ELSEVIER

Medical Image Analysis 6 (2002) 63–75

MEDICAL IMAGE ANALYSIS

www.elsevier.com/locate/media

Estimating the effective degrees of freedom in univariate multiple regression analysis

F. Kruggel[a,*], M. Pélégrini-Issac[b], H. Benali[c]

[a]*Max Planck Institute of Cognitive Neuroscience, Stephanstrasse 1, 04103 Leipzig, Germany*
[b]*Unité 483, INSERM, 9 Quai Saint-Bernard, 75005 Paris, France*
[c]*Unité 494, INSERM, CHU Pitié-Salpêtrière 91, Boulevard de l'Hôpital, 75634 Paris Cedex 13, France*

Received 18 January 2001; received in revised form 4 May 2001; accepted 1 October 2001

Abstract

The general linear model provides the most widely applied statistical framework for analyzing functional MRI (fMRI) data. With the increasing temporal resolution of recent scanning protocols, and more elaborate data preprocessing schemes, data independency is no longer a valid assumption. In this paper, we revise the statistical background of the general linear model in the presence of temporal autocorrelations. First, when detecting the activation signal, we explicitly account for the temporal autocorrelation structure, which yields a generalized F-test and the associated corrected (or effective) degrees of freedom (DOF). The proposed approach is data driven and thus independent of any specific preprocessing method. Then, for event-related protocols, we propose a new model for the temporal autocorrelations (''damped oscillator'' model) and compare this model to another, previously used in the field (first-order autoregressive model, or AR(1) model). In the case of long fMRI time series, an efficient approximation for the number of effective DOF is provided for both models. Finally, the validity of our approach is assessed using simulated and real fMRI data and is compared with more conventional methods.

Keywords: Functional MRI; Effective degrees of freedom; Multiple linear regression; Temporal autocorrelations

1. Introduction

In their seminal papers on analyzing fMRI time series data using multiple regression, Friston et al. (1995a) and Worsley and Friston (1995) elucidated the importance of taking autocorrelations in the statistical tests into account. They suggested modifying the number of degrees of freedom (DOF) associated with the statistical tests performed to detect significantly activated voxels. This correction was found to be necessary, because regressors were convolved with a Gaussian smoothing kernel in the temporal domain, thus correlating the data. The Gaussian kernel was matched to the shape of the hemodynamic response, and introduced in order to obtain a better signal-to-noise ratio.

In the formulation given by Worsley and Friston (1995), the autocorrelations of the filtered fMRI data are approximated by the properties of the matched Gaussian filter, i.e. properties of the unfiltered data are ignored. This approximation is valid if the autocorrelations in the unfiltered data are low, and filter properties are matched to the signal of interest. Since the temporal filter is uniformly applied in the spatial domain to all voxels in the fMRI data set, it follows that a *spatially uniform* assumption about autocorrelations is implied. With increasing knowledge of the signal properties of fMRI data, assumptions about spatially homogeneous hemodynamic responses and spatially homogeneous noise properties no longer hold. More efficient methods for separating the fMRI signal from physiological and system noise were proposed as a preprocessing step (Biswal et al., 1996; Buonocore and

*Corresponding author. Tel.: +49-341-9940-223; fax: +49-341-9940-221.
E-mail address: kruggel@cns.mpg.de (F. Kruggel).

1361-8415/02/$ – see front matter
PII: S1361-8415(01)00052-4

Maddock, 1997; Hu et al., 1995; Kruggel et al., 1999), which renders a second smoothing step unnecessary. However, this raises the issue of how to correct for the enhanced autocorrelations introduced by these filters in the subsequent signal detection step. It appears more attractive to take the properties of the (filtered) data into account directly in the statistical evaluation. Consequently, the number of DOF used for statistical inference must be computed voxelwise and adapted to the properties of the preprocessing filter.

In addition, there are computational issues to discuss. Today's MR scanner hardware make volume acquisition in subsecond intervals possible, and experimental designs of more than 2000 time steps are common. It is therefore desirable to design methods of statistical analysis for which the computation time is *not* a limiting factor to their routine use.

In this paper, we start by presenting a rather scholastic review of the statistical background of handling autocorrelations in the general linear model. While comprehensive treaties of this problem appeared early in the statistical literature (Cochrane and Orcutt, 1949; Durbin and Watson, 1950, 1951), formulations targeted to analyze fMRI data appear only scattered throughout specialized textbooks (Graybill, 1983; Johnson and Wichern, 1988; Rencher, 1995; Lange and Zeger, 1997; Moonen and Bandettini, 1999), and deserve a compact review. We derive a generalized F-test for detecting significantly activated areas and explicitly compute the number of effective DOF associated with this F-test. The test is independent from a smoothing matrix introduced in the design and thus allows the application of any suitable filter or signal restoration algorithm in a preprocessing step. However, it depends on the temporal autocorrelation structure of the data. Therefore, we propose a new autocorrelation model (''damped oscillator'' model) suitable for event-related studies and compare it with the more conventional AR(1) model. Optimizations and approximations are discussed which make a statistically valid handling of autocorrelations computationally tractable. We then validate the proposed autocorrelation model by studying semi-synthetic fMRI data filtered by different preprocessing algorithms. Finally, a realistic example demonstrates the consequences of this approach in the signal detection step of fMRI data analysis.

2. Theoretical background

2.1. General linear model

Univariate multiple regression represents one application of the general linear model (Rencher, 1995), which is a powerful tool for assessing relationships between the fMRI signal and experimental factors (Friston et al., 1995b; see Moonen and Bandettini, 1999, for a thorough discussion). Denote by $\boldsymbol{y}$ an N-vector of fMRI time series data (usually preprocessed data) and by $\boldsymbol{x}$ an (N,P) matrix assumed to have full rank. Each of the P columns of $\boldsymbol{x}$ is called a ''regressor'', which is either determined by the experimental design (''regressors of interest'') or represents confounds (''dummy regressors''). The general linear model relates observations with the model via a P-vector of regression coefficients $\boldsymbol{\beta}$,

$$\boldsymbol{y} = \boldsymbol{x}\boldsymbol{\beta} + \boldsymbol{\epsilon}, \tag{1}$$

where $\boldsymbol{\epsilon}$ corresponds to an N-vector of error terms. Let us assume that the errors are restricted to the observations $\boldsymbol{y}$, i.e. there are no errors in the regressors $\boldsymbol{x}$. The process generating $\boldsymbol{\epsilon}$ is assumed to be stationary in time. Errors are assumed to be normally distributed and correlated, following $\mathcal{N}(0,\sigma^2\boldsymbol{V})$, where σ^2 is a constant and $\boldsymbol{V}$ corresponds to the autocorrelation matrix. In the following discussion, the actual structure of $\boldsymbol{x}$ is not relevant: the stimulus function may have been convolved with a hemodynamic modulation function using pre-defined (Friston et al., 1995b) or data-driven parameters (Rajapakse et al., 1998).

Solving the general linear model consists of deciding whether or not $\boldsymbol{y}$ represents a brain activation signal by estimating the regression coefficients $\hat{\boldsymbol{\beta}}$ and determining, using a statistical test, whether or not they contribute significantly to predicting the signal $\boldsymbol{y}$.

2.2. Solving the general linear model using ordinary least squares

The most widely used approach for solving the general linear model consists of estimating the regression coefficients $\hat{\boldsymbol{\beta}}$ using ordinary least squares (OLS),

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x})^{-1}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{y}, \tag{2}$$

where $\boldsymbol{x}\hat{\boldsymbol{\beta}}$ represents the estimated signal and $\hat{\boldsymbol{\epsilon}} = \boldsymbol{y} - \boldsymbol{x}\hat{\boldsymbol{\beta}}$ is the residual, i.e. the error between the actual and the estimated time series.[1] The estimated regression coefficients $\hat{\boldsymbol{\beta}}$ are unbiased, but do not have minimum variance because the OLS approach assumes that errors are independently distributed.

A number Q of regressors may be tested for having a significant influence on predicting the time series $\boldsymbol{y}$ under the null hypothesis $H_0 : \hat{\boldsymbol{\beta}}_1 = 0, \ldots, \hat{\boldsymbol{\beta}}_Q = 0$. For doing so, a reduced model $\boldsymbol{y} = \boldsymbol{x}_{\mathrm{r}}\boldsymbol{\beta}_{\mathrm{r}} + \boldsymbol{\epsilon}_{\mathrm{r}}$ has to be solved, where $\boldsymbol{x}_{\mathrm{r}}$ represents $\boldsymbol{x}$ in which the Q studied factors have been removed. The $\hat{\boldsymbol{\beta}}_{\mathrm{r}}$ are estimated using OLS and the following sums of squares are computed:

- The sum of squares due to the residuals $\mathrm{SSR} = \hat{\boldsymbol{\epsilon}}^{\mathrm{T}}\hat{\boldsymbol{\epsilon}} = \boldsymbol{y}^{\mathrm{T}}\boldsymbol{y} - \hat{\boldsymbol{\beta}}^{\mathrm{T}}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{y}$, with an associated number of effective DOF ν_{SSR}.
- The sum of squares due to the model $\mathrm{SSH} = \hat{\boldsymbol{\epsilon}}_{\mathrm{r}}^{\mathrm{T}}\hat{\boldsymbol{\epsilon}}_{\mathrm{r}} -$

[1] If $\boldsymbol{x}$ does not have full rank, then $\hat{\boldsymbol{\beta}} = (\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x})^{-}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{y}$, where ''$^{-}$'' denotes the generalized inverse.

$\hat{\boldsymbol{\epsilon}}^{\mathrm{T}}\hat{\boldsymbol{\epsilon}} = \hat{\boldsymbol{\beta}}^{\mathrm{T}}\boldsymbol{x}^{\mathrm{T}}\boldsymbol{y} - \hat{\boldsymbol{\beta}}_{\mathrm{r}}^{\mathrm{T}}\boldsymbol{x}_{\mathrm{r}}^{\mathrm{T}}\boldsymbol{y}$, with an associated number of effective DOF ν_{SSH}.

2.3. F-test in the case of uncorrelated errors

If the errors are not correlated ($\boldsymbol{V} = \boldsymbol{I}_N$), SSH and SSR follow a χ^2 distribution with $\nu_{\mathrm{SSH}} = Q$ and $\nu_{\mathrm{SSR}} = N - P$ degrees of freedom, respectively, and the null hypothesis can be tested using the following Fisher's F-value:

$$F = \frac{(N-P)}{Q}\frac{\mathrm{SSH}}{\mathrm{SSR}}. \tag{3}$$

F then follows $F_{\alpha,Q,N-P}$, where α is the type I error, with Q and $N-P$ degrees of freedom.

2.4. Generalized F-test in the case of correlated errors

If the errors are correlated, the problem consists of assessing the distribution of the sums of squares SSR (Section 2.4.1) and SSH (Section 2.4.2), estimating the number of effective DOF ν_{SSR} and ν_{SSH} and deriving an appropriate way of testing the null hypothesis (Section 2.4.3).

2.4.1. Distribution of SSR

Let $\boldsymbol{R} = \boldsymbol{I}_N - \boldsymbol{x}(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x})^{-1}\boldsymbol{x}^{\mathrm{T}}$ be the projector on the plane perpendicular to the plane spanned by the P columns of $\boldsymbol{x}$. It can be shown that

$$\mathrm{SSR} = \hat{\boldsymbol{\epsilon}}^{\mathrm{T}}\hat{\boldsymbol{\epsilon}} = \boldsymbol{\epsilon}^{\mathrm{T}}\boldsymbol{R}\boldsymbol{\epsilon} = \boldsymbol{\epsilon}^{\mathrm{T}}\boldsymbol{R}\boldsymbol{R}\boldsymbol{\epsilon} = \mathrm{trace}(\boldsymbol{R}\boldsymbol{\epsilon}\boldsymbol{\epsilon}^{\mathrm{T}}\boldsymbol{R})$$
$$= \sum_{i=1}^{N}(\boldsymbol{R}\boldsymbol{\epsilon}\boldsymbol{\epsilon}^{\mathrm{T}}\boldsymbol{R})_i,$$

where $(\boldsymbol{R}\boldsymbol{\epsilon}\boldsymbol{\epsilon}^{\mathrm{T}}\boldsymbol{R})_i$ is the ith diagonal element of $\boldsymbol{R}\boldsymbol{\epsilon}\boldsymbol{\epsilon}^{\mathrm{T}}\boldsymbol{R}$. Since $\boldsymbol{\epsilon}$ is an $(N,1)$ matrix from $\mathcal{N}_N(0,\sigma^2\boldsymbol{V})$, then $(\boldsymbol{R}\boldsymbol{\epsilon}\boldsymbol{\epsilon}^{\mathrm{T}}\boldsymbol{R})_i$ follows a χ_1^2 distribution (Mardia et al., 1979, p. 67). Therefore, SSR follows a linear combination of χ_1^2 distributions, which can in turn be approximated by a Gamma distribution $G(\delta_{\mathrm{SSR}},\rho_{\mathrm{SSR}})$ (the so-called ''Satterthwaite–Welsh'' approximation (Satterthwaite, 1946; Wood, 1989)) with parameters

$$\delta_{\mathrm{SSR}} = \frac{[\mathrm{trace}(\boldsymbol{R}\boldsymbol{V})]^2}{2\,\mathrm{trace}(\boldsymbol{R}\boldsymbol{V}\boldsymbol{R}\boldsymbol{V})} \quad \text{and} \quad \rho_{\mathrm{SSR}} = \frac{\mathrm{trace}(\boldsymbol{R}\boldsymbol{V})}{2\,\sigma^2\,\mathrm{trace}(\boldsymbol{R}\boldsymbol{V}\boldsymbol{R}\boldsymbol{V})}.$$

2.4.2. Distribution of SSH

Let $\boldsymbol{R}_{\mathrm{r}}$ be the projector on the plane perpendicular to the plane spanned by the $P-Q$ columns of $\boldsymbol{x}_{\mathrm{r}}$. It can be shown that $\boldsymbol{R}_{\mathrm{s}} = \boldsymbol{R}_{\mathrm{r}} - \boldsymbol{R}$ is also a projector (on the plane spanned by the Q columns of $\boldsymbol{x}$ one wishes to test). It can also be shown that $\mathrm{SSH} = \boldsymbol{y}^{\mathrm{T}}\boldsymbol{R}_{\mathrm{s}}\boldsymbol{y}$ is equal to $\boldsymbol{\epsilon}^{\mathrm{T}}\boldsymbol{R}_{\mathrm{s}}\boldsymbol{\epsilon}$ under the null hypothesis. Hence, in analogy to the previous section, SSH follows a Gamma distribution $G(\delta_{\mathrm{SSH}},\rho_{\mathrm{SSH}})$ with parameters

$$\delta_{\mathrm{SSH}} = \frac{[\mathrm{trace}(\boldsymbol{R}_{\mathrm{s}}\boldsymbol{V})]^2}{2\,\mathrm{trace}(\boldsymbol{R}_{\mathrm{s}}\boldsymbol{V}\boldsymbol{R}_{\mathrm{s}}\boldsymbol{V})}$$

and

$$\rho_{\mathrm{SSH}} = \frac{\mathrm{trace}(\boldsymbol{R}_{\mathrm{s}}\boldsymbol{V})}{2\,\sigma^2\,\mathrm{trace}(\boldsymbol{R}_{\mathrm{s}}\boldsymbol{V}\boldsymbol{R}_{\mathrm{s}}\boldsymbol{V})}.$$

2.4.3. Test for the null hypothesis

Since $\boldsymbol{R}$ and $\boldsymbol{R}_{\mathrm{s}}$ are orthogonal, SSH and SSR are orthogonal. Since they both follow a Gamma distribution, the null distribution of the ratio

$$\frac{\delta_{\mathrm{SSR}}\rho_{\mathrm{SSR}}}{\delta_{\mathrm{SSH}}\rho_{\mathrm{SSH}}}\frac{\mathrm{SSH}}{\mathrm{SSR}}$$

is well approximated (Worsley and Friston, 1997) by $F_{\alpha,2\delta_{\mathrm{SSH}},2\delta_{\mathrm{SSR}}}$ (Johnson et al., 1995, p. 348) (the so-called generalized F-distribution). The quantity $\delta_{\mathrm{SSR}}\rho_{\mathrm{SSR}}/\delta_{\mathrm{SSH}}\rho_{\mathrm{SSH}}$ is the weighting factor of the generalized F-distribution. Using the results from Sections 2.4.1 and 2.4.2, the test quantity

$$F = \frac{\nu_{\mathrm{SSR}}^2}{\nu_{\mathrm{SSH}}^2}\frac{\mathrm{trace}(\boldsymbol{R}_{\mathrm{s}}\boldsymbol{V})}{\mathrm{trace}(\boldsymbol{R}\boldsymbol{V})}\frac{\mathrm{SSH}}{\mathrm{SSR}} \tag{4}$$

follows an $F_{\alpha,\nu_{\mathrm{SSH}},\nu_{\mathrm{SSR}}}$ distribution with the following number of effective DOF:[2]

$$\nu_{\mathrm{SSH}} = \frac{[\mathrm{trace}(\boldsymbol{R}_{\mathrm{s}}\boldsymbol{V})]^2}{\mathrm{trace}(\boldsymbol{R}_{\mathrm{s}}\boldsymbol{V}\boldsymbol{R}_{\mathrm{s}}\boldsymbol{V})} \quad \text{and} \quad \nu_{\mathrm{SSR}} = \frac{[\mathrm{trace}(\boldsymbol{R}\boldsymbol{V})]^2}{\mathrm{trace}(\boldsymbol{R}\boldsymbol{V}\boldsymbol{R}\boldsymbol{V})}. \tag{5}$$

The expression for ν_{SSR} corresponds to the number of effective DOF given by Worsley and Friston (1995).

2.4.4. Computational issues

Since fMRI time series typically consist of several hundred time steps, the F-test according to Eq. (4) involves voxelwise products of large matrices. For general autocorrelation matrices $\boldsymbol{V}$, this is very costly computationally. Therefore, it is worthwhile deriving efficient computation schemes for ν_{SSR} and ν_{SSH}.

When N becomes large, it can be shown that

$$\nu_{\mathrm{SSR}} \text{ tends to } \nu_{\mathrm{app}} = \frac{N(N-P)}{\mathrm{trace}(\boldsymbol{V}\boldsymbol{V})}. \tag{6}$$

The reader is referred to Appendix A for a complete proof.

Unfortunately, since $\boldsymbol{R}_{\mathrm{s}}$ is a projector on a subspace with *finite* dimension Q, there is no similar approximation of ν_{SSH} for large N. The idea, therefore, is to compute ν_{SSH} efficiently by avoiding the computation of the products $\mathrm{trace}(\boldsymbol{R}_{\mathrm{s}}\boldsymbol{V})$ and $\mathrm{trace}(\boldsymbol{R}_{\mathrm{s}}\boldsymbol{V}\boldsymbol{R}_{\mathrm{s}}\boldsymbol{V})$. It is easy to show that

$$\nu_{\mathrm{SSH}} = \frac{\left[\sum_{i=1}^{Q}\boldsymbol{u}^{i\mathrm{T}}\boldsymbol{V}\boldsymbol{u}^{i}\right]^2}{\sum_{i=1}^{Q}[\boldsymbol{u}^{i\mathrm{T}}\boldsymbol{V}\boldsymbol{u}^{i}]^2}, \tag{7}$$

where $\boldsymbol{u}^i$ corresponds to the Q eigenvectors of $\boldsymbol{R}_{\mathrm{s}}$, which

[2] It is straightforward to check that, in the case of uncorrelated errors ($\boldsymbol{V} = \boldsymbol{I}_N$), ν_{SSH} reduces to $\mathrm{trace}(\boldsymbol{R}_{\mathrm{s}}) = Q$ and ν_{SSR} reduces to $\mathrm{trace}(\boldsymbol{R}) = N - P$. Thus, Eq. (4) corresponds to the simple test from Eq. (3).

can easily be obtained using the singular value decomposition of $\boldsymbol{x}_Q$ (the matrix consisting of the Q regressors of $\boldsymbol{x}$ one wishes to include in the F-test). This implementation is efficient in that it involves only Q vector-by-matrix and vector-by-vector products.

In summary, Eq. (4) can be approximated for large N as

$$F_{\text{app}} = \frac{\nu_{\text{app}}^2}{\nu_{\text{SSH}}^2} \frac{\text{trace}(\boldsymbol{R}_s \boldsymbol{V})}{N-P} \frac{\text{SSH}}{\text{SSR}}, \tag{8}$$

where ν_{app} is computed by Eq. (6), and ν_{SSH} using Eq. (7).

2.5. Estimation of the autocorrelation matrix $\boldsymbol{V}$

2.5.1. Models for the autocorrelation structure

To compute the F-test according to Eqs. (4) and (5), the autocorrelation matrix $\boldsymbol{V}$ has to be known or estimated. It is impossible to determine this (N,N) matrix from a single time series of N steps without introducing a model structure. Most typically, such a model is introduced by adapting an autocorrelation function $\rho(\cdot)$ to the data, which must be non-negative definite (Priestley, 1996, p. 110). This is most easily achieved by choosing a function from a parametric family whose members are known to be at least non-negative definite.

Because stationarity in the temporal domain is assumed, the autocorrelation function depends only on the ''lag'' k between two time points. The entries $\hat{v}_{ij}$ of the estimated autocorrelation matrix $\hat{\boldsymbol{V}}$ are then defined by

$$\begin{cases} \hat{v}_{ii} = \rho(0) = 1, \\ \hat{v}_{ij} = \rho(|i-j|) = \rho(|k|), \quad \text{otherwise.} \end{cases} \tag{9}$$

In event-related fMRI experiments, some (quasi-)periodicity is expected to show up in the residuals (e.g., due to breathing or heart beat if these physiological phenomena are not pre-filtered (Petersen et al., 1998)). This quasi-periodicity in the residuals is also present in the autocorrelation function (see, for instance, Fig. 2). In this case, we propose a new model for $\rho(k)$, which we term the ''damped oscillator'' model,

$$\rho(k) = \exp(a_1 k)\cos(a_2 k), \tag{10}$$

where $a_1 < 0$. This model is fitted in practice as follows:

$$\begin{cases} \rho(0) = 1, \\ \rho(k) = a_0^2 - \exp(a_1 k)\cos(a_2 k), \quad \text{where } a_0^2 < 1,\ a_1 < 0. \end{cases}$$

The parameter a_0^2 reflects the so-called ''nugget effect'' (Cressie, 1993, p. 59), i.e. it accounts for the fact that, due to measurement errors, $\rho(\cdot)$ may not be continuous at the origin in practice. Using (Christakos, 1984), we have shown that ρ is non-negative definite. This condition is necessary and sufficient to guarantee that this autocorrelation function yields a valid autocorrelation matrix $\boldsymbol{V}$. We refer the reader to the proof given in Appendix B.

One interesting particular case of the ''damped oscillator'' model, obtained for $a_2 = 0$, corresponds to the conventional AR(1) process, which has already been used to model temporal autocorrelations in fMRI time series. For instance, an AR(1) autocorrelation in the (spatially averaged) residuals of fMRI data from blocked experiments was first demonstrated by Bullmore et al. (1996).

To sum up, $\boldsymbol{V}$ is estimated as follows. First, temporal autocorrelations present in the data are estimated by the sample autocorrelation function (ACF),

$$\hat{\rho}(k) = \frac{\sum_{i=1}^{N-k} (\hat{\epsilon}_i - \bar{\epsilon})(\hat{\epsilon}_{i+k} - \bar{\epsilon})}{\sum_{i=1}^{N-1} (\hat{\epsilon}_i - \bar{\epsilon})^2}, \tag{11}$$

where $\bar{\epsilon}$ is the sample mean of the residuals. Then, the model function (given by Eq. (10)) is fitted to the sample ACF using any nonlinear minimization algorithm. Powell's algorithm is a suitable choice here (Press et al., 1992, p. 309). Finally, using fitted parameters of the model function, the autocorrelation matrix $\hat{\boldsymbol{V}}$ is computed from Eq. (9).

2.5.2. Computational issues

Once $\hat{\boldsymbol{V}}$ is known, the F-test has to be computed. In some cases, for example if $\hat{\boldsymbol{V}}$ is a symmetric Toeplitz matrix, F may be calculated directly using Eqs. (4) and (5). Indeed, $\hat{\boldsymbol{V}}$ can be represented by a vector and convolution can be used to compute $\boldsymbol{R}\hat{\boldsymbol{V}}$ (Press et al., 1992, p. 92). Note that the model proposed in Eq. (9) yields such a symmetric Toeplitz matrix.

However, $\hat{\boldsymbol{V}}$ may not be a Toeplitz matrix, for example if the process is temporally non-stationary ($\hat{v}_{ij}$ does not depend only on the lag k). Computing ν_{SSR} using Eq. (5) then becomes computationally costly on a voxelwise basis. In this case, Eq. (6) can be used, but it may also be possible to derive a closed form for ν_{app} using the analytical expression for $\hat{v}_{ij}$ and series expansion. This further simplifies the computation of Eq. (6) by avoiding the calculation of trace($\hat{\boldsymbol{V}}\hat{\boldsymbol{V}}$).

To give an example of how such an analytical expression can be obtained, we derived an approximation of trace($\hat{\boldsymbol{V}}\hat{\boldsymbol{V}}$) using the expression for $\hat{v}_{ij}$ given by Eq. (9). When $\boldsymbol{V}$ is modeled using the ''damped oscillator'' model, the approximation is given by

$$\nu_{\text{app}}^{*} = (N-P)\left[\frac{1-\exp(2a_1)}{1+\exp(2a_1)}\right] \times \left[\frac{1+\exp(4a_1)-2\exp(2a_1)\cos(2a_2)}{1+\exp(4a_1)-\exp(2a_1)(1+\cos(2a_2))}\right]. \tag{12}$$

The reader is referred to Appendix C for a detailed derivation of Eq. (12).

The approximation corresponding to the particular case of the AR(1) model is derived straightforwardly from Eq. (12) with $a_2 = 0$,

$$\nu^{*}_{\mathrm{app}} = (N-P)\left[\frac{1-\exp(2a_1)}{1+\exp(2a_1)}\right], \quad (13)$$

where $\exp(2a_1) = \rho(1)^2$ is the squared first-order autocorrelation.

In either case, the approximation for the F-test is given by Eq. (8).

2.6. Algorithm

To sum up, the simplest method to handle autocorrelations in multiple regression is as follows:

1. Estimate the regression coefficients $\hat{\boldsymbol{\beta}}$ using Eq. (2).
2. Compute the residuals $\hat{\boldsymbol{\epsilon}} = \boldsymbol{y} - \boldsymbol{x}\hat{\boldsymbol{\beta}}$.
3. Estimate the sample autocorrelation function in each voxel using Eq. (11).
4. Estimate the parameters of the ''damped oscillator'' model.
5. Determine the number of effective DOF ν^{*}_{app} using Eq. (12).
6. Compute ν_{SSH} using Eq. (7).
7. Compute the F-value using Eq. (8), with ν^{*}_{app} and ν_{SSH} degrees of freedom.

If a different model is selected for $\boldsymbol{V}$, either derive a closed form for it, or replace Step 5 by

5′. Determine the number of effective DOF ν_{app} using Eq. (6).

If N/P is not large enough (typically, $N/P > 50$), it is better to compute $\boldsymbol{R}$ and to resort to the general form for ν_{SSR} as given in Eq. (5). A scheme similar to that described for ν_{SSH} (Section 2.4.4) can be applied in this case for an efficient determination of ν_{SSR}. Alternatively, fast computation using results based on the Hadamard product or Frobenius norm (Golub and Van Loan, 1984, p. 14) may be implemented.

3. Experiments

3.1. Simulations

We are now interested in the performance of the proposed method in relation to the preprocessing methods typically applied for fMRI data. For reasons of simplicity, we restrict ourselves to the case of $Q = 1$, thus $\nu_{\mathrm{SSH}} = 1$. The aims of these simulations were:

1. to study the fit of the ACF to the ''damped oscillator'' model proposed by Eq. (10) for fMRI data preprocessed with different techniques, and to compare this fit with that obtained using the simpler AR(1) model;
2. to compare the weighting factors of the generalized F-test (Eq. (4)), defined as $wf = \nu^2_{\mathrm{SSR}}[\mathrm{trace}(\boldsymbol{R}_s\boldsymbol{V})/\mathrm{trace}(\boldsymbol{RV})]$, obtained using both models (using the general form of Eq. (5) for ν_{SSR});
3. to test the approximations for ν_{SSR} given by the closed forms of Eqs. (12) and (13).

3.1.1. Test bed

For simulation experiments, we resorted to the methods described in detail in conjunction with our fMRI filter test bed (Kruggel et al., 1999). We created simulated fMRI data based on an event-related experiment in language comprehension (TR 2 s, 12 time steps per trial). This experiment was run using an EPI protocol on a Bruker 3 T Medspec 100 system (128×64 voxels, resolution 1.9×3.8 mm, 5 mm slice thickness, 2 mm gap). Three data sets were selected. A prototypical hemodynamic response (HR) was obtained by averaging the responses of a highly activated sample region in space and time (see example waveform in Fig. 1). From each data set, three different patches were chosen from a brain region where no activation was determined using a standard statistical procedure (linear regression using the prototypical HR as a regressor, transformation of F-values into z-scores, determination of the significance of an activated cluster by its spatial extent for $z > 3$ (Friston et al., 1994)). Each of the 9 patches consisted of 10×10 voxels and 500 time steps and contained random proportions of voxels representing grey matter, white matter and cerebrospinal fluid. The prototypical HR was modulated onto the test patches as a single 4-voxel activation (Fig. 1) at a signal-to-noise ratio of 0.20.

In all test patches, the baseline was estimated by low-pass filtering in the temporal domain and subsequently corrected by subtraction of the estimated baseline from the data. As examples of typical preprocessing procedures, patches were either (1) unprocessed (UP) or filtered by (2) low-pass filtering in the temporal domain (LP, cut-off frequency 0.38 times a trial length in time steps), (3) Gaussian filtering in the temporal domain (GT, $\sigma = 2.8$ s), or (4) spatio-temporal signal restoration by a Markov Random Field (Descombes et al., 1998) (MRF, $\beta = 1.0$, $\delta = 1600$). Filter parameters were optimized for a given experiment as determined by the test bed (Kruggel et al., 1999). A total of 36 preprocessed test patches were studied by the proposed algorithm. The design matrix used for the multiple regression analysis was comprised of three regressors: one corresponded to the prototypical HR, the others were used to correct for trends and means.

3.1.2. Results

The autocorrelation structure was studied thoroughly for a large set of voxels by inspecting the sample autocorrelation functions computed according to Eq. (11). Examples from the same sample voxel are compiled in Fig. 2. From these plots it is clear that the autocorrelation structure depends on the preprocessing operation applied. Autocorrelations in the residuals of both LP- and GT-filtered data (Fig. 2(a,b)) follow a periodic model function according to Eq. (10), where the period length is close to

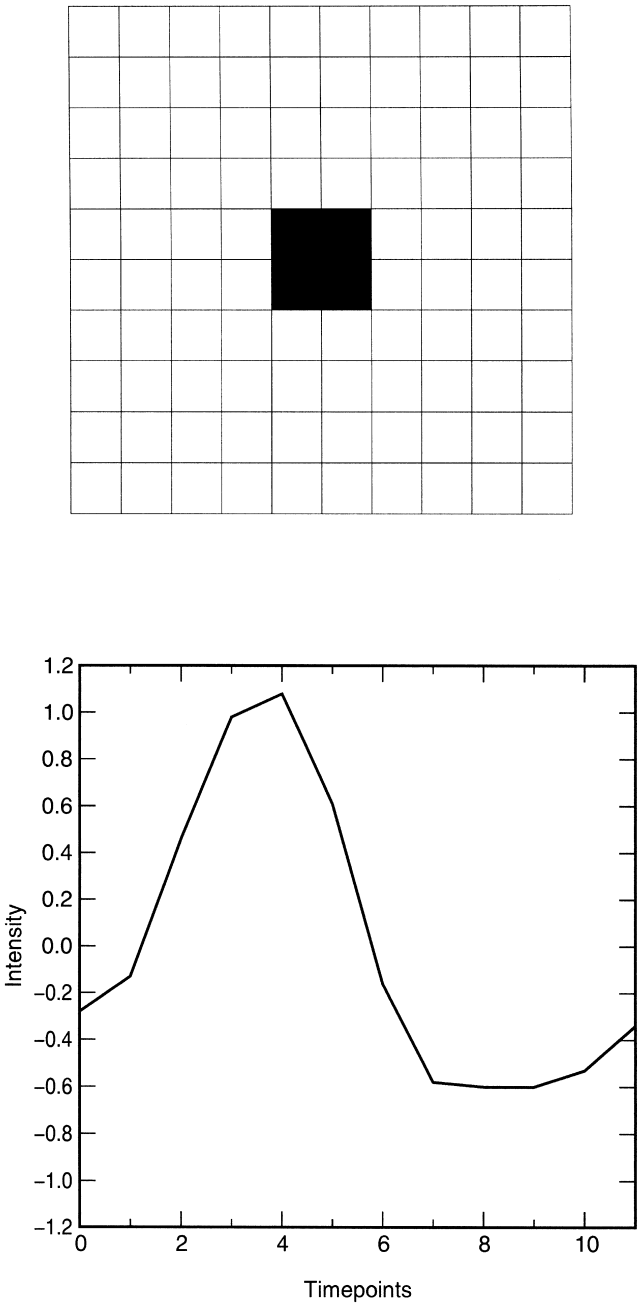

Fig. 1. Scheme of the spatial pattern and the prototypical hemodynamic response used in the simulation experiments. Black pixels correspond to a modulation with the test signal at a signal-to-noise ratio of 0.2, white pixels are untouched.

the trial length of the experiment ($a_2 = 12.4$, for 12 time steps per trial in this experiment). In unfiltered data (Fig. 2(c)), (quasi-)periodic signals unmodeled by the experimental design show up. In the example displayed here, the periodicity was most likely attributable to breathing ($a_2 = 2.87$, corresponding to 10/min). In the case of the MRF filter (Fig. 2(d)), more complex autocorrelation structures dominate in the residuals.

In all cases studied, the ''damped oscillator'' autocorrelation function (Eq. (10)) fitted reasonably well to the sample ACF, and much better in comparison with the AR(1) model.

The normalized ratio $wf/(N-P)$ (computed according to Eq. (5) for $Q = 1$ and $\nu_{SSH} = 1$) was compared between the ''damped oscillator''model and the AR(1) model. As can be seen in Fig. 3, all points lie above the main diagonal, therefore wf was severely overestimated when using an AR(1) model. It is evident that the weighting factor depends on the preprocessing technique and is spatially not uniform. In the background voxels, the ratio $wf/(N-P)$ varied for unfiltered data between 0.9 and 1.0, for GT-filtered data between 0.1 and 0.6, for LP-filtered data between 0.05 and 0.4, and for MRF-filtered data between 0.2 and 0.7. In the foreground voxels, this ratio was typically higher by 0.1–0.2. These variations underline the importance of a voxelwise computation of the DOF. Note that, for the MRF model, autocorrelations were not captured by the AR(1) model.

Finally, we compared ν_{SSR} as computed for the ''damped oscillator'' model by the general form (Eq. (5)) with the closed form (Eq. (12)), and for the AR(1) model (Eq. (5) versus Eq. (13)). Results for all test patches are compiled in Fig. 4, and demonstrate the validity of the approximations for large N for both autocorrelation models.

3.2. A real example

As a final example, we re-evaluate an event-related fMRI experiment in language comprehension, from which we select a single data set. Using a single shot EPI protocol, we acquired on a Bruker 3 T Medspec 100 system, every 2 s, 4 T_2^* slices of 128×64 voxels with a spatial resolution of $1.9 \times 3.8 \times 5$ mm and 2 mm gap. The auditory presentation of a sentence needed approximately 6 s (3 time steps), followed by a pause of 18 s (9 time steps). 76 trials (912 time steps) were recorded during an experiment lasting approximately 30 min. Right-handed test persons were asked to classify a sentence for syntactical correctness and to respond by pressing a button with the right hand.

Standard procedures for the detection of functional activity in this dataset were applied as follows: *signal preprocessing* using motion correction, estimation of the baseline by voxelwise low-pass filtering in the temporal domain using a cut-off frequency of 1.5 times the stimulation frequency and baseline correction by subtracting the estimated baseline from the motion-corrected data, reduction of physiological and system noise by voxelwise Gaussian filtering in the temporal domain ($\sigma = 2.8$ s), *signal detection* by computing a linear regression with a box-car waveform corresponding to the auditory stimulation, which was convolved with a Gaussian function (temporal shift 6 s, $\sigma = 2.8$ s), applying one of the autocorrelation models (see below) to derive the number of effective DOF, conversion of the F-values into z-scores,

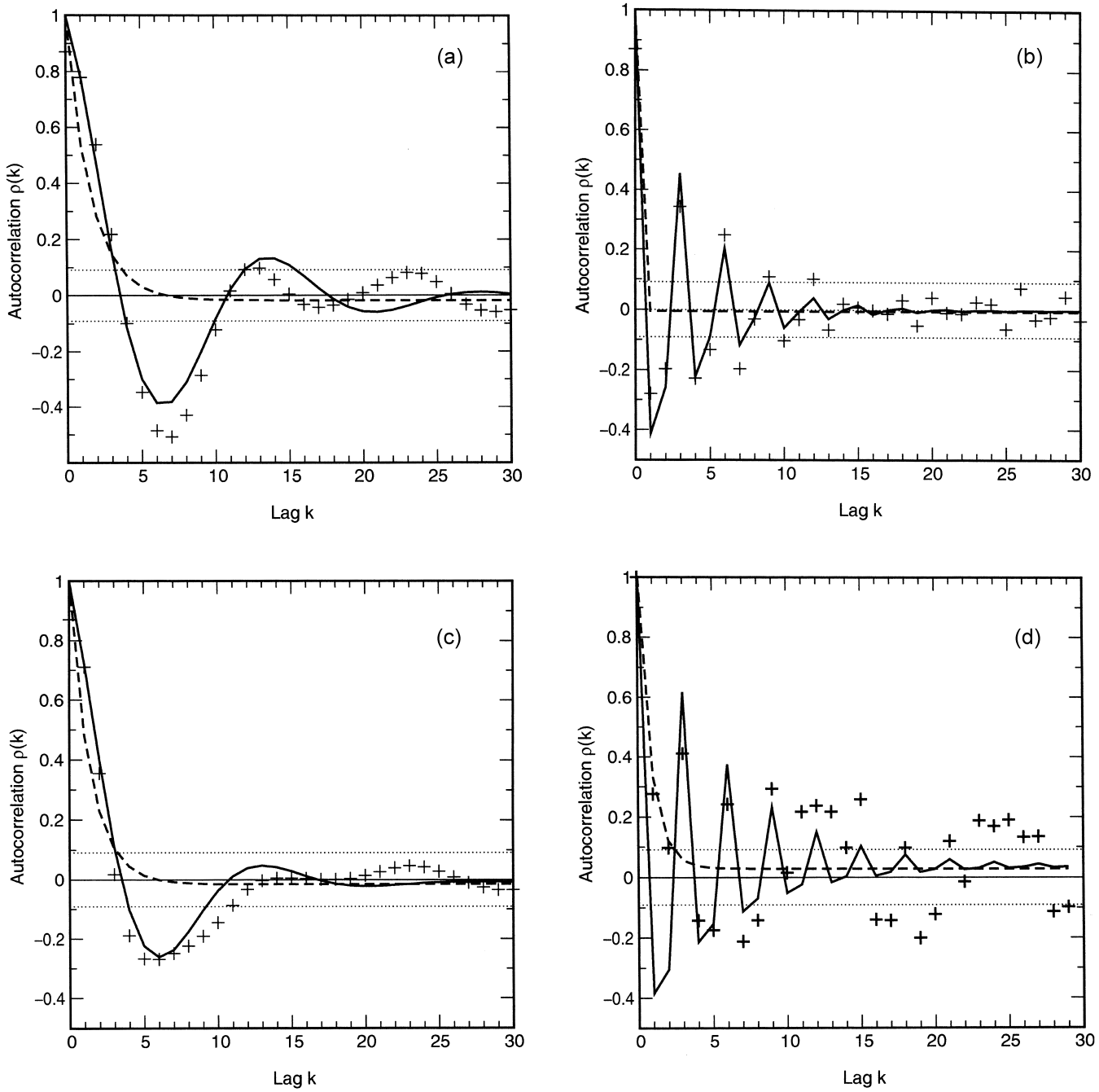

Fig. 2. (a,b) Estimated autocorrelation coefficients (+, Eq. (11)), fitted autocorrelation functions (AR(1) model: - - -, "damped oscillator" model: ———) for a sample background voxel of an LP- (top) and a GT-filtered (bottom) data set. Bartlett's significance interval ($\pm 2/\sqrt{N}$) is shown as a dotted line. (c,d) Estimated autocorrelation coefficients (+, Eq. (11)), fitted autocorrelation functions (AR(1) model: - - -, "damped oscillator" model: ———) for a sample background voxel of an unprocessed (top) and MRF-filtered (bottom) data set. Bartlett's significance interval ($\pm 2/\sqrt{N}$) is shown as a dotted line.

and thresholding of the corresponding z-maps by a score of 3, and, finally, assessment of the activated regions for their significance on the basis of their spatial extent (Friston et al., 1994). The resulting z-score maps were overlaid onto their corresponding anatomical slices and color coded using a hot-iron scale ranging from $z = 8$ (red) to $z = 16$ (white). In the top row of Fig. 5, no correction for temporal autocorrelations was applied ($\nu = N - P = 909$). In the middle row, a spatially stationary correction approach, as implemented in SPM99 (FIL Methods Group, 2000), was used: the autocorrelation matrix $\hat{\boldsymbol{V}}$ was computed using the characteristics of the prototypical HR and the F-test value was computed based on the assumption that both SSR and SSH follow a χ^2 distribution and are divided by their expectation (trace($\boldsymbol{RV}$) and trace($\boldsymbol{R}_s\boldsymbol{V}$), respectively). This yielded the following approximation for the F-value:

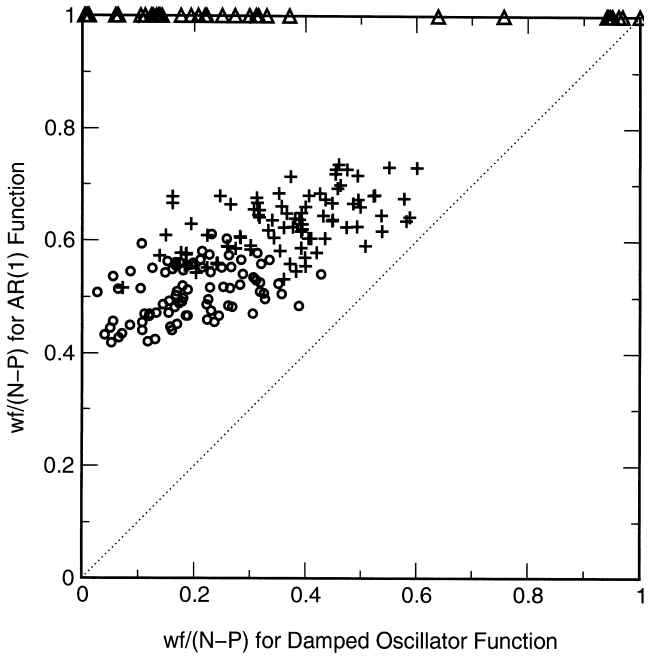

Fig. 3. Comparison of weighting factors *wf*, as computed for the ''damped oscillator'' model and the AR(1) model, for voxels from patches treated by different filtering methods (LP (○), GT (+), MRF (△)).

$$F_{\alpha,\nu_{\mathrm{SSH}},\nu_{\mathrm{SSR}}} = \frac{\mathrm{trace}(\boldsymbol{RV})}{\mathrm{trace}(\boldsymbol{R}_s\boldsymbol{V})}\frac{\mathrm{SSH}}{\mathrm{SSR}}. \tag{14}$$

In the bottom row, a voxelwise correction was used, based on the ''damped oscillator'' model proposed here.

Correcting for the number of effective DOF led to a general reduction of *z*-scores by 5–8 units (see rows A and B), and rendered some of the spurious activations insignificant. Introducing the ''damped oscillator'' model (row C) and using a spatially non-stationary correction led to a much better delineated detection of functional activation with somewhat higher *z*-scores in areas expected to be related to the experiment (i.e., the superior temporal gyrus on both sides, the thalami), and less spurious activations (e.g., in overprojection to the putamen, the inter-hemispheric cleft, and the anterior tips of the temporal lobe). This correlates nicely with the finding in the simulation experiments, where the weighting factor was higher in focal activation.

4. Discussion

We review how the statistical inference may be corrected for temporal autocorrelations present in fMRI time series. Our approach is similar to the method described by Worsley and Friston (1995) and that implemented in SPM99, in that we modify the DOF associated with the statistical test used for detecting significantly activated areas. Obviously, if the number of effective DOF is large

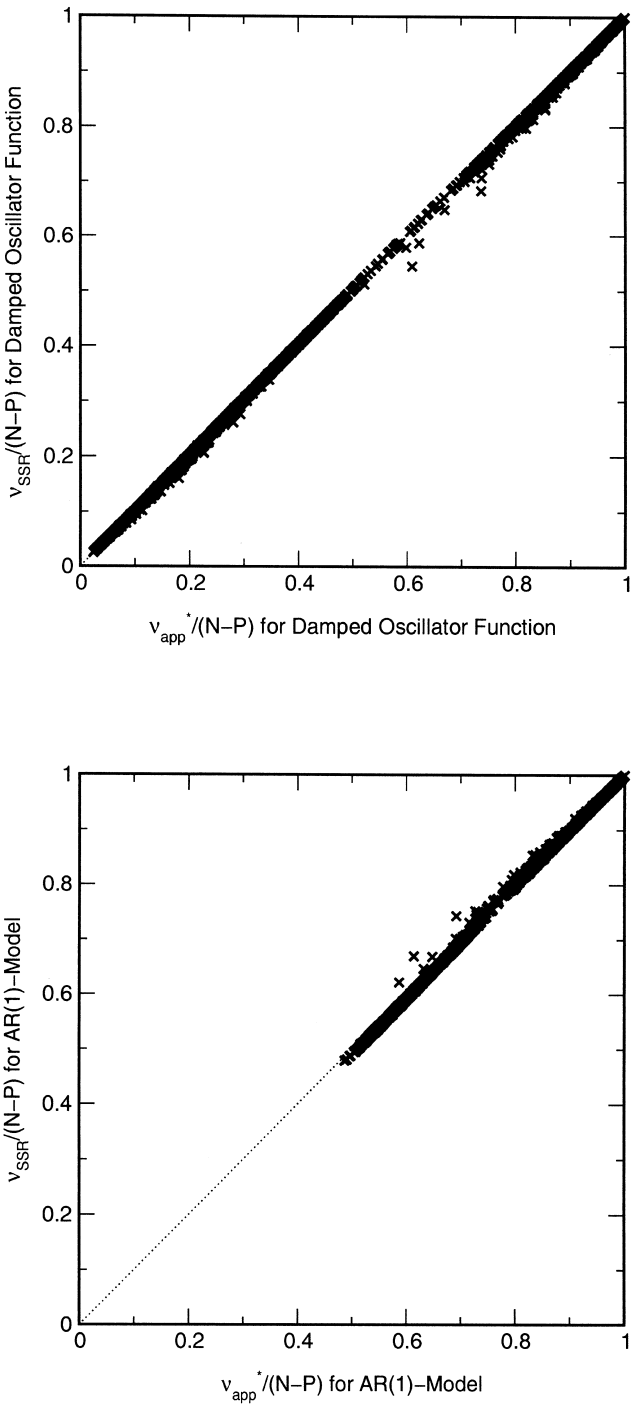

Fig. 4. (Top) Comparison of the number of effective DOF, computed on the basis of the ''damped oscillator'' model by the general form (ν_{SSR}, Eq. (5)) and its approximation (ν^{*}_{app}, Eq. (12)) ($y = 1.00x - 1.10$, $r^2 = 0.9999$). (Bottom) Same, but for the AR(1) model ($y = 1.02x - 10.1$, $r^2 = 0.9999$).

enough (say, >50), then approximate distributions can be used (e.g., normal or χ^2) and ν_{SSR} set to infinity. However, it is impossible to predict the number of effective DOF from N (the number of time steps): large N does not imply large ν_{SSR}. Indeed, if N is very large, but the data are highly correlated, the number of effective DOF may be

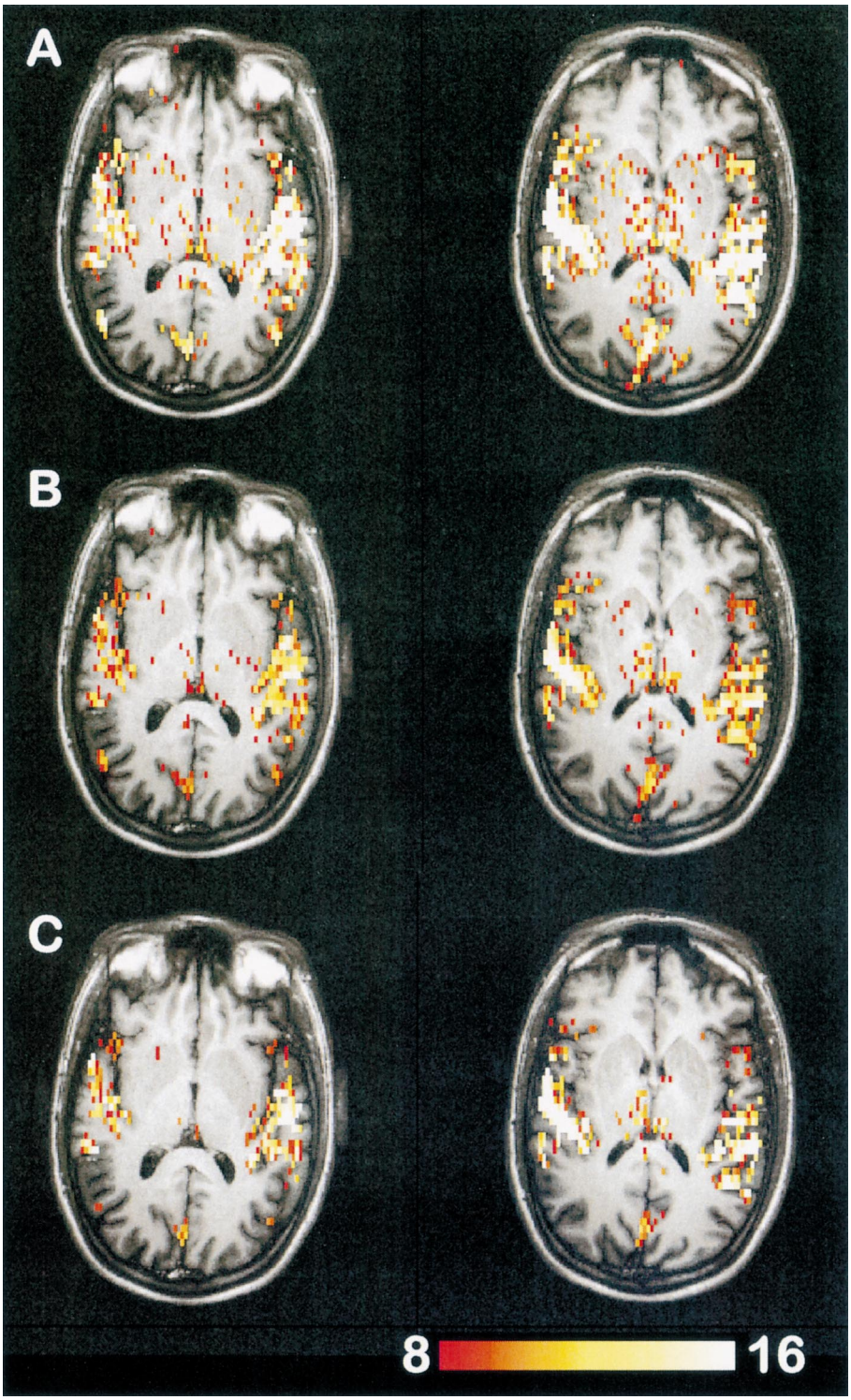

Fig. 5. Two *z*-score maps, overlaid onto the corresponding anatomical slices, from an event-related fMRI experiment in language comprehension. In the top row (A), *z*-scores are uncorrected; in the middle row (B) they are corrected by a spatially stationary correction based on averaged properties of the hemodynamic response (as implemented in SPM99); and in the bottom row (C) they are corrected by the autocorrelation-based approach, using the ''damped oscillator'' function. The same color range was used to visualize *z*-scores.

very low and the approximations may not be valid. Estimating the number of effective DOF is therefore an important issue even for long fMRI time series.

Our method differs from previous methods in three areas:

- In both previous approaches, the implicit assumption

was made that the distribution of SSR and SSH followed a χ^2 distribution and the resulting F-value (Eq. (14)) was an approximation. A more formal treatment of this subject reveals that the more general Gamma distribution is required to describe the distribution of SSR and SSH properly. This leads to a generalized F-test (see Eq. (4)) which differs from the test given by Eq. (14) by the weighting factor associated with the ratio SSH/SSR, but involves the *same* number of DOF (Eq. (5)). If the errors are weakly correlated (i.e., $\boldsymbol{V} \sim \boldsymbol{I}_N$), the weighting factor used in our approach tends to the same value (close to $(N-P)/Q$) as the factor implemented in SPM99.

- While previous methods applied a spatially stationary correction for the number of effective DOF, we perform a voxelwise analysis of the residuals. The number of effective DOF is found to differ substantially between different brain regions and activated versus non-activated sites. Results from simulation studies and an example real experiment suggest that our approach leads to higher F-values in activated areas and lower values in the background, thus improving the detection of experiment-related activation.
- We use the autocorrelation matrix $\boldsymbol{V}$ of the residuals to derive the number of effective DOF. The advantage of this interpretation is given by the fact that $\boldsymbol{V}$ is determined from the data, and not linked to the properties of a pre-defined smoothing matrix. Thus, if data fulfill stationarity assumptions, as required for the general linear model, then Eq. (4) (respectively (5)) is expected to yield an exact result for the F-test (respectively, for the number of effective DOF), independent of any preprocessing that preceded the statistical analysis.

It is impossible to determine the contents of $\boldsymbol{V}$ from a single fMRI time series without introducing an autocorrelation model. For data from event-related experiments preprocessed by filters commonly applied for fMRI time series (LP, GT), we propose a new model, the ''damped oscillator'' model, to account for the autocorrelation structure. We demonstrate that this model yields a valid (i.e., non-negative definite) autocorrelation matrix $\boldsymbol{V}$ and adequately fits the temporal autocorrelations in the residuals. A simple AR(1) autocorrelation model generally does not fit the temporal autocorrelations found in event-related fMRI experiments and therefore generally overestimates ν_{SSR}. More flexible models have been proposed recently, for example AR(p) models (Friston et al., 2000), and these need to be compared with our approach.

An OLS approach is taken for solving the general linear model. A generalized least squares approach may be used instead, which would first consist of transforming (''pre-whitening'') the data

$$\boldsymbol{y}^* = \frac{1}{\sigma}\boldsymbol{V}^{-1/2}\boldsymbol{y}, \quad \boldsymbol{x}^* = \frac{1}{\sigma}\boldsymbol{V}^{-1/2}\boldsymbol{x}, \quad \boldsymbol{\epsilon}^* = \frac{1}{\sigma}\boldsymbol{V}^{-1/2}\boldsymbol{\epsilon},$$

and solving the transformed linear model $\boldsymbol{y}^* = \boldsymbol{x}^*\boldsymbol{\beta}^* + \boldsymbol{\epsilon}^*$ using OLS. Since $\boldsymbol{\epsilon} \sim \mathcal{N}(0,\sigma^2\boldsymbol{V})$ it follows that $\boldsymbol{\epsilon}^* \sim \mathcal{N}(0,\boldsymbol{I}_N)$ and the statistical tests would be conducted using $N-P$ and Q degrees of freedom. While pre-whitening is the most efficient method if $\boldsymbol{V}^{-1/2}$ is known or can be estimated in a robust manner, we do not favor this approach. Estimating $\boldsymbol{V}^{-1/2}$ is usually an ill-conditioned problem (i.e., in some cases, small variations in $\boldsymbol{V}$ may induce large variations in $\boldsymbol{V}^{-1/2}$). Moreover, this approach is computationally more costly. Indeed, if an analytical expression for $\boldsymbol{V}^{-1/2}$ is not known, a preliminary OLS analysis has to be conducted anyway to estimate $\boldsymbol{V}$ from the residuals.

The number of effective DOF is determined on a voxelwise basis. Care is taken to ensure an efficient computation even for the long time series now commonly found in fMRI experiments. Fast computation of Eq. (5) is tractable using results based on the Hadamard product or Frobenius norm. Moreover, we propose approximations for the typical case of $N/P > 50$, which further reduces the computation time. The approximation of ν_{SSR} as defined by Eq. (6) only requires computation of trace($\boldsymbol{VV}$), which is an N^2 operation. For both autocorrelation functions studied here, analytical expressions (see Eqs. (12) and (13)) simplify the computation of ν_{SSR} even further.

The proposed approach can be extended in space, i.e. for the multivariate (spatio-temporal) case. Here, $\boldsymbol{V}$ would be interpreted as a spatio-temporal autocorrelation matrix. Modeling $\boldsymbol{V}$ as a matrix separable in time and space has already been proposed in fMRI (Benali et al., 1997; Kruggel and von Cramon, 1999). Non-separable spatio-temporal autocorrelation models were first proposed in other fields such as meteorology (Cressie and Huang, 1999). We have recently introduced a non-separable model for fMRI data (Benali et al., 2001), which generalizes the ''damped oscillator'' model proposed in this paper. The relevance of non-separable models in fMRI remains to be evaluated. The statistical tests to be used for detecting activation remain to be adapted to the multivariate case and the specificity and the sensitivity of the approach need to be studied, for example using simulated data and ROC analyses as proposed by Burock and Dale (2000).

Acknowledgements

The authors are grateful to J.B. Poline (Service Hospitalier Frédéric Joliot, Orsay, France) for fruitful discussions and for advice about implementation details of SPM99. They also thank anonymous referees for significant input into the theoretical developments.

Appendix A. Matrix approximation of ν_{ssr} for large N

It is easy to show that trace($\boldsymbol{RV}$) can be rewritten as trace($\boldsymbol{V}$) − trace($\boldsymbol{HV}$), with $\boldsymbol{H} = \boldsymbol{x}(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{x})^{-1}\boldsymbol{x}^{\mathrm{T}}$. Similarly,

$$\mathrm{trace}(\boldsymbol{RVRV}) = \mathrm{trace}(\boldsymbol{VV}) - 2\,\mathrm{trace}(\boldsymbol{HVV}) + \mathrm{trace}(\boldsymbol{HVHV}).$$

Using these expressions and $\mathrm{trace}(\boldsymbol{V}) = N$ (by construction, since $\mathrm{trace}(\boldsymbol{V})$ is an autocorrelation matrix), Eq. (5) becomes

$$\nu_{\mathrm{SSR}} = \frac{N^2}{\mathrm{trace}(\boldsymbol{VV})}\left[1 - \frac{\mathrm{trace}(\boldsymbol{HV})}{N}\right]^2 \times \left[1 - 2\frac{\mathrm{trace}(\boldsymbol{HVV})}{\mathrm{trace}(\boldsymbol{VV})} + \frac{\mathrm{trace}(\boldsymbol{HVHV})}{\mathrm{trace}(\boldsymbol{VV})}\right]^{-1}.$$

We wish to prove that $\mathrm{trace}(\boldsymbol{HV})/N$, $\mathrm{trace}(\boldsymbol{HVV})/\mathrm{trace}(\boldsymbol{VV})$ and $\mathrm{trace}(\boldsymbol{HVHV})/\mathrm{trace}(\boldsymbol{VV})$ all tend to 0 when N becomes large. To do this, let us cite the following theorem (Graybill, 1983, p. 94).

Theorem. *Consider any* (N,N) *matrix* $\boldsymbol{A}$. *Then*

1. $\|\boldsymbol{A}\|_{\mathrm{E}}^2 = trace(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})$, *where* $\|\cdot\|_{\mathrm{E}}$ *is the Euclidian norm defined by* $\|\boldsymbol{A}\|_{\mathrm{E}} = [\Sigma_i\,\Sigma_j(|a_{ij}|)^2]^{1/2}$.
2. *If* $\boldsymbol{A}$ *is non-negative, then the spectral norm* $\|\boldsymbol{A}\|_2$ *is defined by* $\|\boldsymbol{A}\|_2 = max_j(\lambda_j)$, *where the* λ_j *are the eigenvalues of* $\boldsymbol{A}$.
3. *For any* (N,N) *matrix* $\boldsymbol{B}$, *it follows that* $\|\boldsymbol{AB}\|_{\mathrm{E}} \leqslant \|\boldsymbol{A}\|_{\mathrm{E}}\|\boldsymbol{B}\|_2$.

Using these propositions, we obtain

$$\mathrm{trace}(\boldsymbol{HV}) = \|\boldsymbol{HV}^{1/2}\|_{\mathrm{E}}^2 \leqslant \|\boldsymbol{H}\|_{\mathrm{E}}^2\|\boldsymbol{V}^{1/2}\|_2^2.$$

It is easy to see that $\|\boldsymbol{H}\|_{\mathrm{E}}^2 = P$, so it follows that

$$\frac{\mathrm{trace}(\boldsymbol{HV})}{N} \leqslant \frac{\|\boldsymbol{H}\|_{\mathrm{E}}^2\|\boldsymbol{V}^{1/2}\|_2^2}{N} = \frac{1}{N}\left(P \cdot \left[\max_j(\lambda_j)\right]^2\right),$$

where the λ_j are the eigenvalues of $\boldsymbol{V}^{1/2}$. It has been shown by Fine and Romain (1984) that the eigenvalues of a correlation matrix converge to a finite value when N becomes large. Therefore, the quantity $P \cdot [\max_j(\lambda_j)]^2$ is bounded when N becomes large, and the quantity $\mathrm{trace}(\boldsymbol{HV})/N$ tends to 0.

We can also show that $\mathrm{trace}(\boldsymbol{HVV}) = \|\boldsymbol{HV}\|_{\mathrm{E}}^2$ and thus

$$\frac{\mathrm{trace}(\boldsymbol{HVV})}{\mathrm{trace}(\boldsymbol{VV})} \leqslant \frac{\|\boldsymbol{H}\|_{\mathrm{E}}^2\|\boldsymbol{V}\|_2^2}{\mathrm{trace}(\boldsymbol{VV})} = \frac{P \cdot [\max_j(\mu_j)]^2}{\mathrm{trace}(\boldsymbol{VV})}, \tag{A.1}$$

where the μ_j are the eigenvalues of $\boldsymbol{V}$. We can also show that

$$\mathrm{trace}(\boldsymbol{HVHV}) = \mathrm{trace}((\boldsymbol{HVH})^{\mathrm{T}}(\boldsymbol{HVH})) = \|\boldsymbol{HVH}\|_{\mathrm{E}}^2$$

and

$$\|\boldsymbol{HVH}\|_{\mathrm{E}}^2 \leqslant \|\boldsymbol{HV}\|_{\mathrm{E}}^2\|\boldsymbol{H}\|_2^2.$$

Since $\boldsymbol{H}$ is a projector, the maximum value of its eigenvalues is 1 (Golub and Van Loan, 1984, p. 20), thus $\|\boldsymbol{H}\|_2^2 = 1$. This yields

$$\|\boldsymbol{HVH}\|_{\mathrm{E}}^2 \leqslant \|\boldsymbol{HV}\|_{\mathrm{E}}^2 \leqslant \|\boldsymbol{H}\|_{\mathrm{E}}^2\|\boldsymbol{V}\|_2^2,$$

thus

$$\frac{\mathrm{trace}(\boldsymbol{HVHV})}{\mathrm{trace}(\boldsymbol{VV})} \leqslant \frac{\|\boldsymbol{H}\|_{\mathrm{E}}^2\|\boldsymbol{V}\|_2^2}{\mathrm{trace}(\boldsymbol{VV})} = \frac{P \cdot [\max_j(\mu_j)]^2}{\mathrm{trace}(\boldsymbol{VV})}. \tag{A.2}$$

On the other hand, using the Cauchy–Schwartz inequality on the norm defined by the ''trace'' operator on non-negative definite matrices, we have

$$[\mathrm{trace}(\boldsymbol{V})]^2 \leqslant \mathrm{trace}(\boldsymbol{I}_N^2)\mathrm{trace}(\boldsymbol{V}^2) = N\,\mathrm{trace}(\boldsymbol{VV}).$$

Since $\mathrm{trace}(\boldsymbol{V}) = N$, it follows that $\mathrm{trace}(\boldsymbol{VV}) \geqslant N$. Coming back to Eqs. (A.1) and (A.2), we then see that, for all N,

$$\frac{\mathrm{trace}(\boldsymbol{HVV})}{\mathrm{trace}(\boldsymbol{VV})} \leqslant \frac{1}{N}\left(P \cdot \left[\max_j(\mu_j)\right]^2\right)$$

and

$$\frac{\mathrm{trace}(\boldsymbol{HVHV})}{\mathrm{trace}(\boldsymbol{VV})} \leqslant \frac{1}{N}\left(P \cdot \left[\max_j(\mu_j)\right]^2\right).$$

As before, the quantity $P \cdot [\max_j(\mu_j)]^2$ is bounded when N becomes large, hence $\mathrm{trace}(\boldsymbol{HVV})/\mathrm{trace}(\boldsymbol{VV})$ and $\mathrm{trace}(\boldsymbol{HVHV})/\mathrm{trace}(\boldsymbol{VV})$ tend to 0 when N becomes large.

Hence

$$\frac{[\mathrm{trace}(\boldsymbol{RV})]^2}{\mathrm{trace}(\boldsymbol{RVRV})} \to \frac{N^2}{\mathrm{trace}(\boldsymbol{VV})}$$

when N becomes large. To refine this approximation, note that when there are no autocorrelations in the data, i.e. $\boldsymbol{V} = \boldsymbol{I}_N$, the number of DOF should be $N - P$ (independent data). We can therefore correct post hoc the previous expression and finally obtain

$$\frac{[\mathrm{trace}(\boldsymbol{RV})]^2}{\mathrm{trace}(\boldsymbol{RVRV})} \to \nu_{\mathrm{app}} = \frac{N(N-P)}{\mathrm{trace}(\boldsymbol{VV})}$$

when N becomes large, which is equal to $N - P$ for $\boldsymbol{V} = \boldsymbol{I}_N$.

Eq. (8) is finally obtained by noticing that

$$\mathrm{trace}(\boldsymbol{RV}) = N\left(1 - \frac{\mathrm{trace}(\boldsymbol{HV})}{N}\right)$$

tends to N when N becomes large. Since $\mathrm{trace}(\boldsymbol{RV})$ should be equal to $N - P$ when $\boldsymbol{V} = \boldsymbol{I}_N$, we correct the limit value to $N - P$. It is straightforward to check that Eq. (8) reduces to Eq. (3) when $\boldsymbol{V} = \boldsymbol{I}_N$.

Appendix B. Validity of the ''damped oscillator'' correlation model

According to Christakos (1984, p. 255), a continuous function $\rho(r)$ ($r = |k|$, where k is the lag), defined in $\mathbb{R}^1$, is a permissible autocorrelation model, i.e. it is non-negative definite, if and only if its associated spectral function $\mathscr{R}_1(\omega)$ satisfies

$$\mathscr{R}_1(\omega) \geqslant 0,$$

where the spectral function $\mathscr{R}_1(\omega)$ is defined as

$$\mathcal{R}_1(\omega) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \cos(\omega r)\rho(r)\,\mathrm{d}r.$$

For the ''damped oscillator'' model, we have $\rho(r) = \exp(a_1 r)\cos(a_2 r)$, $a_1 < 0$, hence

$$\begin{aligned}\mathcal{R}_1(\omega) &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} \exp(a_1 r)\cos(a_2 r)\cos(\omega r)\,\mathrm{d}r \\ &= \frac{1}{\pi}\int_{0}^{+\infty} \exp(a_1 r)\cos(a_2 r)\cos(\omega r)\,\mathrm{d}r, \text{ since } r \geqslant 0 \\ &= \frac{1}{2\pi}\left(\int_{0}^{+\infty} \exp(a_1 r)\cos((a_2+\omega)r)\,\mathrm{d}r \right. \\ &\qquad \left. + \int_{0}^{+\infty} \exp(a_1 r)\cos((a_2-\omega)r)\,\mathrm{d}r\right).\end{aligned}$$

Using the integral

$$\int_{0}^{+\infty} \exp(-\alpha x)\cos(\beta x)\,\mathrm{d}x = \frac{\alpha}{\alpha^2+\beta^2},$$

and taking $\alpha = -a_1$ and $\beta = a_2 \pm \omega$, we have

$$\begin{aligned}\mathcal{R}_1(\omega) &= \frac{1}{2\pi}\left(\frac{-a_1}{a_1^2+(a_2+\omega)^2} + \frac{-a_1}{a_1^2+(a_2-\omega)^2}\right) \\ &= -\frac{a_1}{\pi}\frac{a_1^2+a_2^2+\omega^2}{(a_1^2+(a_2+\omega)^2)(a_1^2+(a_2-\omega)^2)}.\end{aligned}$$

It is straightforward to see that $\rho(r)$ is a permissible autocorrelation model if and only if $a_1 \leqslant 0$.

Appendix C

Closed form for ν_{app}

To find the closed form for ν_{app}, we need to calculate trace($\boldsymbol{VV}$) for the ''damped oscillator'' autocorrelation structure. From Eq. (9), we have

$$\mathrm{trace}(\boldsymbol{VV}) = \sum_{i=1}^{N}\sum_{j=1}^{N} \hat{v}_{ij}^{\,2} = \sum_{i=1}^{N}\sum_{j=1}^{N} \rho(|i-j|)^2.$$

For the ''damped oscillator'' autocorrelation model given by Eq. (10),

$$\begin{aligned}\rho(|i-j|) &= \exp(a_1|i-j|)\frac{\exp(\mathrm{i}a_2|i-j|) + \exp(-\mathrm{i}a_2|i-j|)}{2} \\ &= \frac{1}{2}(A^{|i-j|} + B^{|i-j|}),\end{aligned}$$

where $\mathrm{i}^2 = -1$, $A = \exp(a_1 + \mathrm{i}a_2)$ and $B = \exp(a_1 - \mathrm{i}a_2)$. trace($\boldsymbol{VV}$) then becomes

$$\begin{aligned}&\mathrm{trace}(\boldsymbol{VV}) \\ &= \frac{1}{4}\sum_{i=1}^{N}\sum_{j=1}^{N}[A^{|i-j|} + B^{|i-j|}]^2 \\ &= \frac{1}{4}\sum_{i=1}^{N}\sum_{j=1}^{N} A^{2|i-j|} + \frac{1}{4}\sum_{i=1}^{N}\sum_{j=1}^{N} B^{2|i-j|} + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} AB^{|i-j|}.\end{aligned}$$

Since we impose $a_1 < 0$, it follows that $|A^2| = \exp(2a_1) < 1$, $|B^2| = \exp(2a_1) < 1$ and $|AB| = \exp(2a_1) < 1$.

On the other hand, we have the following general result for $\alpha < 1$:

$$\begin{aligned}&\sum_{i=1}^{N}\sum_{j=1}^{N} \alpha^{|i-j|} \\ &= N + 2\sum_{i=1}^{N} \alpha^{i}\sum_{j=1}^{i-1} \alpha^{-j} \\ &= \frac{N}{(1-\alpha)^2}\left[(1-\alpha^2) - \frac{2\alpha}{N} + \frac{2\alpha^{N+1}}{N}\right], \\ &\text{using } \sum_{i=0}^{n-1} r^i = \frac{1-r^n}{1-r}.\end{aligned}$$

Since α/N and α^{N+1} tend to 0, $\Sigma_{i=1}^{N}\Sigma_{j=1}^{N}\alpha^{|i-j|}$ is equivalent, for large N, to $N[(1+\alpha)/(1-\alpha)]$.

Applying this result to the terms of trace($\boldsymbol{VV}$), we have

$$\begin{aligned}&\mathrm{trace}(\boldsymbol{VV}) \\ &\approx \frac{N}{4}\left[\frac{1+A^2}{1-A^2}\right] + \frac{N}{4}\left[\frac{1+B^2}{1-B^2}\right] + \frac{N}{2}\left[\frac{1+AB}{1-AB}\right] \\ &= \frac{N}{2}\left[\frac{1-(AB)^2}{1-(A^2+B^2)+(AB)^2}\right] + \frac{N}{2}\left[\frac{1+AB}{1-AB}\right].\end{aligned}$$

Since

$$\begin{aligned}A^2 + B^2 &= \exp(2a_1)(\exp(\mathrm{i}2a_2) + \exp(-\mathrm{i}2a_2)) \\ &= 2\exp(2a_1)\cos(2a_2),\end{aligned}$$

this yields

$$\begin{aligned}&\mathrm{trace}(\boldsymbol{VV}) \\ &\approx \frac{N}{2}\left[\frac{1-\exp(4a_1)}{1-2\exp(2a_1)\cos(2a_2)+\exp(4a_1)}\right] \\ &\quad + \frac{N}{2}\left[\frac{1+\exp(2a_1)}{1-\exp(2a_1)}\right] \\ &= N\left[\frac{1+\exp(2a_1)}{1-\exp(2a_1)}\right] \\ &\quad \left[\frac{1+\exp(4a_1)-\exp(2a_1)(1+\cos(2a_2))}{1+\exp(4a_1)-2\exp(2a_1)\cos(2a_2)}\right],\end{aligned}$$

and from Eq. (6),

$$\begin{aligned}\nu_{\mathrm{app}} &\approx (N-P)\left[\frac{1-\exp(2a_1)}{1+\exp(2a_1)}\right] \\ &\quad \times \left[\frac{1+\exp(4a_1)-2\exp(2a_1)\cos(2a_2)}{1+\exp(4a_1)-\exp(2a_1)(1+\cos(2a_2))}\right].\end{aligned}$$

When the data are totally correlated, i.e. $\exp(a_1)\cos(a_2) \to 1$ ($a_1 \to 0$ and $a_2 \to 0$), the number of DOF should tend to 0. The above equation is not defined for $a_1 = 0$ and $a_2 = 0$, but it can be shown using series expansion that $\nu_{\text{app}} \underset{a_1 \to 0,\ a_2 \to 0}{\to} 0$.

When there are no autocorrelations in the data, i.e. $\exp(a_1)\cos(a_2) \to 0$, which is only possible with $a_1 \to -\infty$, the number of DOF is $N - P$ (independent data). This completes the proof for Eq. (12).

References

Benali, H., Buvat, I., Anton, J.L., Pélégrini, M., Di Paola, M., Bittoun, J., Burnod, Y., Di Paola, R., 1997. Space–time statistical model for functional MRI image sequences. In: Duncan, J., Gindi, G. (Eds.), Information Processing in Medical Imaging. Springer, Berlin, pp. 285–298.

Benali, H., Pélégrini-Issac, M., Kruggel, F., 2001. Spatio-temporal covariance model for medical images sequences: application to functional MRI data. In: Insana, M., Leahy, R.M. (Eds.), Information Processing in Medical Imaging. Springer, Berlin, pp. 197–203.

Biswal, B., DeYoe, E.A., Hyde, J.S., 1996. Reduction of physiological fluctuations in fMRI using digital filters. Magn. Reson. Med. 35, 107–113.

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

Buonocore, M.H., Maddock, R.J., 1997. Noise suppression digital filter for functional magnetic resonance imaging based on image reference data. Magn. Reson. Med. 38, 456–469.

Burock, M.A., Dale, A.M., 2000. Estimation and detection of event-related fMRI signals with temporally correlated noise: a statistically efficient and unbiased approach. Hum. Brain Mapp. 11, 249–260.

Christakos, G., 1984. On the problem of permissible covariance and variogram models. Water Resources Res. 20, 251–265.

Cressie, N.A.C., 1993. Statistics for Spatial Data. Wiley, New York.

Cressie, N.A.C., Huang, H.C., 1999. Classes of nonseparable, spatio-temporal stationary covariance functions. J. Am. Statist. Assoc. 44, 1330–1340.

Cochrane, D., Orcutt, G.H., 1949. Application of least squares regression to relationships containing autocorrelated errors. J. Am. Statist. Assoc. 44, 32–61.

Descombes, X., Kruggel, F., von Cramon, D.Y., 1998. FMRI signal restoration using an edge preserving spatio-temporal Markov random field. NeuroImage 8, 340–349.

Durbin, J., Watson, G.S., 1950. Testing for serial correlations in least squares regression: I. Biometrika 37, 409–428.

Durbin, J., Watson, G.S., 1951. Testing for serial correlations in least squares regression: II. Biometrika 38, 159–178.

FIL Methods Group, 2000. http//www.fil.ion.ucl.ac.uk/spm/spm99.html.

Fine, J., Romain, Y., 1984. Reduced principal component analysis. Mathematische Operationsforschung und Statistik 15, 493–512.

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

Friston, K.J., Holmes, A.P., Poline, J.B., Grasby, P.J., Williams, S.C.R., Frackowiak, R.S.J., Turner, R., 1995a. Analysis of fMRI time-series revisited. NeuroImage 2, 45–53.

Friston, K.J., Holmes, A.P., Worsley, K.J., Poline, J.B., Frith, C.D., Frackowiak, R.S.J., 1995b. Statistical parametric maps in functional imaging, a general linear approach. Hum. Brain Mapp. 2, 189–210.

Friston, K.J., Josephs, O., Zarahn, E., Holmes, A.P., Rouquette, S., Poline, J.B., 2000. To smooth or not to smooth? Bias and efficiency in fMRI time-series analysis. NeuroImage 12, 196–208.

Golub, G.H., Van Loan, C.F., 1984. Matrix Computations. The Johns Hopkins University Press, Baltimore.

Graybill, F.A., 1983. Matrices with Applications in Statistics. Wadsworth, Belmont.

Hu, X., Le, T.H., Parrish, T., Erhard, P., 1995. Retrospective estimation and correction of physiological fluctuation in functional MRI. Magn. Reson. Med. 35, 201–212.

Johnson, N.L., Kotz, S., Balakrishnan, N., 1995. 2nd Edition. Continuous Univariate Distributions, Vol. 2. Wiley, New York.

Johnson, R.A., Wichern, D.W., 1988. Applied Multivariate Statistical Analysis. Prentice Hall, Englewood Cliffs.

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

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

Lange, N., Zeger, S.L., 1997. Non-linear Fourier time series analysis for human brain mapping by functional magnetic resonance imaging. J. R. Statist. Soc., Appl. Statist. 46, 1–19.

Mardia, K.V., Kent, J.T., Bibby, J.M., 1979. Multivariate Analysis. Academic Press, London.

Moonen, C., Bandettini, P.A., 1999. Functional MRI. Springer, Berlin.

Petersen, N.V., Jensen, J.L., Burchhardt, J., Stødkilde-Jørgensen, H., 1998. State space models for physiological noise in fMRI time series. NeuroImage 7, S592.

Press, W.H., Flannery, B.P., Teukolsky, S.A., Vetterling, W.T., 1992. Numerical Recipes in C. Cambridge University Press, Cambridge.

Priestley, M.B., 1996. Spectral Analysis and Time Series, 9th Edition. Academic Press, London.

Rajapakse, J.C., Kruggel, F., Maisog, J., von Cramon, D.Y., 1998. Modeling hemodynamic response for analysis of functional MRI time series. Hum. Brain Mapp. 6, 283–300.

Rencher, A.C., 1995. Methods of Multivariate Analysis. Wiley, New York.

Satterthwaite, F.E., 1946. An approximate distribution of estimates of variance components. Biometrika 33, 110–114.

Wood, A.T.A., 1989. An *F* approximation to the distribution of a linear combination of Chi-squared variables. Commun. Statist. – Simulations 18, 1439–1456.

Worsley, K.J., Friston, K.J., 1995. Analysis of fMRI time-series revisited – again. NeuroImage 2, 173–181.

Worsley, K.J., Friston, K.J., 1997. Characterizing the response of PET and fMRI data using multivariate linear models (MLM). NeuroImage 6, 305–319.