ELSEVIER

Pattern Recognition Letters 22 (2001) 1247–1252

Pattern Recognition Letters

www.elsevier.com/locate/patrec

Nonlinear regression analysis of the hemodynamic response in functional MRI

F. Kruggel *, D.Y. von Cramon

Max-Planck-Institute of Cognitive Neuroscience, Stephanstraße 1, 04103 Leipzig, Germany

Abstract

Functional magnetic resonance imaging is a non-invasive technique to study brain activity in humans. Among other effects, the stimulus pattern influences the timecourse of activations. A non-linear regression context is proposed to quantitatively compare stimulation parameters with activation timecourses. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Functional magnetic resonance imaging; Hemodynamic response; Non-linear regression; Event-related experimental paradigm

1. Introduction

Functional magnetic resonance imaging (fMRI, Belliveau et al., 1991) has become one of the major methods for investigating brain function in cognitive science. Most fMRI studies apply the blood-oxygen-level-dependent (BOLD) effect, where a metabolic correlate of brain activation, the socalled hemodynamic response (HR) of the vascular system, is measured. Activations are found as stimulus-linked and time-dependent local intensity changes in two-dimensional images which are acquired during stimulation at regular time intervals. Data analysis is complicated by the following circumstances: (1) the effect is rather small, (2) data are noisy and overlaid by artifacts, (3) the HR due to a brief stimulus is typically delayed by 3–5 s and dispersed by 1.5–2.5 s, and (4) perhaps also displaced with respect to the activation site.

Recently, event-related experimental designs were introduced (ER-fMRI, Zarahn et al., 1997). Such designs allow the randomized presentation of behavioural trials and the study of responses to a specific stimulation context separately. As a consequence, there is an increasing interest in describing the timecourse of the HR in relation to experimental stimulation parameters, or: how much can be inferred from the HR shape characteristics about the underlying neuronal activation?

This paper proposes a flexible non-linear regression context for modeling the HR in fMRI data. We explain the framework, the background model, and inferential statistics for the model in the following section. Three examples illustrate the use of this context.

2. A modeling context for the hemodynamic response

Typically, fMRI data are collected as a set of 2D image slices recorded at different positions of the brain (say, every 7 mm) and at regular time intervals (say, every 1 s). Activated brain regions

* Corresponding author.

PII: S0167-8655(01)00056-3

may be detected by any viable choice of fMRI signal detection methods (for a review, see Lange, 1996). We define regions-of-interest (ROIs) by selecting a few highly activated voxels from a brain region and then focus on modeling the timecourse of the signal in this ROI. We will now formally introduce this modeling context.

2.1. Model layout

We consider a subset $\boldsymbol{y}$ of the fMRI data, collected from k 4-connected voxel sites at l discrete timesteps (i.e. $n = kl$ data points) of a single experimental period (a trial). Data are modeled as a sum of a deterministic function $g(\cdot)$ and a stochastic part $\boldsymbol{\epsilon}$:

$$\boldsymbol{y} = g(\boldsymbol{t}, \boldsymbol{\beta}) + \boldsymbol{\epsilon}, \quad \boldsymbol{\epsilon} \sim N_n(0, \boldsymbol{V}), \tag{1}$$

where $\boldsymbol{t}$ denotes the vector of discrete timesteps and $\boldsymbol{\beta}$ the p-dimensional vector of shape parameters of a HR model function. $g(\cdot)$ is bounded in time within the experimental period, and undefined otherwise. The elements ϵ_i of the stochastic part are uncorrelated with $g(\cdot)$. In the following section, we discuss how an estimate of the covariance matrix $\hat{\boldsymbol{V}}$ is determined from the data. The ML estimate $\hat{\boldsymbol{\beta}}$ of our model parameters is found as the vector $\boldsymbol{\beta}$ from non-linear minimization in p dimensions:

$$\arg\min_{\beta}(\hat{\boldsymbol{\epsilon}}^{\mathrm{T}} \hat{\boldsymbol{V}}^{-1} \hat{\boldsymbol{\epsilon}}), \quad \text{where} \quad \hat{\boldsymbol{\epsilon}} = \boldsymbol{y} - g(\boldsymbol{t}, \boldsymbol{\beta}). \tag{2}$$

2.2. Covariance matrix

The spatio-temporal covariance matrix $\boldsymbol{V}$ contains $n(n+1)/2$ unknowns and thus may not be determined from the data without making certain simplifications. The following assumptions are approximately fulfilled for preprocessed fMRI data (for a full discussion, see Kruggel and von Cramon, 1999): (1) elements ϵ_i are normally distributed, (2) their correlation is described by an AR(1) model in space and time (Bullmore et al., 1996, Benali et al., 1997), (3) the covariance matrix is separable in space and time: $\boldsymbol{V} = \boldsymbol{S} \odot \boldsymbol{T}$.

The determination of the spatial correlation matrix is somewhat involved, because residuals are only available on an irregular configuration of sites. The AR(1) correlation function of a stationary process on a regular spatial grid is given by:

$$\rho_{\mathrm{S}}(s_0, s_1) = \rho_X^h \rho_Y^v \quad \text{using } h = H(s_0, s_1),\ v = V(s_0, s_1), \tag{3}$$

where $H(\cdot)$ and $V(\cdot)$ return the absolute distance between two sites in the x and y direction, and ρ_X, ρ_Y denote the spatial autocorrelations. To determine ρ_X and ρ_Y from the data, we first form subsets of all pairs of sites, which are located the same absolute distance h, v:

$$S_{h,v} = \{s_0,\ s_1 \,|\, s_0 \in S,\ s_1 \in S,\ H(s_0, s_1) = h,\ V(s_0, s_1) = v\} \tag{4}$$

then compute the maximum likelihood estimate $\hat{\rho}_{h,v}$ for this subset. Now, we find $\hat{\rho}_X, \hat{\rho}_Y$ by LS estimation:

$$\arg\min_{\rho_X, \rho_Y} \left\{ \sum_{h,v} (\hat{\rho}_{h,v} - \rho_X^h \rho_Y^v)^2 \right\}. \tag{5}$$

Using these estimates, we can compute the correlation $\hat{\rho}_{\mathrm{S}}$ between any sites s_i, s_j in the ROI by Eq. (3), and thus set up the spatial correlation matrix $\hat{\boldsymbol{S}}$. Similarly, a matrix $\hat{\boldsymbol{T}}$ is formed for the temporal domain and composed with the spatial matrix as indicated above.

2.3. Goodness-of-fit and confidence limits

A simple measure for conformance of the time series with a given model (GOF) is given by:

$$\mathrm{GOF} = 1 - \frac{\hat{\boldsymbol{\epsilon}}^{\mathrm{T}} \hat{\boldsymbol{V}}^{-1} \hat{\boldsymbol{\epsilon}}}{\boldsymbol{y}^{\mathrm{T}} \hat{\boldsymbol{V}}^{-1} \boldsymbol{y}}, \tag{6}$$

which ranges in $[0, 1]$, with 1 denoting a perfect fit. Using a first-order linear model, we can derive confidence limits for $\boldsymbol{\beta}$ from the inverse of the Fisher information matrix $\boldsymbol{F}$:

$$\hat{\boldsymbol{\beta}} \sim N(\boldsymbol{\beta}, \boldsymbol{F}_{\boldsymbol{\beta}}^{-1}), \quad \text{where} \quad \boldsymbol{F}_{\boldsymbol{\beta}} = \boldsymbol{G}_{\boldsymbol{\beta}} \hat{\boldsymbol{V}}^{-1} \boldsymbol{G}_{\boldsymbol{\beta}}^{\mathrm{T}}, \tag{7}$$

and $\boldsymbol{G}_{\boldsymbol{\beta}}$ denotes the Jacobian matrix of $g(\cdot)$ with respect to $\boldsymbol{\beta}$. Exact confidence limits on the estimated parameters may be derived using Hartley's suggestion (see Seber and Wild, 1989, p. 236):

$$\left\{ \boldsymbol{\beta} : \hat{\boldsymbol{\epsilon}}^{\mathrm{T}} \boldsymbol{P} \hat{\boldsymbol{\epsilon}} \leqslant \hat{\boldsymbol{\epsilon}}^{\mathrm{T}} \hat{\boldsymbol{V}}^{-1} \hat{\boldsymbol{\epsilon}} \frac{f}{1+f} \right\}$$

$$\text{where} \quad f = F^{\alpha}_{p,n-p} \frac{p}{n-p}$$

$$\text{and} \quad \boldsymbol{P} = \hat{\boldsymbol{V}}^{-1} \boldsymbol{G}_{\boldsymbol{\beta}}^{\mathrm{T}} \boldsymbol{F}_{\boldsymbol{\beta}}^{-1} \boldsymbol{G}_{\boldsymbol{\beta}} \hat{\boldsymbol{V}}^{-1}, \tag{8}$$

which is a $100(1-\alpha)\%$ confidence region for $\boldsymbol{\beta}$.

3. Examples

Three different examples will now illustrate the use of this modeling context. In the first example, we show how HR shape parameters may be linked to experimental stimulation parameters using a simple model function. The second example introduces a more elaborate model, by which we try to separate "neuronal" from "vascular" shape parameters. A complex model function for multiple HRs within a trial is discussed as a third example.

3.1. Gaussian function

The most simple suitable choice for $g(\cdot)$ in Eq. (1) is given by a 4-parameter Gaussian function:

$$g(t, \boldsymbol{\beta}) = \beta_0 \exp\left[-\frac{(t-\beta_2)^2}{2\beta_1^2} \right] + \beta_3, \tag{9}$$

where we denote the components of $\boldsymbol{\beta}$ as β_0: gain (the "height" of the HR), β_1: dispersion (proportional to the duration of the HR), β_2: lag (the time delay from stimulation onset to the HR peak), and β_3: baseline.

To illustrate results using this function, we selected a fMRI dataset from a recent study in language comprehension (Meyer et al., 2000). Single sentences were presented orally, and subjects were asked to classify the sentence grammatically for correctness. The presentation of a sentence lasted between 2.56 and 4.46 s during a 6 s interval, followed by a 18 s pause. Seventy-six trials were recorded during an approximately 30 min experiment. During this time, we acquired every 2 s four slices of 128×64 voxels with a spatial resolution of $1.9 \times 3.8 \times 5$ mm^3.

The standard procedures to detect functional activation in this dataset were performed: (1) *preprocessing* using motion correction, baseline correction and lowpass-filtering to reduce physiological and system noise (Kruggel et al., 1999), (2) *statistical analysis* for activated regions by Pearson correlation with a time-shifted box-car waveform ($\Delta = 6$ s), and conversion of the correlation coefficients into z-scores, (3) *assessment of significance* to the activated regions on the basis of their spatial extent (Friston et al., 1994). We focus on results of two ROIs (see Fig. 1): the primary auditory cortex on the left side $\mathrm{PAC_L}$, and an

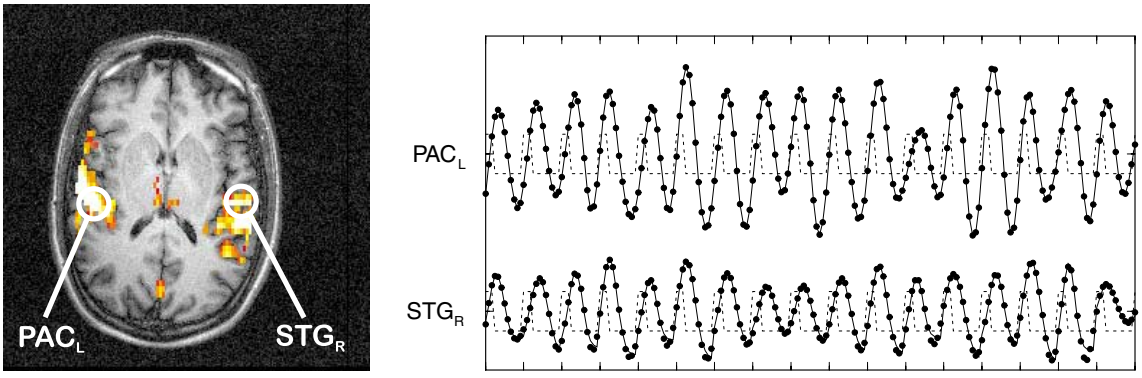

Fig. 1. Activated brain regions in a fMRI experiment on language comprehension (left). A few trials from the (spatially averaged) timecourse of the signal in two sample regions are shown on the right. Dotted lines denote the acoustic stimulation, the thick dots denote the measurements, and the black line the estimated waveforms.

auditory association cortex on the right superior temporal gyrus (STG_R).

Using our model, we obtained 76 estimates of gain, lag and dispersion for each ROI. These estimates were compared with the experimental parameters (sentence presentation length, correctness manipulation) using linear regression. When evaluating over all ROIs, we found that per second of increasing sentence length, the lag increased by 880 ms and the dispersion by 210 ms. So within this temporal stimulus range, there is a proportionality between stimulus and HR duration. Likewise, the gain increased by 21% per second stimulus duration. A significant interaction ($P < 0.05$) between the gain and the correctness manipulation was found for the association cortex STG_R, but not for the primary auditory cortex PAC_L.

3.2. Convolved asymmetric Gaussian function

An approach to closer model to processes underlying the HR is to define the model function $g(\cdot)$ as a convolution of a neuronal stimulation function $n(t)$ with a hemodynamic modulation function $f(t)$:

$$g(t, \boldsymbol{\beta}) = n(t) \otimes f(t) + b, \tag{10}$$

where $\otimes$ denotes the convolution operator and b is a baseline term. We simply assume a square-wave function for the neuronal stimulation $n(t)$:

$$n(t) = \begin{cases} a & \text{if } t \geqslant t_0 \text{ and } t < t_0 + t_1, \\ 0 & \text{otherwise.} \end{cases} \tag{11}$$

A Gaussian function is introduced for hemodynamic modulation function $f(\cdot)$, here with different dispersions (d_0, d_1) for the rising and the falling edge:

$$f(t) = \begin{cases} \exp(-t^2/(2d_0^2)) & \text{if } t < 0, \\ \exp(-t^2/(2d_1^2)) & \text{if } t \geqslant 0. \end{cases} \tag{12}$$

In this model, $\boldsymbol{\beta}$ consists of six parameters (d_0: dispersion on the rising edge, d_1: dispersion on the falling edge, a: gain, t_0: response onset, t_1: response duration, b: offset). Modeling of the convolution process allows to address the meaning of a and t_1 as neuronal parameters, resp. t_0, d_0, d_1 as vascular parameters.

Applying this model to data from region PAC_L, we found that the response onset t_0 was independent of the sentence length, and varied only within relatively small margins (3.49 ± 0.30 s), while the response duration t_1 was proportional to the sentence length (+964 ms per second sentence length). In addition, the onset dispersion d_0 was independent of the sentence length, while d_1 increased by 320 ms per second stimulation length. d_0 was always shorter than d_1 (3.18 vs. 3.62 s), indicating a slightly asymmetric HR. Thus, it appears feasible to separate neuronal from vascular parameters by this estimation context. The model applied here, however, is a gross simplification of physiological processes underlying the HR, and needs to be enhanced whenever more information about the neurono-vascular coupling becomes available.

3.3. Modeling multiple responses

As a final example, we selected an experiment in human working memory, where multiple HRs per trial are expected. Subjects were shown a cue set of 3–6 letters. After a variable delay length (2.0–7.0 s) a probe letter appeared, and the subject was asked to respond via a button press whether the probe letter belonged to the previously presented set. An experimental run consisted of 48 different combinations, subjects completed 4 runs.

As a model function $g(\cdot)$ we set up of a sum of two Gaussian functions:

$$g(t, \boldsymbol{\beta}) = \beta_0 \exp\left[-\frac{(t-\beta_2)^2}{2\beta_1^2} \right] + \beta_3 \exp\left[-\frac{(t-\beta_5)^2}{2\beta_4^2} \right]. \tag{13}$$

Because of the HR dispersion, cue and probe phase responses merge for short delay times. Modeling each trial separately would lead to biased estimates, so here, we included the whole timeseries in the model. Unlike previous examples with separate modeling and analysis stages, parameters $\boldsymbol{\beta}$ were now directly set up as functions of the stimulation context of a given trial.

Results for two sample regions are shown in Fig. 2. Because we modeled the whole timeseries,

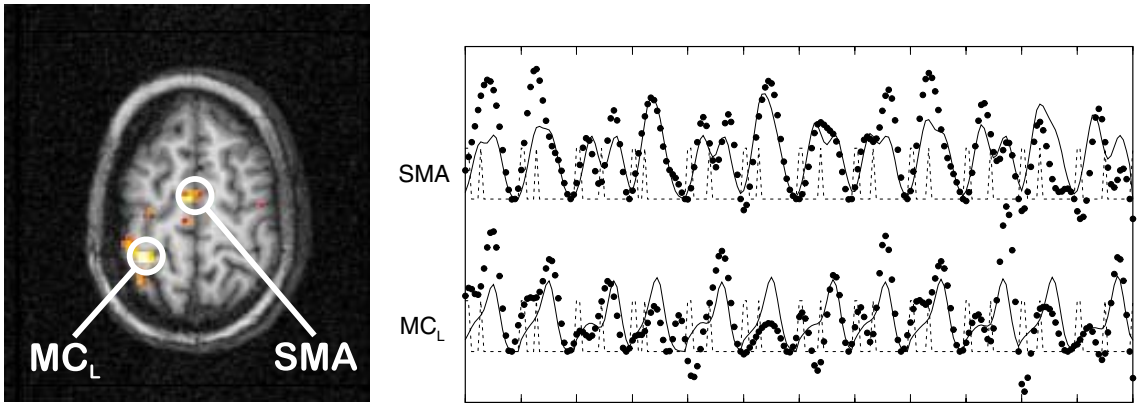

Fig. 2. Activated brain regions in a fMRI experiment on working memory (left). A few trials from the (spatially averaged) timecourse of the signal in two sample regions are shown on the right. Dotted lines denote the visual stimulation. Note the variable delay time between cue and probe phase. The thick dots correspond to the measurements, and the black lines to the estimated waveforms.

GOF values were not as high (0.70–0.95) as in the first example, however, confidence limits as derived in Section 2.3 still proved to be small enough to allow drawing several conclusions. Both regions were active in both phases: for the cue phase in the supplementary motor area (SMA), a marked dependency of the gain on the set size was determined, while the probe phase response was independent of the set size manipulation in both areas. Not surprisingly, the motor cortex (MC_L) was predominantely active during the probe phase, were a motor response was requested. The time difference between cue and probe phase response matched the delay time manipulation with a slope of 1.015.

4. Conclusion

We proposed a non-linear regression context to model the hemodynamic response in ER–fMRI. A Gaussian function fits well to single responses, and parameters of this function (i.e. gain, lag and dispersion) may be assigned a physiological meaning, which are readily compared with experimental stimulation parameters. The application of this model was illustrated in three variants using recent fMRI experiments.

The choice of the Gaussian function is heuristic: whenever a deeper understanding about the physiological properties underlying the hemodynamic response is available, any functional relation may be embedded within this context. However, in closely timed event-related experiments, usually a few timepoints per trial (6–20) are recorded, so the number of parameters of a model function are limited. Note that we modeled linear relations between stimuli and HR parameters only. First results about a non-linear response behaviour are available (Robson et al., 1998, Vazquez and Noll, 1998), but for longer stimulus durations than those used in our experiments. All of the three variants proposed here are readily extendible to detect and model non-linear behaviour.

Increasingly complex experimental designs require more elaborate statistical procedures to quantitatively compare stimulus and response. Our non-linear modeling context offers a high degree of flexibility. It may serve as another tool to face the challenge of understanding brain function.

Acknowledgements

We would like to thank the anonymous reviewers for helpful comments on this manuscript.

References

Belliveau, J.W., Kennedy, D.N., McKinstry, R.C., Buchbinder, B.R., Weisskoff, R.M., Cohen, M.S., Vevea, J.M., Brady, T.J., Rosen, B.R., 1991. Functional mapping of the human

visual cortex by magnetic resonance imaging. Science 254, 716–719.

Benali, H., Buvat, I., Anton, J.L., Pelegrini, M., Di Paola, M., Bittoun, J., Burnod, Y., Di Paola, R. 1997. Space–time statistical model for functional MRI image sequences. In: Inf. Process. Medical Imaging, Lecture Notes in Computer Science, Vol. 1230, Springer, Heidelberg, pp. 285–298.

Bullmore, E., Brammer, M., Williams, S.C.R., Rabe-Hesketh, S., Janoth, N., David, A., Mellers, J., Howard, R., Sham, P., 1996. Statistical methods of estimation and inference for functional MR image analysis. Magn. Reson. Med. 35, 261–277.

Friston, K.J., Worsley, K.J., Frackowiak, R.S.J., Mazziotta, J.C., Evans, A.C., 1994. Assessing the significance of focal activations using their spatial extent. Hum. Brain Mapp. 1, 210–220.

Kruggel, F., Cramon, D.Y., 1999. Modeling the hemodynamic response in single-trial functional MRI experiments. Magn. Reson. Med. 42, 787–797.

Kruggel, F., Descombes, X., von Cramon, D.Y., 1999. Comparison of filtering methods for fMRI datasets. Neuroimage 10, 530–543.

Lange, N., 1996. Tutorial in biostatistics: Statistical approaches to human brain mapping by functional resonance imaging. Stat. Med. 15, 389–428.

Meyer, M., Friederici, A.D., von Cramon, D.Y., 2000. Neurocognition of auditory sentence comprehension: event-related fMRI reveals sensitivity to syntactic violations and task demands. Cogn. Brain Res. 9, 19–33.

Robson, M.D., Dorosz, J.L., Gore, J.C., 1998. Measurement of the temporal fMRI response of the human auditory cortex to trains of tones. Neuroimage 7, 185–198.

Seber, G.A.F., Wild, C.J., 1989. Nonlinear Regression. Wiley, New York.

Vazquez, A.L., Noll, D.C., 1998. Nonlinear aspects of the BOLD Response in functional MRI. Neuroimage 7, 108–118.

Zarahn, E., Aguirre, G., D'Esposito, M., 1997. A trial-based experimental design for fMRI. Neuroimage 6, 122–138.