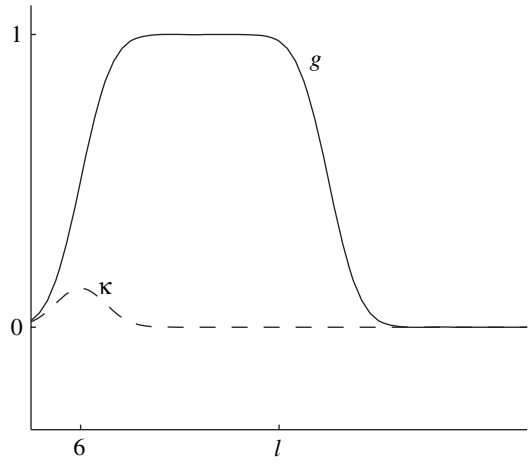

Fig. 3. Gaussian response function κ (dashed) and the corresponding integrated response function g (solid).

3.1.2. HRF as an integral of Gamma functions

Other empirical studies (Glover, 1999) have shown that Gamma functions may be more appropriate than Gaussian densities to capture the shape of the HRF. Glover uses the difference of two Gamma functions, one to capture the main response and the other to capture the late undershoot; that is, the HRF is modelled by

$$\kappa(t) = \left[\left(\frac{t}{p_1} \right)^{a_1} \exp\left(-\frac{t-p_1}{b_1} \right) - c \left(\frac{t}{p_2} \right)^{a_2} \exp\left(-\frac{t-p_2}{b_2} \right) \right] \mathbb{1}\{t>0\},$$

where t is the time in seconds and $p_j = a_j b_j$ is the time to the peak. In repeated stimulus experiments, $\kappa(t)$ is then convolved with the time course of the stimuli. This model can be made more flexible by expanding $\kappa(t)$ as a Taylor series and convolve the time course with $-\kappa(t) - t\partial\kappa(t)/\partial t$ instead (Worsley, 2000).

In our formulation, this means that the mark m^1 is now given by $m^1 = \{a_1, a_2, b_1, b_2, c, l\}$, where l describes the duration of the activation. The number of unknown parameters in the mark can be reduced by using the results from Glover (1999). For auditory response, the parameters were fit to $a_1 = 6, a_2 = 12, b_1 = b_2 = 0.9$ and $c = 0.35$. Motor response gave the result $a_1 = 5, a_2 = 12, b_1 = 1.1, b_2 = 0.9$ and $c = 0.4$. An example is shown in Fig. 4.

3.2. Spatial activation

The simplest model for the spatial activation is a symmetric Gaussian bell function

$$h(y; m^2) = \theta_1 \exp\left(-\frac{\|y\|^2}{2\theta_2} \right), \tag{5}$$

where $m^2 = (\theta_1, \theta_2)$, $\theta_1, \theta_2 > 0$ and $\|\cdot\|$ is the Euclidean norm in $\mathcal{X}$.

This can be extended as follows. Let $m^2 = (\theta_1, \Theta_2)$ where $\theta_1 > 0$ and Θ_2 is a $p \times p$ positive definite matrix ($p = 2$ or 3). The spatial activation function now becomes

$$h(y; m^2) = \theta_1 \exp\left(-\frac{1}{2} y^{\mathrm{T}} \Theta_2^{-1} y \right), \tag{6}$$

where y is assumed to be a column vector and $(\cdot)^{\mathrm{T}}$ stands for transpose (see also Hartvig, 2002).

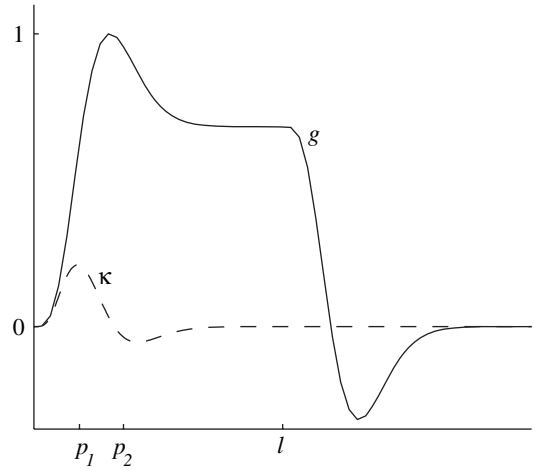

Fig. 4. Gamma response function κ (dashed) and the corresponding integrated response function g (solid).

4. The underlying spatio-temporal point process

The unmarked point process will be denoted by $\Phi = \{[t_i, x_i]\}$ and its intensity measure by Λ. For $A \in \mathcal{B}(\mathbb{R} \times \mathcal{X})$, the Borel σ-algebra on $\mathbb{R} \times \mathcal{X}$, $\Phi(A)$ is the number of unmarked points $[t_i, x_i]$ in A. Then,

$$\Lambda(A) = \mathbb{E}\Phi(A).$$

If $\Psi(A \times B)$ denotes the number of marked points $[t_i, x_i; m_i]$ with $[t_i, x_i] \in A$ and $m_i \in B$, $A \in \mathcal{B}(\mathbb{R} \times \mathcal{X})$ and $B \in \mathcal{B}(\mathcal{M})$, the intensity measure of the marked point process is defined by

$$\Lambda_m(A \times B) = \mathbb{E}\Psi(A \times B).$$

As $\Lambda_m(\cdot \times B) << \Lambda$, there exists for each $(u, y) \in \mathbb{R} \times \mathcal{X}$ a probability distribution $P_{u,y}$ on $(\mathcal{M}, \mathcal{B}(\mathcal{M}))$ such that

$$\Lambda_m(A \times B) = \int_A P_{u,y}(B)\Lambda(\mathrm{d}u, \mathrm{d}y),$$

see also Stoyan *et al.* (1995, p. 108). Note that $P_{u,y}$ can be interpreted as the distribution of the mark at (u, y).

Example 1 (the repeated stimulus experiment). The standard repeated stimulus experiment has earlier been described, using this framework (cf. Hartvig, 2002). In such an experiment we have k activation periods with known starting times t_i and known durations l_i, $i = 1, \ldots, k$ (see Fig. 5). The activation centres in the brain is described in Hartvig (2002) by a marked spatial point process $\{[x_j; m_j^2]\}$, where m_j^2 represents the spatial extension of the activation around x_j. The activation profile is specified as

$$\sum_{j=1}^{n} h(x - x_j; m_j^2)\varphi_t, \tag{7}$$

where φ_t is of the form

$$\varphi_t = \int_{-\infty}^{\infty} \pi_u \kappa(t - u)\,\mathrm{d}u,$$

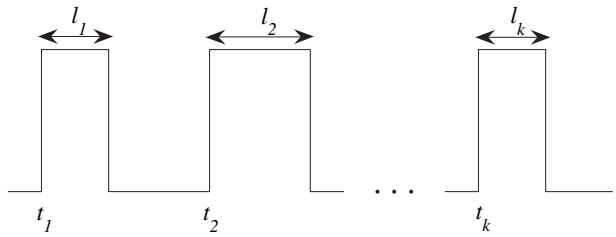

Fig. 5. Repeated stimulus experiment.

$\pi_u = 1$ if $u \in \cup_{i=1}^{k}[t_i, t_i + l_i]$ and κ is the response function for an activation at time 0. The expression (7) can be rewritten as

$$\sum_{i=1}^{k} \sum_{j} f_{tx}(t_i, x_j; m_i^1, m_j^2),$$

where f_{tx} satisfies (3) with g of the form (2). It follows that the model can be described by a marked spatio-temporal process

$$\Psi = \{[t_i, x_j; (m_i^1, m_j^2)]\},$$

where t_i and $m_i^1 = l_i$ are known. Note that for this process the intensity measure Λ satisfies

$$\Lambda = \Lambda_1 \times \Lambda_2,$$

where Λ_1 and Λ_2 are measures on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ and $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$, respectively, and Λ_1 is a discrete measure with weight 1 in t_i, $i = 1, \ldots, k$.

Example 2 (the non-stimulus experiment). During a non-stimulus experiment, the brain is not subjected to systematic stimuli but activated at random time points unknown to the experimenter. We will formalize this in the following fashion. The marked point process Ψ is assumed to be time stationary in the sense that

$$\Psi_t = \{[t_i + t, x_i; m_i]\}$$

has the same distribution as Ψ for all $t \in \mathbb{R}$. As a consequence, the intensity measure Λ is of the form

$$\Lambda = c\nu^1 \times \Lambda_2,$$

where $c > 0$ and ν^1 is the Lebesgue measure on $\mathbb{R}$. Furthermore, time stationarity implies that the mark distribution $P_{u,y}$ does not depend on the time point u.

5. Moment relations

In this section, we derive moment relations for the observed MR signal Z_{tx}, under various assumptions on the spatio-temporal point process $\Psi = \{[t_i, x_i; m_i]\}$.

5.1. The mean value relation

Using the Campbell–Mecke theorem for marked point processes, we find

$$\mathbb{E}Z_{tx} = \mu_x + \int_{\mathbb{R}\times\mathcal{X}} \int_{\mathcal{M}} f_{tx}(u, y; m) P_{u,y}(\mathrm{d}m) \Lambda(\mathrm{d}u, \mathrm{d}y).$$

The mean value relation can be further simplified if Ψ is separable.

Definition 1
The spatio-temporal point process Ψ *is called separable if the activation profile is on the product form (3),*

$$\Lambda = \Lambda_1 \times \Lambda_2 \tag{8}$$

and

$$P_{u,y} = P_u^1 \times P_y^2. \tag{9}$$

Here, Λ_1 *and* Λ_2 *are measures on* $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ *and* $(\mathcal{X}, \mathcal{B}(\mathcal{X}))$ *while* P_u^1 *and* P_y^2 *are probability measures on* $(\mathcal{M}_1, \mathcal{B}(\mathcal{M}_1))$ *and* $(\mathcal{M}_2, \mathcal{B}(\mathcal{M}_2))$ *respectively.*

Note that (8) is satisfied for the repeated stimulus and non-stimulus experiments discussed in examples 1 and 2 respectively. The assumption (9) is trivially satisfied if the marks are non-random.

For a separable model, we have

$$\mathbb{E}Z_{tx} = \mu_x + \alpha_t \beta_x, \tag{10}$$

where

$$\alpha_t = \int_{\mathbb{R}} \int_{\mathcal{M}_1} g(t-u; m^1) P_u^1(\mathrm{d}m^1) \Lambda_1(\mathrm{d}u)$$

and

$$\beta_x = \int_{\mathcal{X}} \int_{\mathcal{M}_2} h(x-y; m^2) P_y^2(\mathrm{d}m^2) \Lambda_2(\mathrm{d}y).$$

The parameters α_t can be further simplified for repeated stimulus and non-stimulus experiments respectively.

Example 1 (continued). The measure Λ_1 is here a discrete measure with weight 1 in $t_i, i = 1, \ldots, k$, and

$$\alpha_t = \sum_{i=1}^{k} \int_{\mathcal{M}_1} g(t-t_i; m^1) P_{t_i}^1(\mathrm{d}m^1). \tag{11}$$

In particular, if $P_{t_i}^1$ is concentrated in l_i, the known duration of the ith activation, then

$$\alpha_t = \sum_{i=1}^{k} g(t-t_i; m_i^1) \tag{12}$$

is known. The mean value specification (10) is a linear regression.

Example 2 (continued). As $\Lambda_1 = cv^1$ and P_u^1 does not depend on $u \in \mathbb{R}$, we have

$$\begin{aligned}
\alpha_t &= c \int_{\mathbb{R}} \int_{\mathcal{M}_1} g(t-u; m^1) P_u^1(\mathrm{d}m^1)\, \mathrm{d}u \\
&= c \int_{\mathcal{M}_1} \int_{\mathbb{R}} g(t-u; m^1)\, \mathrm{d}u P^1(\mathrm{d}m^1) \\
&= c\mathbb{E}\alpha_1(M^1),
\end{aligned}$$

where

$$\alpha_1(m^1) = \int_{\mathbb{R}} g(v; m^1)\, \mathrm{d}v$$

and M^1 is a random mark, distributed according to P^1. Accordingly, the parameter α_t does not depend on t and the same is true for $\mathbb{E}Z_{tx}$.

5.2. The covariance structure

In contrast to first-order properties, the covariance structure of Z_{tx} depends on the specific choice of point process model. The covariance can be expressed in terms of the so-called second-order factorial moment measure (see Stoyan *et al.*, 1995, p. 111 and onwards).

Let us here study the case of a marked point process $\Psi = \{[t_i, x_i; m_i]\}$ with conditional independent marks, such that conditionally on $\Phi = \{[t_i, x_i]\}$, $\{m_i\}$ are independent and $m_i \sim P_{t_i, x_i}$. Then,

$$\begin{aligned}
&\mathbb{E}\left(\sum_{i,i'} f_{tx}(t_i, x_i; m_i) f_{t'x'}(t_{i'}, x_{i'}; m_{i'})\right) \\
&\quad = \int_{\mathbb{R}\times\mathcal{X}} \int_{\mathcal{M}} f_{tx}(u, y; m) f_{t'x'}(u, y; m) P_{u,y}(\mathrm{d}m) \Lambda(\mathrm{d}u, \mathrm{d}y) \\
&\qquad + \int_{\mathbb{R}\times\mathcal{X}} \int_{\mathbb{R}\times\mathcal{X}} \int_{\mathcal{M}} \int_{\mathcal{M}} f_{tx}(u, y; m) f_{t'x'}(u', y'; m') P_{u,y}(\mathrm{d}m) P_{u',y'}(\mathrm{d}m') \\
&\qquad\qquad \times \alpha^{(2)}(\mathrm{d}u, \mathrm{d}y, \mathrm{d}u', \mathrm{d}y'),
\end{aligned}$$

where $\alpha^{(2)}$ is the second-order factorial moment measure for Φ, which is defined for $A, A' \in \mathcal{B}(\mathbb{R}\times\mathcal{X})$ by

$$\alpha^{(2)}(A \times A') = \mathbb{E}\sum_{i\neq i'} \mathbb{1}\{[t_i, x_i] \in A, [t_{i'}, x_{i'}] \in A'\}.$$

It follows that

$$\begin{aligned}
&\mathrm{cov}(Z_{tx}, Z_{t',x'}) \\
&\quad = \int_{\mathbb{R}\times\mathcal{X}} \int_{\mathcal{M}} f_{tx}(u, y; m) f_{t'x'}(u, y; m) P_{u,y}(\mathrm{d}m)\ \Lambda(\mathrm{d}u, \mathrm{d}y) \\
&\qquad + \int_{\mathbb{R}\times\mathcal{X}} \int_{\mathbb{R}\times\mathcal{X}} \int_{\mathcal{M}} \int_{\mathcal{M}} f_{tx}(u, y; m) f_{t',x'}(u', y'; m') P_{u,y}(\mathrm{d}m) P_{u',y'}(\mathrm{d}m') \\
&\qquad \times [\alpha^{(2)}(\mathrm{d}u, \mathrm{d}y, \mathrm{d}u', \mathrm{d}y') - \Lambda(\mathrm{d}u, \mathrm{d}y)\Lambda(\mathrm{d}u', \mathrm{d}y')] \\
&\qquad + \mathbb{1}\{(t, x) = (t', x')\}\sigma_x^2. \qquad (13)
\end{aligned}$$

The second-order factorial moment measure $\alpha^{(2)}$ is equal to $\Lambda \times \Lambda$ if Φ is a Poisson point process (see Stoyan *et al.*, 1995, p. 44). If

$$\alpha^{(2)}(\mathrm{d}u, \mathrm{d}y, \mathrm{d}u', \mathrm{d}y') - \Lambda(\mathrm{d}u, \mathrm{d}y)\Lambda(\mathrm{d}u', \mathrm{d}y') > 0,$$

then pairs of activations are more likely to occur jointly at (u, y) and (u', y') than for a Poisson point process with intensity measure Λ.

6. Specific point process models

In this section, we give two examples of spatio-temporal point process models that can exhibit the desired long-distance dependence. Under such a model, brain regions far apart may interact in the sense that if one of the regions is activated at time t then it is likely that the other regions are also activated at time t.

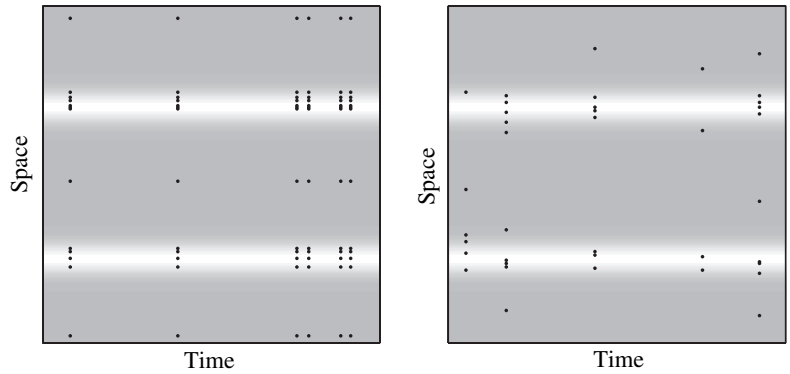

Fig. 6. Independent spatial and temporal Poisson processes (left) and conditionally independent Poisson processes (right). The associated intensity functions are shown in grey scale.

Example 3 (independent spatial and temporal point patterns). Suppose that $\Psi = \{[t_i, x_j; m_i^1, m_j^2]\}$ where $\Psi_1 = \{[t_i; m_i^1]\}$ and $\Psi_2 = \{[x_j; m_j^2]\}$ are independent. This model is separable. An example with Ψ_1 and Ψ_2 Poisson is shown in Fig. 6, left. We have

$$Z_{tx} = \mu_x + A_t B_x + \sigma_x \varepsilon_{tx},$$

where

$$A_t = \sum_i g(t - t_i; m_i^1) \quad \text{and} \quad B_x = \sum_j h(x - x_j; m_j^2) \tag{14}$$

are independent. The covariance is of the form

$$\begin{aligned}\operatorname{cov}(Z_{tx}, Z_{t'x'}) = {} & \operatorname{cov}(A_t, A_{t'})\ \operatorname{cov}(B_x, B_{x'}) + \operatorname{cov}(A_t, A_{t'})\beta_x\beta_{x'} \\ & + \alpha_t\alpha_{t'}\ \operatorname{cov}(B_x, B_{x'}) + \mathbb{1}\{(t,x) = (t',x')\}\sigma_x^2.\end{aligned}$$

For a repeated stimulus experiment, A_t is deterministic and the expression for the covariance reduces to

$$\operatorname{cov}(Z_{tx}, Z_{t'x'}) = \alpha_t\alpha_{t'}\ \operatorname{cov}(B_x, B_{x'}) + \mathbb{1}\{(t,x) = (t',x')\}\sigma_x^2,$$

where α_t takes the form (11) or (12), depending on the specific assumption on the HRF. A model of this type has already been considered in Hartvig (2002).

In a non-stimulus experiment, Ψ is time stationary and $\operatorname{cov}(A_t, A_{t'})$ only depends on $|t - t'|$. In particular, if the temporal process $\{t_i\}$ is Poisson and conditionally on $\{t_i\}$, $\{m_i^1\}$ are independent and $m_i^1 \sim P^1$, we have

$$\operatorname{cov}(A_t, A_{t'}) = \int_{\mathbb{R}} \int_{\mathcal{M}_1} g(t - u; m^1) g(t' - u; m^1) P^1(\mathrm{d}m^1) \Lambda_1(\mathrm{d}u) = \rho_{t,t'},$$

say. As $\Lambda_1(\mathrm{d}u) = c\ \mathrm{d}u$, we obtain

$$\begin{aligned}\rho_{t,t'} &= c \int_{\mathbb{R}} \int_{\mathcal{M}_1} g(t - u; m^1) g(t' - u; m^1) P^1(\mathrm{d}m^1)\, \mathrm{d}u \\ &= c \int_{\mathcal{M}_1} \int_{-\infty}^{\infty} g(v; m^1) g(v + |t' - t|; m^1)\, \mathrm{d}v P^1(\mathrm{d}m^1) \\ &= c\, \mathbb{E}\alpha_2(|t' - t|; M^1),\end{aligned}$$

say, where

$$\alpha_2(t;m^1)=\int_{-\infty}^{\infty} g(v;m^1)g(v+t;m^1)\,\mathrm{d}v$$

and M^1 is a random mark distributed according to P^1. If the spatial process is Poisson with conditionally independent marking

$$\mathrm{cov}(B_x,B_{x'})=\int_{\mathcal{X}}\int_{\mathcal{M}_2} h(x-y;m^2)h(x'-y;m^2)P_y^2(\mathrm{d}m^2)\Lambda_2(\mathrm{d}y)=\tau_{x,x'},$$

say. The parameter $\tau_{x,x'}$ will be small if Λ_2 is concentrated around x and x' but the distance between x and x' is large. If both processes are Poisson with conditionally independent marking, we thus have

$$\mathrm{cov}(Z_{tx},Z_{t'x'})=\rho_{t,t'}\tau_{x,x'}+\rho_{t,t'}\beta_x\beta_{x'}+\alpha_t\alpha_{t'}\tau_{x,x'}+\mathbb{1}\{(t,x)=(t',x')\}\sigma_x^2. \tag{15}$$

More generally, if both processes $\{t_i\}$ and $\{x_j\}$ have conditionally independent marking, $\{t_i\}$ is Poisson and $\{x_j\}$ is a general point process with second-order factorial moment measure $\alpha^{(2)}$, then

$$\begin{aligned}\mathrm{cov}(Z_{tx},Z_{t'x'})&=[\rho_{t,t'}+\alpha_t\alpha_{t'}][\tau_{x,x'}+\delta_{x,x'}]\\&\quad+\rho_{t,t'}\beta_x\beta_{x'}+\mathbb{1}\{(t,x)=(t',x')\}\sigma_x^2,\end{aligned} \tag{16}$$

where

$$\begin{aligned}\delta_{x,x'}=\int_{\mathcal{X}}\int_{\mathcal{X}}\int_{\mathcal{M}_2}\int_{\mathcal{M}_2} &h(x-y;m^2)h(x'-y';m^{2'})P_y^2(\mathrm{d}m^2)P_{y'}^2(\mathrm{d}m^{2'})\\&\times[\alpha^{(2)}(\mathrm{d}y,\mathrm{d}y')-\Lambda_2(\mathrm{d}y)\Lambda_2(\mathrm{d}y')].\end{aligned}$$

Note that $\delta_{x,x'}=0$ if $\{x_j\}$ is Poisson.

Example 4 (conditional independent spatial processes). The spatio-temporal process is given by $\Psi=\{[t_i,x_{ij};m_i^1,m_{ij}^2]\}$. Conditionally on the temporal process $\Psi_1=\{[t_i;m_i^1]\}$, the spatial processes $\Psi_{2i}=\{[x_{ij};m_{ij}^2]\}$ are independent and identically distributed with second-order factorial moment measure $\alpha^{(2)}$. It is not difficult to show that if (3) is satisfied, then Ψ is separable.

Under this model, the covariance is of the form

$$\begin{aligned}\mathrm{cov}(Z_{tx},Z_{t'x'})&=\mathrm{cov}(A_t,A_{t'})\beta_x\beta_{x'}+\rho_{t,t'}\ \mathrm{cov}(B_x,B_{x'})\\&\quad+\mathbb{1}\{(t,x)=(t',x')\}\sigma_x^2,\end{aligned}$$

with the notation of the previous example. For a repeated stimulus experiment (see example 1), A_t is deterministic and the expression for the covariance reduces to

$$\mathrm{cov}(Z_{tx},Z_{t',x'})=\rho_{t,t'}\ \mathrm{cov}(B_x,B_{x'})+\mathbb{1}\{(t,x)=(t',x')\}\sigma_x^2.$$

If instead the temporal process is Poisson, we have an example of a non-stimulus experiment, cf. example 2, and the covariance is of the form

$$\mathrm{cov}(Z_{tx},Z_{t'x'})=\rho_{t,t'}[\tau_{x,x'}+\delta_{x,x'}+\beta_x\beta_{x'}]+\mathbb{1}\{(t,x)=(t',x')\}\sigma_x^2, \tag{17}$$

again with the notation of the previous example. An example with Ψ_1 and Ψ_2 Poisson is shown in Fig. 6, right.

7. Statistical inference

In this section, we discuss statistical inference based on moment relations. We also briefly touch upon Bayesian inference.

7.1. Inference based on the mean value relation

In this section, we will discuss within the framework of a separable model, the estimation of the intensity measure Λ_2 of the spatial point process, using the general mean value relation (10). We will assume that the marks are identical for all points in which case

$$\mathbb{E}Z_{tx} = \mu_x + \alpha_t \beta_x,$$

where

$$\alpha_t = \int_{\mathbb{R}} g(t-s; m^1)\Lambda_1(\mathrm{d}s)$$

and

$$\beta_x = \int_{\mathcal{X}} h(x-y; m^2)\Lambda_2(\mathrm{d}y).$$

In what follows, we let $\tilde{\alpha}_t = \alpha_t$ in a repeated stimulus experiment [see (12)] while $\tilde{\alpha}_t = 1$ in a non-stimulus experiment. Note that for fixed m^1 the parameters $\tilde{\alpha}_t$ are known. Likewise, we let $\tilde{\Lambda}_2 = \Lambda_2$ in a repeated stimulus experiment and $\tilde{\Lambda}_2 = c\alpha_1(m^1)\Lambda_2$ in a non-stimulus experiment where

$$\alpha_1(m^1) = \int_{\mathbb{R}} g(u; m^1)\,\mathrm{d}u.$$

The method to be described can be applied if the baseline intensity μ_x can be regarded as known. The baseline intensity can vary by a factor of 2 to 3 across the brain, due to variations in the brain tissue as well as variations in the scanner. The baseline μ_x is well determined from data in repeated stimulus experiments, otherwise additional data is needed.

If μ_x can be regarded as known, we can let $\mu_x = 0$. The mean value relation can then be written as

$$\mathbb{E}Z_{tx} = \tilde{\alpha}_t \tilde{\beta}_x,$$

where

$$\tilde{\beta}_x = \int_{\mathcal{X}} h(x-y; m^2)\tilde{\Lambda}_2(\mathrm{d}y).$$

We will consider the estimation of Λ_2 (or equivalently $\tilde{\Lambda}_2$) under the assumption that Λ_2 is a discrete measure concentrated in $y_j, j = 1, \ldots, N$, with masses $\lambda_2(y_j) = \Lambda_2(\{y_j\})$, $j = 1, \ldots, N$. Here, N may be chosen as the number of voxels. Let us suppose that we have discretely observed data in time with spacing Δ

$$\{Z_{i\Delta, x} : i = i_0 + 1, \ldots, i_0 + n, x \in \mathcal{X}\},$$

where all time points are free of edge effects. A simple estimation procedure is to estimate $\tilde{\beta}_x$ by the regression estimate

$$\tilde{Z}_x = \frac{\sum_{i=1}^n \tilde{\alpha}_{(i_0+i)\Delta} Z_{(i_0+i)\Delta, x}}{\sum_{i=1}^n \tilde{\alpha}^2_{(i_0+i)\Delta}}$$

and for each m^1 and m^2 minimize

$$\sum_{i=1}^N \left[\tilde{Z}_{y_i} - \sum_{j=1}^N h(y_i - y_j; m^2)\tilde{\lambda}_2(y_j)\right]^2 \tag{18}$$

with respect to $\{\tilde{\lambda}_2(y_j)\}$, subject to the condition $\tilde{\lambda}_2(y_j) \geq 0$ for all j. Note that in a non-stimulus experiment, Λ_2 and c cannot be separated, using this estimation procedure.

The variance of $\tilde{Z}_x$ may, however, depend on x. As an example, let us consider a repeated stimulus experiment with Poisson-distributed activation centres as described in example 3. Then,

$$\mathbb{V}\tilde{Z}_x = \tau_{x,x} + \frac{1}{\sum_1^n \alpha^2_{(i_0+i)\Delta}}\sigma_x^2.$$

An unbiased estimate of σ_x^2 is

$$\hat{\sigma}_x^2 = \frac{1}{n-1}\sum_{i=1}^n (Z_{(i_0+i)\Delta,x} - \alpha_{(i_0+i)\Delta}\tilde{Z}_x)^2.$$

Furthermore, a discrete version of $\tau_{x,x}$ is

$$\tau_{x,x} = \sum_{j=1}^N h(x-y_j;m^2)^2\lambda_2(y_j).$$

The unweighted sum of squares may then be replaced by

$$\frac{\sum_{i=1}^N \left[\tilde{Z}_{y_i} - \sum_{j=1}^N h(y_i-y_j;m^2)\lambda_2(y_j)\right]^2}{\mathbb{V}\tilde{Z}_{y_i}},$$

where we insert the derived form of $\mathbb{V}\tilde{Z}_{y_i}$ and the estimate $\hat{\sigma}^2_{y_i}$. This sum of squares should be minimized with respect to λ_2 for fixed m^1 and m^2.

As another example, let us consider a non-stimulus experiment with independent temporal and spatial Poisson point processes. Then,

$$\tilde{Z}_x = \bar{Z}_{\cdot x} = \frac{1}{n}\sum_{i=1}^n Z_{(i_0+i)\Delta,x}$$

and, using (15), we find

$$\mathbb{V}\bar{Z}_{\cdot x} = \frac{c}{n^2}\left[n\alpha_2(0;m^1) + 2\sum_{i=1}^{n-1}(n-i)\alpha_2(i\Delta;m^1)\right]\left[\tau_{x,x}+\beta_x^2\right] + c^2\alpha_1(m^1)^2\tau_{x,x} + \frac{1}{n}\sigma_x^2.$$

The empirical variance

$$\hat{\sigma}_{x,x} = \frac{1}{n-1}\sum_{i=1}^n (Z_{(i_0+i)\Delta,x} - \bar{Z}_{\cdot x})^2$$

can be used to estimate σ_x^2 but it is important to correct for bias in this estimate caused by correlations inside the time series. We will now derive the bias of the estimate.

Generally, if $\mathrm{cov}(Z_{tx}, Z_{tx'})$ only depend on t and t' via $|t-t'|$,

$$\mathrm{cov}(Z_{tx}, Z_{t',x'}) = \sigma_{x,x'}(|t-t'|),$$

say, then the estimate

$$\hat{\sigma}_{x,x'} = \frac{1}{n-1}\sum_{i=1}^n (Z_{(i_0+i)\Delta,x} - \bar{Z}_{\cdot x})(Z_{(i_0+i)\Delta,x'} - \bar{Z}_{\cdot x'}), \tag{19}$$

is a biased estimate of $\sigma_{x,x'} = \sigma_{x,x'}(0)$. We thus have

$$\mathbb{E}(\hat{\sigma}_{x,x'}) = \sigma_{x,x'} - \frac{2}{n(n-1)} \sum_{i=1}^{n-1} (n-i)\sigma_{x,x'}(i\Delta). \tag{20}$$

Using (15), (19) and (20), we find that

$$\mathbb{E}(\hat{\sigma}_{x,x}) = c\left[\alpha_2(0;m^1) - \frac{2}{n(n-1)} \sum_{i=1}^{n-1} (n-i)\alpha_2(i\Delta;m^1)\right]\left[\tau_{x,x} + \beta_x^2\right] + \sigma_x^2.$$

The variance of $\bar{Z}_{\cdot x}$ can therefore be written as

$$\mathbb{V}\bar{Z}_{\cdot x} = \frac{1}{n}\mathbb{E}(\hat{\sigma}_{x,x}) + c^2\alpha_1(m^1)^2\tau_{x,x} + \frac{2c}{n(n-1)} \sum_{i=1}^{n-1} (n-i)\alpha_2(i\Delta;m^1)[\tau_{x,x} + \beta_x^2].$$

The unweighted sum of squares may be replaced by

$$\frac{\sum_{i=1}^{N}\left[\bar{Z}_{\cdot y_i} - c\alpha_1(m^1)\sum_{j=1}^{N} h(y_i - y_j;m^2)\lambda_2(y_j)\right]^2}{\mathbb{V}\bar{Z}_{\cdot y_i}}$$

and minimized with respect to $\{\lambda_2(y_j)\}$ for each fixed c, m^1 and m^2.

7.2. Inference based on covariances

The method described in the previous section is simple but requires, for a non-stimulus experiment, that μ_x is known from external sources. If this is not feasible, one may try to get information about the intensity measure Λ_2 of the spatial point process from $\mathrm{cov}(Z_{tx}, Z_{t',x'})$ instead. The covariances do not depend on the μ_xs.

This approach depends on a specific point process model. As an example, let us consider the model for a non-stimulus experiment with both temporal and spatial processes Poisson. Irrespectively of whether the processes are independent or conditionally independent (example 3 or 4), the mean value of the empirical covariance estimate (19) can be approximated for x, x' with large mutual distance by

$$\mathbb{E}(\hat{\sigma}_{x,x'}) \approx c\gamma(m^1)\beta_x\beta_{x'},$$

where

$$\gamma(m^1) = \alpha_2(0;m^1) - \frac{2}{n(n-1)} \sum_{i=1}^{n-1} (n-i)\alpha_2(i\Delta;m^1),$$

cf. (15), (17), (19) and (20). Assume that an activation centre $\mathcal{X}_0 \subset \mathcal{X}$ with N_0 points is known. Then, for x' with large mutual distance from all points $x \in \mathcal{X}_0$,

$$\mathbb{E}\left(\frac{1}{N_0}\sum_{x\in\mathcal{X}_0}\hat{\sigma}_{x,x'}\right) \approx c\gamma(m^1)\bar{\beta}_{\cdot}\sum_{i=1}^{N} h(x' - x_i;m^2)\lambda_2(x_i), \tag{21}$$

where

$$\bar{\beta}_{\cdot} = \frac{1}{N_0}\sum_{x\in\mathcal{X}_0}\beta_x.$$

This expression is linear in λ_2 if we regard $\bar{\beta}_{\cdot}$ as an unknown constant. We can thus use least squares methods to estimate $\lambda_2(x)$ for $x \in \mathcal{X} \setminus \mathcal{X}_0$ up to a constant, as in the previous section.

Another relevant question is what kind of information about the model parameters can be gained from the covariances under a less specified model, for instance, if we relax the assumption that the spatial point process is Poisson. Let us concentrate on the conditional-

independent processes, presented in example 4. We consider a non-stimulus experiment with a Poisson process as temporal process. Then, cf. (17),

$$\begin{aligned}\operatorname{cov}(Z_{tx}, Z_{t',x'}) = c\alpha_2(|t'-t|; m^1) \\ \times \left[\tau_{x,x'} + \int_{\mathcal{X}}\int_{\mathcal{X}} h(x-y; m^2)h(x'-y'; m^2)\alpha^{(2)}(\mathrm{d}y, \mathrm{d}y')\right] \\ + \mathbb{1}\{(t,x)=(t',x')\}\sigma_x^2.\end{aligned}$$

In particular, for x, x' with large mutual distance

$$\operatorname{cov}(Z_{tx}, Z_{t',x'}) \approx c\alpha_2(0; m^1)\int_{\mathcal{X}}\int_{\mathcal{X}} h(x-y; m^2)h(x'-y'; m^2)\alpha^{(2)}(\mathrm{d}y, \mathrm{d}y').$$

The slope of the regression of $Z_{tx'}$ on Z_{tx},

$$\frac{\operatorname{cov}(Z_{tx}, Z_{tx'})}{\mathbb{V}Z_{tx}},$$

is thus for fixed x and varying x' proportional to

$$\int_{\mathcal{X}}\int_{\mathcal{X}} h(x-y; m^2)h(x'-y'; m^2)\alpha^{(2)}(\mathrm{d}y, \mathrm{d}y').$$

If $h(u; m^2)$ is concentrated around $u=0$, a plot of the slopes will reveal $x' \in \mathcal{X}$ for which $\alpha^{(2)}(\mathrm{d}x, \mathrm{d}x')$ is large. Recall that $\alpha^{(2)}(\mathrm{d}x, \mathrm{d}x')$ can be interpreted as the probability of having simultaneously an activation at x and x'.

In Greicius *et al.* (2003), the average time series from one brain region is used as explanatory variable in the analysis of the time variation in other regions of the brain. Under the model specified above, Greicius' analysis leads to a study of the second-order factorial moment measure of the spatial point process.

7.3. Bayesian inference

In this section, we will briefly discuss the Bayesian inference. A more complete treatment of this approach is planned to appear elsewhere. As earlier, μ_x requires a special treatment. When considering Bayesian methods we may simply replace Z_{tx} by $Z_{tx} - \bar{Z}_{\cdot x}$ and f_{tx} by $f_{tx} - \bar{f}_{\cdot x}$. The new data have $\mu_x = 0$ and the same correlation structure as the original data if T is large. For brevity, we write $Z = \{Z_{tx}\}$.

7.3.1. Prior distributions

We concentrate on the case where $m_i = m$ and $\sigma_x = \sigma$ are known. We then need to specify a prior density of the point process Φ and its parameters. We assume that the intensity function of Φ is of the following form

$$\lambda(t,x) = \sum_{l=1}^{k} \lambda_l \mathbf{1}\{x \in \mathcal{X}_l\},$$

where the sets $\mathcal{X}_l \subseteq \mathcal{X}$ are disjoint. Their union may be the whole brain $\mathcal{X}$ but need not be. The sets $\mathcal{X}_l$ should be specified by the experimenter while the parameters λ_l are unknown.

It turns out to be a good idea to transform the parameters. We let $c = \sum \lambda_l|\mathcal{X}_l|$ and $\pi_l = \lambda_l|\mathcal{X}_l|/\sum \lambda_l|\mathcal{X}_l|$. Note that the π_l values satisfy

$$\pi_l \geq 0, \quad \sum_{l=1}^{k} \pi_l = 1.$$

The new parameters have nice interpretations. Thus,

$$c = \frac{1}{T}\int_{[0,\,T]\times\mathcal{X}} \lambda(t,x)\,\mathrm{d}t\,\mathrm{d}x$$
$$= \frac{1}{T}\mathbb{E}\Phi([0,T]\times\mathcal{X})$$

is the expected number of activations per time unit while

$$\pi_l = \frac{\displaystyle\int_{[0,\,T]\times\mathcal{X}_l} \lambda(t,x)\,\mathrm{d}t\,\mathrm{d}x}{\displaystyle\int_{[0,\,T]\times\mathcal{X}} \lambda(t,x)\,\mathrm{d}t\,\mathrm{d}x}$$
$$= \frac{\mathbb{E}\Phi([0,T]\times\mathcal{X}_l)}{\mathbb{E}\Phi([0,T]\times\mathcal{X})}$$

is the expected fraction of all activations that occur in $\mathcal{X}_l$. Note that

$$\lambda(t,x) = c\lambda_2(x), \tag{22}$$

where

$$\lambda_2(x) = \sum_{l=1}^{k} \pi_l \frac{\mathbf{1}\{x\in\mathcal{X}_l\}}{|\mathcal{X}_l|}, \tag{23}$$

satisfies

$$\int_{\mathcal{X}} \lambda_2(x)\,\mathrm{d}x = 1.$$

The prior distribution of Φ will be chosen as Poisson with intensity function λ. Note that there is no interaction between points in the prior distribution. Interaction found in the posterior distribution of the point process will therefore be 'caused' by the data z. We consider the restriction

$$\Phi_0 = \Phi\cap([T_{0-}, T_{0+}]\times\mathcal{X})$$

of Φ to a time interval $[T_{0-}, T_{0+}]$ containing $[0,T]$. The interval $[T_{0-}, T_{0+}]$ is chosen such that it is very unlikely that a point from $\Phi\backslash\Phi_0$ will affect an MR signal observed in $[0,T]$. Using (22) and (23), the density of Φ_0 with respect to the distribution ν of a unit rate Poisson process on $[T_{0-}, T_{0+}]\times\mathcal{X}$ becomes

$$p(\phi_0\,|\,c,\pi) = \exp\left(-\int_{[T_{0-},\,T_{0+}]\times\mathcal{X}} [\lambda(t,x)-1]\,\mathrm{d}t\;\mathrm{d}x\right)\prod_{[u,\,y]\in\phi_0}\lambda(u,y)$$
$$= \exp(-(T_{0+}-T_{0-})(c-|\mathcal{X}|))c^{n(\phi_0)}\prod_l\left(\frac{\pi_l}{|\mathcal{X}_l|}\right)^{n_l(\phi_0)},$$

where $n(\phi_0)$ is the number of points in ϕ_0 and $n_l(\phi_0)$ is the number of these points falling in $\mathcal{X}_l$.

We will use non-informative priors for c and $\pi=(\pi_1,\ldots,\pi_k)$. The prior density of c will be specified as

$$p(c) = \frac{1}{c_{\max}}\mathbf{1}\{c<c_{\max}\},$$

where $c_{\max}$ is a large known constant while the prior density of π is

$$p(\pi) = \frac{1}{\mathrm{vol}(D)}\mathbf{1}\{\pi\in D\},$$

where

$$D = \{\pi : \pi_l > 0, \sum \pi_l = 1\}.$$

7.3.2. Posterior simulation

The complete posterior density is

$$p_{m,\sigma^2}(c, \pi, \phi_0 \mid z) \propto p(c)p(\pi)p(\phi_0 \mid c, \pi)p_{m,\sigma^2}(z \mid \phi_0), \tag{24}$$

as the conditional density of z given c, π, ϕ_0

$$p_{m,\sigma^2}(z \mid \phi_0) = [2\pi\sigma^2]^{-NT/2} \exp\left(-\frac{1}{2\sigma^2}\|z - f(\phi_0; m)\|^2\right)$$

only depends on ϕ_0. Here,

$$\|z - f(\phi_0; m)\|^2 = \sum_{t,x}\left(z_{tx} - \sum_{[t_i, x_i] \in \phi_0} f_{tx}(t_i, x_i; m)\right)^2.$$

For simulation from the posterior density we use a fixed scan Metropolis within Gibbs algorithm where in each scan c, π and ϕ_0 are updated in turn. The full conditional for c is a Gamma distribution

$$c \mid \pi, \phi_0, z \sim \Gamma(n(\phi_0) + 1, T_{0+} - T_{0-}), \tag{25}$$

with the constraint that $c < c_{\max}$. We use (25) as proposal where ϕ_0 is the current state, and sample from $\Gamma(n(\phi_0) + 1, T_{0+} - T_{0-})$ until the constraint is satisfied.

The density of the full conditional distribution of π takes the form

$$\begin{aligned} p(\pi \mid c, \phi_0, z) &\propto p(\pi)p(\phi_0 \mid c, \pi) \\ &\propto \prod \pi_l^{n_l(\phi_0)} \end{aligned}$$

with the constraint $\pi \in D$. For $k = 1$, this step can be omitted as $D = \{1\}$. For $k = 2$,

$$\pi \mid c, \phi_0, z \sim \text{Beta}\big(n_1(\phi_0) + 1, n_2(\phi_0) + 1\big),$$

while for $k > 2$ we have

$$\pi \mid c, \phi_0, z \sim \text{Dirichlet}\big(n_1(\phi_0) + 1, n_2(\phi_0) + 1, \ldots, n_k(\phi_0) + 1\big).$$

We sample a π from the appropriate distribution, using the current value of ϕ_0.

The last step is to simulate from

$$p(\phi_0 \mid c, \pi, z) \propto c^{n(\phi_0)} \prod_l \pi_l^{n_l(\phi_0)} \exp\left(-\frac{1}{2\sigma^2}\|z - f(\phi_0; m)\|^2\right).$$

The point process is simulated using a birth, death, and move algorithm as described in chapter 7 in Møller and Waagepetersen (2004). The starting value of the simulation is a Poisson process. In each iteration we then propose one of the following steps: deleting a uniformly chosen point from the current point pattern, moving a uniformly chosen point in the current point pattern to a new random location (in time and space) or adding a new point in a random location. The different proposals are selected with equal probability. We use the current values of c and π to calculate the acceptance ratio for the suggested change in the point pattern.

8. A simulation study

We have simulated data from the model in (1) with independent spatial and temporal Poisson point patterns as in example 3. The object of the simulation study was the analysis of a non-stimulus experiment. Thus, we gave the temporal intensity function a constant value, $\lambda_1(t) = c$ for all $t \in [0, 100]$, while the spatial activation pattern comprised activated areas of various sizes, shapes and peak intensity. The HRF was given by an integral (sum) of Gaussian densities as in section 3.1.1 with $m^1 = l = 5$ and the spatial activation was modelled by a symmetric Gaussian bell function as in section 3.2 with $m^2 = (\theta_1, \theta_2) = (4, 4)$. Furthermore, the errors were standard Gaussian distributed, $\varepsilon_{tx} \sim N(0, 1)$, and we set σ_x^2 to be equal to 40% of the baseline signal. The standard deviation σ_x is roughly three times larger than the maximum intensity of an activation centre.

The activation pattern is shown in Fig. 7, with the realization of the temporal activity left and the arranged spatial activity right. Two time series from the simulation are shown in Fig. 8, one is from an activated area and one from an area with no activation. The activation pattern in the former clearly follows the temporal activation pattern shown in Fig. 7 (left). The development of the activation over time has also been shown in Fig. 1.

We have estimated the spatial intensity function λ_2, using the three different methods outlined in section 7. In section 8.1, we used the method based on the general mean value relation as described in section 7.1. In section 8.2, we assumed one of the activated areas, $\mathcal{X}_0 \subset \mathcal{X}$, to be known and we searched for other areas in $\mathcal{X}$, functionally connected to $\mathcal{X}_0$.

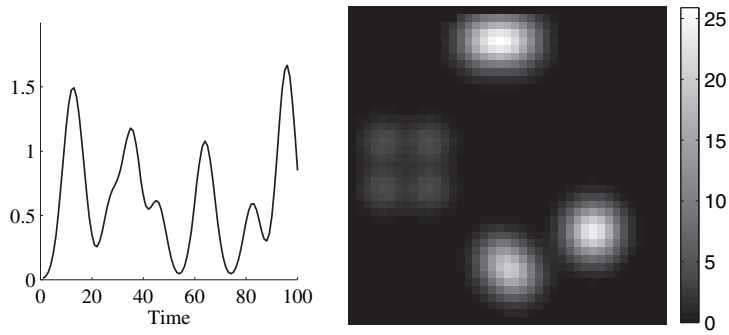

Fig. 7. The realization of the temporal activity used in the simulation (left) and the spatial activity (right). The haemodynamic response function was modelled by a sum of Gaussian functions with mark $m^1 = 5$ and the spatial activation factor was modelled by a Gaussian bell function with $m^2 = (4, 4)$. See the main text for more details.

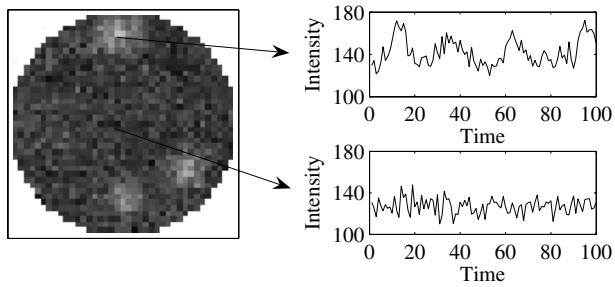

Fig. 8. Illustration of time-series data from the simulation. Left: simulated data at time $t = 20$. Right: time series of respectively an active (top) and a non-active (bottom) voxel.

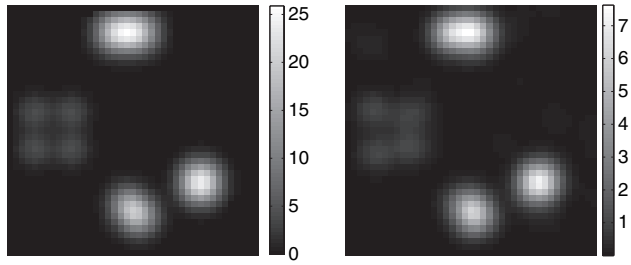

Fig. 9. The true spatial activation pattern (left) and the estimated spatial activation pattern (right) with marks $m^1 = 5$ and $m^2 = (4, 4)$.

That is, we estimated λ_2 in $\mathcal{X} \setminus \mathcal{X}_0$ using covariances. This method is similar to the inference discussed in section 7.2. Finally, in section 8.3, we use Bayesian inference.

8.1. Estimation of λ_2 using mean value relations

We used the method described in section 7.1 and, for fixed m^1 and m^2, minimized (18). This method gave us an estimate $\hat{\lambda}_2$ of λ_2 up to a constant of proportionality. We scaled $\hat{\lambda}_2$ such that $0 \leq \hat{\lambda}_2(y_i) \leq 1$ for all $i = 1, \ldots, N$, with y_i being the midpoint of each voxel (pixel). The estimated activation pattern was determined at each $x \in \mathcal{X}$ as

$$\sum_{i=1}^{N} h(x - y_i; m^2)\hat{\lambda}_2(y_i).$$

Figure 9 shows the estimated activation pattern for $m^1 = 5$ and $m^2 = (4, 4)$ (right) together with the true activation pattern (left). The method gives an estimate of the correct activation pattern, up to multiplication with a constant.

8.2. Estimation of λ_2 using covariances

We assume that we have given an activated area, $\mathcal{X}_0$, in $\mathcal{X}$ and wish to find other areas with functional connection to $\mathcal{X}_0$, using analysis based on covariances. Following section 7.2 we calculate the slope of the regression of $Z_{tx'}$ on Z_{tx} for the simulated data, where x is the point in $\mathcal{X}$ with maximum intensity. This approach gives a first estimate of the spatial intensity function (see Fig. 10).

We can also estimate the spatial activation, using the covariances. We supposed the upper middle activation centre in Fig. 7 (right) to be known. We then used (21) to obtain an

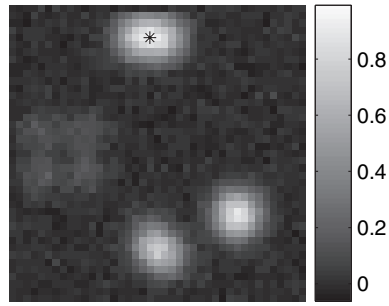

Fig. 10. The slope of the regression of $Z_{tx'}$ on Z_{tx} for a fixed point $x \in \mathcal{X}$ and all $x' \in \mathcal{X}$. The point x, shown as an asterisk in the figure, is the point in $\mathcal{X}$ with maximum intensity.

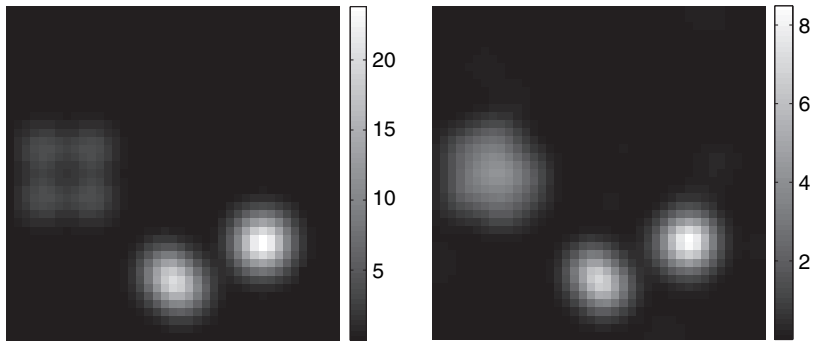

Fig. 11. The true spatial activation pattern (left) and the estimated spatial activation pattern (right) for the marks $m^1 = 5$ and $m^2 = (4, 4)$. The upper middle activation centre in Fig. 7, denoted by $\mathcal{X}_0$ in the text, is not shown, as it is assumed known and thus not estimated.

estimate $\hat{\lambda}_2(x)$ of $\tilde{\lambda}_2(x) = c\gamma(m^1)\bar{\beta}.\lambda_2(x)$ for all $x \in \mathcal{X} \setminus \mathcal{X}_0$. Given the estimate of the spatial intensity function, the spatial activation was reconstructed as in the previous section. The results for $m^1 = 5$ and $m^2 = (4, 4)$ are shown in Fig. 11. As before, the method finds the correct activation areas, but the intensities are only known up to a multiplication with a constant.

8.3. Bayesian inference

We analysed the simulated data in two different ways with the method described in section 7.3. In the first case, we have restricted our attention to two regions of the brain, the high intensity region $\mathcal{X}_1$ and the low intensity region $\mathcal{X}_2$. The areas of these two regions are 51 and 49 pixels respectively. Their position is indicated in Fig. 7 (right) where $\mathcal{X}_1$ is the three separate bright parts while $\mathcal{X}_2$ is the low intensity region in the mid-left part of the illustration. We assumed there is no activation in $\mathcal{X} \setminus (\mathcal{X}_1 \cup \mathcal{X}_2)$, and so we used the method in section 7.3 with $k = 2$. In the second case, we have set $k = 1$ with the region of interest equal to $\mathcal{X}$. The area of $\mathcal{X}$ is 1245 pixels. Figure 12 shows plots of the log posterior density, $\log p_{m,\sigma^2}(c, \pi, \phi_0 \mid z)$, as a function of iteration number. The normalization constant for the posterior density is unknown, cf. (24), so we only know the log posterior density up to

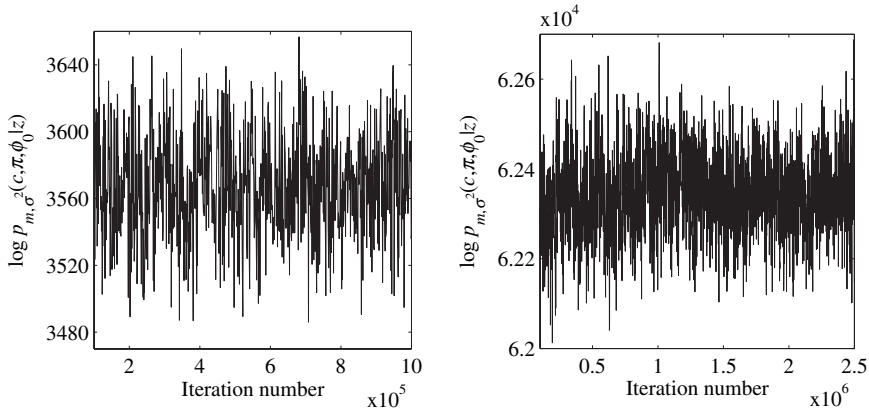

Fig. 12. The log-posterior density as a function of iteration number for $k = 2$ (left) and $k = 1$ (right). Note that the log-posterior density is only known up to a constant, cf. (24). The values on the y-axis are thus only correct up to an additive constant.

addition with a constant. This is, however, irrelevant when the plots are used to study the convergence of the algorithm.

When the algorithm had converged, we sampled point processes $\{\Phi_j\}_{j=1}^M$ from the posterior and used them to estimate the activation pattern. The estimated spatial activation pattern was determined at each $x \in \mathcal{X}$ as

$$\frac{1}{M}\sum_{j=1}^{M}\sum_{y_i\in\Phi_j} h(x-y_i;m^2).$$

In Fig. 13, the resulting estimates of the activation pattern are shown together with the true activation pattern. Note that for comparison the true activation pattern shown in Fig. 13 (left), has been multiplied with 13, the number of time points of activation in the true pattern.

The simulated point processes give us information about the temporal activation as well as the spatial activation. We estimate the temporal activation in the same manner as the spatial activation by

$$\frac{1}{M}\sum_{j=1}^{M}\sum_{s_i\in\Phi_j} g(t-s_i;m^1)$$

for each $t \in [0, T]$. The resulting estimates are shown in Fig. 14 together with the true activation pattern. Note that for comparison the true activation pattern shown in Fig. 14 (solid line) has been multiplied with 30, the number of activation centres in space in the true pattern.

To explain the higher level of activation obtained for $k=1$, we have divided $\mathcal{X}$ into two parts and estimated the temporal activation pattern for $k=1$ in each part separately. The

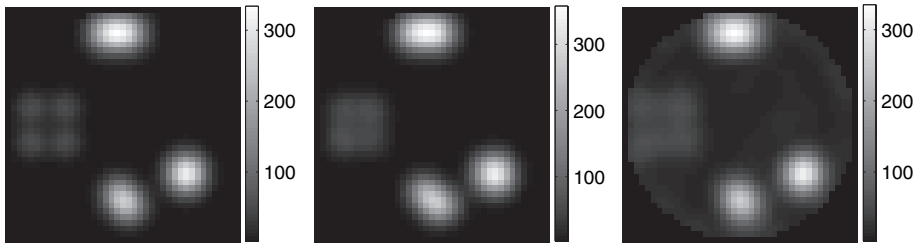

Fig. 13. The true spatial activation pattern (left), the estimated spatial activation pattern for $k=2$ (middle), and the estimated spatial activation pattern for $k=1$ (right). The marks have values $m^2=(4,4)$ in all the figures.

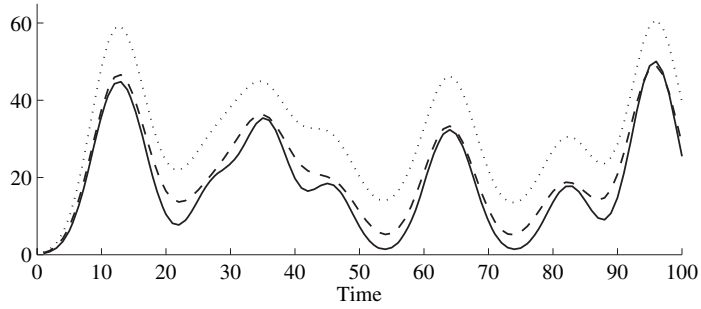

Fig. 14. The true temporal activation pattern (solid), the estimated temporal activation pattern for $k=2$ (dashed), and the estimated temporal activation pattern for $k=1$ (dotted). The mark has value $m^1=5$.

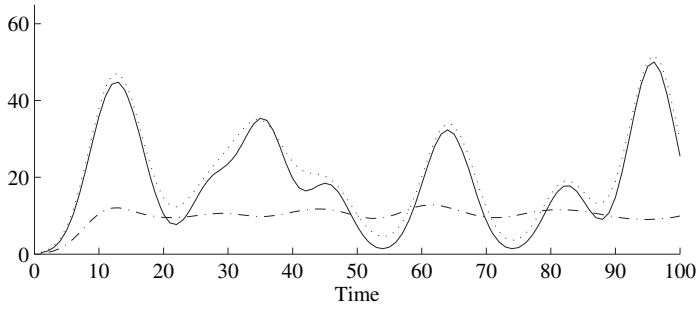

Fig. 15. The true temporal activation pattern (solid), the estimated temporal activation pattern for $k=1$ in activated areas (dotted), and the estimated temporal activation pattern for $k=1$ in non-activated areas (dash-dotted). The mark has value $m^1=5$.

first part, $\tilde{\mathcal{X}}_1$, consists of the areas that were used in the analysis for $k=2$, along with those voxels in $\mathcal{X}\setminus(\mathcal{X}_1\cup\mathcal{X}_2)$ influenced by voxels in $\mathcal{X}_1\cup\mathcal{X}_2$. Thus,

$$\tilde{\mathcal{X}}_1=\{x\in\mathcal{X}\,|\,\exists y\in\mathcal{X}_1\cup\mathcal{X}_2: h(x-y;m^2)\geq 2\}.$$

The threshold chosen equals half the maximum of the function h. The estimated temporal activation pattern for $\tilde{\mathcal{X}}_1$ is shown as a dotted line in Fig. 15. The second part, $\tilde{\mathcal{X}}_2$, consists of the remaining voxels in $\mathcal{X}$, $\tilde{\mathcal{X}}_2=\mathcal{X}\setminus\tilde{\mathcal{X}}_1$. The estimated temporal activation pattern for $\tilde{\mathcal{X}}_2$ is shown by a dash-dotted line in Fig. 15. The estimated temporal activation pattern for $\tilde{\mathcal{X}}_1$ is close to the true pattern, while the estimated pattern for $\tilde{\mathcal{X}}_2$ is approximately uniform in time apart from an edge effect in the beginning. The higher level of the temporal activation pattern for $k=1$ in Fig. 14 can thus be explained by extra points, uniformly distributed in time, in the simulated point process outside the activated areas.

9. An example of analysis of fMRI data

The data we consider here are from a larger investigation of the resting state network (cf. Beckmann *et al.*, 2005). The data have an exceptional high time resolution with 120 ms between neighbouring time points. Such high resolution enables the investigator to distinguish neural effects from non-neural physiological effects such as aliased cardiac or respiratory cycles. An independent component analysis of these data revealed a resting state network involving the sensory-motor cortices bilaterally (cf. Beckmann *et al.*, 2005). A Fourier analysis of the estimated temporal activation pattern of the network showed a dominating period of approximately 15–20 s.

The data are from a single slice through an axial plane that intersects the sensory-motor cortices bilaterally. The number of time points is 2000, corresponding to a total duration of the experiment of 4 min. After masking the data, in order to remove non-brain voxels, it consists of 932 voxels. We analyse this data under our model using Bayesian inference for one area, the whole slice, with $l=5s$, $\theta_1=100$ and $\theta_2=2$. Initially, we performed band-pass filtration with limits of 1 and 60 s so as to remove low-frequency drift as well as some of the effects relating to cardiac and respiratory cycles. We have used the program FSL for the preprocessing of the data (see Smith *et al.*, 2004 for an overview of FSL).

The resting state network found in Beckmann *et al.* (2005) involves three regions of interest: the middle region $\mathcal{X}_1$, the left motor cortex $\mathcal{X}_2$, and the right motor cortex $\mathcal{X}_3$ as shown in Fig. 16. A Bayesian analysis supports such a network. In Fig. 17 samples from the posterior density of time points of activation in the three regions of interest are shown. Clearly,

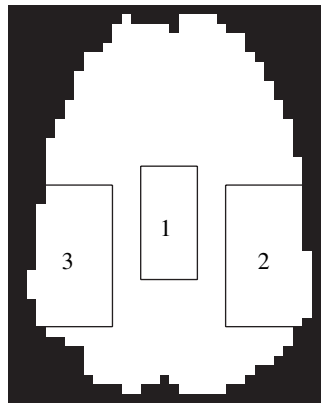

Fig. 16. Delineation of the three regions of interest: $\mathcal{X}_1$, a middle region; $\mathcal{X}_2$, includes the left motor cortex; and $\mathcal{X}_3$, includes the right motor cortex.

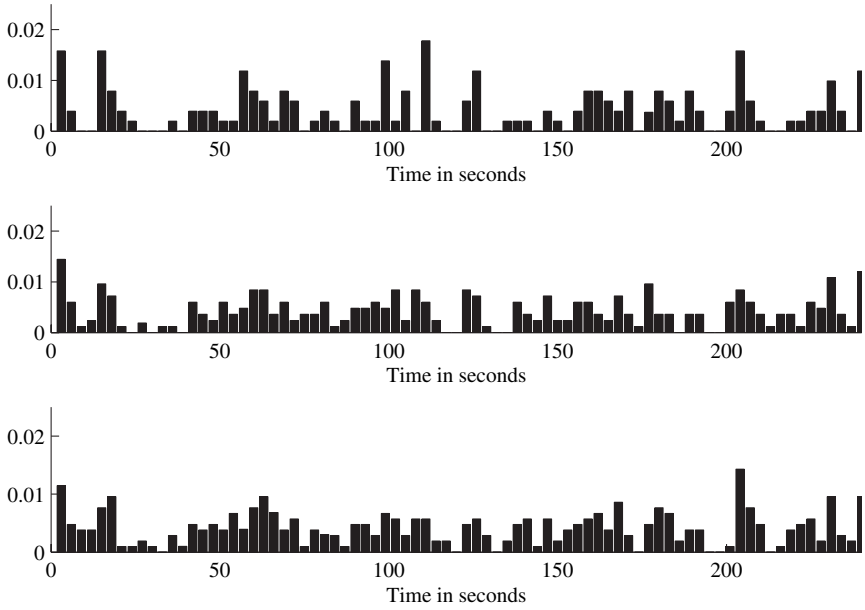

Fig. 17. Samples from the posterior density of time points of activation for region $\mathcal{X}_1$ (top), $\mathcal{X}_2$ (middle) and $\mathcal{X}_3$ (bottom). Each bar represents three seconds.

the three temporal point processes are positively correlated. Examples of observed and estimated time series are shown in Fig. 18 while the estimated spatial activation pattern in the three regions can be found in Fig. 19.

10. Discussion

In fMRI experiments, data may have a more complicated structure than the one predicted by our model (cf. e.g. Hartvig, 2002). An extended model will most likely include a drift component d_{tx}

$$Z_{tx} = \mu_x + d_{tx} + \sum_i f_{tx}(t_i, x_i; m_i) + \sigma_x \varepsilon_{tx}, \tag{26}$$

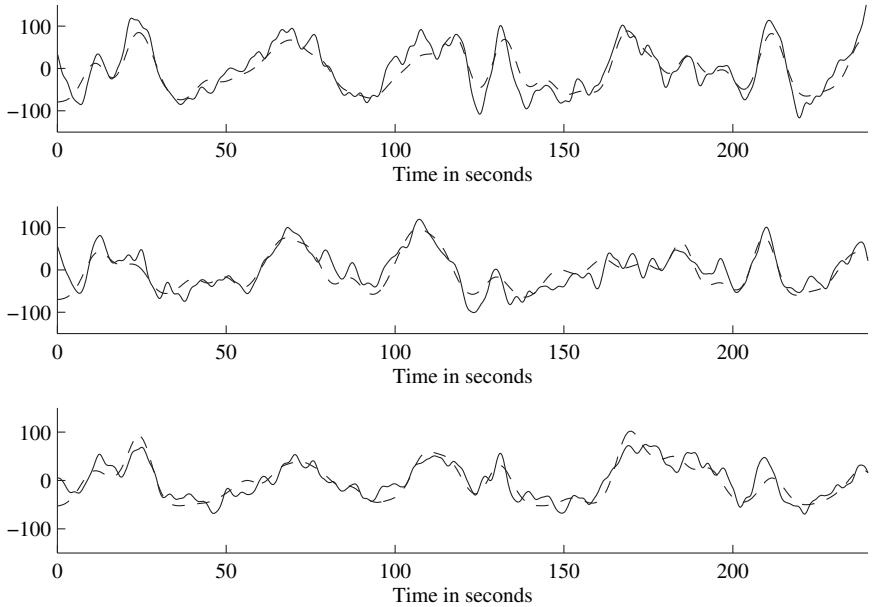

Fig. 18. Examples of time series from the three regions after pre-processing (solid) together with the estimated temporal activation (dashed). The top figure shows a time series from $\mathcal{X}_1$, the middle figure shows a time series from $\mathcal{X}_2$ and the bottom figure shows a time series from $\mathcal{X}_3$.

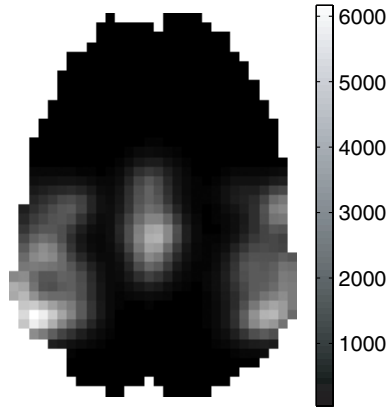

Fig. 19. The estimated spatial activation pattern in the three regions cumulated over time.

cf. Genovese (2000). This component describes the slow drifts in the static magnetic field during the experiment and residual motion not accounted for by prior motion correction. Often, the drift is removed using filtering, before any further analysis of the data (cf. Friston *et al.*, 2000), or included in a general linear model (cf. Friston *et al.*, 1995). It should also be a part of an initial analysis to examine whether the data should be transformed. In Hartvig (2002), log-transformed signal intensities are analysed by a model as in (26) with $\sigma_x^2 = \sigma^2$. Note that the variance of the untransformed intensities will then depend on t and x.

In the present paper, we have mainly used the simple model described in section 3.1.1 for the temporal activity, one reason being that we need to focus on the spatial modelling. In Genovese (2000), models for the HRF are reviewed, including a model based on splines. In

Purdon *et al.* (2001), a new model for a physiologically based haemodynamic response is described.

We have assumed that the errors $\{\varepsilon_{tx}\}$ are mutually independent. It is important to consider more general error models. In particular, the noise is often autocorrelated in time, as emphasized in Worsley (2000). A more general model for the errors is the multivariate Gaussian model,

$$\varepsilon \sim N_{|\mathcal{X}|\times T}(0, \Sigma). \tag{27}$$

For a standard whole brain analysis, the covariance model Σ will be very large, e.g. $|\mathcal{X}| \times T = 10{,}000 \times 100 = 10^6$. It is therefore necessary to make some simplifications of the model to make it computationally feasible. In Woolrich *et al.* (2004), this type of noise model is investigated in Bayesian settings. The authors propose the use of a space-time simultaneously specified autoregressive model,

$$\varepsilon_{tx} = \sum_{y\in\mathcal{N}_x} \beta_{xy}\varepsilon_{(t-1)y} + \sum_{s=1}^{3} \alpha_{sx}\varepsilon_{(t-s)x} + \eta_{tx},$$

where $\mathcal{N}_x$ is a neighbourhood of the voxel x, β_{xy} is the spatial autocorrelation between voxel x and y at a time lag of one with $\beta_{xy} = \beta_{yx}$, α_{sx} is the temporal autocorrelation between time point t and $t-s$ at voxel x, and $\{\eta_{tx}\}$ are independent noise variables with distribution

$$\eta_{tx} \sim N(0, \sigma_\eta^2).$$

It still remains to study more systematically explicit point process models that can describe how activities in different regions of the brain are related. The regions may either be activated simultaneously or some regions may be activated with delay compared with other regions. For modelling the spatial point process $\{x_i\}$, Taskinen (2001) has suggested a cluster point process. Note also that the point process term

$$\sum_i f_{tx}(t_i, x_i; m_i)$$

of (26) has the form of the random intensity field of a shot-noise Cox process if Ψ is Poisson (see e.g. Møller, 2003).

In the analysis of the real fMRI data, we have concentrated on Bayesian analysis because this type of analysis is expected to be the most interesting in practice. Inference based on the mean value relation will only be possible if the baseline signal μ_x is known from external sources. Inference based on covariances requires that one of the activation centers in the network is selected *a priori*.

An ICA analysis of the fMRI data resulted in a spatial activation pattern of the network similar to that provided by our Bayesian analysis in Fig. 19 (cf. Beckmann *et al.*, 2005, Fig. 4c). But the Bayesian analysis presented in our paper yields, in addition, information about the correlation between time points of activation in the different regions that are classified as one component in the ICA analysis (cf. Fig. 17).

From an applied point of view, an important next step is to design non-stimulus experiments along the lines described in Greicius *et al.* (2003) and analyse the data, using the modelling framework presented in this paper. In the examples considered until now, we have assumed that the marks are identical for all activations. The common value m has been treated as an unknown parameter. It will be interesting to include the distribution of the temporal duration of the activation in a Bayesian analysis.

One further possibility for extending the model is to consider K independent marked spatio-temporal point processes Ψ_k, $k = 1, \ldots, K$, instead of just one spatio-temporal point process. In the particular case of independent spatial and temporal point processes, where

$$\Psi_k = (\Psi_{k1}, \Psi_{k2}),$$

and $\Psi_{k1} = \{[t_{ki}; m^1_{ki}]\}$ and $\Psi_{k2} = \{[x_{kj}; m^2_{kj}]\}$ are independent, we obtain the following model equation

$$Z_{tx} = \mu_x + \sum_{k=1}^{K} A_{kt} B_{xk} + \sigma_x \varepsilon_{tx},$$

where

$$A_{kt} = \sum_i g(t - t_{ki}; m^1_{ki}) \quad \text{and} \quad B_{xk} = \sum_j h(x - x_{kj}; m^2_{kj}).$$

Note that $\{A_{k*}\}$ are independent corresponding to a temporal ICA model and $\{B_{*k}\}$ are independent corresponding to a spatial ICA model respectively (cf. McKeown *et al.*, 2003). The resulting model may be analysed by first performing an ICA analysis and then analysing the estimated components, using point process theory.

11. Summary

In the present paper, we have suggested a new modelling framework for non-stimulus experiments, using point process theory. The key idea is to replace the controlled on–off activation times from repeated stimulus experiments with *random* activation times from a stationary point process. Bayesian analysis of the model provides an estimate of the posterior distribution of the spatio-temporal point process of activations. Such dynamic output (including the posterior density of time points of activation) is not provided by the standard methods available for analysis. The model may be used for an exploratory analysis of the whole brain or a more detailed analysis of parts of the brain that have been spotted as regions of special interest in earlier analyses. As knowledge of the activation profile is used in the model, it is expected that the new approach will give a more clear picture of what is going on in the brain than non-parametric analysis like Greicius' regression method. The model offers a way of further analysing output from ICA analysis.

Acknowledgements

This work was supported by the Danish Natural Science Research Council. The authors thank Christian F. Beckmann for sharing his data. Further, the authors are grateful for fruitful discussions with Klaus B. Bærentsen, Anders C. Green and Hans Stødkilde-Jørgensen.

References

Beckmann, C. F. & Smith, S. M. (2005). Tensorial extensions of independent component analysis for multisubject fMRI analysis. *NeuroImage* **25**, 294–311.

Beckmann, C. F., DeLuca, M., Devlin, J. T. & Smith, S. M. (2005). Investigations into resting-state connectivity using independent component analysis. *Philos. Trans. R. Soc. Lond. Ser. B, Biol. Sci.* **360**, 1001–1013.

Buxton, R. B., Uludağ, K., Dubowitz, D. J. & Liu, T. T. (2004). Modelling the haemodynamic response to brain activation. *NeuroImage* **23**, 220–233.

Cao, J. & Worsley, K. (1999). The geometry of correlation fields with an application to functional connectivity of the brain. *Ann. Appl. Probab.* **9**, 1021–1057.

Friston, K. J. (1998). Modes or models: a critique on independent component analysis for fMRI. With discussion. *Trends Cogn. Sci.* **2**, 373–375.

Friston, K. J. (2002). Bayesian estimation of dynamical systems: an application to fMRI. *NeuroImage* **16**, 513–530.

Friston, K. J., Holmes, A. P., Poline, J.-B., Grasby, P. J., Williams, S. C. R., Frackowiak, R. S. J. & 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. & Turner, R. (1998). Event-related fMRI: characterizing differential responses. *NeuroImage* **7**, 30–40.

Friston, K. J., Josephs, O., Zarahn, E., Holmes, A. P., Rouquette, S., & Poline, J.-B. (2000). To smooth or not to smooth? *NeuroImage* **12,** 196–208.

Friston, K. J., Penny, W., Philips, C., Kiebel, S., Hinton, G. & Ashburner, J. (2002a). Classical and Bayesian inference in neuroimaging: theory. *NeuroImage* **16**, 465–483.

Friston, K. J., Glaser, D. E., Henson, R. N. A., Kiebel, S., Phillips, C. & Ashburner, J. (2002b). Classical and Bayesian inference in neuroimaging: applications. *NeuroImage* **16**, 484–512.

Genovese, C. R. (2000). A Bayesian time-course model for functional magnetic resonance imaging data. With discussion and a reply by the author. *J. Amer. Statist. Assoc.* **95**, 691–719.

Glover, G. H. (1999). Deconvolution of impulse response in event related fMRI. *NeuroImage* **9**, 416–429.

Greicius, M. D., Krasnow, B., Reiss, A. L. & Menon, V. (2003). Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. *Proc. Natl Acad. Sci. USA* **100**, 253–258.

Hartvig, N. V. (2002). A stochastic geometry model for functional magnetic resonance images. *Scand. J. Statist.* **29**, 333–353.

McKeown, M. J., Hansen, L. K. & Sejnowski, T. J. (2003). Independent component analysis of functional MRI: what is signal and what is noise? *Curr. Opin. Neurobiol.* **13**, 620–629.

Møller, J. (2003). Shot noise Cox processes. *Adv. Appl. Probab.* **35**, 614–640.

Møller, J. & Waagepetersen, R. P. (2004). *Statistical inference and simulation for spatial point processes*. Chapman & Hall/CRC, New York.

Ogawa, S., Tank, D. W., Menon, R., Ellermann, J. E., Kim, S., Merkle, H. & Ugurbil, K. (1992). Intrinsic signal changes accompanying sensory stimulation: functional brain mapping with magnetic resonance imaging. *Proc. Natl Acad. Sci. USA* **89**, 5951–5955.

Purdon, P. L., Solo, V., Weisskoff, R. M. & Brown, E. N. (2001). Locally regularized spatiotemporal modeling and model comparison for functional MRI. *NeuroImage* **14**, 912–923.

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., Bannister, P. R., De Luca, M., Drobnjak, I., Flitney, D. E., Niazy, R. K., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J. M. & Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage* **23**, S208–S219.

Stone, J. V. (2002). Independent component analysis: an introduction. *Trends Cogn. Sci.* **6**, 59–64.

Stoyan, D., Kendall, W. S. & Mecke, J. (1995). *Stochastic geometry and its applications*, 2nd edn. John Wiley & Sons, Chichester, UK.

Taskinen, I. (2001). *Cluster priors in the Bayesian modelling of fMRI data*. PhD Thesis. Department of Statistics, University of Jyväskylä, Jyväskylä.

Tonini, G., McIntosh, A. R., Russell, D. P. & Edelman, G. M. (1998). Functional clustering: identifying strongly interactive brain regions in neuroimaging data. *NeuroImage* **7**, 133–149.

Woolrich, M. W., Jenkinson, M., Brady, J. M. & Smith, S. M. (2004). Fully Bayesian spatio-temporal modeling of fMRI data. *IEEE Trans. Med. Imaging* **23**, 213–231.

Worsley, K. J. (2000). Comment on A Bayesian time-course model for functional magnetic resonance imaging by C. Genovese. *J. Amer. Statist. Assoc*. **95**, 691–719.

Received November 2004, in final form November 2006

Eva B. Vedel Jensen, Department of Mathematical Sciences, The T. N. Thiele Centre, University of Aarhus, Building 530, Ny Munkegade, DK-8000 Aarhus C, Denmark.
E-mail: eva@imf.au.dk