STATISTICS IN MEDICINE
Statist. Med. 2007; **26**:3845–3861
Published online 5 June 2007 in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/sim.2936

A comparative study of one-level and two-level semiparametric estimation of hemodynamic response function for fMRI data

C. M. Zhang*,†, Y. Jiang and T. Yu

Department of Statistics, University of Wisconsin-Madison, 1300 University Avenue, Madison, WI 53706, U.S.A.

SUMMARY

Functional magnetic resonance imaging (fMRI) is emerging as a powerful tool for studying the process underlying the working of the many regions of the human brain. The standard tool for analyzing fMRI data is some variant of the linear model, which is restrictive in modeling assumptions. In this paper, we develop a semiparametric approach, based on the cubic smoothing splines, to obtain statistically more efficient estimates of the underlying hemodynamic response function (HRF) associated with fMRI experiments. The hypothesis testing of HRF is conducted to identify the brain regions which are activated when a subject performs a particular task. Furthermore, we compare one-level and two-level semiparametric estimates of HRF in significance tests for detecting the activated brain regions. Our simulation studies demonstrate that the one-level estimates combined with a bias-correction procedure perform best in detecting the activated brain regions. We illustrate this method using a real fMRI data set and compare it with popular methods offered by AFNI and FSL. Copyright © 2007 John Wiley & Sons, Ltd.

KEY WORDS: false discovery rate; multiple comparison; semiparametric test; smoothing splines; stimuli; time resolution

1. INTRODUCTION

The past decades have witnessed a big boom in medical imaging technologies. Functional magnetic resonance imaging (fMRI) is emerging as a powerful tool for studying the process underlying the working of the many regions of the human brain. Though merely a decade old, fMRI has already revolutionized neuroscience research.

With the growing flood of brain imaging data, statistical methods play an increasingly critical role in extracting knowledge from experimental data, characterizing patterns of brain activity,

*Correspondence to: C. M. Zhang, Department of Statistics, University of Wisconsin, 1300 University Avenue, Madison, WI 53706, U.S.A.
†E-mail: cmzhang@stat.wisc.edu

Contract/grant sponsor: Wisconsin Alumni Research Foundation
Contract/grant sponsor: National Science Foundation; contract/grant numbers: DMS-03-53941, DMS-07-05209

Received 18 April 2007
Accepted 19 April 2007

and drawing inference about brain dynamics. Recent reviews of the statistical issues in fMRI for brain imaging and the statistical methods for analyzing fMRI data can be found in Lange [1], Lazar *et al*. [2], and Worsley *et al*. [3], among others. Several significantly challenging issues arise in statistical modeling aspects. First, a typical fMRI data set for a single scan on a single subject consists of a temporally highly correlated time series of measurements, taken every 2 s or so for about an hour, on each of, say, $64 \times 64 \times 30$ voxels throughout the brain. Consequently, the total amount of data is enormous, and proper accommodation of both temporal and spatial correlations is needed. Second, models relating fMRI signals to neural changes are complex. The standard tool for analysis done within scans is a variant of the linear model, usually fitted separately by least squares to each voxel [4]. Columns of the design matrix are typically obtained from the schedule of applied stimuli, convolved with an assumed relationship between these stimuli and the blood oxygenation level-dependent (BOLD) response. After that, perform the tests of significance of the model parameters and draw colors on top of significant voxels. This comprises the major idea of statistical parametric mapping (SPM) [5], popularly used in neuroimage study.

In this paper, we intend to incorporate a novel semiparametric model that relaxes the above parametric modeling assumptions. Unlike the general linear model approach in previous studies, our approach has the advantage that we do not assume any *a priori* shape of the hemodynamic response function (HRF), and that we do not assume a particular form of the temporal drift function. Taking full advantage of these flexibilities will help reduce the bias due to model misspecification and enhance the power of detection.

In Section 2, we introduce the one-level (in Section 2.3) and two-level (in Section 2.4) semiparametric estimations of HRF, based on the cubic smoothing spline approach. Section 3 discusses semiparametric F-test statistics for identifying regions of the brain that are activated when a subject performs certain motor, sensory, or cognitive tasks. In nonparametric regression literature, it is customary to apply two-level estimation to enhance the smoothness of the raw estimates from one-level estimation; however, it is not convincingly clear whether two-level estimates really improve upon the discriminating power of one-level estimates in semiparametric tests. Of particular importance to fMRI data analysis is to evaluate whether two-level estimates really gain efficiency over one-level estimates in statistical estimation and inference of HRF. Interestingly, our simulation studies in Sections 4 and 5 provide empirical evidence for the opposite point of view and examine the extent to which one-level estimates outperform two-level estimates for the estimation of HRF and that one-level estimates combined with a bias-correction procedure perform best in detecting the activated brain regions. In Section 6, we apply the semiparametric F-test associated with bias corrected one-level estimates to a real fMRI data set, and compare the activated brain regions using popular imaging analysis tools AFNI (at `http://afni.nimh.nih.gov/afni/`) [6] and FSL (at `http://www.fmrib.ox.ac.uk/fsl/`) [7, 8]. Section 7 concludes this paper with a discussion.

2. SEMIPARAMETRIC ESTIMATION OF HRF

2.1. Semiparametric model for single-voxel and single-run fMRI

The BOLD signal response to neuronal activity is heavily lagged and damped by the hemodynamic response. Following Ward [9] and Worsley *et al*. [3], a single-voxel fMRI time-series data, for a

DOI: 10.1002/sim

given scan and a given subject, can be captured by the convolution model

$$y(t) = \int_0^t s(t-\tau)h(\tau)\,\mathrm{d}\tau + d(t) + \varepsilon(t) \tag{1}$$

where $y(t)$ is the measured noisy fMRI signal, $s(t)$ is the external input stimulus at time t, which could be from a design either block or event related and $s(t)=1$ or 0 indicates the presence or absence of a stimulus, $h(\tau)$ is the HRF at time τ after neural activity, $d(t)$ is a slowly drifting baseline of time t (usually a polynomial trend of at most third order), and $\varepsilon(t)$ is a zero-mean error process, consisting of nonneural noise (due to respiration and blood flow pulsations through the cardiac cycle) and 'white noise' (from random/thermal currents in the body and the scanner).

The HRF $h(t)$ is the part of primary interest to neuroscientists. It should be a smooth function and not identically zero if a voxel responds to the stimuli. We will propose a semiparametric method for characterizing properties of the hemodynamic response in the presence of an unknown smooth drift. Such characterization is essential for accurate prediction of the time-course behavior of neuronal responses.

In practice, the output $y(t)$ and the input $s(t)$ are measured at discrete time points t_i, $i=1,\ldots,n$. The previous model (1) assumes that the fMRI signal, $y(t)$, will be determined by the time-varying stimuli for time points up to t. After discretizing (1), we get

$$y(t_i) = \sum_{j:0\leqslant\tau_j\leqslant t_i} s(t_i-\tau_j)h(\tau_j)\Delta_j + d(t_i) + \varepsilon(t_i), \quad i=1,\ldots,n \tag{2}$$

where $\Delta_j = \tau_j - \tau_{j-1}$. Note that the peak response of HRF $h(t)$ is reached after a short delay of the stimulus and drops quickly to zero. See an example of $h(t)$ given in Glover [10]. We thus suppose $h(t)=0$ for $t>m$ and focus only on the values of $h(t)$ at the first m time points. In practice, n far exceeds m. For instance, m is 18 and n is 185 for each run in the real fMRI data set described in Section 6. Assume that y and s have equal time resolutions of 1 s. Using matrix representation, (2) becomes

$$\mathbf{y} = \mathbf{S}\mathbf{h} + \mathbf{d} + \boldsymbol{\varepsilon} \tag{3}$$

where

$$\mathbf{y} = \begin{bmatrix} y(1) \\ \vdots \\ y(n) \end{bmatrix}, \quad \mathbf{S} = \begin{bmatrix} s(0) & 0 & \cdots & 0 \\ s(1) & s(0) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ s(m-1) & s(m-2) & \cdots & s(0) \\ s(m) & s(m-1) & \cdots & s(1) \\ \vdots & \vdots & \ddots & \vdots \\ s(n-1) & s(n-2) & \cdots & s(n-m) \end{bmatrix}, \quad \mathbf{h} = \begin{bmatrix} h(1) \\ \vdots \\ h(m) \end{bmatrix}, \quad \mathbf{d} = \begin{bmatrix} d(1) \\ \vdots \\ d(n) \end{bmatrix}$$

and $\boldsymbol{\varepsilon} = (\varepsilon(1),\ldots,\varepsilon(n))^{\mathrm{T}}$. A sufficient condition for the identifiability of model (3) is discussed in Appendix A. It is easy to see that $\mathbf{S}$ is a Toeplitz matrix. Handling unequal time resolutions can similarly be done.

DOI: 10.1002/sim

To incorporate the combined effects of multiple types of stimuli, $s_1(t), \ldots, s_r(t)$, associated with HRFs, $h_1(t), \ldots, h_r(t)$, we build an extended model of (1) *via*

$$y(t) = s_1 * h_1(t) + \cdots + s_r * h_r(t) + d(t) + \varepsilon(t)$$

in which $*$ is the convolution operator as in (1). Analogous to (2), the discretized version of this extended model is

$$\mathbf{y} = \mathbf{S}_1\mathbf{h}_1 + \cdots + \mathbf{S}_r\mathbf{h}_r + \mathbf{d} + \boldsymbol{\varepsilon}$$

Apparently, for this extended model, model (3) continues to hold, with

$$\mathbf{S} = [\mathbf{S}_1, \ldots, \mathbf{S}_r] \quad \text{and} \quad \mathbf{h} = \begin{bmatrix} \mathbf{h}_1 \\ \vdots \\ \mathbf{h}_r \end{bmatrix} \tag{4}$$

Additionally, for multiple run fMRI data, we only need to supplement the matrix $\mathbf{S}$ by adding rows composed of the Toeplitz design matrix for each run.

Model (3) is conceivably a semiparametric regression model, with a vector $\mathbf{h}$ of length rm for *parametric* components and a vector $\mathbf{d}$ of length n for *nonparametric* components. The parametric components (related to the unknown HRF) are of our primary interest, whereas the nonparametric components (related to the unknown temporal drift) serve as nuisance effects, and the noise components $\boldsymbol{\varepsilon}$ are serially correlated.

2.2. Cubic smoothing splines

Before we discuss the semiparametric estimation of HRF, we start with a brief overview of the cubic smoothing spline, which is a popular nonparametric curve-fitting method. For a sample of random pairs $\{(X_i, Y_i)\}_{i=1}^n$ from a population (X, Y), suppose that the design variables, $X_1, \ldots, X_n$, are points in an interval $[a, b]$ satisfying $a < X_1 < \cdots < X_n < b$ and we are interested in estimating the regression curve $g(x) = E(Y|X = x)$. The cubic smoothing spline estimator, denoted by $\widehat{g}$, minimizes the penalized sum of squared errors

$$\frac{1}{n}\sum_{i=1}^{n}\{Y_i - g(X_i)\}^2 + \lambda\int_a^b \{g''(x)\}^2\,\mathrm{d}x, \quad \lambda > 0 \tag{5}$$

over the class of all twice-differentiable functions g. The smoothing parameter or the penalty factor λ, upon which the smoothing spline estimator depends, regulates the 'rate of exchange' between fidelity to the observed data and smoothness of the fitted curve. Define $h_i = X_{i+1} - X_i$, $1 \leqslant i \leqslant n-1$, and two band matrices: $\mathscr{R}$, which is $(n-2) \times (n-2)$ with elements given by

$$\begin{aligned} \mathscr{R}(i, i) &= (h_i + h_{i+1})/3 \quad \text{if } i = 1, \ldots, n-2 \\ \mathscr{R}(i, i+1) &= \mathscr{R}(i+1, i) = h_{i+1}/6 \quad \text{if } i = 1, \ldots, n-3 \\ \mathscr{R}(i, j) &= 0 \quad \text{if } |i - j| \geqslant 2 \end{aligned}$$

DOI: 10.1002/sim

and $\mathcal{Q}$, which is $n \times (n-2)$, with entries given by

$$\begin{aligned}
\mathcal{Q}(i,i) &= 1/h_i \quad \text{if } i = 1, \ldots, n-2 \\
\mathcal{Q}(i+1,i) &= -1/h_i - 1/h_{i+1} \quad \text{if } i = 1, \ldots, n-2 \\
\mathcal{Q}(i+2,i) &= 1/h_{i+1} \quad \text{if } i = 1, \ldots, n-2 \\
\mathcal{Q}(i,j) &= 0 \quad \text{if } |i-j| \geqslant 3
\end{aligned}$$

Define $\mathbf{g} = (g(X_1), \ldots, g(X_n))^{\mathrm{T}}$. According to Theorem 2.1 of Green and Silverman [11], $\int_a^b \{g''(x)\}^2 \, \mathrm{d}x = \mathbf{g}^{\mathrm{T}} K \mathbf{g}$, where $K = \mathcal{Q}\mathcal{R}^{-1}\mathcal{Q}^{\mathrm{T}}$ is a symmetric matrix. Thus, the criterion function (5) can be rewritten as a quadratic form,

$$(\mathbf{y} - \mathbf{g})^{\mathrm{T}}(\mathbf{y} - \mathbf{g})/n + \mathbf{g}^{\mathrm{T}} \lambda K \mathbf{g}$$

Its minimizer gives the cubic smoothing spline estimator of $\mathbf{g}$, represented by

$$\widehat{\mathbf{g}} = (\mathbf{I} + n\lambda K)^{-1}\mathbf{y}$$

where $\mathbf{I}$ is an identity matrix. We call

$$S = (\mathbf{I} + n\lambda K)^{-1} \tag{6}$$

the smoothing matrix of cubic smoothing splines. This matrix carries information about the design points and smoothing parameter, but does not rely on the configuration of the response variables. Refer to Section 2.3 of Zhang [12] for further discussion on finite sample and asymptotic properties of the smoothing matrices for higher-order smoothing splines.

2.3. One-level estimation of HRF

We first describe a one-level estimation of HRF in the semiparametric model (3). For a known value of $\mathbf{h}$, a cubic smoothing spline estimate of $\mathbf{d}$ is given by

$$\widehat{\mathbf{d}}(\mathbf{h}) = S_d(\mathbf{y} - \mathbf{S}\mathbf{h})$$

where S_d is an $n \times n$ smoothing matrix as defined in (6) for the cubic smoothing splines, associated with the design points $\{1, 2, \ldots, n\}$, and the subscript d emphasizes the goal of estimating the temporal drift. Substituting $\mathbf{d}$ in (3) by $\widehat{\mathbf{d}}(\mathbf{h})$, we get

$$\widetilde{\mathbf{y}} \approx \widetilde{\mathbf{S}}\mathbf{h} + \boldsymbol{\varepsilon} \tag{7}$$

where $\widetilde{\mathbf{y}} = (\mathbf{I} - S_d)\mathbf{y}$ and $\widetilde{\mathbf{S}} = (\mathbf{I} - S_d)\mathbf{S}$. Treating (7) as a parametric regression model, the weighted least-squares estimate of $\mathbf{h}$ is produced by

$$\widehat{\mathbf{h}}_{\text{rough}} = (\widetilde{\mathbf{S}}^{\mathrm{T}}\widehat{R}^{-1}\widetilde{\mathbf{S}})^{-1}\widetilde{\mathbf{S}}^{\mathrm{T}}\widehat{R}^{-1}\widetilde{\mathbf{y}} \tag{8}$$

where $\widehat{R}$ estimates the correlation matrix R of $\boldsymbol{\varepsilon}$, namely, $\mathrm{cov}(\boldsymbol{\varepsilon}, \boldsymbol{\varepsilon}) = \sigma^2 R$, with variance σ^2. This $\widehat{\mathbf{h}}_{\text{rough}}$ supplies the one-level estimate of HRF.

DOI: 10.1002/sim

There is an extensive literature on the estimation of covariance matrices for correlated outcomes. See Anderson [13], Diggle and Verbyla [14], Boik [15], and many others. To estimate the large covariance matrix for the current fMRI data, we apply a computationally more feasible method studied in Zhang *et al*. [16], which assumes the stationarity of the noise process and that the autocovariance sequence vanishes for time lag exceeding 2.

2.4. Two-level estimation of HRF

We now describe a two-level estimation of HRF. Note that the one-level estimates of HRF at m points are usually rough, since the smoothness of the underlying HRF is not taken into account. To obtain a smooth estimate of the hemodynamic response curve $h(t)$, the two-level estimation simply applies the cubic spline smoothing to the one-level estimates $\widehat{\mathbf{h}}_{\text{rough}}$ to produce refined estimates

$$\widehat{\mathbf{h}}_{\text{smooth}} = S_h \widehat{\mathbf{h}}_{\text{rough}} \tag{9}$$

where S_h denotes an $m \times m$ cubic spline smoothing matrix, associated with the design points $\{1, 2, \ldots, m\}$, and the subscript h pinpoints the use of this smoothing matrix for estimating the HRF $\mathbf{h}$ in (3). We call $\widehat{\mathbf{h}}_{\text{smooth}}$ the two-level estimate of HRF.

Following the estimation of HRF, the one-level and two-level estimates of the drift components are formed by

$$\widehat{\mathbf{d}}_{\text{rough}} = \widehat{\mathbf{d}}(\widehat{\mathbf{h}}_{\text{rough}}) = S_d(\mathbf{y} - \mathbf{S}\widehat{\mathbf{h}}_{\text{rough}}) \quad \text{and} \quad \widehat{\mathbf{d}}_{\text{smooth}} = \widehat{\mathbf{d}}(\widehat{\mathbf{h}}_{\text{smooth}}) = S_d(\mathbf{y} - \mathbf{S}\widehat{\mathbf{h}}_{\text{smooth}})$$

respectively.

We would like to make some remarks on the computational aspect of smoothing matrices S_d and S_h. Note that for each fixed smoothing parameter, they do not vary from one session to another session and thus each matrix only needs to be computed once. This feature will significantly reduce the computational burden, particularly when n is large.

3. DETECTION OF ACTIVATED BRAIN REGIONS

Identification of a particular brain region with a specific function has become a central theme in neuroscience. Traditional methods use SPM to identify activation foci or characterization of distributed changes with spatial modes. Statistical inferences about the resulting SPM are made using distributional approximations from the theory of Gaussian random fields [17, 18]. In order to generate activation maps, statistical inference must be drawn from the estimated hemodynamic responses. However, since in our approach the HRFs are estimated semiparametrically, existing statistical inference procedures are not immediately applicable to our one-level and two-level semiparametric estimates of HRF.

3.1. Hypothesis test for HRF

Determining whether the fMRI time series on a particular voxel responds to a given signal waveform can be expressed in terms of a statistical test for the hypothesis

$$H_0 : h_1(t) = \cdots = h_r(t) = 0$$

DOI: 10.1002/sim

Testing HRF also plays an important role in comparing shapes of HRF across different tissue types, corresponding to the hypothesis

$$H_0 : h_A(t) \equiv h_B(t) \quad \text{for all } t$$

where $h_A(t)$ and $h_B(t)$ are the corresponding HRFs at two different voxels (for e.g. the left and right amygdalae), and in detecting contrast in the stimuli, i.e. testing for the hypothesis

$$H_0 : h_k(t) = 0 \quad \text{for a given } k = 1, \ldots, r$$

Under our semiparametric model (3), all these testing problems can be formulated in a more general form

$$H_0 : A\mathbf{h} = \mathbf{0} \quad \textit{versus} \quad H_1 : A\mathbf{h} \neq \mathbf{0} \tag{10}$$

where A is a full row rank matrix.

3.2. Semiparametric test statistics

We now discuss a semiparametric F-test. Recall that approximation (7) takes the form of a general linear model. Suppose that the rank of A in (10) is k and that $\widehat{\mathbf{h}}$ is an estimate of $\mathbf{h}$. The semiparametric F-test statistic for (10) can be constructed by

$$\frac{(A\widehat{\mathbf{h}})^{\mathrm{T}}\{A(\widetilde{\mathbf{S}}^{\mathrm{T}}\widehat{R}^{-1}\widetilde{\mathbf{S}})^{-1}A^{\mathrm{T}}\}^{-1}(A\widehat{\mathbf{h}})/\mathsf{k}}{\widehat{\mathbf{r}}^{\mathrm{T}}\widehat{R}^{-1}\widehat{\mathbf{r}}/(n-rm)} \tag{11}$$

where $\widehat{\mathbf{r}} = \widetilde{\mathbf{y}} - \widetilde{\mathbf{S}}\widehat{\mathbf{h}}$, and the denominator in (11) serves as a variance estimator. The asymptotic distribution of this test statistic under the null hypothesis can be approximated by an F distribution with k and $n - rm$ degrees of freedom. The results in Fan *et al.* [19] demonstrate that a class of the generalized likelihood ratio statistics based on some appropriate nonparametric function estimators is asymptotically distribution free and follows χ^2-distributions under null hypotheses for a number of useful hypotheses and a variety of useful models including Gaussian white noise models, nonparametric regression models, varying coefficient models and generalized varying coefficient models. These results, when translated to the current setup, provide insights into the distributional approximation.

The above semiparametric F-test statistic usually works reasonably well. Nonetheless, from the identity, $\widetilde{\mathbf{y}} = \widetilde{\mathbf{S}}\mathbf{h} + \boldsymbol{\varepsilon} + \{\mathbf{d} - \widehat{\mathbf{d}}(\mathbf{h})\}$, and numerical evidences in Figure 3, we observe that this test may not be very accurate when the noise level is small. The loss of accuracy is mainly caused by approximation (7), which entirely ignores an additive term $\mathbf{d} - \widehat{\mathbf{d}}(\mathbf{h})$. This term is negligible when the noise level is relatively large, but may become a significant source of bias when the noise level is comparatively small. To remedy the bias problem, we adopt the following bias-corrected semiparametric F-test statistic:

$$\frac{(A\widehat{\mathbf{h}}_{\mathrm{bc}})^{\mathrm{T}}\{A(\widetilde{\mathbf{S}}^{\mathrm{T}}\widehat{R}^{-1}\widetilde{\mathbf{S}})^{-1}A^{\mathrm{T}}\}^{-1}(A\widehat{\mathbf{h}}_{\mathrm{bc}})/\mathsf{k}}{\widehat{\mathbf{r}}_{\mathrm{bc}}^{\mathrm{T}}\widehat{R}^{-1}\widehat{\mathbf{r}}_{\mathrm{bc}}/(n-rm)} \tag{12}$$

where $\widehat{\mathbf{h}}_{\mathrm{bc}} = \widehat{\mathbf{h}} - (\widetilde{\mathbf{S}}^{\mathrm{T}}\widehat{R}^{-1}\widetilde{\mathbf{S}})^{-1}\widetilde{\mathbf{S}}^{\mathrm{T}}\widehat{R}^{-1}\widetilde{\widehat{\mathbf{d}}}$ and $\widehat{\mathbf{r}}_{\mathrm{bc}} = \widetilde{\mathbf{y}} - \widetilde{\mathbf{S}}\widehat{\mathbf{h}} - \widetilde{\widehat{\mathbf{d}}}$, with $\widetilde{\widehat{\mathbf{d}}} = (\mathbf{I} - S_d)\widehat{\mathbf{d}}$ and $\widehat{\mathbf{d}} = S_d(\mathbf{y} - \mathbf{S}\widehat{\mathbf{h}})$. Note that for a large noise level, $\widetilde{\widehat{\mathbf{d}}}$ is negligible and thus the bias-corrected and uncorrected versions will be similar.

DOI: 10.1002/sim

Table I. One-level and two-level semiparametric F-test statistics.

F1: $\widehat{\mathbf{h}} = \widehat{\mathbf{h}}_{\text{rough}}$, without bias correction	F2: $\widehat{\mathbf{h}} = \widehat{\mathbf{h}}_{\text{rough}}$, with bias correction
F3: $\widehat{\mathbf{h}} = \widehat{\mathbf{h}}_{\text{smooth}}$, without bias correction	F4: $\widehat{\mathbf{h}} = \widehat{\mathbf{h}}_{\text{smooth}}$, with bias correction

For both test statistics in (11) and (12), the HRF estimate, $\widehat{\mathbf{h}}$, could utilize either the one-level estimate $\widehat{\mathbf{h}}_{\text{rough}}$ or the two-level estimate $\widehat{\mathbf{h}}_{\text{smooth}}$. Therefore, we have four options for constructing semiparametric F-test statistics as listed in Table I. (There, $\widehat{\mathbf{d}} = \widehat{\mathbf{d}}_{\text{rough}}$ in F1 and F2, and $\widehat{\mathbf{d}} = \widehat{\mathbf{d}}_{\text{smooth}}$ in F3 and F4.) In Sections 4 and 5, we will use simulation studies to explore whether two-level estimates have efficiency over one-level estimates in detecting activated brain regions.

4. SIMULATION STUDY OF fMRI AT A SINGLE VOXEL

4.1. Estimation of HRF and drift

We simulate an fMRI experiment with a single run and a single type of stimulus. In the simulation, the length of time series is 200, and 1000 realizations are conducted. For the reader's convenience, Appendix B summarizes the implementation details of semiparametric analysis for single-voxel fMRI data.

- The time-varying stimuli are generated from independent Bernoulli trials such that $P\{s(t) = 1\} = 0.5$.
- The HRF is $h(t) = g_1(1.5(t-1))/a_1 - g_2(1.5(t-1))/a_2$, $t = 1, \ldots, 18$ (so that $m = 18$), where $g_1(t) = (t-5.5)^5 \exp\{-(t-5.5)/0.9\}$ and $g_2(t) = 0.4(t-5.5)^{12} \exp\{-(t-5.5)/0.7\}$, $a_1 = \max\{g_1(t)\}$ and $a_2 = \max\{g_2(t)\}$. Following Glover [10], the HRF here takes the difference of two gamma functions, the canonical form implemented in FSL.
- The drift function is $d(t) = 10 \sin\{\pi(t/200 - 0.21)\}$, $t = 1, \ldots, 200$.
- The noise process $\varepsilon(t)$ is the sum of independent noise processes ε_1 and ε_2 (see Purdon *et al.* [20]); $\{\varepsilon_1(t)\}$ are i.i.d. normal, with mean zero and variance 0.5216^2, 0.3689^2, 0.2608^2 and 0.1844^2, respectively; ε_2 is AR(1), i.e. $\varepsilon_2(t) = \rho\varepsilon_2(t-1) + z(t)$, with $\rho = 0.638$, and $z(t)$ follows the normal distribution, with mean zero and variance 0.5216^2, 0.3689^2, 0.2608^2 and 0.1844^2, respectively. These choices give the noise lag-one auto-correlation equal to 0.4 and the signal-to-noise ratio (SNR) about 1, 2, 4, and 8, respectively where $\text{SNR} = \text{var}(\mathbf{Sh})/\text{var}(\boldsymbol{\varepsilon})$.

The one-level and two-level estimates of HRF and temporal drift are displayed in Figure 1 with $\text{SNR} = 1$ and in Figure 2 with $\text{SNR} = 8$. Evidently, compared with the two-level estimates, the one-level estimates are more centered around the actual values of HRF. On the other hand, the drift function estimates are less affected by one-level and two-level estimations of HRF. Besides, the HRF and drift estimates by our method outperform those by AFNI.

4.2. Hypothesis test for HRF

As an illustration, the hypothesis test for $H_0 : \mathbf{h} = \mathbf{0}$ *versus* $H_1 : \mathbf{h} \neq \mathbf{0}$ is considered. This is used for testing whether the brain activity in a voxel is triggered or not. To check the agreement between the F distribution with null distributions of the semiparametric F-test statistics, the fMRI data are simulated in the same way as above, except that $\mathbf{h} = \mathbf{0}$.

DOI: 10.1002/sim

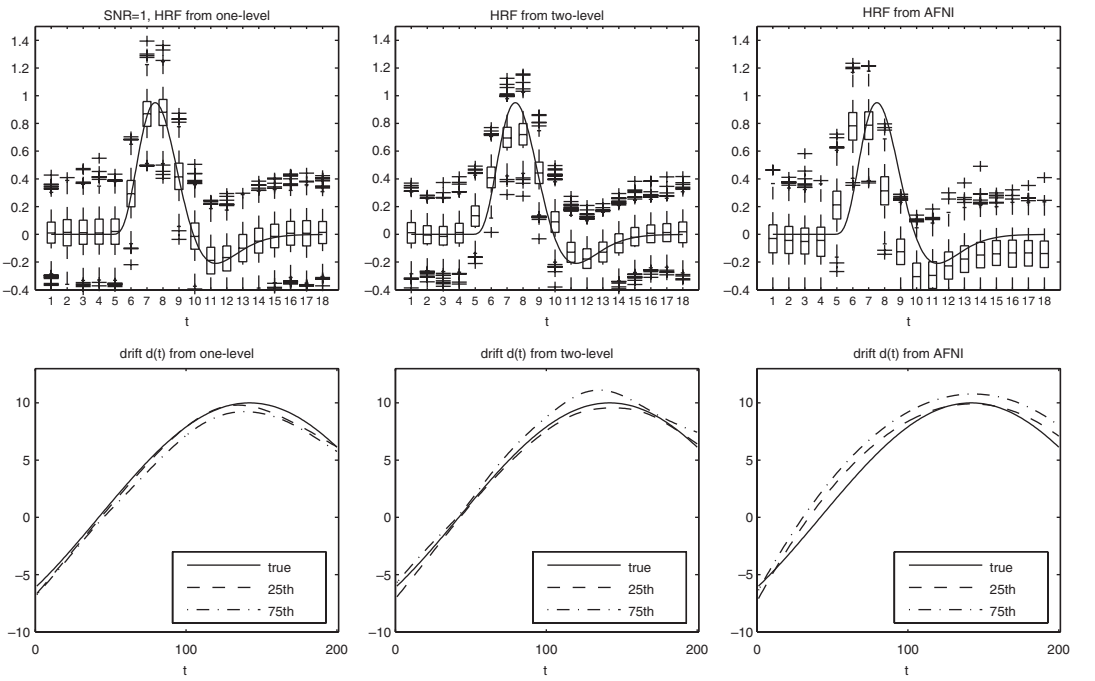

Figure 1. Comparison of one-level, two-level and AFNI estimates of HRF and drift. Top panel: boxplots of $\widehat{\mathbf{h}}_{\text{rough}}$ and $\widehat{\mathbf{h}}_{\text{smooth}}$; the solid curves are the true function $h(t)$. Bottom panel: solid curves denote the true drift $d(t)$, and the estimated curves from two typical samples are presented corresponding to the 25th (the dashed curve), and the 75th (the dash-dotted curve) percentiles among the ASE-ranked values, where $\text{ASE} = \sum_{i=1}^{n} \{\widehat{d}(i) - d(i)\}^2/n$.

The QQ plots of the (1st up to 99th) percentiles of F-test statistics against those of the $F_{m,n-m}$ distribution are presented in Figure 3. In the top panel, the semiparametric F-test statistics use the true covariance matrix and fix the smoothing parameters at their theoretically optimal values (minimizing the mean-squared errors of estimators) for estimating HRF and drift in each simulation. For the sake of clarity, only the cases of SNR equal to 1 and 8 are presented; the former is the 'large noise level' case, whereas the latter is the 'small noise level' case. In either case, we observe that the finite sampling distributions of `F1` and `F2`, based on the one-level estimation, nearly perfectly agree with the F distribution. In contrast, the sampling distributions of `F3` and `F4`, based on the two-level estimation, depart further from the F distribution. The QQ plots also lend support that the bias-corrected versions, `F2` and `F4`, are better than or at least as good as their bias uncorrected counterparts, `F1` and `F3`, respectively.

For a more realistic comparison, the semiparametric F-test statistics in the bottom panel of Figure 3 use the estimated covariance matrices and data-driven smoothing parameters, following Steps 1–7 in Appendix B. The results are similar in spirit to the ones in the top panel and continue to support the bias-correction procedure when applied to either one-level or two-level estimation. In this regard, the detection accuracy based on the test `F3` is lower than any of the other three semiparametric F-tests.

DOI: 10.1002/sim

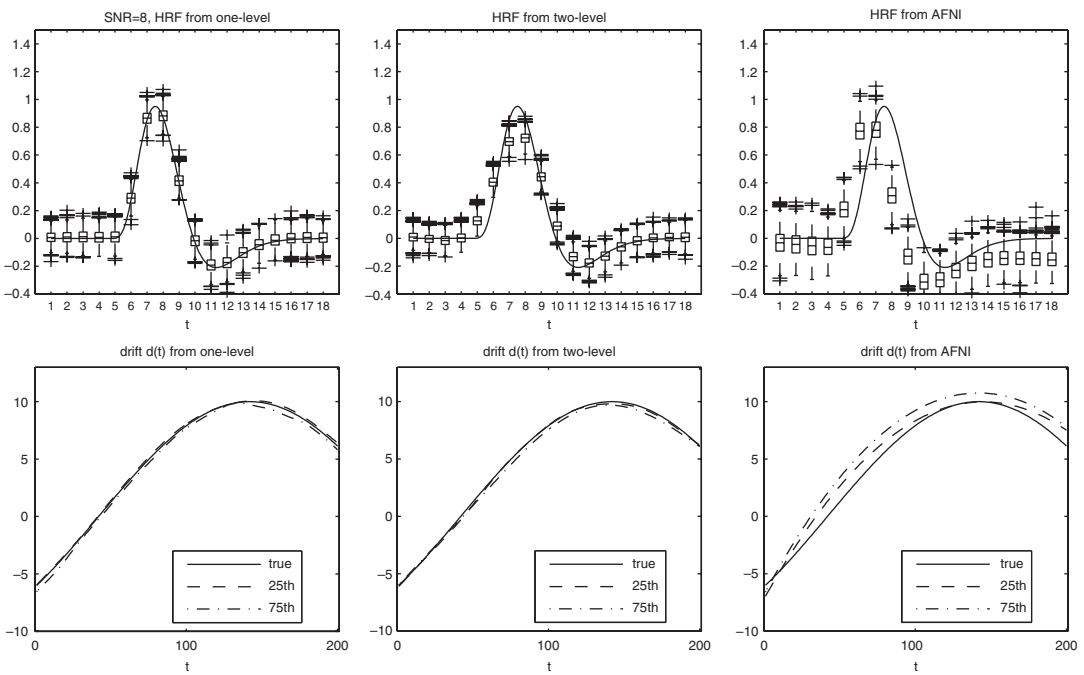

Figure 2. Same as in Figure 1, except SNR = 8.

5. SIMULATION STUDY OF fMRI OVER AN ENTIRE BRAIN

5.1. Simulating the fMRI data

We simulate a whole brain fMRI data set, with the aim of mimicing true brain activity, to the maximum extent feasible. The experimental design, timings and size are exactly the same as those of the real data set in Section 6. An HRF profile is extracted from a voxel which shows the strongest responses in the real data set. For each voxel, the simulated drift is obtained from an adequate smoothing to the time series for the corresponding voxel of the real data set. The simulated noise variance profile is determined from a variance map, which is made by a $5 \times 5 \times 5$ spatial median smoothing on median values of squared residuals of the real time series by subtracting the simulated drift profile as mentioned before. The noise process $\varepsilon(t)$ is generated in a fashion similar to that explained in Section 4.1. Specifically, the variances of $\varepsilon_1(t)$ and $z(t)$ are chosen to be equal such that $\text{var}\{\varepsilon(t)\}$ is one-fifth of the variance map. The HRF profiles, in accordance with the stimuli in the experiment, are added to two regions that are postulated to be truly active. In these two zones, the HRFs have been rescaled to about 17 and 12 per cent of the amplitude of the original HRF profiles. The purpose of rescaling HRFs and noise variance is to amplify the drift effect and weaken the HRF response so that the estimation of HRF is more challenging. Plate 1 gives nine different slices which highlight the two activated brain regions. Note that, throughout the paper, we apply the same registration transform from the real brain data to the $T1$ high-resolution image of the subject's brain.

DOI: 10.1002/sim

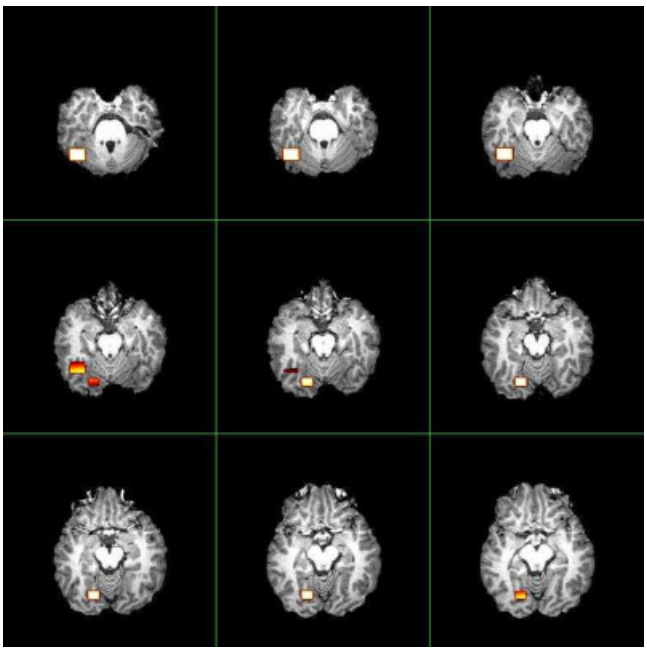

Plate 1. True activated regions (denoted by hot color) for the simulated fMRI data set.

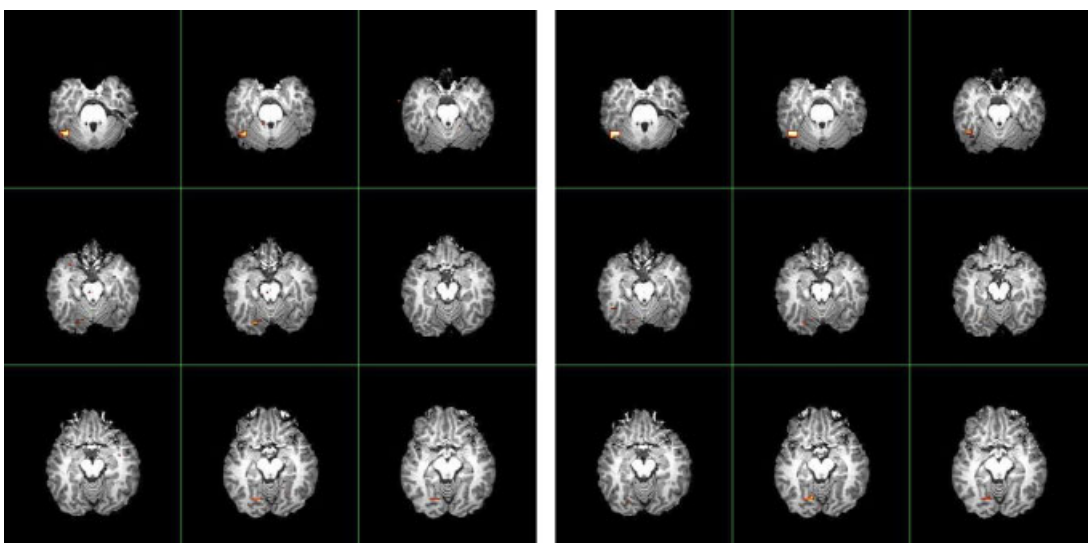

Plate 2. Activated regions discovered by AFNI (on the left) and FSL (on the right) for the simulated fMRI data set. The FDR level is 0.05.

DOI: 10.002/sim

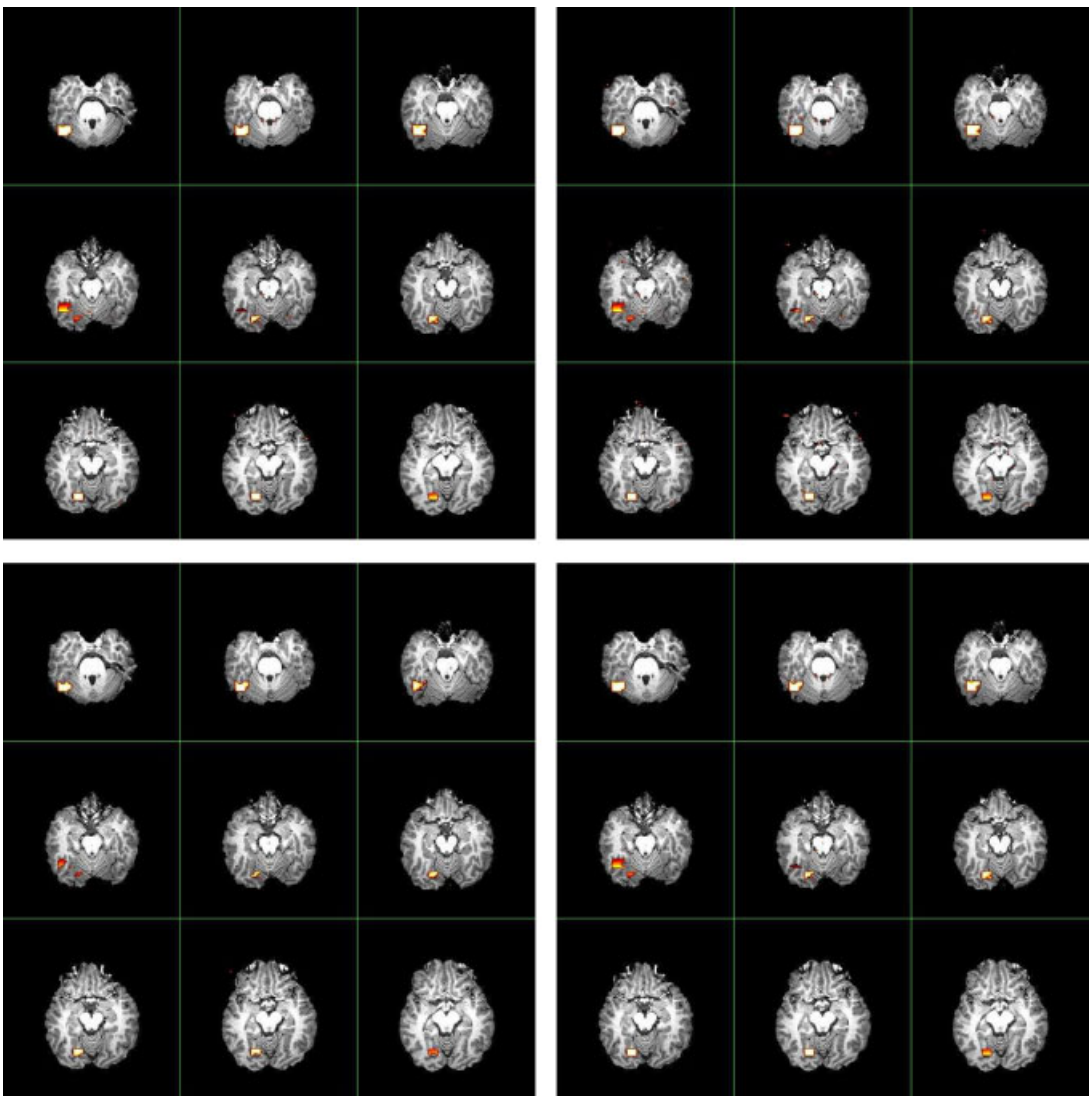

Plate 3. Activated regions discovered by our methods for the simulated fMRI data set. Top panel: `F1` (on the left) and `F2` (on the right). Bottom panel: `F3` (on the left) and `F4` (on the right). The FDR level is 0.05.

DOI: 10.002/sim

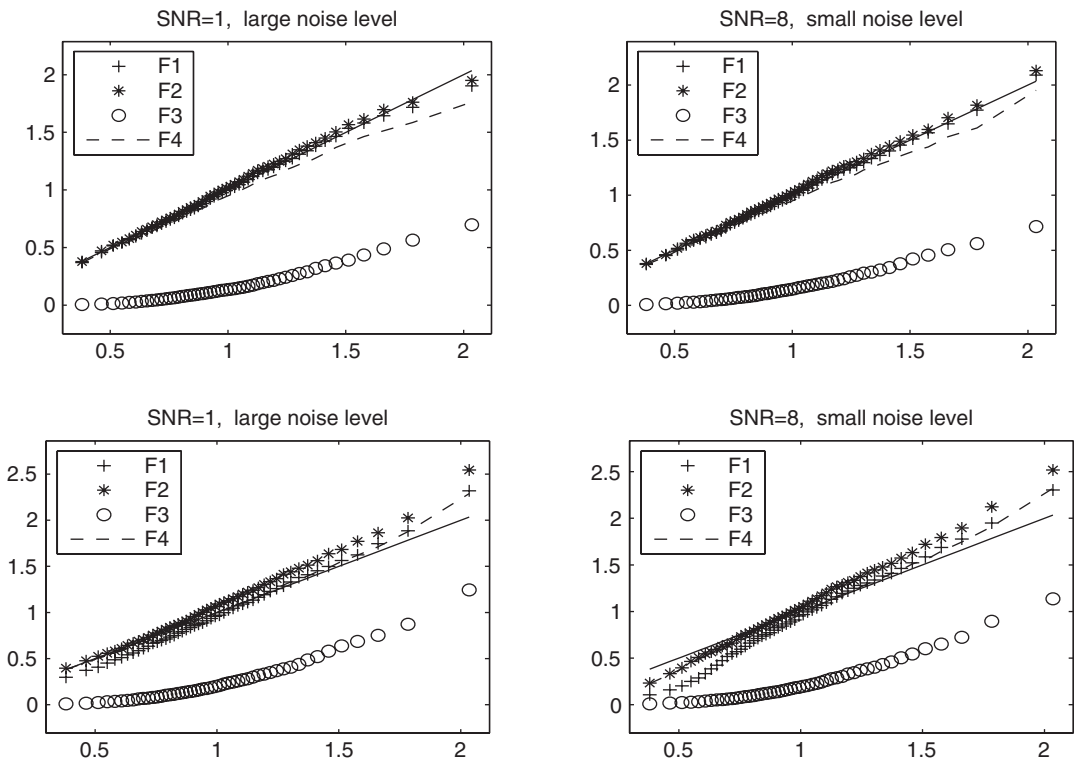

Figure 3. Empirical quantiles (on the y-axis) of semiparametric F-test statistics (where the top panel uses the true covariance matrix and the optimal smoothing parameters, and the bottom panel uses the estimated covariance matrices and data-driven smoothing parameters) *versus* quantiles (on the x-axis) of $F_{m,n-m}$ distribution. Solid line: the 45-degree reference line.

5.2. Detection of activated brain regions

We compare the activated brain regions identified by our approach with methods offered by popular softwares AFNI and FSL. The false discovery rate (FDR) level 0.05 is used for performing the multiple comparison. To facilitate comparison and to help practitioners better view possible differences when applying different methods, Table II tabulates the HRF, drift and error implemented in AFNI and FSL. Inspection of Plate 2 reveals that both AFNI and FSL fail to locate an activated brain area, and that the other region, although correctly detected, has an appreciably reduced size relative to the actual size. This detection bias suggests that the stringent modeling assumptions be relaxed to ameliorate the effects of misspecification.

The brain regions detected using our approach are shown in Plate 3. All four versions of F-test statistics are capable of locating both active regions, and `F1` and `F2` are superior to `F3` and `F4` in terms of the volume of the detected regions. Moreover, to eliminate tiny and scattered regions which are potentially falsely discovered, we apply a new multiple comparison procedure (at FDR level 0.05) proposed in Zhang *et al*. [16] to all four F-test statistics. The theoretical justification for that procedure (controlling the FDR) can be found in Zhang *et al*. [16]. As evidenced in Plate 4, all four versions achieve much more accurate detections than their counterparts in Plate 3, with

DOI: 10.1002/sim

Table II. HRF, drift, and error implemented in AFNI and FSL.

	AFNI (tool `3dDeconvolve`)	FSL (tool `FEAT`)
$h(t)$	Finite impulse response filter	Difference of two gamma functions, which is the canonical form
$d(t)$	Quadratic polynomial	Removed in the preprocessing, using high-pass temporal filtering (Gaussian-weighted LSF straight line fitting)
$\varepsilon(t)$	i.i.d.	Autocorrelation estimated by Tukey tapering of the spectrum of the residuals

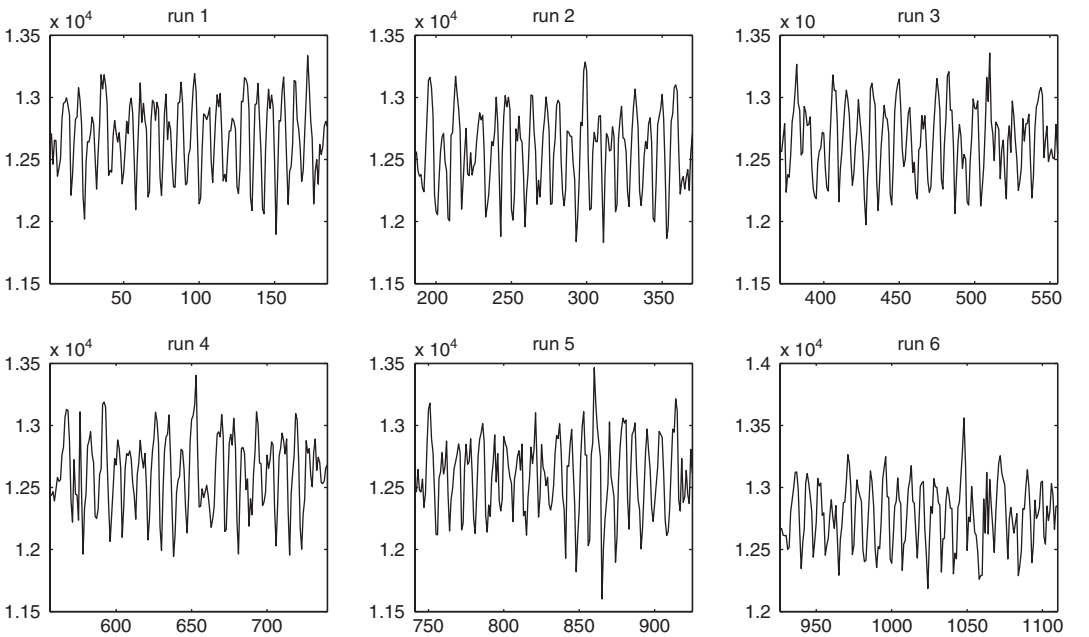

Figure 4. Raw fMRI time series from an activated brain voxel (49, 41, 10).

`F1` and `F2` continuing to outperform `F3` and `F4`. Therefore, for applications to the real fMRI data set in Section 6, we will employ only `F1` and `F2`.

6. APPLICATION TO REAL fMRI DATA

6.1. Experiment design, data description, and analysis aim

In an emotional control study, subjects saw a series of negative or positive emotional images, and were asked to either suppress or enhance their emotional responses to the image, or to simply attend to the image. Therefore, there were six types of trials (i.e. six types of stimuli): negative–enhance (`neg-enh`), negative–attend (`neg-att`), negative–suppress (`neg-sup`),

DOI: 10.1002/sim

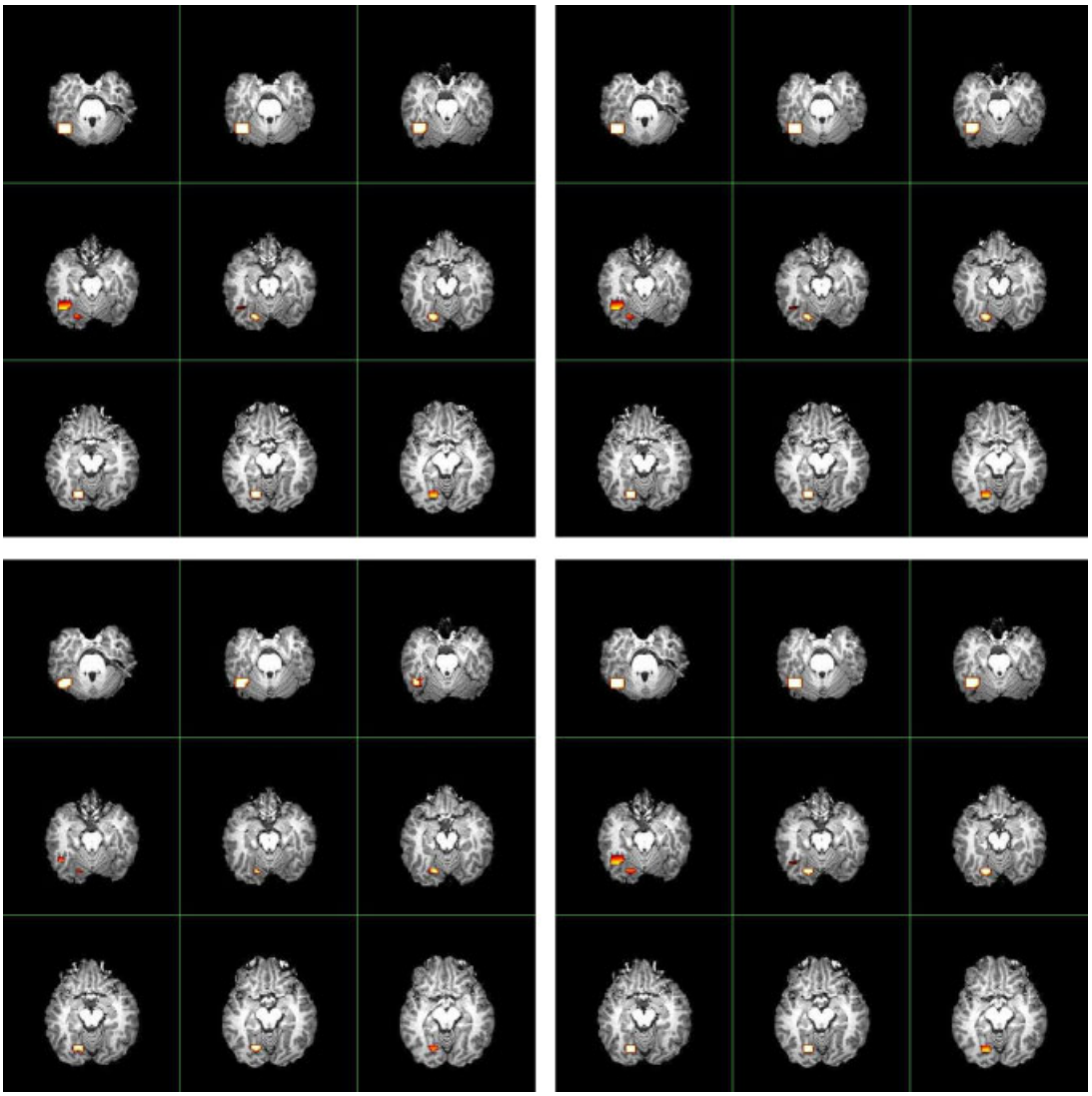

Plate 4. Same as in Plate 3, except the multiple comparison procedure uses Zhang *et al*. [16].

DOI: 10.002/sim

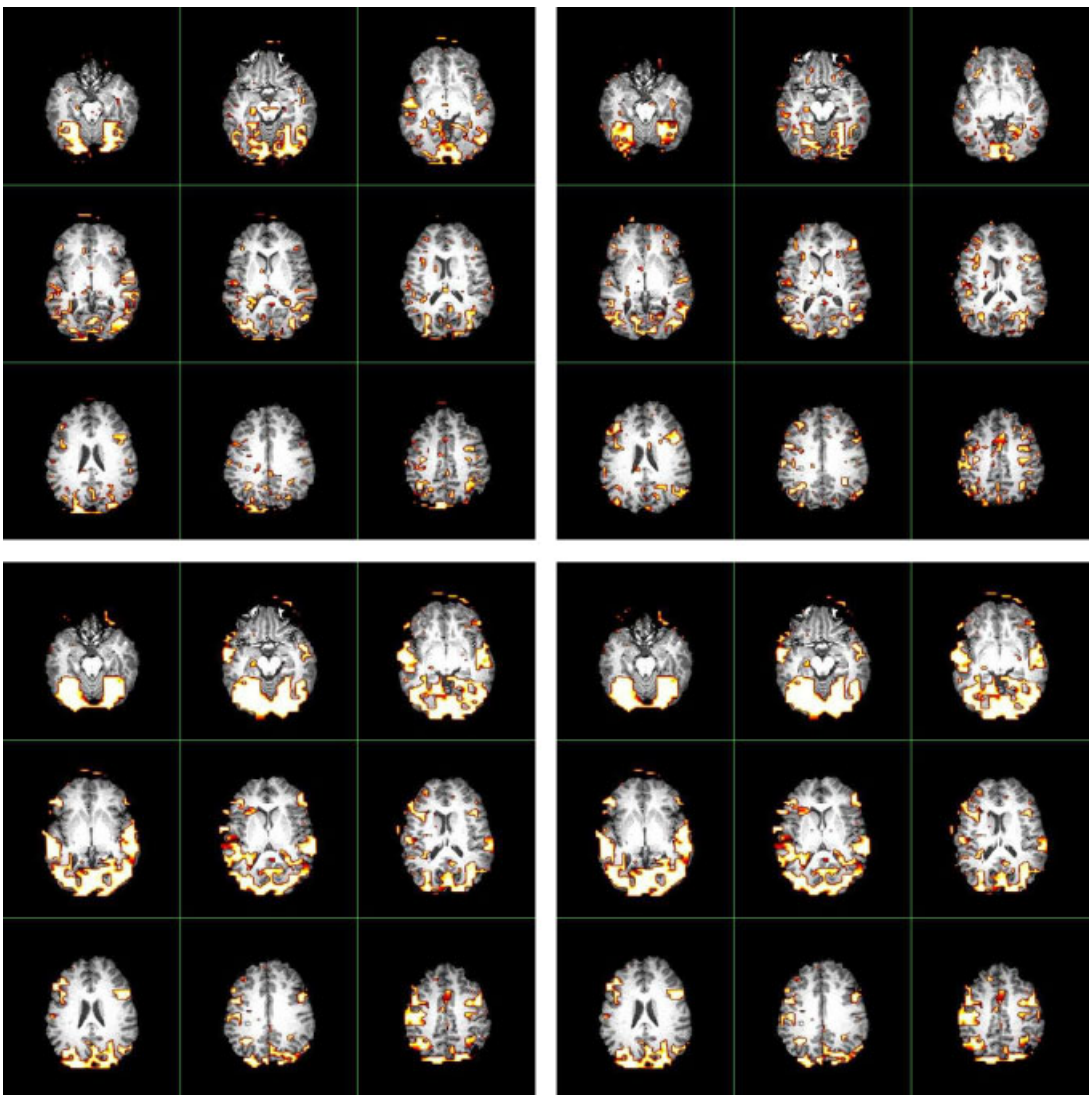

Plate 5. Activated regions discovered for the real fMRI data set. Top panel: AFNI (on the left) and FSL (on the right). Bottom panel: `F1` (on the left) and `F2` (on the right), using the multiple comparison procedure in Zhang *et al*. [16]. The FDR level is 0.001.

DOI: 10.002/sim

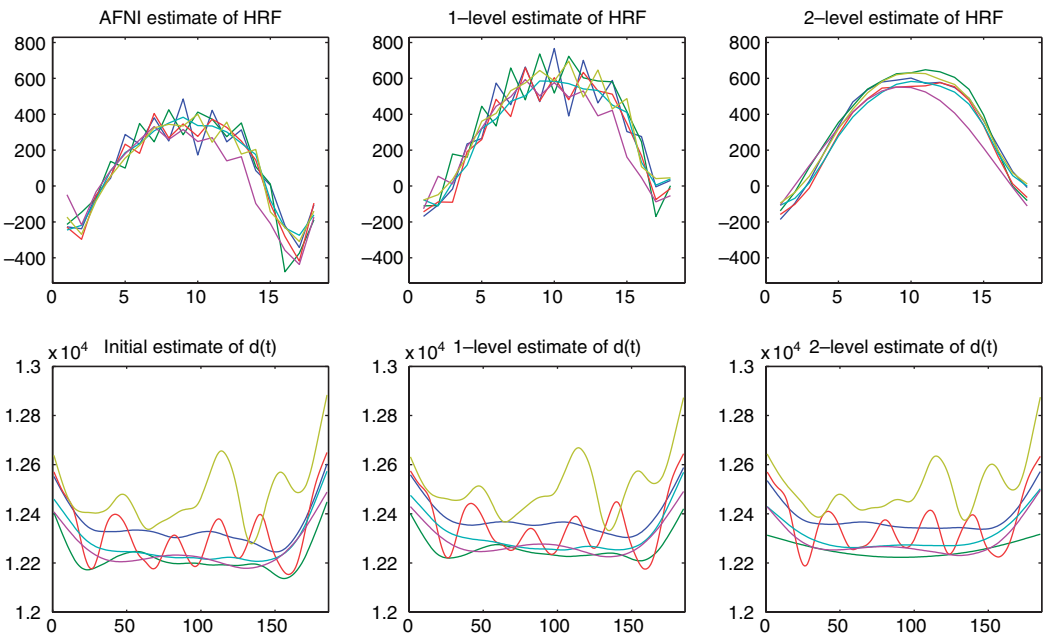

Figure 5. One-level and two-level estimates of HRF (for each of six stimuli) and drift (for each of six runs) at an activated voxel (49, 41, 10). The initial estimate of drift follows Step 3 in Appendix B.

positive–enhance (`pos-enh`), positive–attend (`pos-att`), and positive–suppress (`pos-sup`). The sequence of trials was randomized. The time between successive trials also varied. There were 24 trials each of the `neg-enh`, `neg-sup`, `pos-enh`, and `pos-sup`, and 11 trials each of the `neg-att` and `pos-att`.

The size of the whole brain data set is $64 \times 64 \times 30$. At each voxel, the time series has six runs, each containing 185 observations with a time resolution of 2 s, thus TR = 2 s and the total length is 1110. In contrast, the length of stimuli is 2220; the timing of the stimuli has a time resolution of 1 s, and thus each HRF output will also be sampled at 1 s. Hence, the odd rows of the design matrix $\mathbf{S}$ in (4) suffice for analysis.

The study aims to estimate the BOLD response to each of the trial types for 1–18 s following the image onset. We analyze the fMRI data set containing one subject. The length of the estimated HRF is set equal to 18.

6.2. Single-voxel fMRI analysis

Figure 4 plots the fMRI time series of six runs from an activated brain voxel (49, 41, 10). The estimates of BOLD response to each stimulus type, as given by AFNI analysis, the one-level and two-level semiparametric methods, are compared in the top panel of Figure 5. Note that since the HRF is common to all runs, the HRF is estimated based on data from the combined runs. In contrast, as the drift is specific to each separate run, data from that particular run are used in the estimation of the drift.

DOI: 10.1002/sim

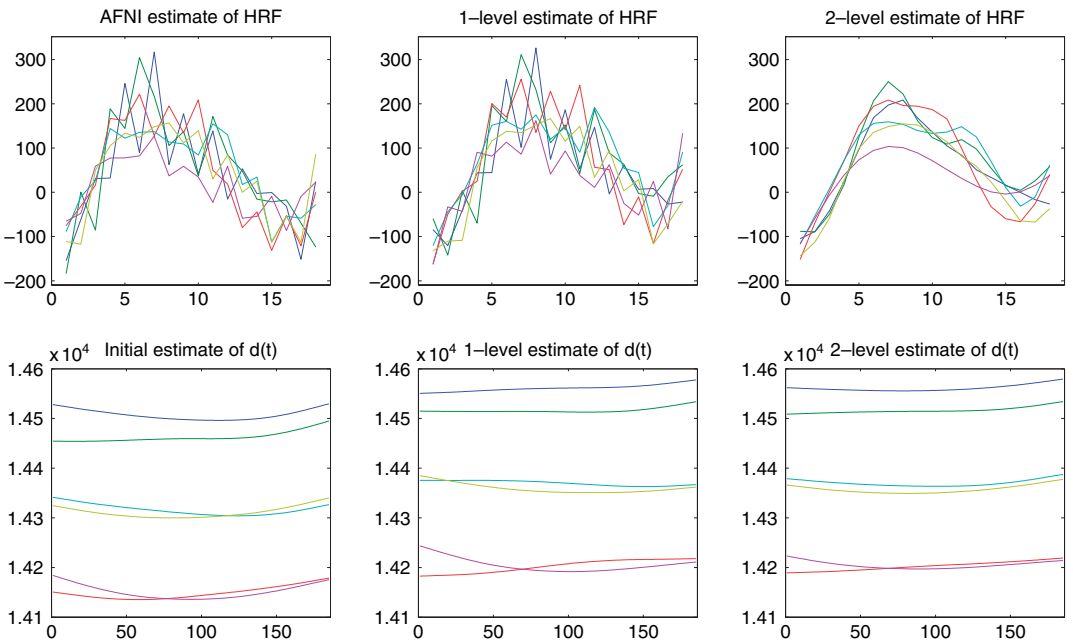

Figure 6. Same as in Figure 5, except from an activated voxel (24, 32, 7).

The activation effect detected by our methods appears to be stronger than that of AFNI. Meanwhile, the drift curve estimated by our methods exhibits a highly nonlinear pattern, which suggests that AFNI misspecifies the drift to some degree. Henceforth, the one-level and two-level estimates of HRF are expected to suffer less from model bias than AFNI estimates.

For an alternative activated voxel (24, 32, 7), the bottom panel of Figure 6 shows low-order polynomial trends of the estimated drift functions using our methods. This tends to support the specification of drift in AFNI. Figure 6 manifests that even for models ideally suited for AFNI's analysis of a particular voxel, our method yields HRF estimates favorably comparable with or nearly indistinguishable from those calculated by AFNI.

6.3. Detection of activated brain regions

We compare the brain regions detected by our methods (based on `F1` and `F2`) with AFNI and FSL. Again, HRF in FSL is specified as the difference between two gamma functions, and the drift term in AFNI is specified as a quadratic polynomial. We use FDR at level 0.001 to carry out the multiple comparisons. This level is set to avoid excessive discoveries, most of which are thought to be false. A comparison of the detected regions is visualized in Plate 5. Our detected regions are closer to those afforded by AFNI, but our methods find activation in much more clustered regions of the brain. For example, our results do not have the holes as seen in the detected regions on the first slice of AFNI and FSL. AFNI gives more tiny scattered findings, which are more likely to be false discoveries. FSL detects very scattered regions which are difficult to interpret. In addition, the volumes of the detected regions by FSL are substantially smaller than those of the regions deleted by AFNI and our methods.

DOI: 10.1002/sim

7. DISCUSSION

In this paper, we present a flexible semiparametric procedure based on the smoothing spline approach for the modeling, estimation, and inference of HRF for brain fMRI data. We demonstrate that, contrary to some conventional thought, the two-level estimation of HRF does not have efficiency over the one-level estimation. Furthermore, in the ideal case that the true covariance matrix structure of the error process is known, the proposed semiparametric F-test based on the one-level estimation of HRF outperform its two-level counterpart in the detection of activated brain areas. For realistic applications, we recommend the bias-corrected semiparametric F-test combined with the one-level estimation. This test compares favorably with some popular imaging analysis tools in which the modeling assumptions are more restrictive than those in our semiparametric model. We intend to explore issues like more efficient estimation of large covariance matrix and more rigorous investigation of the sampling properties of the semiparametric F-test statistics in future work.

APPENDIX A: IDENTIFIABILITY OF MODEL (3)

Before showing the identifiability of model (3), we need the proposition.

Proposition A1
If the matrix $\mathrm{cov}(\mathbf{S}^{\mathrm{T}}, \mathbf{S}^{\mathrm{T}})$ is positive definite, then model (3) is identifiable.

Proof
Suppose that there exist coefficient vectors $\mathbf{h}_1$, $\mathbf{h}_2$, $\mathbf{d}_1$, and $\mathbf{d}_2$ such that

$$\mathbf{S}\mathbf{h}_1 + \mathbf{d}_1 = \mathbf{S}\mathbf{h}_2 + \mathbf{d}_2 \tag{A1}$$

or, equivalently, $(\mathbf{h}_1 - \mathbf{h}_2)^{\mathrm{T}}\mathbf{S}^{\mathrm{T}} + (\mathbf{d}_1 - \mathbf{d}_2)^{\mathrm{T}} = \mathbf{0}$. This, along with the fact that $\mathbf{S}$ is a stochastic matrix and $\mathbf{d}_1$ and $\mathbf{d}_2$ are deterministic, implies that

$$0 = \mathrm{cov}\{(\mathbf{h}_1 - \mathbf{h}_2)^{\mathrm{T}}\mathbf{S}^{\mathrm{T}}, (\mathbf{h}_1 - \mathbf{h}_2)^{\mathrm{T}}\mathbf{S}^{\mathrm{T}}\} = (\mathbf{h}_1 - \mathbf{h}_2)^{\mathrm{T}}\,\mathrm{cov}(\mathbf{S}^{\mathrm{T}}, \mathbf{S}^{\mathrm{T}})(\mathbf{h}_1 - \mathbf{h}_2)$$

From the assumption $\mathrm{cov}(\mathbf{S}^{\mathrm{T}}, \mathbf{S}^{\mathrm{T}}) > \mathbf{0}$, we directly deduce that $\mathbf{h}_1 = \mathbf{h}_2$. Consequently, (A1) yields $\mathbf{d}_1 = \mathbf{d}_2$. This completes the proof of Proposition A1. □

For model (3), we now check the validity of $\mathrm{cov}(\mathbf{S}^{\mathrm{T}}, \mathbf{S}^{\mathrm{T}}) > \mathbf{0}$ in Proposition A1. Let $\mathbf{s}_i$ denote the ith column vector of the matrix $\mathbf{S}^{\mathrm{T}}$. For $i = 1, \ldots, m-1$, since certain entries in $\mathbf{s}_i$ are always 0 (due to the fact that $\mathbf{S}$ is a Toeplitz matrix), we can conclude $\mathrm{cov}(\mathbf{s}_i, \mathbf{s}_i) \geqslant \mathbf{0}$ (but not necessarily $\mathrm{cov}(\mathbf{s}_i, \mathbf{s}_i) > \mathbf{0}$). In contrast, for $i = m, \ldots, n$, noting that the vector

$$\mathbf{s}_i = \begin{bmatrix} s(i-1) \\ \vdots \\ s(i-m) \end{bmatrix}$$

DOI: 10.1002/sim

does not contain entries which are always 0, we observe that

$$\mathrm{cov}(\mathbf{s}_i, \mathbf{s}_i) = \mathrm{diag}[\mathrm{var}\{s(i-1)\}, \ldots, \mathrm{var}\{s(i-m)\}] > \mathbf{0}$$

provided that $0 < P\{s(i) = 1\} < 1$, and hence $\mathrm{cov}(\mathbf{S}^{\mathrm{T}}, \mathbf{S}^{\mathrm{T}}) = \sum_{i=1}^{n} \mathrm{cov}(\mathbf{s}_i, \mathbf{s}_i) > \mathbf{0}$. The identifiability of (3) is thus verified.

APPENDIX B: SINGLE-VOXEL ANALYSIS OF fMRI UNDER (3)

Step 1: Obtain an initial estimate of $\mathbf{h}$ using a difference-based method. The difference-based estimate $\widehat{\mathbf{h}}_0$ is based on the lag-one difference of $\mathbf{y}$ and $\mathbf{S}$ simultaneously to remove the drift term, followed by an ordinary least-squares estimation of HRF.

Step 2: Estimate the noise variance and correlation matrix. Remove the estimated parametric component; apply the second-order difference to the residual $\mathbf{y} - \mathbf{S}\widehat{\mathbf{h}}_0$; obtain estimates $\widehat{R}$ of the correlation matrix and $\widehat{\sigma}^2$ of the variance using the method in Zhang *et al.* [16], which assumes the stationarity of the noise process and that the autocovariance sequence vanishes for time lag exceeding 2.

Step 3: Get an initial estimate $\widehat{\mathbf{d}}_0$ of the drift, by applying a cubic smoothing spline to $\mathbf{y} - \mathbf{S}\widehat{\mathbf{h}}_0$, i.e. $\widehat{\mathbf{d}}_0 = S_d(\mathbf{y} - \mathbf{S}\widehat{\mathbf{h}}_0)$. The smoothing parameter in S_d can be selected *via* generalized cross-validation and is searched over grid points which are logarithmically equally spaced between decades $10^{2.7 \times 10^{-7}}$ and $10^{1.1 \times 10^{-3}}$.

Step 4: Select the optimal data-driven smoothing parameter in S_d for the estimator $\widehat{\mathbf{h}}_{\mathrm{rough}}$. The optimal parameter minimizes the plug-in version of the mean-squared error of $\widehat{\mathbf{h}}_{\mathrm{rough}}$, where the unknown quantities, R, σ^2, and $\mathbf{d}$, are replaced by $\widehat{R}$, $\widehat{\sigma}^2$, and $\widehat{\mathbf{d}}_0$ obtained in Steps 2 and 3 above. The grid search is the same as in Step 3.

Step 5: Obtain the one-level estimate $\widehat{\mathbf{h}}_{\mathrm{rough}}$ of HRF, as in (8). Use the selected smoothing parameter in Step 4 and estimate the covariances.

Step 6: Obtain the two-level estimate $\widehat{\mathbf{h}}_{\mathrm{smooth}}$ of HRF, as in (9), where the smoothing parameter in S_h is selected based on the Mallow's C_p criterion. The smoothing parameter is searched over grid points, which are logarithmically equally spaced between decades $10^{4.7 \times 10^{-6}}$ and $10^{9.1 \times 10^{-4}}$.

Step 7: Calculate the semiparametric F-test statistics defined in (11) and (12) for each brain voxel.

Step 8: If needed, obtain one-level and two-level drift estimates $\widehat{\mathbf{d}}_{\mathrm{rough}}$ and $\widehat{\mathbf{d}}_{\mathrm{smooth}}$ as in Step 3, except that $\widehat{\mathbf{h}}_0$ is replaced by $\widehat{\mathbf{h}}_{\mathrm{rough}}$ and $\widehat{\mathbf{h}}_{\mathrm{smooth}}$, respectively.

ACKNOWLEDGEMENTS

We thank Richard Davidson, Tom Johnstone, and Terry Oakes at the Waisman Laboratory for Brain Imaging and Behavior, University of Wisconsin-Madison, for helpful comments and providing the fMRI data set. We are also grateful to Professor Emery N. Brown and an anonymous referee for their constructive comments and suggestions that greatly improved the presentation of this paper.

DOI: 10.1002/sim

REFERENCES

1. Lange N. Tutorial in biostatistics: statistical approaches to human brain mapping by functional magnetic resonance imaging. *Statistics in Medicine* 1996; **15**:389–428.
2. Lazar NA, Eddy WF, Genovese CR, Welling J. Statistical issues in fMRI for brain imaging. *International Statistical Review* 2001; **69**:105–127.
3. Worsley KJ, Liao CH, Aston J, Petre V, Duncan GH, Morales F, Evans AC. A general statistical analysis for fMRI data. *NeuroImage* 2002; **15**:1–15.
4. Worsley KJ, Friston KJ. Analysis of fMRI time-series revisited-again. *NeuroImage* 1995; **2**:173–181.
5. Friston *et al*. SPM course notes. http://www.fil.ion.ucl.ac.uk/spm/ (1997).
6. Cox RW. AFNI: software for analysis, visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research* 1996; **29**:162–173.
7. Smith S, Jenkinson M, Woolrich M, Beckmann CF, Behrens TEJ, Johansen-Berg H, Bannister PR, De Luca M, Drobnjak I, Flitney DE, Niazy RK, Saunders J, Vickers J, Zhang Y, De Stefano N, Brady JM, Matthews PM. Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage* 2004; **23**:208–219.
8. Woolrich MW, Ripley BD, Brady M, Smith SM. Temporal autocorrelation in univariate linear modelling of FMRI data. *NeuroImage* 2001; **14**:1370–1386.
9. Ward BD. Deconvolution analysis of fMRI time series data. *Technical Report*, Biophysics Research Institute, Medical College of Wisconsin, 2001.
10. Glover GH. Deconvolution of impulse response in event-related BOLD fMRI. *NeuroImage* 1999; **9**:416–429.
11. Green PJ, Silverman BW. *Nonparametric Regression and Generalized Linear Models*: *A Roughness Penalty Approach*. Chapman & Hall: London, 1994.
12. Zhang CM. Calibrating the degrees of freedom for automatic data smoothing and effective curve checking. *Journal of the American Statistical Association* 2003; **98**:609–628.
13. Anderson TW. Asymptotically efficient estimation of covariance matrices with linear structure. *Annals of Statistics* 1973; **1**:135–141.
14. Diggle PJ, Verbyla AP. Nonparametric estimation of covariance structure in longitudinal data. *Biometrics* 1998; **54**:401–415.
15. Boik R. Spectral models for covariance matrices. *Biometrika* 2002; **89**:159–182.
16. Zhang CM, Lu Y, Johnstone T, Davidson R. Nonparametric modeling and inference for event-related FMRI data (preliminary report). *Technical Report #1125*, Department of Statistics, University of Wisconsin-Madison, 2006.
17. Adler RJ. *The Geometry of Random Fields*. Wiley: New York, 1981.
18. Worsley KJ, Marrett S, Neelin P, Vandal AC, Friston KJ, Evans AC. A unified statistical approach for determining significant signals in images of cerebral activation. *Human Brain Mapping* 1996; **4**:58–73.
19. Fan J, Zhang CM, Zhang J. Generalized likelihood ratio statistics and Wilks phenomenon. *Annals of Statistics* 2001; **29**:153–193.
20. Purdon PL, Solo V, Weissko, RM, Brown E. Locally regularized spatiotemporal modeling and model comparison for functional MRI. *NeuroImage* 2001; **14**:912–923.

DOI: 10.1002/sim