Magnetic Resonance in Medicine 42:787–797 (1999)

Modeling the Hemodynamic Response in Single-Trial Functional MRI Experiments

Frithjof Kruggel* and D. Yves von Cramon

Today, most studies of cognitive processes using functional MRI (fMRI) experiments adopt a single-trial design. Highly flexible stimulation paradigms require new statistical models in which not only the activation amount but also the time course of the measured hemodynamic response is analyzed. Most previous approaches have been based on a linear regression context and have introduced hemodynamic model functions to improve the signal detection. In this report a nonlinear regression context is derived, from which shape parameters for the hemodynamic response are obtained per trial and per region of interest. These parameters allow the investigation of stimulus-induced shape variations of the hemodynamic response. By embedding the estimation into a robust statistical framework and rigorously analyzing the spatiotemporal interactions in the fMRI data, it is possible to derive statistically valid descriptions of single hemodynamic responses. The model, estimation algorithm, validation, and an example analysis from a single-trial fMRI study are reported. Magn Reson Med 42:787–797, 1999. © **1999 Wiley-Liss, Inc.**

Key words: fMRI; single-trial design; hemodynamic response; Ornstein-Uhlenbeck process; functional data analysis; nonlinear regression

Within the last few years, functional MRI (fMRI) has gained enormous interest in the study of cognitive processes in humans. Primary reasons for its success are found in its high spatial and (moderate) temporal resolution, its noninvasiveness, and its ease of application. However, the blood oxygen level–dependent (BOLD) effect (1–3) measured in fMRI is only indirectly linked to the neuronal activation via mechanisms of the neuronovascular coupling, which currently are not fully understood. It is now commonly accepted that the hemodynamic response (HR) after neuronal activation consists of several phases: 1) Within milliseconds after stimulus onset, the local oxygen consumption elicits a transient shortening of the T_2^* time and thus causes a signal decrease (4,5). 2) The initial hypooxygenation in the active region is oversupplied by an increased inflow of oxygenated hemoglobin (HbO_2), leading to a signal increase at a maximum of 5–6 sec after stimulus onset (6). 3) Limitations of the vascular regulation and/or transport of the HbO_2 excess induce a dispersion of the signal increase by 3–4 sec. 4) a final negative dip has been reported but not yet interpreted.

It is currently under discussion to what extent properties of the underlying neuronal activation can be inferred from the shape of the HR. It is not unreasonable to hypothesize that 1) a stronger neuronal activation leads to an increased bold response; 2) a prolonged neuronal activation is accompanied by a prolonged HR; and 3) a time difference in the activation onset (i.e., between a sensory and a response-related area) is reflected by a temporal shift in the HRs of these areas. However, whereas a cognitive task under study takes a few hundred milliseconds to complete, the time scale of the vascular response is much longer. Furthermore, the blood supply is not uniform throughout the brain, and besides regional differences other physiological variables (e.g., heart rate, blood pressure, slope of perfusion changes) might influence the shape of the HR. Nevertheless, recent fMRI studies have shown that experimentally introduced delays led to temporal shifts of the HR (7,8) that were indeed on the order of the neuronal time scale.

Recently, single-trial fMRI designs (STD) have been introduced (9–11). A trial-based behavior is understood as a ''self-contained behavioural/perceptual unit which is temporally limited'' (9). Such designs allow the randomization of the presentation of behavioral trials and the study of the response to a specific stimulation separately.

The focus of this report is to describe a general statistical context for modeling the HR in single-trial fMRI experiments. Four concepts are important in the context of this model and deserve discussion beforehand:

1. We separate the signal detection problem from the signal description problem. Throughout this report, we assume that the locus of a functional activation is known. This knowledge may arise from a previous determination by well-established signal detection procedures or defined as regions of neurofunctional interest (ROI). In these regions we determine a trial- and ROI-wise description of the HR by neurophysiologically meaningful parameters. The results of such an estimation procedure are given as a three-dimensional table, where the first dimension corresponds to separate ROIs, the second dimension to the individual experimental trials, and the third dimension to the set of estimated parameters.
2. We introduce a nonlinear regression context which allows a wide variety of functional models for the HR. Because the signal-to-noise ratio in fMRI data is low, special care must be taken to ensure the stability, validity and accuracy of the estimation procedures.
3. We separate the statistical evaluation of the experiment from the signal description. The data table as a result from the estimation process is a very compact model representation of the brain activation elicited in this experiment and may be imported into any statistical package for further evaluation.
4. Regarding trials as behavioral entities, we now have a single set of HR parameters per region available for combination with parameters of the experimental design and behavioral data (responses, reaction times,

Max-Planck-Institute of Cognitive Neuroscience, Leipzig, Germany.
*Correspondence to: Frithjof Kruggel, Max-Planck-Institute of Cognitive Neuroscience, Stephanstrasse 1, D-04103 Leipzig, Germany.
E-mail: kruggel@cns.mpg.de

Received 3 August 1998; revised 16 June 1999; accepted 7 July 1999.

etc.). This allows the setting up of very clearly defined statistical models without the need of handling baseline conditions, signal drift, or questions of activation significance. The compact representation makes quick testing of different statistical models and alternate hypotheses possible.

The remainder of this report is organized as follows. In the next section, we introduce a nonlinear regression model suitable for the estimation of the HR in fMRI data. We then report on the model validation using simulated data and a comparison of confidence intervals determined by the mathematical model with a Monte Carlo simulation. We then underline the usage of this model with results from an example fMRI experiment. In the Discussion, we weight advantages and disadvantages of this model. This report is primarily concerned with theoretical aspects of the model. Detailed neurobiological results will be presented separately.

A NONLINEAR REGRESSION MODEL FOR fMRI

In a quite general model of focal brain activation we consider a measured event $E(e, s, t, \cdot\cdot\cdot)$, which is dependent on the stimulation event e, a brain site s, and the time t. The dots denote a number of other influences (aperiodic and quasiperiodic body motion, scanner instabilities, noise, motivation, etc.), which are not discussed here. These influences have either been reduced by a preprocessing step (12) or are collectively modeled as an (additive) stochastic part $\boldsymbol{\epsilon}$. Let us assume that a significantly activated brain region s has been determined by any of the well-established methods (see ''Example fMRI Analysis'' and ''Importance of Temporal Characterization''), which is constant in position and extent. These assumptions simplify the event model to $E(e, s, t) + \boldsymbol{\epsilon}$. The estimation problem is now the adaptation of a model function to the time course of E.

Description of the Model

This section focuses on the maximum likelihood (ML) estimation of hemodynamic model parameters from fMRI data (13,14). Let $\mathbf{S}$ be an ROI of k 8-connected voxels, and $\mathbf{T}$ an l-dimensional vector of discrete time steps in a single experimental trial. Denote a subset of the fMRI data $\mathbf{Y} = \{y(s, t),\ s \in \mathbf{S},\ t \in \mathbf{T}\}$. We model the data as a sum of a deterministic function $g(t, \boldsymbol{\beta})$, where $\boldsymbol{\beta}$ denotes a p-dimensional vector of model parameters, and a stochastic part $\boldsymbol{\epsilon} = \{\epsilon(s, t), s \in \mathbf{S}, t \in \mathbf{T}\}$:

$$y(s, t) = g(t, \boldsymbol{\beta}) + \epsilon(s, t)\ \forall\ s \in \mathbf{S},\ t \in \mathbf{T}. \qquad [1]$$

Thus, we have $n = k\,l$ equations to estimate the p model parameters $\boldsymbol{\beta}$, with $p < n$. To analyze the problem using standard results from multivariate calculus, it is required that $g(t, \boldsymbol{\beta})$ is differentiable with respect to $\boldsymbol{\beta}$. Furthermore, we assume that the stochastic part is independent of the signal and stationary with respect to time, and its elements are normally distributed with a nonsingular covariance matrix $\mathbf{V}$:

$$\boldsymbol{\epsilon} \sim N_n(0, \mathbf{V}), \qquad \text{then} \qquad \mathbf{Y} \sim N_n(\mathbf{g}(\mathbf{T}, \boldsymbol{\beta}), \mathbf{V}), \qquad [2]$$

where $\mathbf{g(T, \beta)}$ denotes the vector-valued equivalent of $g(\cdot)$. We will indicate a way to determine the covariance structure from experimental data later in this section.

The ML estimate $\hat{\boldsymbol{\beta}}$ is obtained as the vector $\boldsymbol{\beta}$ that minimizes the quantity:

$$\arg\min_{\boldsymbol{\beta}} \{(\mathbf{Y} - \mathbf{g}(\mathbf{T}, \boldsymbol{\beta}))^{\mathrm{T}} \mathbf{V}^{-1} (\mathbf{Y} - \mathbf{g}(\mathbf{T}, \boldsymbol{\beta}))\}. \qquad [3]$$

Any nonlinear optimization algorithm, such as the downhill simplex method (15), may be used to obtain $\hat{\boldsymbol{\beta}}$.

Model Functions for the Hemodynamic Response

A number of functions have been proposed to model the HR: the Poisson function (16), the γ function (17,18) and the Gaussian function (19), which all are viable choices for $g(\cdot)$. In the context of this report, we concentrate on the Gaussian function:

$$g(t, \boldsymbol{\beta}) = \frac{\beta_0}{\beta_1} \exp\left(-\frac{(t - \beta_2)^2}{2{\beta_1}^2}\right) + \beta_3, \qquad \text{where} \qquad t \in \mathbf{T}. \qquad [4]$$

We propose this function because it offers the best goodness of fit (19), and its model parameters are interpretable in terms of physiological variables. We denote the components of $\boldsymbol{\beta}$ as $\boldsymbol{\beta}_0$: gain (the ''height'' of the HR), $\boldsymbol{\beta}_1$: dispersion (in seconds, proportional to the duration of the HR), $\boldsymbol{\beta}_2$: lag (in seconds, the time delay from stimulation onset to the HR peak), and $\boldsymbol{\beta}_3$: baseline. Parameters $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$ are constrained to be positive and within the length of a trial period: $0 < \boldsymbol{\beta}_1 \leq t_{l-1}$ and $0 < \boldsymbol{\beta}_2 \leq t_{l-1}$. As an independent measure derived from the data, we define the norm of an HR as the difference between a data point and the baseline value, summed over all points and time steps in a trial: $\Sigma_{s \in \mathbf{S}} \Sigma_{t \in \mathbf{T}} (y(s, t) - \boldsymbol{\beta}_3)$.

The Stochastic Part

The spatiotemporal covariance matrix V, as introduced in Eq. [2], contains n^2 unknowns and thus may not be determined directly without making certain assumptions. Covariance structures in fMRI data were first examined in the temporal domain by Bullmore et al. (20) and in the spatial domain by Benali et al. (21). Both found an AR(1) model in good accordance with the data. To ensure that the ML procedure (see Eq. [3]) yields an unbiased estimate, it is necessary to show that the residuals are (asymptotically) normally distributed and stationary in time and space. Any process which is normally distributed, stationary and has an AR(1) covariance structure may be denoted as a discrete realization of an Ornstein-Uhlenbeck process (22,23), and hence is Markovian. We now demonstrate how 1) normality, 2) stationarity, and 3) the covariance structure are determined from the data. These assumptions also assure that the covariance matrix is invertible.

Normality

The normality of the residuals is assessed by the analysis of the first four moments. Let us denote γ_1 as the mean, γ_2 as the dispersion, γ_3 as the skewness, and γ_4 as the kurtosis of

the residuals ϵ, which are computed by using standard textbook procedures (24). We collect all n residuals from all m trials (i.e., $q = n\ m$ data points) in a check for normality. According to our model in Eq. [2], we require that $\gamma_1 \approx 0$. The normality assumption will be rejected at a significance level of 5% if

$$T_1 = \frac{|\gamma_3|}{2\sqrt{\nu_1}} > 1 \quad \text{using} \quad \nu_1 = \frac{6(q-2)}{(q+1)(q+3)}, \quad \text{or} \quad [5]$$

$$T_2 = \frac{|\gamma_4 - 3 + (6/(q+1)|}{2\sqrt{\nu_2}} > 1$$

$$\text{using} \quad \nu_2 = \frac{24q(q-2)(q-3)}{(q+1)^2(q+3)(q+5)}. \quad [6]$$

Stationarity

Weak second-order stationarity is tested across all $r = l\ m$ time steps within a ROI. For the temporal domain, we apply Anscombe's test and check the correlation of the residuals with time t (24):

$$c_t = \frac{\sum_{t=0}^{r-1}\left[\sum_{s\in\mathbf{S}}(\epsilon(s,t)-\bar{\epsilon})(t-\bar{t})\right]}{\sqrt{\sum_{t=0}^{r-1}\left[\sum_{s\in\mathbf{S}}(\epsilon(s,t)-\bar{\epsilon})^2(t-\bar{t})^2\right]}}, \quad [7]$$

where bars denote the average of a variable across all items. The stationarity assumption will be rejected at a significance level of 5% if $|c_t|\sqrt{kr-1}/2 > 1$.

For the spatial domain, we use the Goldfeldt-Quandt procedure (24). Residuals from any two sites s_0, s_1 in the ROI are collected across all trials and then used to compute the ratio

$$c_{0,1} = \frac{\sum_{t=0}^{r-1}[\epsilon(s_0,t)-\bar{\epsilon}(s_0)]^2}{\sum_{t=0}^{r-1}[\epsilon(s_1,t)-\bar{\epsilon}(s_1)]^2} \ \forall\ s_0, s_1 \in \mathbf{S}, \quad [8]$$

where $c_{0,1}$ has a Fisher distribution with $(r-1,\ r-1)$ degrees of freedom. The stationarity assumption in space will be rejected at a significance level of 5% if:

$$c_{0,1} > F_{0.05}(r-1, r-1) \quad \text{or} \quad c_{0,1} < F_{0.95}(r-1, r-1). \quad [9]$$

Covariance Structure

For a weakly stationary process, the correlation between residuals $\epsilon(s_0, t_0)$ and $\epsilon(s_1, t_1)$ depends only on the "distance" between sites in space and time. Since we assume AR(1) correlations, the covariance matrix $\mathbf{V}$ is defined completely by the variance ${\gamma_2}^2$ and the lag-1 autocorrelation coefficients ρ in the x, y, and t directions. For reasons of simplicity, we estimate ρ separately in space and time, thus separating $\mathbf{V}$ into a spatial and a temporal correlation matrix:

$$\mathbf{V} = {\gamma_2}^2(\mathbf{V}_S \otimes \mathbf{V}_T), \quad [10]$$

where $\otimes$ denotes the Kronecker product. For the temporal domain, we simply estimate the lag-1 sample autocorrelation:

$$\hat{\rho}_T = \frac{\sum_{s\in\mathbf{S}}\left[\sum_{t=1}^{l-1}\epsilon(s,t)\epsilon(s,t-1)\right]}{\sum_{s\in\mathbf{S}}\left[\sum_{t=0}^{l-1}\epsilon(s,t)^2\right]} \quad [11]$$

and set up the inverse of the temporal correlation matrix directly:

$$\mathbf{V}_T{}^{-1} = \frac{1}{(1-{\rho_T}^2)}\begin{pmatrix} 1 & -\rho_T & 0 & \cdots & 0 \\ -\rho_T & 1+{\rho_T}^2 & -\rho_T & \cdots & 0 \\ 0 & -\rho_T & 1+{\rho_T}^2 & \cdots & 0 \\ & & & \cdots & \\ 0 & & & \cdots & 1 \end{pmatrix}. \quad [12]$$

The determination of the spatial correlation matrix is more 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 (13, p. 97):

$$\rho_S(s_0, s_1) = {\rho_X}^h{\rho_Y}^v,$$

$$\text{using} \quad h = H(s_0, s_1),\ v = V(s_0, s_1), \quad [13]$$

where H(·) and V(·) return the absolute distance between two sites in the x and y directions, and ρ_X, ρ_Y denote the spatial autocorrelations. The problem is now to determine ρ_X and ρ_Y from the data. First, we form subsets of all pairs of sites, which are located the same absolute distance h, v:

$$S_{h,v} = \{s_0, s_1 | s_0 \in S, s_1 \in S, H(s_0, s_1) = h, V(s_0, s_1) = v\}, \quad [14]$$

and then estimate $\hat{\rho}_{h,v}$ for this subset:

$$\hat{\rho}_{\boldsymbol{h,v}} = \frac{\sum_{t\in\mathbf{T}}\left[\sum_{s_0,s_1\in\mathbf{S}_{h,v}}\epsilon(s_0,t)\epsilon(s_1,t)\right]}{\sum_{t\in\mathbf{T}}\left[\sum_{s_0\in\mathbf{S}_{h,v}}\epsilon(s_0,t)^2\right]}. \quad [15]$$

Now, we find $\hat{\rho}_X$, $\hat{\rho}_Y$ by LS estimation:

$$\arg\min_{\rho_X\rho_Y}\left\{\sum_{h,v}(\hat{\rho}_{h,v} - {\rho_X}^h{\rho_Y}^v)^2\right\}. \quad [16]$$

By using these estimates, we can compute the correlation ρ_S between any sites s_i, s_j in the ROI by Eq. [13], and thus set up the spatial correlation matrix:

$$\mathbf{V}_S = \begin{pmatrix} 1 & \rho_S(s_0,s_1) & \rho_S(s_0,s_2) & \cdots & \rho_S(s_0,s_{k-1}) \\ \rho_S(s_0,s_1) & 1 & \rho_S(s_1,s_2) & \cdots & \rho_S(s_1,s_{k-1}) \\ \rho_S(s_0,s_2) & \rho_S(s_1,s_2) & 1 & \cdots & \rho_S(s_2,s_{k-1}) \\ & & & \cdots & \\ \rho_S(s_0,s_{k-1}) & & & \cdots & 1 \end{pmatrix}, \quad [17]$$

which is symmetric and always invertible. Finally, we compose the inverse of the spatiotemporal covariance matrix: $\mathbf{V}^{-1} = 1/\gamma_2{}^2\,(\mathbf{V}_s{}^{-1} \otimes \mathbf{V}_T{}^{-1})$.

Confidence Intervals

Confidence intervals may be determined from Monte Carlo simulations. This requires prior knowledge about the properties of brain activation, and a model for noise and artifacts. We report on the results of such a model in the experimental section. Deriving confidence intervals for the estimates $\boldsymbol{\beta}$ from the model equations is more involved (13,26,27). To simplify the notation, we drop the dependence of $\mathbf{g}(\cdot)$ on $\mathbf{T}$. We expand $g(\boldsymbol{\beta})$ by a first-order Taylor-series about the true parameter vector $\tilde{\boldsymbol{\beta}}$:

$$g(\hat{\boldsymbol{\beta}}) \approx \mathbf{g}(\hat{\boldsymbol{\beta}}) + \mathbf{G}_{\hat{\boldsymbol{\beta}}}{}^{\mathrm{T}}(\hat{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}}), \qquad [18]$$

where $\mathbf{G}_{\hat{\boldsymbol{\beta}}}$ denotes the Jacobian matrix of $\mathbf{g}(\cdot)$ with respect to $\hat{\boldsymbol{\beta}}$. By using an equivalent formulation of Eq. [3]:

$$\mathbf{G}_{\hat{\boldsymbol{\beta}}}\mathbf{V}^{-1}\,[\mathbf{Y} - \mathbf{g}(\boldsymbol{\beta})] = 0, \qquad [19]$$

and insertion of Eq. [18] yields:

$$\mathbf{G}_{\hat{\boldsymbol{\beta}}}\mathbf{V}^{-1}\,[\mathbf{Y} - \mathbf{g}(\hat{\boldsymbol{\beta}}) - \mathbf{G}_{\hat{\boldsymbol{\beta}}}{}^{\mathrm{T}}(\tilde{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}})] = 0. \qquad [20]$$

Solving for $\tilde{\boldsymbol{\beta}}$, we find:

$$\tilde{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}} + \mathbf{F}_{\hat{\boldsymbol{\beta}}}{}^{-1}\mathbf{G}_{\hat{\boldsymbol{\beta}}}\mathbf{V}^{-1}(\mathbf{Y} - \mathbf{g}(\hat{\boldsymbol{\beta}})), \qquad [21]$$

where $\mathbf{F}$ is called the Fisher Information Matrix $\mathbf{F}_{\boldsymbol{\beta}} = \mathbf{G}_{\boldsymbol{\beta}}\mathbf{V}^{-1}\mathbf{G}_{\boldsymbol{\beta}}{}^{\mathbf{T}}$. Because $\tilde{\boldsymbol{\beta}}$ is a linear function of the data, its estimator follows a multivariate normal distribution:

$$\hat{\boldsymbol{\beta}} \sim \mathrm{N_p}(\tilde{\boldsymbol{\beta}}, \mathbf{F}_{\tilde{\boldsymbol{\beta}}}{}^{-1}). \qquad [22]$$

Thus, under this first-order linear model, the estimator is unbiased, and its covariance, $\mathbf{F}\hat{\boldsymbol{\beta}}^{-1}$, is the Cramer-Rao lower bound for the nonlinear model of Eq. [1].

It is possible to improve the confidence bands by introducing higher order approximations (28). However, this requires Monte Carlo simulations to derive the confidence intervals for each parameter set, which is not feasible for routine estimation.

Estimation Algorithm

Given a set of regions of interest and a set of trials, we repeat the following estimation procedure for each ROI and each trial. First, we collect data from the ROI in a specific trial, set $\mathbf{V} = \mathbf{I}_n$ and starting parameters for $\boldsymbol{\beta}$. The optimization problem is given by minimizing Eq. [3] in the p parameters of the model function $g(\cdot)$. From the residuals, the spatiotemporal covariance matrix is computed, and the optimization process repeated. A fixed number of iterations (typically 5) is used. For a region of interest this procedure needs about 10 sec for 100 trials. A typical fMRI data set is evaluated within 5 min on a conventional workstation. At length, we summarize our algorithm as follows:

1. Collect data points $y(s,t)\ \forall\ s \in S, t \in T$.
2. Set the spatiotemporal covariance matrix as the identity matrix: $\mathbf{V} = \mathbf{I}_n$.
3. Select starting values for $\boldsymbol{\beta}$.
4. Repeat for N iterations:
   Optimize Eq. [3] by the downhill simplex method
   Compute γ_2, ρ_X, ρ_Y, and ρ_T from the residuals (Eqs. [11] and [16])
   Set up the spatiotemporal covariance matrix $\mathbf{V}$ (Eqs. [10], [12], and [17]).
5. Find confidence intervals for model parameters $\hat{\boldsymbol{\beta}}$.
6. Examine normality and stationarity in the residuals from all trials in a ROI.

MODEL VALIDATION

For the validation of the model explained in the previous section, we have to show that 1) the parameter estimation is stable with respect to noise levels and distributions typically found in fMRI data sets, and 2) the confidence intervals for the model parameters derived analytically (see ''Confidence Intervals'' in previous section) comply with experimentally obtained results. We refer to the HR parameters as defined in ''Model Functions for the Hemodynamic Response.''

Performance of Parameter Estimation

To assess the stability of the parameter estimation, we used simulated fMRI data sets. We modulated a prototypical HR with known properties (spatial distribution, temporal shape; Fig. 1) onto a brain patch with unknown background characteristics. A recent single-trial fMRI experiment with a TR = 2 sec and a trial length of 12 time points was chosen, and a prototypical HR was collected from a significantly activated cortical area and averaged in time and space (Fig. 1a). A patch of 10×10 voxels was selected from the same experiment which did not contain any significant task-related activation. The prototypical HR was modulated onto the patch in a connected region of 6 voxels (Fig. 1b) over 76 trials, using an amplitude range of 20 arbitrary units (AU) to 500 AU in 25 steps, corresponding to z scores between 4.4 for an amplitude of 80 AU to 14.2 for an amplitude of 500 AU. Patches with a smaller amplitude were rendered insignificant by a z score threshold of 3.5. To determine the true (noiseless) parameters for the prototypical HR, we modulated the signal with an amplitude of 500 onto an artificial noise-free patch. We determined HR parameters for the 26 amplitude settings and the 76 trials.

The results of the tests were summarized as follows:

- The gain was proportional to the HR amplitude by a factor of 2.48. For an amplitude of 500 AU, the gain was determined as 1241 AU ± 217 AU, which compares favorably with the true value of 1245 AU in the noiseless case.
- The lag was slightly underestimated for cases of a low signal-to-noise ratio (Fig. 2a), which is a consequence of the asymmetric distribution of the lag values. The lag, again for an amplitude of 500 AU, is 7.09 sec ± 0.40 sec (true: 7.09 sec). As expected, the distribution

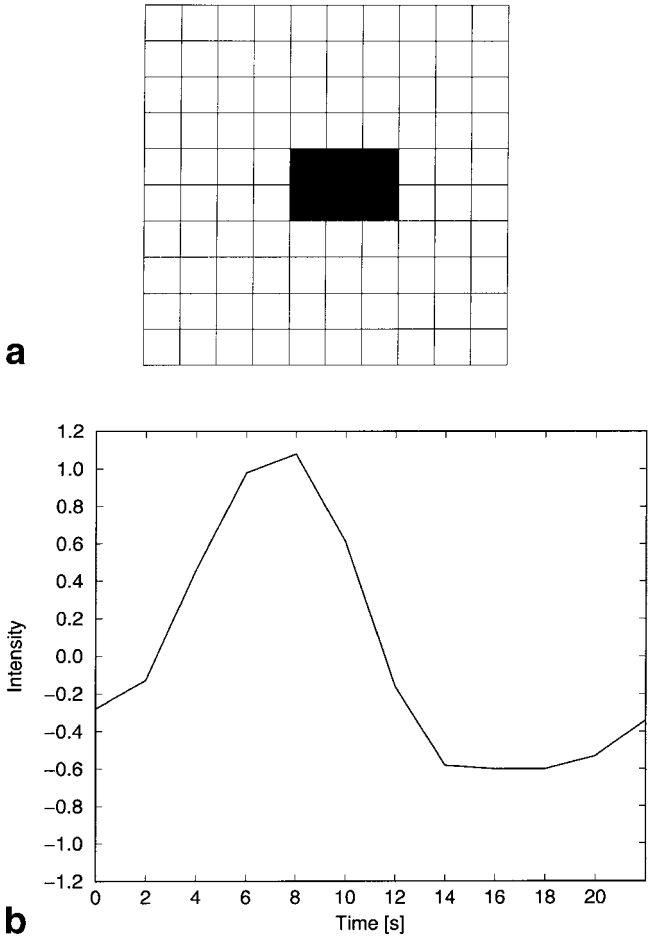

FIG. 1. Spatial pattern (**a**) and time course (**b**) of the hemodynamic response used in the simulation study. Parameters of this prototypic response (at an amplitude factor of 500 AU) are: gain 1245 AU, lag 7.09 sec, dispersion 2.97 sec.

of the lag narrowed with increasing signal-to-noise ratio.

- The dispersion was constant, independent of the amplitude (Fig. 2b), and estimated as 2.90 sec ± 0.43 sec (true: 2.97 sec). As expected, the distribution of the dispersion narrowed with increasing signal-to-noise ratio.
- The norm was proportional to the amplitude by a factor of 6.19. The norm for an amplitude of 500 AU was 3097 AU ± 532 AU, which agrees with the true value of 3113.

We conclude that this nonlinear regression procedure is stable against noise levels typically found in fMRI data sets.

Confidence Intervals

To learn more about the confidence regions for the estimation, we ran a Monte Carlo simulation. We modulated the prototypical HR onto an artificial patch, using an amplitude of 100 AU and a Gaussian noise level of $\sigma = 40$, which compares with the usual signal to noise ratio found in preprocessed fMRI data. 25 activated areas of 25 voxel size and 100 periods were generated, thus yielding 2500 sets of parameter estimates. Their distribution and confidence regions (99.73%, 99.0%, 95.4%, 90.0%, and 68.3%) are plotted in Figure 3. The 95% confidence intervals at this noise level are determined as gain: 252 AU ± 74 AU, lag: 7.11 sec ± 0.71 sec, dispersion: 2.97 sec ± 0.48 sec, norm: 620 AU ± 182 AU. The confidence intervals, as determined from the diagonal elements of the inverted Fisher matrix (see Eq. [22], are found for the gain: 258 AU ± 68 AU, lag: 7.20 sec ± 0.40 sec, dispersion: 3.07 sec ± 0.42 sec, norm: 652 AU ± 165 AU, which agree well with the simulation.

EXAMPLE fMRI ANALYSIS

We randomly selected an fMRI data set from a recent experimental study on language comprehension. Using a single-shot EPI protocol, we acquired every 2 sec 4 T_2^*-weighted slices of 128 × 64 voxel with a spatial resolution of 1.9 × 3.8 × 5mm and 2 mm gap. The auditory presentation of a sentence needed approx. 6 sec (3 time steps), followed by a pause of 18 sec (9 time steps). 76 trials (912 time steps) were recorded during an approximately 30 min experiment. The right-handed test persons were asked to classify a sentence for syntactical correctness and to respond via a button press with the right hand. For consistency, a single data set was chosen to demonstrate the use

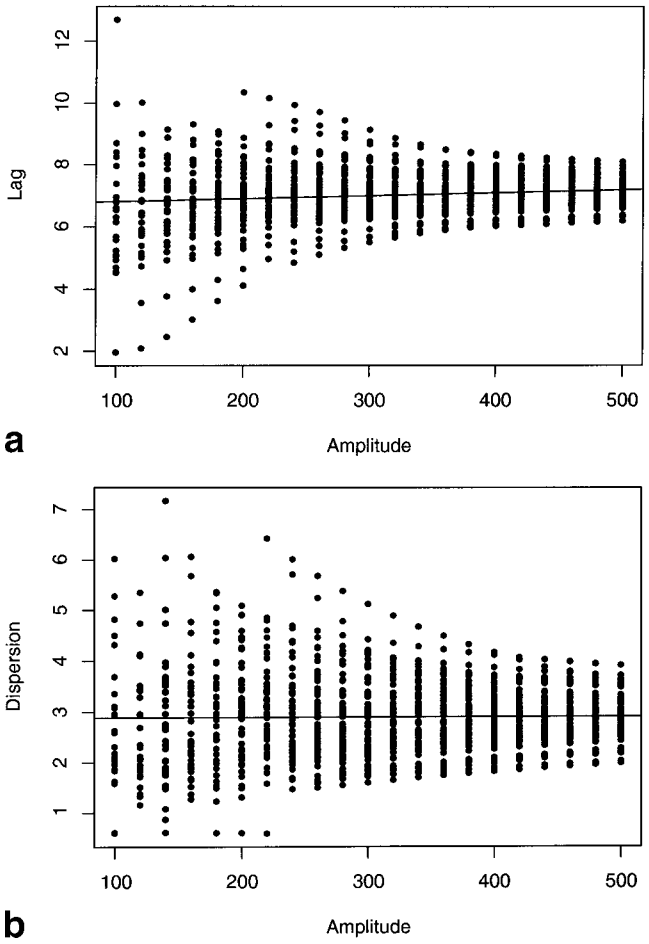

FIG. 2. Dependency of the HR parameters on the test function amplitude, shown here for the lag (**a**) and the dispersion (**b**).

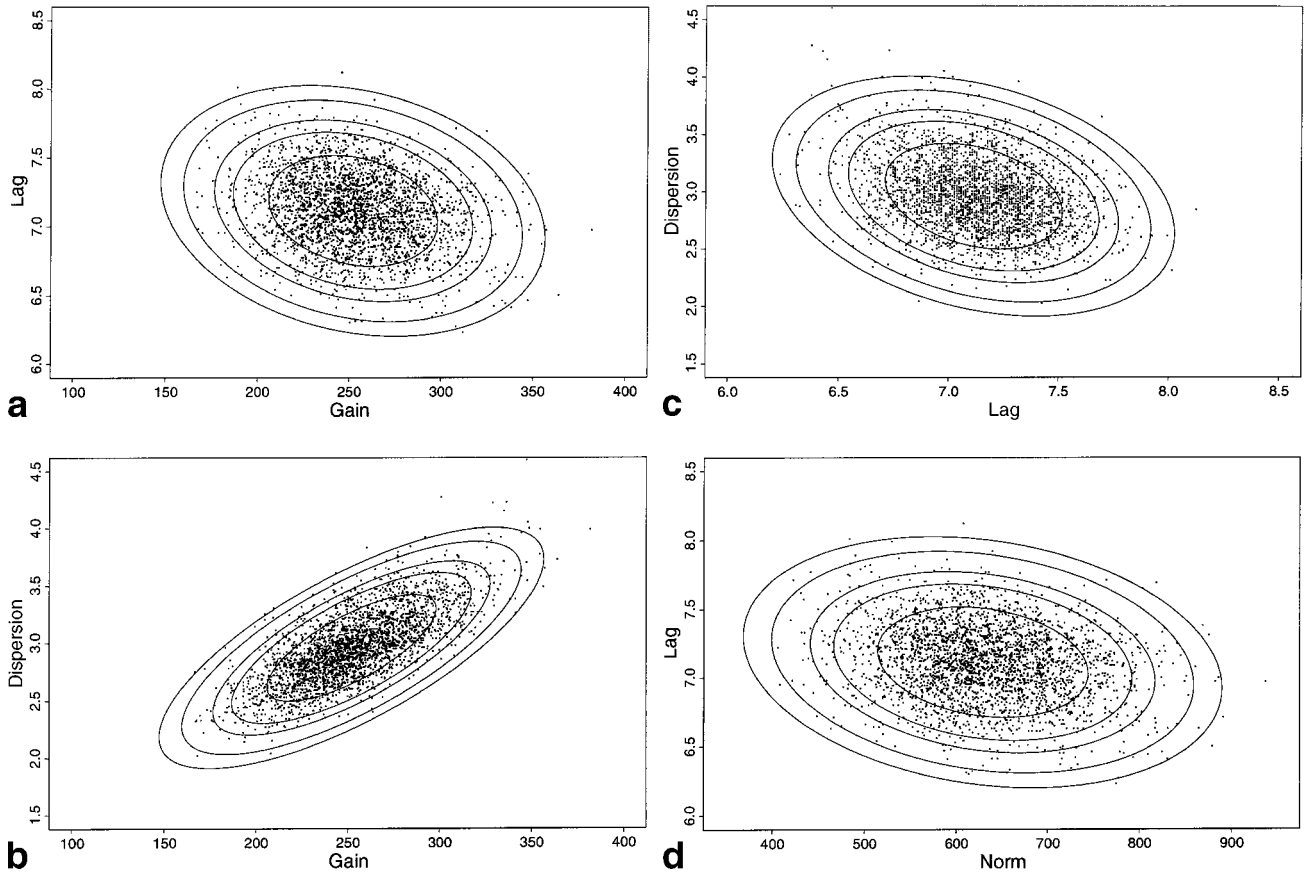

FIG. 3. **a–d**. Results of a Monte Carlo simulation to study the confidence regions for HR parameters. Confidence ellipsoids are given for levels of 99.73%, 99.0%, 95.4%, 90.0%, and 68.3%.

of the procedures discussed in this report. Similar evaluations were performed on a large number of data sets and different experimental designs.

The standard procedures to detect functional activity in this data set were applied as follows: *signal preprocessing* (12) using 1) motion correction, 2) estimation of the baseline by voxelwise low-pass filtering in the temporal domain using a cutoff frequency of 1.5 times of the stimulation frequency, 3) baseline correction by subtracting the estimated baseline from the motion-corrected data, and 4) signal restoration using spatiotemporal Markov-Random-Field model (29) to reduce physiological and system noise; *signal detection* by 1) analysis for activated regions by Pearson correlation with a time-shifted box-car waveform ($\Delta = 6$ sec), 2) conversion of the correlation coefficients into *z* scores, 3) thresholding of the corresponding SPM[z] by a score of 8, and finally 4) assessment of the activated regions for their significance on the basis of their spatial extent (30) (Fig. 4). For *signal quantification* by the model proposed in this report, 44 ROIs were defined by their 6 most highly activated voxels from the SPM[z]. We modeled the HR trial- and ROI-wise in preprocessed data using a Gaussian function by the estimation context explained in ''A Nonlinear Regression Model for fMRI'' and yielded 44 by 76 estimates of gain, norm, lag, and dispersion. Timings were corrected for the acquisition delay of separate slices in the EPI protocol. Nine neurofunctionally interesting regions were marked in Figure 5.

Before we actually report on the HR estimation results, the estimation performance and model validity (by analysis of the residuals) shall be discussed.

Estimation Performance

Results of an example region in Heschl's gyrus on the right side (HG_R in Fig. 5), where a satisfying fit to the Gaussian model function was found, are shown in Figure 6a. In other regions, especially with a lower signal-to-noise ratio, some HR are crippled or too deviant in the sample region to warrant a good fit. Figure 5b displays an example from the precuneus (*PC* in Fig. 5), where the second and the seventh period are not well adapted. By applying certain heuristics on the estimated parameters and examining the goodness of fit, these misfits can be detected and marked in the output. Typically 0–15% of the trials are marked as outliers.

Analysis of the Residuals

From the sample signal in Heschl's gyrus (HG_R in Fig. 5), we computed the distribution of the residuals as $N(-0.022, 40.9^2)$. A Q-Q plot revealed a high conformance with a normal distribution; however, a deviance on extreme cases ($<2\gamma_2$ or $>2\gamma_2$) was found. Normality in the distribution according to Eq. [5], with $T_1 = 0.243$ and $T_2 = 0.437$, was confirmed for this region. Normality criteria were not fulfilled for all ROIs, however. For most regions,

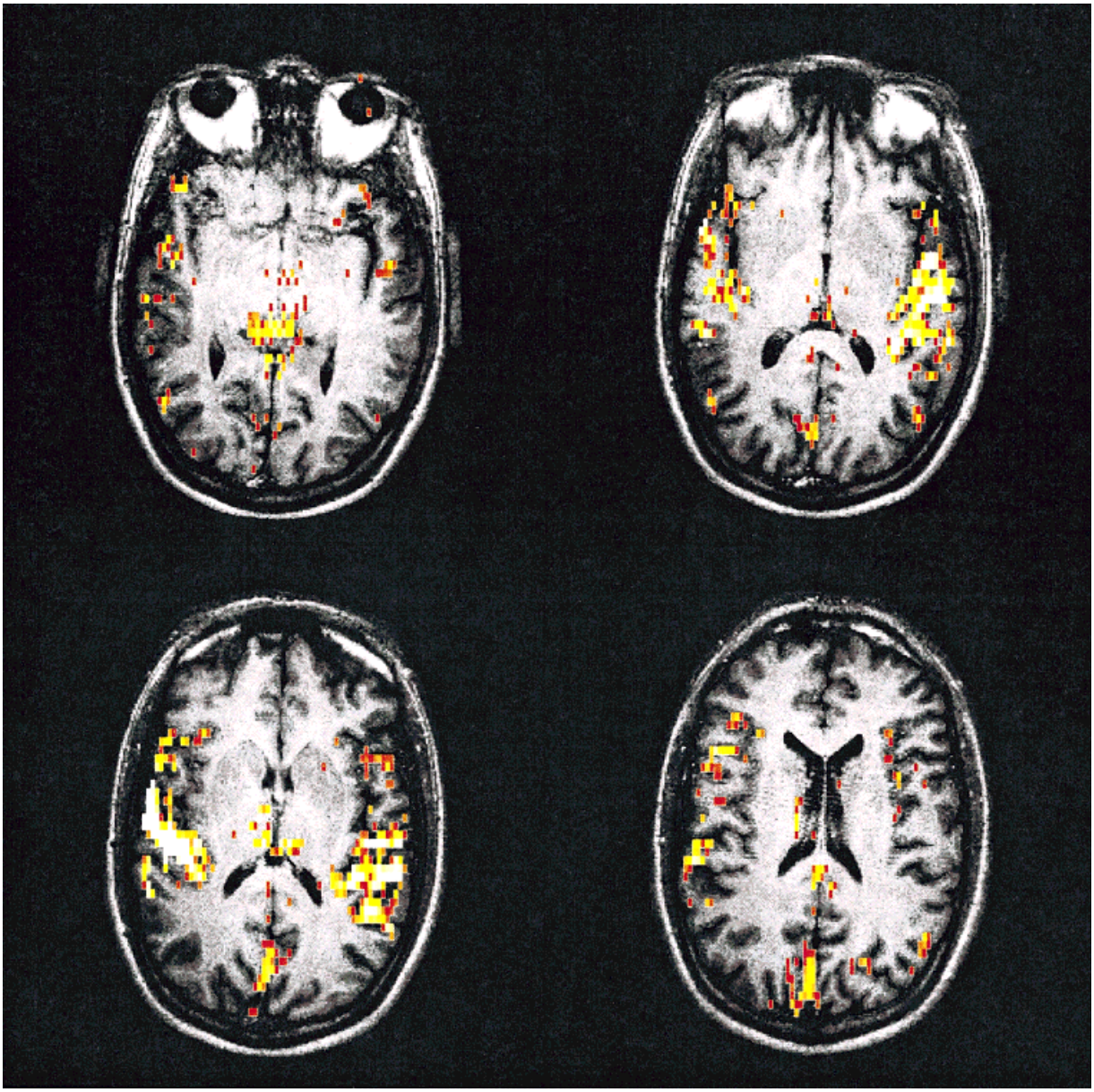

FIG. 4. Significantly activated regions from a study in language comprehension, overlaid onto their corresponding brain slice.

T_1 had small positive values, indicating a distribution slightly skewed to the left. In some regions, normality was rejected due to $T_2 > 1$. We found distributions with longer tails here, mostly as a consequence of unmodeled components and (physiological) artifacts in the data. Such violations of the distributional assumptions are not considered harmful to the ML estimation process (20). These criteria indicate regions (or temporal intervals) disturbed by artifacts, so the corresponding HR parameters might be regarded with caution or excluded from further analysis. Stationarity as defined by Anscombe's test (see Eq. [7]) and the Goldfeld-Quandt procedure (see Eq. [8]) was accepted for all ROIs in preprocessed data. If a spatial or spatiotemporal filter was applied during preprocessing, we found $\rho_{X,Y}$ values in the range of 0.40–0.60, in comparison with 0.10–0.30 in unfiltered data. If spatial correlations are low ($\rho \sim 0.10$), the spatial matrix $\mathbf{V}_S$ is well approximated by the identity matrix and might be substituted as such in Eq. [10]. Correlations in the temporal domain were determined in a similar range. As indicated from the performance study, a good signal-to-noise ratio improves the confidence bands of the estimation. Since filtering for system and physiological noise in fMRI data is effective in the temporal domain (12), we always preprocess the raw data and consider the temporal autocorrelations in the estimation.

Estimation Results

HR parameters were estimated for all the 44 ROIs in 76 trials. Summary statistics for 9 sample regions from Fig. 5 are compiled in Table 1.

The inter-trial variation of the lag values is considerably higher in comparison with the simulated data. This is easily explained by considering that sentences were au-

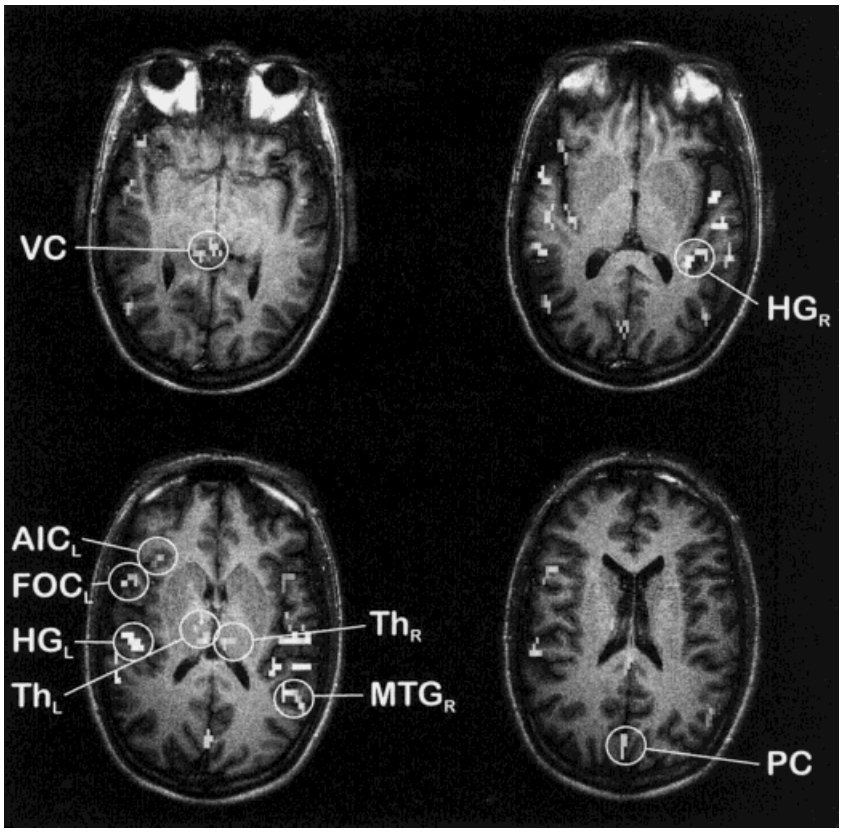

FIG. 5. Selection of highly activated regions (see Fig. 4) with an extent of 6 voxels. Nine neurofunctionally interesting regions were labeled.

rally presented and lag times were computed from the onset of the presentation. In addition, we ran t-tests (single sided, paired, unequal variance, $p < 0.05$) on the lag values for single trials. We found a significant temporal ordering as follows:

$$HG_L \sim HG_R < MTG_R \sim Th_R \sim Th_L < FOC_L \sim PC < VC \quad [23]$$

Interestingly, a temporal ordering can also be found along a trajectory starting from the mesial part of Heschl's gyrus, running outwards on the superior temporal gyrus and to the anterior temporal lobe.

Comparison With Stimulation Data

Since the duration of the aurally presented sentence and the sentence type (correct or incorrect) are known they may be compared with the HR parameters.

For all ROIs, norm, lag, and dispersion increase with sentence length, indicating that a longer activation time corresponds to a higher energy consumption and thus a more increased BOLD effect. We fitted linear models to test the dependency of norm, lag, and dispersion on the sentence length (Fig. 7). All fits were highly significant ($p < 1e-4$). Per second of increasing sentence length, the norm increases by 188 AU (corresponding to an increase of +21% for a step from 3 to 4 sec), the lag by +880 msec, and the dispersion by +210 msec.

In addition, the norm depends on the correctness manipulation: presentation of an incorrect sentence leads to an increase of the lag (+245 msec), the norm (+17%), and the dispersion (+100 msec). During the experiment, norm and dispersion are constant, whereas the lag decreases slightly (−360 msec). This may be attributed to practicing the stimulus.

When analyzing single ROIs, significant interactions between the norm and the sentence length are found in a number of ROIs on the superior temporal gyrus on both sides. The correctness manipulation interacts with the norm only in the left thalamus, the precuneus, and the right superior temporal gyrus.

DISCUSSION

This report deals with a method to analyze fMRI data targeted to STDs. The hemodynamic response is modeled in a nonlinear regression context trial- and ROI-wise. We will now focus on a number of issues that require discussion.

Choice of Model Function

Describing the HR by a model function is a major advantage because it provides a compact and concise description of the HR shape. Currently, the discussion about the ''best'' model function is open, and any choice is clearly arbitrary, because none is based on a comprehensive physiological model. Cohen (31) and Lange and Zeger (17) proposed the Gamma function and were followed by Friston et al. (18), who used a set of two fixed-parameter Gamma functions to model the HR. In accordance with Lange, we found—especially in STD—that the Gamma function yielded a rather high number of unsatisfactory fits. The Poisson function offers only a single-shape parameter and is not

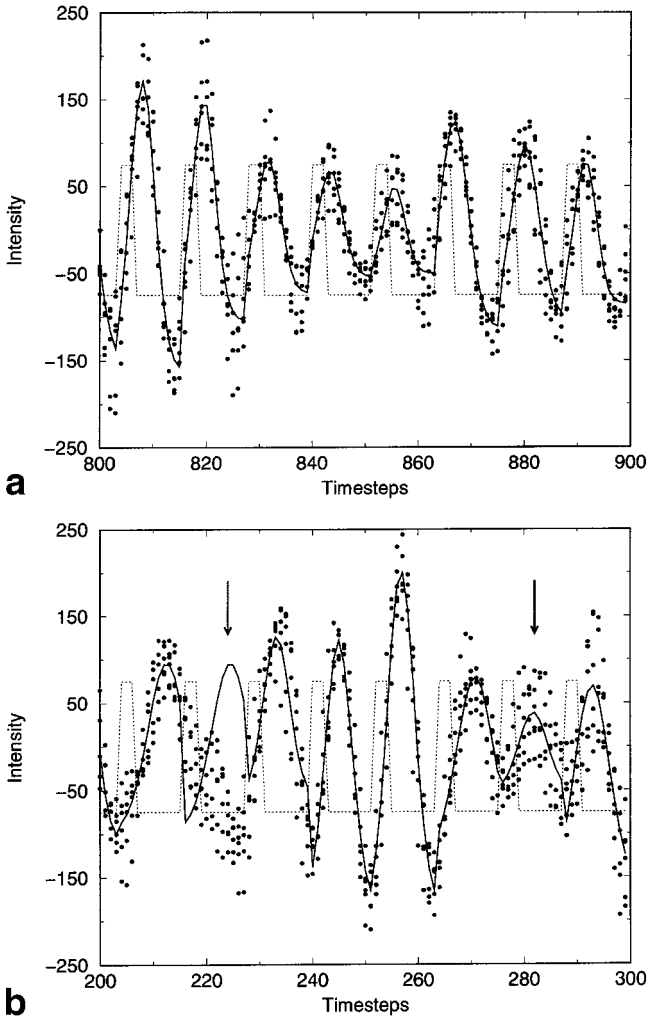

FIG. 6. From the experiment in sentence comprehension, 8 trials were selected of 12 time steps each. The auditory stimulus is shown as a dotted line, together with data (shown as dots) sampled from two regions, Heschl's gyrus on the right side (HG_R, **a**) and the precuneus (PC, **b**) and their trialwise modeled response using a Gaussian function (solid line). Arrows mark trials were a satisfying adaption could not be achieved.

considered a suitable model function. Judging from goodness-of-fit criteria, the Gaussian function models the HR in STDs best. This finding is in agreement with findings of a standard linear regression context by Rajapakse et al. (19). An argument for using asymmetrical model functions is given by the expectation that the slope of the rising edge of the HR is steeper than the falling edge. In our closely timed single-trial experiments this is rarely found. When using spatially or temporally averaged responses, a Gaussian function is expected to yield the best fit. Other choices of model functions are sinusoidal orthogonal basis functions (20) or splines (32). However, they require more parameters (and thus more time steps per trial in the data), and the determined parameters do not allow an easy derivation of physiological meaningful variables (such as maximum, slope, etc.).

From calculus we know that at least as many time steps per trial as parameters of the model function are required.

Table 1
Characterization of the HR Parameters in Sample Regions (see Fig. 5) of the Example fMRI Experiment

Location	Norm (AU)	Lag (sec)	Dispersion (sec)
Vein complex	1680 ± 877	10.25 ± 2.36	5.01 ± 1.27
Precuneus	861 ± 389	9.66 ± 2.30	4.86 ± 1.39
Heschl's gyrus left	1241 ± 934	6.98 ± 1.28	4.82 ± 1.08
Heschl's gyrus right	1353 ± 515	7.30 ± 1.21	4.75 ± 0.93
Middle temporal gyrus right	1341 ± 576	8.24 ± 2.02	4.77 ± 1.18
Thalamus left	893 ± 451	8.49 ± 1.80	4.97 ± 1.44
Thalamus right	956 ± 395	8.34 ± 1.97	4.96 ± 1.38
Anterior insular cortex left	766 ± 366	9.80 ± 2.97	4.87 ± 1.50
Frontal opercular cortex left	926 ± 402	9.66 ± 2.47	4.85 ± 1.32

For the three- and four-parameter functions discussed here we found sufficient fits in STDs with a period length down to 8 time steps. We did not attempt to model the multiphasic nature of the hemodynamic response. Trials are tightly spaced in standard STDs, so the terminal undershoot runs over into the next rising edge. The early dip is only detectable in highly time-resolved fMRI experiments at

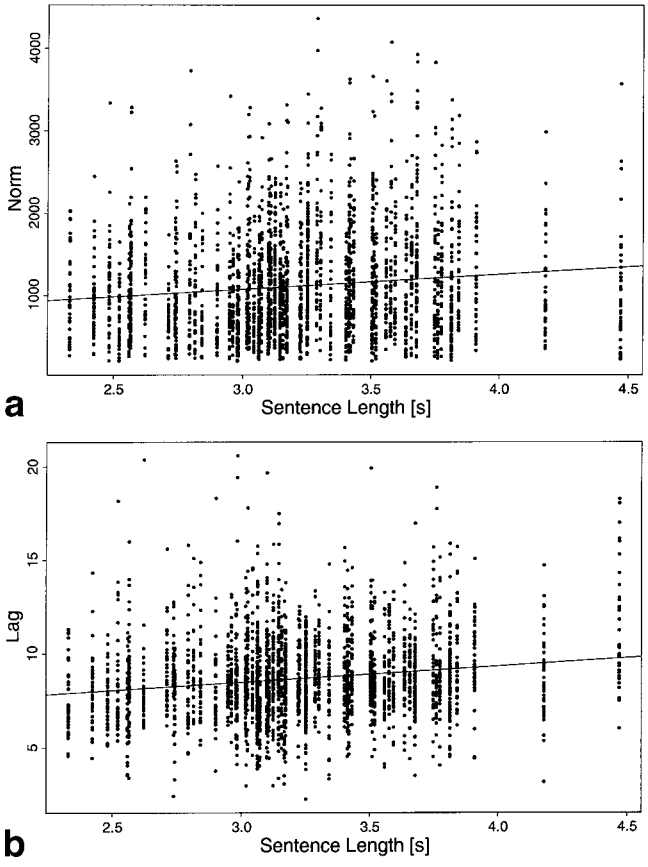

FIG. 7. Dependency of the HR model parameters norm (**a**) and lag (**b**) on the auditory presentation time of a sentence. The solid lines indicate linear regression models fitted to the data ($p < 1e\text{-}4$). Per second of increasing sentence length, the norm increases by 188 AU, the lag by +880 msec and the dispersion (not shown) by +210 msec.

higher fields. It is a question at issue what should be considered as baseline in these experiments. In our estimation context we were ignorant to these findings and modeled only the signal increase. We do not assign any physical meaning to offset parameter β_3 in Eq. [4] and regard a trialwise constant offset just as a first-order approximation.

Of course, the Gaussian model function here is not well suitable to model prolonged activation. Even if it is possible in principle to use this estimation context with block designs, none of these functions would obtain a good fit. Therefore, ideally, the stimulus presentation should be short. In a random selection of several experiments we found acceptable and reliable fits with presentation times of up to 4 sec (i.e., aurally presented sentences).

Interpretation of the HR Parameters

Another advantage of using model functions is the ability to derive physiologically meaningful parameters. The lag as the time to the HR maximum can be understood in terms of the experiment and easily be integrated into statistical models (i.e. for comparison with stimulation length, reaction times, learning etc.). The lag is useful to introduce a temporal ordering of the arising HR. To better interpret the reason for a temporal ordering, the dispersion information must be taken into account: a comparatively high dispersion corresponds to a temporally long HR, whereas a low dispersion reflects a rather short HR, which is possibly shifted in time with respect to the stimulation onset. The usefulness of lag and dispersion in inferring about the sequence of neuronal activations is currently under investigation.

Gain and norm may be used interchangeably according to our experiences. In the majority of activations analyzed, they are linked with a constant factor with relatively tight limits. The norm might be given preference because this parameter is ''closer to the data.'' In our experience, norm and gain reflect the energy consumption of an activated area. The dispersion is proportional to the duration of a brain activation and thus interpretable as an indicator of the ''computation time'' of a cognitive process. The quotient of gain and lag is an approximation of the HR rise time and may be linked to the energy consumption of an activated area. Finally, the residual variance γ_2^2 is an important quality indicator of the fit. The quotient $\gamma_2^2/gain$ is trivially dependent on the signal-to-noise level, but in addition, an indicator of a possible misfit. This quotient is found to be in the order of 0.05–0.20.

Importance of Temporal Characterization

It is a major advantage of this procedure that the temporal properties of a HR are now accessible to statistical analysis. Rather than the conventional computation of averaged waveforms and (visual) comparison of their shape characteristics the method includes shape properties as statistical parameters. First results show that conclusions about time differences of less than 300 msec seem to be statistically significant. This confirms similar reports by Friston et al. (18). By no means is this estimation context restricted to regular designs in time. With respect to the discussion of estimation restrictions above, the length of a trial may vary in wide margins.

When ROIs are considered as connected, significantly activated, regions, this does not necessarily imply that these region are homogeneous with respect to their temporal characteristics. In a number of experiments we found confluent activated regions, for example along the superior temporal gyrus or the intraparietal sulcus, which showed significantly different properties when split up into distinct ROIs.

Of course, it is necessary to alleviate the current requirement of predefined ROIs. This is straightforward in this modeling context: first, we estimate the HR parameters voxelwise and then decide on the basis of some similarity criteria which voxels may be merged. The center and extent of such a region is not necessarily constant from trial to trial. As much as the signal-to-noise ratio allows, it is necessary to introduce stricter criteria for region homogeneity and add location and extent to the parameter list in this estimation context.

CONCLUSION

We have extended the current approaches of modeling the HR in functional MRI into the context of explicit STDs. By combining the model with a rigorous stochastical analysis of the spatiotemporal interactions in fMRI data and embedding the estimation into a robust framework, we are able to derive statistically sound descriptions of HRs in a certain brain area. This approach achieves a much finer level of description of a fMRI experiment and opens new perspectives for functional interpretation.

ACKNOWLEDGMENTS

The authors thank Angela Friederici, Martin Meyer, and Christopher Wiggins for their support.

REFERENCES

1. Belliveau JW, Kennedy DN, McKinstry RC, Buchbinder BR, Weisskoff RM, Cohen MS, Vevea JM, Brady TJ, Rosen BR. Functional mapping of the human visual cortex by magnetic resonance imaging. Science 1991;254:716–719.
2. Kwong KK, Belliveau JW, Chesler DA, Goldberg IE, Kennedy DN, Weisskoff RM, Poncelet BP, Hoppel BE, Cohen MS, Turner R, Cheng H, Brady TJ, Rosen BR. Dynamic magnetic resonance imaging of the human brain activity during sensory stimulation. Proc Natl Acad Sci USA 1992;89:5675–5679.
3. Ogawa S, Tank DW, Menon R, Ellermann JM, Kim SG, Merkle H, Ugurbil K. Intrinsic signal changes accompanying sensory stimulation: functional brain mapping with magnetic resonance imaging. Proc Natl Acad Sci USA 1992;89:5951–5955.
4. Ernst T, Henning J. Observation of fast response in functional MR. Magn Reson Med 1994;32:146–149.
5. Menon RS, Ogawa S, Tank DW, Ugurbil K. 4 Tesla gradient recalled echo characteristics of photic stimulation-induced signal changes in the human primary visual cortex. Magn Reson Med 1993;30:380–386.
6. Vilringer A, Dirnagl U. Coupling of brain activity and cerebral blood flow: basis of functional neuroimaging. Cereb Brain Metab Rev 1995;7: 240–276.
7. Luknowsky DC, Thomas CG, Gati JS, Menon RS. Millisecond sequencing of neural activation in simple tasks determined by the BOLD fMRI neurovascular response. Neuroimage 1998;6:S280.
8. Richter W, Tegeler C, Georgopoulos A, Ugurbil K, Kim SG. Time-resolved fMRI of motor area activity during mental rotation. Neuroimage 1998;6:S58.

9. Zarahn E, Aguirre G, D'Esposito M. A trial-based experimental design for fMRI. Neuroimage 1997;6:122–138.
10. Josephs O, Turner R, Friston K. Event-related fMRI. Hum Brain Mapp 1997;5:243–248.
11. Richter W, Ugurbil K, Georgopoulos A, Kim SG. Time-resolved fMRI of mental rotation. Neuroreport 1997;8:3697–3702.
12. Kruggel F, Descombes X, von Cramon DY. Preprocessing of fMRI data sets. In: Workshop on biomedical image analysis (Santa Barbara). Los Alamitos: IEEE Computer Press; 1998.
13. Rosenfeld A, Kak AC. Digital picture processing. San Diego: Academic; 1982.
14. Seber GAF, Wild CJ. Nonlinear regression. New York: Wiley; 1989.
15. Nelder JA, Mead R. The simplex method. Computer Journal 1965;7:308.
16. Friston KJ, Jezzard P, Turner R. Analysis of functional MRI time-series. Hum Brain Mapp 1994;1:153–171.
17. Lange N, Zeger SL. Nonlinear Fourier time series analysis for human brain mapping by functional magnetic resonance imaging. J R Stat Soc Appl Stat 1997;46:1–29.
18. Friston KJ, Josephs O, Rees G, Turner R. Nonlinear event-related responses in fMRI. Magn Reson Med 1998;39:41–52.
19. Rajapakse JC, Kruggel F, von Cramon DY. Modeling hemodynamic response for analysis of functional MRI time-series. Hum Brain Mapp 1998;6:283–300.
20. Bullmore E, Brammer M, Williams SCR, Rabe-Hesketh S, Janoth N, David A, Mellers J, Howard R, Sham P. Statistical methods of estimation and inference for functional MR Image analysis. Magn Reson Med 1996;35:261–277.
21. Benali H, Buvat I, Anton JL, Pelegrini M, Di Paola M, Bittoun J, Burnod Y, Di Paola R. Space-time statistical model for functional MRI image sequences. In: Duncan J, Gindi G, eds. Information processing in medical imaging. Lecture notes in computer science 1230. Berlin: Springer Verlag; 1997.
22. Gardiner CW. Handbook of stochastic methods for physics, chemistry and the natural sciences. Berlin: Springer; 1985.
23. Neumaier A, Schneider T. Multivariate autoregressive and Ornstein-Uhlenbeck processes: estimates for order, parameters, spectral information, and confidence regions. ACM Trans Math Soft 1998 http://www.aos.princeton.edu/WWWPUBLIC/tapio/papers/arfit.html.
24. Walter E, Ponzatto L. Identification of parametric models. Heidelberg: Springer; 1997.
25. Cressie NAC. Statistics for spatial data. New York: Wiley; 1993.
26. Bates D, Watts D. Nonlinear regression analysis and its applications. New York: Wiley; 1988.
27. Cox C, Ma G. Asymptotic confidence bands for generalized nonlinear regression models. Biometrics 1995;51:142–150.
28. Qu RP, Kosorok MR. Improving confidence bands for nonlinear regression. Technical Report 114. Madison, Wisconsin: Department of Biostatistics, University of Wisconsin-Madison; 1997.
29. Descombes X, Kruggel F, von Cramon DY. FMRI signal restoration using a spatio-temporal Markov random field preserving transitions. Neuroimage 1998;8:340–349.
30. Friston KJ, Worsley KJ, Frackowiak RSJ, Mazziotta JC, Evans AC. Assessing the significance of focal activations using their spatial extent. Hum Brain Mapp 1994;1:210–220.
31. Cohen MS. Parametric analysis of fMRI data using linear systems methods. Neuroimage 1997;6:93–103.
32. Venables WN, Ripley BD. Modern applied statistics with S-Plus. Heidelberg: Springer; 1997.