Signal Processing in Functional Magnetic Resonance Imaging (fMRI) of the Brain

by

Patrick L. Purdon

A.B. Engineering Sciences
Harvard University, 1996

Submitted to the Department of Electrical Engineering and Computer Science in Partial Fulfillment of the Requirements for the Degree of

Master of Science in Electrical Engineering and Computer Science
at the
Massachusetts Institute of Technology

June 1998

©1998 Patrick L. Purdon. All Rights Reserved.

The author hereby grants to MIT permission to reproduce and to distribute publicly paper and electronic copies of this thesis document in whole or in part.

Signature of Author: ____________________________

Department of Electrical Engineering and Computer Science
May 21, 1998

Certified By: ____________________________

Robert M. Weisskoff
Associate Professor of Radiology
Harvard-MIT Division of Health Sciences and Technology
Thesis Supervisor

Certified By: ____________________________

Alan V. Oppenheim
Distinguished Professor of Electrical Engineering
Thesis Supervisor

Accepted by: ____________________________

Arthur C. Smith
Chair, Committee on Graduate Students
Department of Electrical Engineering and Computer Science

JUL 23 1998
Signal Processing in Functional Magnetic Resonance Imaging (fMRI) of the Brain

by

Patrick L. Purdon

Submitted to the Department of Electrical Engineering and Computer Science on May 21, 1998 in Partial Fulfillment of the Requirements for the Degree of Master of Science in Electrical Engineering and Computer Science

ABSTRACT

Functional Magnetic Resonance Imaging (fMRI) is an important new tool for studying the function of the human brain. The fMRI signal is sensitive to local changes in blood oxygenation and blood volume brought about by neural activity and provides a means of observing dynamical behavior within and among brain regions at a unique combination of high spatial and temporal resolution unmatched by other imaging modalities.

Analysis of fMRI data has focused on detecting regions of functional activation within a hypothesis testing framework. In such analyses the noise is often assumed to be uncorrelated in time, yet it is well-known that the noise are correlated. Simulations were conducted to demonstrate that the presence of correlated noise can bias statistical inference in a way that depends on both the shape of the noise and the experimental paradigm chosen.

Correcting for the presence of colored noise depends on our knowledge of the noise. The degree of variability in the noise between subjects, trials, and across space dictates the most appropriate method for estimating the noise. Analysis of the variability in the noise from resting-state data suggest that noise estimation must be done in a spatially-varying manner concomitantly with estimation of activation parameters.

An alternative to hypothesis testing of fMRI data is to estimate parameters from physiologically-inspired models. For such analysis to be meaningful, the models must accurately account for interesting dynamical features in the data, and robust estimates of the noise must be made for accurate statistical inference. A method employing a physiologically-realistic signal model within a local averaging regularization scheme for noise estimation was implemented and shown to be successful in estimating activation parameters with high spatial resolution while robustly estimating the noise.

Thesis Advisors:

Robert M. Weisskoff
Title: Associate Professor of Radiology, Harvard-MIT Division of Health Sciences and Technology

Alan V. Oppenheim
Title: Distinguished Professor of Electrical Engineering
Chapter 1: Introduction

Non-invasive functional imaging methods have become an important technique for studying the macroscopic function of the human brain. Prior to the mid-1990's, the vast majority of functional imaging experiments were done using nuclear medicine techniques such as positron emission tomography (PET), but currently, well-over half of all such experiments are done with Functional Magnetic Resonance Imaging (fMRI) [personal observations from the 1997 Human Brain Mapping conference in Copenhagen, Denmark]. In fMRI, rapid images of the brain are taken in a way that captures changes in contrast due to changes in blood oxygenation and blood volume brought about by neural activity [Kwong, et al., 1992; Ogawa, et al., 1992]. The signal changes observed through this blood oxygen level dependent (BOLD) contrast mechanism occur on a time scale of seconds, providing a means for observing dynamical behavior within and among brain regions at a unique combination of high spatial and temporal resolution unmatched by other individual imaging modalities [Cohen and Bookheimer, 1994]. The spatio-temporal richness of fMRI data, when combined with its low cost and experimental convenience relative to competing techniques such as PET, make fMRI the method of choice in many brain mapping experiments.

In a typical fMRI experiment, a subject is presented with a sensory stimulus or cognitive task in a periodic "off" - "on" pattern while images of the whole brain are taken in rapid succession. For instance, a subject might be presented with a flickering checkerboard (or some other stimulus) for 10 seconds, repeated every 25 seconds for several trials; this experimental paradigm resembles a square-wave with a 25 second period and 40% duty cycle (referred to as a "block paradigm" in the fMRI literature). In visual experiment of this kind, the MR signal from the visual cortex is observed to fluctuate 2-3% from its baseline level due to the BOLD effect, at the same frequency as the stimulus. Experiments of this general type provide a powerful method for identifying regions of functional specialization and have been used to delineate the organization of the visual cortex [Sereno, et al., 1995] and auditory cortex [Talavage, et al., 1997], and have also provided insight into mechanisms for higher-level cognitive processes such as memory [Buckner and Koutstaal, 1998], attention [O’Craven, et al., 1997], and addiction [Breiter, et al., 1997], to name a few. In clinical settings, experiments of this kind can be used for pre-operative planning in brain tumor surgery [FitzGerald, et al., 1997].

The most widespread analytical approach in fMRI is to perform a binary hypothesis test at each spatial location (or voxel) for the presence of an "activation" signal, creating a spatially distributed map of significance levels (P-values) based on some test statistic (e.g., Student's T-test, Fourier F-Test, Kolmogorov-Smirnov Test,
etc.). Often, the test is carried out as a linear regression against a hypothetical activation signal [Friston, et al., 1995; Friston, et al., 1994]. A typical approach for synthesizing this purported activation signal is to assume that the stimulus paradigm (e.g., the "square-wave" described earlier) is an input into the hemodynamic mechanism which produces the fMRI signal [Cohen, 1997; Friston, et al., 1994; Lange and Zeger, 1997]. In most cases, the hemodynamic mechanism is assumed to be linear and time-invariant (LTI) and the impulse response function of this system (referred to in the brain mapping literature as the "hemodynamic response function") is thought to be some positive, decaying function (e.g., Poisson, gamma, or exponential), with parameters that are either assumed or estimated from data.

Analytical methods of this kind raise a number of important issues that will be the focus of this thesis:

1. Inherent in the hypothesis tests against fMRI time series is the assumption that the data are independent and identically distributed in time. In fact, the noise in fMRI data can have significant temporal correlations due to underlying physiological fluctuations [Biswal, et al., 1995; Weisskoff, et al., 1993], and the effect of this temporal correlation on the accuracy of hypothesis testing methods remains an open question. Chapter 2 of this thesis will describe work done to quantify the effect of this colored noise on false positive rates for several commonly used statistical tests, and will demonstrate that actual false positive rates can be greatly biased from their assumed values, irrespective of which statistical test is used, in a way that depends both on the imaging parameters and the experimental paradigm chosen.

2. With appropriate choice of noise models, it is simple enough to account for the effect of this correlated noise within a linear regression framework by "whitening" or weighted least-squares [Bullmore, et al., 1996]: With a given noise model, the key issue then is how to estimate the noise in a manner which is appropriate to the data and also experimentally and computationally efficient. Three reasonable approaches might be: A) Obtain a noise estimate averaged across many subjects to be applied globally for all subjects; B) Obtain a noise estimate for each individual subject during a "baseline" pre-scan, for use in subsequent experiments with that subject; C) Obtain noise estimates simultaneously with signal parameter estimation [Bullmore, et al., 1996; Locascio, et al., 1997; Solo, et al., 1997]. The first two options are the most simple computationally, but require that the noise is stationary across subjects or trials, respectively, as well as across regions of gray matter; The third option requires the fewest assumptions but can be computationally expensive. Chapter 3 of this thesis will explore noise variability within and across subjects and trials in order to determine if it is possible to create meaningful averages of
noise behavior as described above. We will show that the most appropriate option, for the given noise model and data, is to estimate the noise in a spatially-varying manner concomitantly with estimates of the activation signal.

3. As our knowledge of the mechanisms which produce the BOLD fMRI signal improve, we will be able to gain more detailed knowledge about the underlying physiology of the brain by estimating relevant parameters from physiologically-inspired models, as opposed to simply identifying whether a given region of the brain is "On" or "Off" in response to some experimental treatment. For instance, one problem with the LTI system models described earlier is that they are unable to reproduce "undershoot" dynamics of the fMRI signal under block paradigm conditions-- i.e., during block paradigm experiments, the fMRI signal is often observed to dip below the initial baseline signal level during the "off" phase of the experiment. Recent physiological experiments suggest that the BOLD contrast mechanism actually has dynamical components with two different time scales that cannot be fully represented by the unimodal LTI models described earlier [Mandeville, et al., 1996; Mandeville, et al., 1998; Marota, et al., 1996]. A more complete analysis of the fMRI dynamics using a model which accounts for these mechanisms may provide important physiological information. Chapter 4 describes the implementation of a more realistic model for the hemodynamic response, within a signal processing framework that can account for spatial variations in temporally correlated noise across the brain [Solo, et al., 1997]. We find that this model and implementation are able to represent a greater variety of fMRI dynamics than previous methods, and that it is able to robustly estimate the underlying noise without loss of spatial resolution in the activation parameters.

In summary, the work in this thesis addresses a number of important issues in the analysis of fMRI data that will improve the accuracy of current analytical methods and open the door for development of new methods that incorporate physiological information to ultimately enhance our understanding of the underlying neural activity. Our improved understanding of the effects of colored noise on statistical inference has motivated the need to develop methods which properly estimate and account for this noise in order to achieve more meaningful statistical analysis. The results of the analysis of noise variability across subjects, trials, and space have suggested that the most appropriate method for estimating the noise is to do so in a spatially-varying manner, simultaneously with estimates of the activation parameters. Estimating the noise in this way is difficult to do in a robust manner, particularly given the short fMRI time-series. However, we have implemented a method which incorporates a physiologically inspired signal model within a framework that achieves robust noise estimation. This method provides a solid starting point for estimation methods that seek to quantify and understand underlying physiological parameters, as opposed to simply detecting changes due to experimental treatments, with accurate confidence intervals for these
parameters made possible by our knowledge of the noise. As our knowledge of the mechanisms for the fMRI BOLD response improve, the specific signal models used here may change, but the overall method for robustly estimating the signal and the noise together will remain applicable. Furthermore, as we gain a greater understanding of the relationship between neuronal activation and the BOLD response, estimation methods such as these will become increasingly important.

References


Chapter 2: Un-True False Positives Due to Temporal Autocorrelation in fMRI

INTRODUCTION

Statistical mapping within a binary hypothesis testing framework is the most widely used analytical method in functional MRI of the brain. In combination with experimental paradigms organized around a purported functional principle, this method seeks to identify regions of functional specialization within the brain [Friston, et al., 1995]. Typically, a hypothesis test is applied on a voxel-by-voxel basis to fMRI time series, creating a spatially distributed map of significance levels (P-values) based on some test statistic (e.g., Student’s T-test, Fourier F-Test, Kolmogorov-Smirnov Test, etc.). Inherent in these univariate hypothesis tests against fMRI time series is the assumption that the data are independent and identically distributed in time. In fact, the noise in fMRI data can have significant temporal correlations (or noise "coloration") due to underlying physiological fluctuations [Biswal, et al., 1995; Weisskoff, et al., 1993]. Furthermore, since the signal-to-noise ratio will vary with imaging rate, we should expect that the degree of correlation due to physiological fluctuation will vary with imaging rate as well. The effect of this temporal correlation on the accuracy of hypothesis testing methods remains an unresolved topic of debate. Some authors argue that correlation due to temporal smoothing outweighs any intrinsic physiological correlation, essentially ignoring the physiological correlation in the fMRI time-series [Friston, et al., 1995; Worsley and Friston, 1995]. Other authors have attempted to quantify how such intrinsic correlations might bias false positive rates (Type I error) above the assumed P-value, but have been limited to studies focused on a single imaging rate and experimental paradigm [Aguirre, et al., 1997; Xiong, et al., 1996]. However, because the correlation structure may vary with imaging rate, and because a given correlation structure may impose a characteristic time scale upon the data set, we should expect that the accuracy of P-values used for inference will depend on both imaging rate and paradigm choice. In this study, using empirically obtained noise data, we numerically estimated the actual false positive rates brought about by these conditions and determined the accuracy of P-values for the Student’s T, Kolmogorov-Smirnov, and Fourier-based F tests.

METHODS

In order to investigate the effect of temporal correlation on false positive rates in a systematic way, we approached the problem as follows: 1) We used a simple noise model to quantify the nature of the temporal correlations in fMRI data sets over a range of imaging rates; 2) We then synthesized activation-free noise data using
the parameters determined above and performed statistical analysis on these data to estimate the effect of the noise coloration on false positive rates.

Quantifying the Temporal Autocorrelations in fMRI Time Series: Noise Modeling

Noise in fMRI time series have been modeled in a variety of ways. Some authors have taken the view that the noise is the result of underlying white noise shaped by the brain's hemodynamic response [Friston, et al., 1994], while others have taken a more empirical view in which the noise is thought to come from a mechanism separate from the hemodynamic response of activation. To this end, some authors have used traditional time-series methods [Bullmore, et al., 1996; Locascio, et al., 1997], while others have used more ad-hoc power spectral modeling methods [Lange and Zeger, 1997; Zarahn, et al., 1997]. In this paper, we take the empirical view and employ a first-order auto-regressive (AR) plus white noise model. The AR component represents the excess low frequency noise observed in fMRI time series [Biswal, et al., 1995; Weisskoff, et al., 1993], while the white noise component represents scanner noise. Mathematically, a noise time-series of this type can be represented as the output of a linear, time-invariant (LTI) system with additive white noise

\[ x[n] = w[n] * h[n] + v[n], \]

where \( x[n] \) is the noise time series, \( * \) denotes the convolution operation, \( h[n] \) is the impulse response of the LTI system with Discrete-Time Fourier Transform (DTFT) \( H(e^{j\omega}) = (1 - q)/(1 - q e^{j\omega}) \), and \( w[n] \) and \( v[n] \) are white noises of variance \( A_w \) and \( A_c \), respectively. The power spectrum of this AR plus white noise model is given by

\[ S_{xx}(e^{j\omega}) = A_w + \frac{A_c(1-q)^2}{(1-q e^{-j\omega}) (1-q e^{j\omega})} = A_w + \frac{A_c}{(1+4q(1-q)^2 \sin^2(\omega/2))}. \]

The \( A_w \) and \( A_c \) parameters represent the amounts of white and colored noise in the spectrum, respectively, while \( q \) represents the degree of correlation between adjacent samples of the AR process. Figure 1 provides a plot of what this spectrum looks like for parameter values of \( q=0.65, A_w=1.8, \) and \( A_c=3.6 \).

Noise Estimation and False-Positive Rate Analysis

We estimated the power spectrum on a voxel-wise basis in a visual stimulus data set using a single periodogram at each spatial location. Subjects were
presented a full-field flickering checker-board stimulus in a "box-car" paradigm, alternating between 10 seconds of stimulus and 15 seconds without stimulus for 5 cycles. Data were collected by Kathy O’Craven and Robert Savoy at the MGH-NMR Center (Charlestown, MA) using a GE Signa 1.5 T scanner modified by ANMR for EPI. Gradient Echo images, TE = 50 ms, were acquired using a quadrature head coil in an oblique plane passing through the visual cortex at TR’s of 200, 500, 1000, 2500, and 5000 ms. An ROI was drawn over cortical gray matter, avoiding regions which contained an obvious activation signal. The average power spectrum over this ROI was then fitted to the equation in (2) using an iterative non-linear least squares method.

Using the estimates of \( q \), \( A_c \), and \( A_w \) from the empirical studies, sixty-four by sixty-four images were synthesized with noise parameters analogous to TR’s of 5000, 2500, 1250, 625, and 312.5 ms, without spatial correlation. Parameters for TR’s of 1250, 625, and 312.5 ms were obtained by linear interpolation between the actual measured values for \( A_c \) and \( A_w \), and by direct evaluation of \( q = \exp(-\tau/\tau) \) with \( \tau = 15 \) s, as obtained from the empirical data (see Results). The TR times were chosen so that the overall simulated experiment time would be the same (160 seconds) for each TR value at power-of-two data lengths (facilitating use of the Fourier F-test based on the commonly used power-of-two FFT algorithm [Press, et al., 1992]). These “null” data sets were then analyzed with the T, Kolmogorov-Smirnov (KS), and Fourier F tests with assumed off-on stimulus paradigms of 0.025 Hz and 0.05 Hz, corresponding to assumed off-on periods of 40 seconds and 20 seconds, respectively. Specifically, for the T and KS tests, the "off" and "on" samples were considered as two separate groups and then compared using commonly available "C" computer code for the T and KS tests [Press, et al., 1992]. This calculation was done on a voxel-by-voxel basis to generate a P-value for each voxel.

While the Fourier F test is not as common as the T or KS tests, we describe it in this paper both because of its potential usefulness in fMRI analysis in general and because it has an intuitive interpretation that will be useful in understanding the results to follow. The Fourier F test is used to detect a periodic signals in a background of white noise. In a repeating block designed fMRI experiment, most of the signal power in the BOLD activation signal will be contained in the fundamental paradigm frequency, so in this context we use the Fourier F test to detect a sinusoidal signal at the paradigm frequency. The intuition behind this test is that we are comparing the power in the frequency of interest to the average power in the other frequencies, rejecting the null hypothesis when the signal at the frequency of interest has much more power than the average power of the other frequencies. The Fourier F-test is performed by computing the periodogram \( \hat{I}(\omega) \) (i.e., a simple FFT-based estimate of the power spectrum) of a single-voxel time series and constructing the following F-statistic,
\[
F = \frac{(n-3)I(\omega_p)}{\sum_{n=0}^{N-1} x^2[n] - I(0) - 2I(\omega_p)} \approx F_{1-\alpha}(2, n-3)
\]

\[
I(\omega_k) = N^{-1} \left| \sum_{n=0}^{N-1} x[n] \exp(-j \omega_k n) \right|^2, \omega_k = 2\pi k / N
\]

where \( N \) is the total number of data points, \( F(2, N-3) \) is an \( F \) distribution with 2 and \( N-3 \) degrees, and \( \omega_p \) is the frequency of the paradigm, which in this expression must be a multiple of \( 2\pi/N \) (although a more general expression does exist for cases where this is not true) [Brockwell and Davis, 1991].

The result of the individual tests is an expected false positive probability; that is, the probability that the test statistic could have been that large or larger by chance just from noisy data. Following the brain mapping tradition, we will refer to this value as the “P-value” images. If we simulate data for \( N \) pixels, for an arbitrary significance level, \( \alpha \), one would expect \( \alpha N \) of those pixels to have P-values less than or equal to \( \alpha \). Thus for an activation-free data set, once we choose \( \alpha \), all pixels with \( P < \alpha \) are false-positive activations. So, to test the accuracy of these statistics when the noise is not white, we inspected the resulting P-value maps for each test for false positives over a continuum of assumed significance levels. At each assumed significance level \( \alpha \), the number of pixels with \( P < \alpha \) were counted and divided by the total number of pixels in the image to estimate the false positive rate. Ten separate 64 by 64 data sets for each TR were synthesized and analyzed in this fashion with each of the above statistical tests. The false positive rates for each TR and statistical test were averaged and plotted against the assumed significance level \( \alpha \) to create the False Positive Characteristic (FPC). Data that meet the assumptions of the given statistical test would have false positives rates equal to \( \alpha \), resulting in a linear FPC with unity slope, while those which violate the assumptions of the test would have a nonlinear relationship between false positive rate and \( \alpha \) (i.e., the false positive rate is biased).

RESULTS

In this section we present the results of the noise analysis, illustrating how the noise characteristics change with TR. We then present the results of the statistical analysis, demonstrating that the false positive rates deviate from the assumed significance levels in a way which depends on both the TR and the paradigm frequency.

Noise Estimation

Parameter estimates for \( q, A_C, \) and \( A_w \) are shown in Table 1 for the original TR values. The \( q \) parameter (and hence the degree of correlation) increased with the imaging rate, behaving as if \( q = \exp(-\text{TR}/\tau) \) with \( \tau = 15 \) s for all TR’s, suggesting that there may be an underlying continuous-time decay process that creates the noise
correlation, consistent across imaging rates. In addition to the above noise parameters, Table 1 also shows the calculated AR power, the variance in the noise attributable to the AR process (i.e., $1/(2\pi) \times$ the integral over one cycle of the AR term in equation (2)) and compared this to the white noise variance $A_w$. The ratio of AR power to white noise power increased with TR, consistent with the notion that the AR noise estimated is related to an actual physiological process whose signal-to-noise ratio increases with the TR.

<table>
<thead>
<tr>
<th>TR</th>
<th>$A_w$</th>
<th>$A_c$</th>
<th>$q$</th>
<th>ARpow</th>
<th>ARpow/$A_w$</th>
</tr>
</thead>
<tbody>
<tr>
<td>200 ms</td>
<td>0.1</td>
<td>0.3</td>
<td>0.98</td>
<td>0.003</td>
<td>0.030</td>
</tr>
<tr>
<td>500 ms</td>
<td>0.2</td>
<td>1.0</td>
<td>0.96</td>
<td>0.020</td>
<td>0.102</td>
</tr>
<tr>
<td>1000 ms</td>
<td>0.4</td>
<td>2.3</td>
<td>0.93</td>
<td>0.083</td>
<td>0.209</td>
</tr>
<tr>
<td>2500 ms</td>
<td>1.0</td>
<td>3.5</td>
<td>0.85</td>
<td>0.284</td>
<td>0.284</td>
</tr>
<tr>
<td>5000 ms</td>
<td>1.8</td>
<td>3.6</td>
<td>0.65</td>
<td>0.764</td>
<td>0.424</td>
</tr>
</tbody>
</table>

False-Positive Analysis

The FPCs for each TR and statistical test are shown below in Figures 2-7, constructed by averaging the FPCs for each of the 10 trials conducted under each TR and paradigm period. The sample variance for each of the FPC curves was less than $8 \times 10^{-3}$ over all values of $\alpha$ in all cases. Note that since the Kolmogorov-Smirnov statistic $D$ has a discrete-valued probability distribution, its FPC is piecewise constant. In general, at the low frequency (40 s) paradigm, the FPC tends to bow upwards, indicating that there are more false positives than expected from the assumed significance level $\alpha$. For instance, for the Fourier F-test, at an assumed significance level of $\alpha = 0.05$ and at a TR of 625 ms, the actual false positive rate is 0.16, three times greater than the expected value given by $\alpha$. For the T-test and KS-test under the same situation, the actual false positives are 0.12 and 0.1, respectively, roughly twice as great as the assumed alpha. As we move to smaller alpha, the bias in the false positive rate becomes much worse relative to the assumed $\alpha$. For instance, at an $\alpha$ of 0.02 for the Fourier F-test at a TR of 625 ms, the false positive rate is approximately 0.095, nearly five times the expected value, with similar results for the KS-test and T-test. At the high frequency (20 s) paradigm, there are uniformly fewer false positives than in the lower frequency paradigm, resulting in fewer false positives than the assumed $\alpha$ in some cases. For instance, for the Fourier F-test at a TR of 5000 ms and at $\alpha = 0.05$, the low frequency paradigm gives a false positive rate of 0.06, whereas for the high frequency case the false positive rate is 0.03, a bit more than half the assumed $\alpha$. At a given paradigm frequency, false positives tend to decrease with increasing TR, with the exception of TR = 312 ms at the low frequency paradigm.
Across both TR and paradigm frequencies, all three tests show both similar trends and quite similar actual biases. That is, the FPC curves for the three different tests were very similar at each paradigm/TR combination.

DISCUSSION

Our simulations demonstrate that the disparity between the actual false positive rate and the assumed significance level depends on both the imaging rate and the paradigm frequency. In this section, we (1) provide some interpretations for the behavior of the FPC curves, (2) comment on aspects of the noise modeling, (3) describe a simple way of correcting the false positive bias, and (4) relate this method to an existing method based on the general linear model of Worsley and Friston (1995).

Interpreting the Behavior of the FPC

The varying behavior of the FPC at a given paradigm frequency due to different TR's depends upon the specific noise present at a given TR. Conceptually, however, we can describe the dependence of the FPC on the imaging rate in terms of the effective number of degrees of freedom [Worsley and Friston, 1995]. As the imaging rate and the degree of correlation increase (since the AR correlation parameter $q$ increases with imaging rate), the effective number of degrees of freedom decrease relative to that of the assumed sampling distribution, resulting in an increase in the actual false positive rate with imaging rate by this account.

The influence of the paradigm frequency can be understood in an intuitive way by considering the two limiting cases at a given TR: 1) A paradigm where the first half of the experiment has no stimulus, while the stimulus is applied continuously in the second half (i.e., one period of a box-car paradigm), 2) A paradigm where the stimulus is "off" for one time sample and "on" for the next (i.e., the highest frequency paradigm that we are able to sample). Suppose we attempt to analyze activation-free data for differences in the mean (i.e., a T-test), with known noise parameters. In the first scenario, we will see more false positives than the assumed significance value, due to a reduction in the effective degrees of freedom, as described earlier (i.e., an FPC bowing upwards). However, for the second scenario, the high degree of correlation means that neighboring samples differ by only a small fraction plus a (relatively) small white noise component, and hence the resulting "off" and "on" data sets will be very similar. The difference in the sample means of the "off" and "on" populations will thus be very small, leading us to detect fewer false positives than expected (i.e., an FPC bowing down).

For paradigms of arbitrary frequency, we can develop an approximate frequency domain relationship with a similar interpretation (See Appendix A for derivation). The end result is that the variance of the difference in sample means will be determined principally by the noise power at the fundamental paradigm frequency:
\[ \text{var} \{ \mu_{on} - \mu_{off} \} \propto S_{xx}(e^{i\omega}) . \]  

(4)

An assumption of white noise corresponds to assuming that the power spectrum of the noise is flat, with a value equal to the average value of \( S_{xx}(e^{i\omega}) \) over any interval of \( 2\pi \). The power spectrum of the first order AR noise plus white noise described in equation (2) has an approximate \( 1/\omega^2 \) dependence, seen by taking a small-angle approximation on \( \sin(\omega/2) \), so at low frequencies it is larger than average and at high frequencies it is smaller than average. Hence, for low frequencies we will tend to underestimate the variance of the difference in means, resulting in more false positives than expected. As we increase the paradigm frequency, the actual variance approaches and slides below the average value, so we will tend to detect fewer and fewer false positives, at some point detecting fewer false positives than expected. Figure 8 provides an illustration of this.

**Role of Paradigm Frequency in P-value Distortion**

![Diagram](image)

**Figure 8**

The paradigm dependence of the Fourier-based F-test can be seen in a similar way by directly examining the formula for the F-statistic from equation (3). Its denominator can be interpreted as the average of the power
spectral components minus those at DC (zero frequency) and $\pm \omega_p$, while the numerator can be thought of as the power spectral density at the paradigm frequency. Thus, the Fourier-based F-statistic compares the power at the paradigm frequency to the average power in all frequencies (except DC). As with the T-test, the $1/\omega^2$ power spectrum of first order AR noise, low frequency paradigms will result in more false positives than expected, since power at low frequency will be greater than average, while high frequency paradigms will result in fewer false positives than expected, since power at high frequency will be lower than average.

Comments on Noise Modeling

Modeling the power spectrum as a rational function as in Equation (2), while arbitrary, has a number of advantages over other (also arbitrary) methods proposed previously for fMRI time series [Friston, et al., 1994; Zarahn, et al., 1997]. First, we can easily synthesize simulated data of the desired shape by using a linear, constant-coefficient difference equation implementation of Equation (2) [Oppenheim and Schafer, 1989], facilitating a Monte Carlo study like this one aimed at understanding the basic relationships between temporal correlation, experimental paradigm, and P-values in statistical maps. Second, this model, written in the discrete time domain, properly accounts for the fact that the fMRI time series are sampled data subject to aliasing. To contrast, other authors series [Friston, et al., 1994; Zarahn, et al., 1997] have essentially chosen to model the power spectrum in the Continuous-Time domain (i.e., ignoring the inherent periodicity present in the frequency spectra of any sampled data), an approximation which would require additional steps or assumptions before such a model could be applied properly to the power spectrum of sampled data (from an FFT, for instance).

As described in the previous section, the noise model used in this study has an approximate $1/\omega^2$ dependence in the power spectrum. This $1/\omega^2$ dependence in the power spectrum is analogous to the "1/f" dependence in the magnitude spectrum described by Zarahn, et al. (1997) and is characteristic of noise spectra produced by filtering white noise with some form of first-order linear low-pass filter. To contrast, the term "1/f," as used in the Nonlinear Dynamics and Complex Systems literature, refers to the 1/f power spectrum observed in complex systems exhibiting self-similar dynamics. Since the brain certainly qualifies as a complex system, such self-similar dynamics may well exist when observed over very long time scales, but for now we should make a distinction between the "1/f" noise in complex systems and the simpler $1/\omega^2$ LTI low-pass filtered noise described here and in Zarahn, et al. (1997).

Correcting False-Positive Bias: A "Whitening" Filter

We can see from the previous discussion that a simple degrees-of-freedom correction is not sufficient to compensate for all the effects of this correlated noise, since the paradigm itself also affects the underlying null
distribution. However, it may be possible to correct the false-positive bias by removing the temporal correlation with a "whitening filter." The basic idea is to filter the fMRI time series in such a way that the power spectrum of the noise becomes flat—i.e., we are forcing the noise to become "white." For example, if the noise in fMRI time series are well described by a rational power spectrum, such as equations (1) and (2), for instance, we can apply a whitening filter consisting of the inverse of the minimum-phase spectral factor of the noise power spectrum in (1) [Papoulis, 1991; See Appendix B for a detailed development]:

$$S_{xx}(z) = \frac{A_e(1-q)^2 + A_w (1-qz^{-1})(1-qz)}{(1-qz^{-1})(1-qz)} = \frac{1}{H_w(z)H_w(z^{-1})}$$

(5)

For data containing an activation signal, the resulting "whitened" signal is given by

$$x[n] = \xi[n] + a[n]$$

$$\bar{x}[n] = x[n]*h_w[n] = \xi[n]*h_w[n] + a[n]*h_w[n] = \xi[n]*h_w[n] + \bar{a}[n]$$

$$\bar{a}[n] = a[n]*h_w[n]$$

(6)

where $\xi[n]$ is the colored noise, $a[n]$ is the activation signal (i.e., the assumed neuronal response smoothed by the hemodynamic response), $h_w[n]$ is the impulse response of the whitening filter, $\xi[n]*h_w[n]$ is the whitened noise, and $\bar{a}[n]$ is the modified activation signal. Thus, following the whitening step, signal detection (i.e., correlation analysis or regression) would be done based on the $\bar{a}[n]$ signal. For instance, a T-test is essentially equivalent to regression against a square-wave of the appropriate frequency and duty cycle, so in this whitening framework one would assume a square-wave for $a[n]$ and use the appropriate $\bar{a}[n]$ for regression. In the case of the Fourier-based F-test, this correction corresponds simply to normalizing the periodogram components by the power spectral density of the underlying noise. Note that in a linear regression context, this whitening filter is equivalent to a weighted least-squares estimate, using the (temporal) covariance matrix implied by the autocorrelation function corresponding to (2) (i.e., the inverse DTFT of $S_{xx}(e^{j\omega})$) as our weighting matrix. The advantage here is that, again because of the choice of a rational DTFT domain model of the power spectrum, we can easily implement (6) with a linear, constant-coefficient difference equation which is $O(N)$ in computational complexity, compared to the general $O(N^3)$ complexity of the matrix inversion inherent in weighted least-squares. A plot comparing the FPC's of whitened and unwhitened data for the TR=312 ms data set is given in Figure 9. Note how the FPC of the whitened data closely matches that of the ideal FPC.
The actual feasibility of such a method will depend on our ability to estimate the noise accurately, a topic which will be discussed in greater detail in Chapter 2. Nonetheless, the conceptual framework used in this development is important because it provides a simple and direct way of correcting the bias in the false-positive rate. In the following section we compare our method for correcting false-positives to one based on the work of Worsley and Friston (1995).

Comparisons and Connections with the Extended General Linear Model of Worsley and Friston (1995)

The Extended General Linear Model (E-GLM) of Worsley and Friston (1995) has been used by some authors (Zarahn, et al, 1997) as a means of correcting P-value distortions. The E-GLM method, taken as explicitly stated in Worsley and Friston (1995), does not address the issue of correcting distortions due to intrinsic physiological correlation, though it does provide the correct expressions to account for temporal smoothing imposed during post-processing. However, following the work by Zarahn, et al. (1997), we can make a simple modification to the GLM method which will account for the temporal autocorrelation. In what follows we briefly review the
original E-GLM, describe the modification and demonstrate its ability to correct for P-value distortion, and, finally, we compare this modified E-GLM method to that of the whitening filter method described earlier.

The original E-GLM postulates a linear model $X = G\beta + e$, where $X$ are the data, $G$ is a matrix of postulated covariate waveforms, $\beta$ is the covariate coefficient vector, and $e$ is Gaussian white noise of variance $\sigma^2$. The data $X$ are then smoothed with a matrix $K$ whose rows consist of the hemodynamic response function, yielding an equation $KK = G^*\beta + Ke$, where $G^* = KG$, which yields a non-optimal least squares solution for the estimator of $\beta$, $b = (G^*G^*)^{-1}G^*TKX$ (Worsley and Friston, 1995). The modification consists of (1) replacing the hemodynamic response function with the "shaping filter" of the noise (the inverse of the whitening filter or, equivalently, the impulse response of $1/H_w(e^{i\omega})$) in $K$, and then (2) replacing $KK$ with $X_c$, the observed physiologically correlated data, in the expressions for the estimator and residual vectors (yielding $b = (G^*G^*)^{-1}G^*T X_c$ and $r = RX_c$). Their expression for the effective degrees of freedom is then used to obtain the voxel-wise P-values. The rationale here is that the observed data $X_c$ are already temporally correlated—i.e., it has already been operated on by the shaping filter $K$, so we need not operate on it again. The FPC generated by this modified Extended General Linear Model (ME-GLM), applied to TR = 312 ms data, is shown in Figure 10. This FPC is nearly identical to the ideal FPC, with performance comparable to the whitening filter method. In this scheme, we are regressing against covariates $G$ filtered by the shaping filter $K$ (i.e., $G^* = KG$), a by-product of the original E-
GLM method which may or may not be desired.

While both the ME-GLM and whitening filter produce good corrections for false-positive bias, with a speed advantage in favor of the whitening filter, the ME-GLM method does not provide an optimal solution to the problem posed (Friston and Worsley, 1995). This non-optimal formulation was chosen by Worsley and Friston (1995) to increase robustness of the solution, since the functional form for the K matrix used in their framework would have resulted in an ill-conditioned matrix inversion if a fully-optimal weighted least-squares solution had been chosen. However, the AR plus white noise model used in this study does not suffer from this problem, since the white noise term prevents the spectrum from ever decaying to zero, and can provide the optimal weighted least-squares solution. In preliminary simulations, we have found nearly identical robustness between these two methods.

There is an important conceptual difference between the original E-GLM and the whitening and ME-GLM methods presented here. The former considers noise correlations to come from the hemodynamic response function, while the latter methods take a more empirical view of the underlying noise. While it would be convenient if the underlying noise in the MR imaging process were temporally shaped by the hemodynamic response function, in practice the noise includes both white components and correlated components with time constants that seem longer than the hemodynamic response function. As shown above, the strategy of using the ME-GLM model can be effective in correcting distortions in the false positive rate, but the appropriate K matrix must be determined from the actual noise, not the hemodynamic response. Zarahn et al. (1997) have described further elaborations to the E-GLM which use empirical estimates of noise to account for temporal autocorrelation along with a separate hemodynamic smoothing kernel to shape the covariate waveforms—we omit a detailed analysis of such methods for brevity, but we suggest that it should yield results similar to the ME-GLM. Table 2 provides a summary comparison of how the various methods described here handle the issue of noise modeling and hemodynamic smoothing of covariate waveforms.
Table 2. Comparison of Features in Various fMRI Analysis Methods

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Smoothing of covariate waveforms</td>
<td>Smoothing by Hemodynamic Response</td>
<td>Smoothing by Noise Shaping Filter</td>
<td>Smoothing by Hemodynamic Response</td>
<td>Smoothing by Hemodynamic Response</td>
</tr>
<tr>
<td>Uses Optimal Estimator?</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
</tr>
</tbody>
</table>

Finally, it is important to point out that the specific coefficients for the AR process, which characterize an underlying biological variation and not instabilities in the scanner, depend on spatial resolution, pulse sequence, field strength, etc. and may even vary strongly between individual subjects and regions of the brain. As a result, the specific distortions in false positive rates should not be generalized beyond the examples described.

CONCLUSIONS

We have shown that when temporal autocorrelation in fMRI data sets is ignored, the voxel-level P-values which result are seriously distorted, in a way that depends upon both imaging rate and paradigm choice. Furthermore, we have developed a simple method for correcting these P-value distortions and have used the intuition behind this method to suggest a modification to the Modified General Linear Model that can also correct the P-value distortions. Because the accuracy of P-value assignment is essential to the technique of statistical mapping, it is important that the effects of this temporal autocorrelation are properly accounted for while creating such maps.

REFERENCES


Appendix A

We wish to derive an approximate frequency domain relationship to illustrate the relationship between stimulus paradigm frequency and FPC distortion. Let us define a paradigm waveform \( p[n] \) consisting of a square wave that has a value of -1 for samples corresponding to no stimulus (call this group “Off”) and a value of 1 for samples corresponding to a stimulus (call this group “On”). For simplicity, let \( p[n] \) have zero mean (i.e., the number of samples in the Off group is equal to that in On group). The expected variance of the difference in the sample means of the Off and On groups can be expressed in the frequency domain as an integral over the noise spectrum, \( S_{XX}(\omega) \), weighted by the frequency content of the paradigm:

\[
\mu_{on} - \mu_{off} = \frac{1}{N/2} \sum_{n=0}^{N-1} x[n]p[n] = \frac{1}{\pi N} \int_{-\pi}^{\pi} X(e^{j\omega})P^*(e^{j\omega}) d\omega \quad \text{(Parseval's Relation)}
\]

\[
\left| \mu_{on} - \mu_{off} \right|^2 = \frac{1}{(\pi N)^2} \int_{-\pi}^{\pi} X(e^{j\omega})P^*(e^{j\omega}) d\omega \int_{-\pi}^{\pi} X^*(e^{j\lambda})P(e^{j\lambda}) d\lambda \tag{7}
\]

\[
E\{X(e^{j\omega})X^*(e^{j\lambda})\} = S_{XX}(e^{j\omega}) \delta(\omega - \lambda) \quad \text{for large N}
\]

\[
\therefore \text{var}\{\mu_{on} - \mu_{off}\} = E\{\left| \mu_{on} - \mu_{off} \right|^2\} = \frac{1}{(\pi N)^2} \int_{-\pi}^{\pi} S_{XX}(e^{j\omega}) \left| P(e^{j\omega}) \right|^2 d\omega
\]

where \( \mu_{on} \) and \( \mu_{off} \) are defined as before, \( X(e^{j\omega}) \) and \( P(e^{j\omega}) \) are Discrete-Time Fourier Transforms of \( x[n] \) and \( p[n] \), respectively, \( * \) denotes the complex conjugate operation, \( S_{XX}(e^{j\omega}) \) is the power spectrum of \( x[n] \), and \( \delta(\omega) \) is the Dirac Delta function. Since the paradigm waveform is a square wave, for large \( N \) it can be approximated as a series of delta functions whose fundamental component corresponds to the paradigm frequency. If we use the fact that the higher order Fourier coefficients of \( p[n] \) are much smaller than the fundamental, we can approximate the result of equation (7) as

\[
\text{var}\{\mu_{on} - \mu_{off}\} \propto \int_{-\pi}^{\pi} S_{XX}(e^{j\omega}) \delta(\omega - \omega_p) d\omega = S_{XX}(e^{j\omega_p}) \tag{8}
\]
Appendix B

Expressing (2) as a Z-transform, factoring it, and inverting it, we have

\[ S_{z}(z) = \frac{A_c (1-q)^2 + A_w (1-qz^{-1})(1-qz)}{(1-qz^{-1})(1-qz)} = \frac{K(1-\gamma z^{-1})(1-\gamma z)}{(1-qz^{-1})(1-qz)} = \frac{1}{H_w(z)H_w(z^{-1})} \]

\[ H_w(z) = \frac{\sqrt{K(1-qz^{-1})}}{(1-\gamma z^{-1})} \]

\[ \gamma = \frac{D}{2} \pm \left( \frac{D^2}{4} - 1 \right)^{1/2}, \quad D = \left( \frac{(A_c + A_w) - 2A_c q + (A_c + A_w)q^2}{A_w q} \right) \]

\[ K = A_w q / \gamma \]

where \( H_w(z) \) is the whitening filter and \( \gamma, D, \) and \( K \) are derived variables as given above. Note that \( \gamma \) is a pole of \( H_w(z) \) and must be chosen to be stable (i.e., magnitude less than one) in this context.
Chapter 3. Noise Estimation for Control of False-Positive Rates: How Much Does the Noise Vary Within and Between Subjects and Trials

INTRODUCTION

Control of false-positive rates is an important problem in the statistical analysis of fMRI data. The noise in fMRI time series is known to be correlated in time [Biswal, et al., 1995; Bullmore, et al., 1996; Weisskoff, et al., 1993], with substantial components from underlying physiological fluctuations [Mitra, et al., 1997]. This temporal correlation can bias the false positive rates of hypothesis testing procedures applied to fMRI data. Early attempts to correct this problem focused on using an assumed form of the hemodynamic response function (i.e., within a linear systems framework) to represent these correlations, and calculating an effective degrees of freedom to obtain the correct P-values [Worsley and Friston, 1995]. Zarahn et al. (1997) have found that while this method can work, its effectiveness in correcting false-positive bias is highly dependent upon the filter structure chosen, and thus may not be generally applicable. More recently, authors have focused on using actual estimates of the noise to obtain the correct the false positives. Some have chosen to make local voxel-level estimates of the noise power spectrum while simultaneously estimating activation parameters [Bullmore, et al., 1996; Lange and Zeger, 1997; Solo, et al., 1997], while others have attempted to use global averages of the noise spectra over space and across large populations of subjects [Zarahn, et al., 1997]. Finally, some authors have attempted to account for correlated noise by co-varying out low frequency sinusoids during analysis [Holmes, et al., 1997], but recent evidence [Mitra, pers. comm] indicates that the correlated noise in fMRI is actually broad-band, overlapping with the activation response, suggesting that this method may not be appropriate.

The idea of obtaining noise estimates prior to analysis for activation is an attractive one, provided that the noise estimates meet certain assumptions, because it is easier computationally and is not affected by bias in activation parameter estimates in the way that simultaneous noise estimates might be. However, this kind of analysis cannot work unless there is some form of stationarity in the data, either within individual brains and between trials, or between different subjects. For instance, if we obtained a noise estimate for a subject by scanning "resting-state" data before the experiment, analysis based on this noise estimate would be valid only if we could be sure that this estimate is a reasonable representation of the noise in the subsequent experimental scans. Similarly, if the noise structure were found to be stationary between subjects, it might be possible to use a global noise estimate obtained by averaging the estimates from several subjects. Zarahn et al. (1997) attempted both of these ideas, using
averages taken over the entire brain and across subjects. Averages of this kind were unsuccessful in controlling the omnibus false-positive rate because the averages of noise spectra taken over the entire brain would systematically underestimate regions with highly correlated noise (such as cortical gray matter), while overestimating colored noise in regions that contain mostly white noise or periodic components (such as white matter, the sinuses, or the ventricles) [Zarahn, et al., 1997]. However, since many fMRI studies focus mainly on cortical gray matter regions, it may make more sense to restrict such noise averages to gray matter regions. Spurious activations in non-gray matter regions brought about by using the gray matter noise estimates would be easy to recognize from anatomical considerations and could be ignored. If the noise exhibits some kind of stationarity within gray matter regions, then noise estimation schemes based on pre-scanning or averaging across subjects may be a reasonable approach for controlling false-positive bias in gray matter.

In this chapter, we investigate how the noise in fMRI data sets vary within gray matter regions, and between subjects and trials, in order to gain insight into the problem of noise estimation for controlling false-positive rates.

METHODS

Data Acquisition and Registration

Resting-state BOLD fMRI images were taken using a GE Signa 1.5 T scanner modified by ANMR for echo planar imaging, at a TR (the sampling interval between images) of 2 s over 256 s, yielding a total of 128 images in time. The data were collected by Michael Rotte and Randy Buckner at the MGH-NMR Center (Charlestown, MA). The TR and number of images chosen here are typical of many brain mapping studies. Sixteen slices at an oblique angle were taken, using seven subjects, with 2 consecutive trials per subject. Subjects were instructed to fixate on a centered cross-hair over a dark screen. Tissue regions from a single slice roughly through the center of the brain were segmented by hand using high-resolution T1 images. In identifying gray and white matter regions, care was taken to avoid the ventricles and the venous sinuses. These regions are strongly coupled to heartbeat and respiration, and would yield inherently poor noise estimates that might confound the analysis of variability. The resulting regions of interest were then scaled down for use with the lower resolution BOLD images.
Noise Model

There are many sources of noise in fMRI data sets. The MRI scanner produces noise that is uncorrelated across space and time. There are also fluctuations due to heartbeat and respiration, occurring at roughly 1.0-1.2 Hz and 0.5 Hz, respectively, particularly in the ventricles and venous sinuses. Finally, there are low frequency fluctuations (< 0.3-0.5 Hz) [Biswal, et al., 1995; Weisskoff, et al., 1993] attributable to motion artifact [Zarahn, et al., 1997] and local spontaneous fluctuations in blood flow and volume (i.e., vasomotor fluctuations) [Mitra, et al., 1997]. Figure 1 shows the magnitude spectrum of fMRI noise in representative brain regions. In most regions of functional interest, scanner and low-frequency physiological noise comprise the bulk of the noise power, with relatively small contributions from respiratory and cardiac sources. Hence, we focus here on modeling the scanner noise and low-frequency physiological noise. We represent the scanner as white noise, and use an AR(1) process to represent the low-frequency noise [Bullmore, et al., 1996; Locascio, et al., 1997; Weisskoff, et al., 1993]. The overall noise signal $x[n]$ is then modeled as

$$x[n] = w[n] + c[n]$$

$$c[n] = q c[n-1] + v[n],$$

where $w[n]$ is the scanner noise, with variance $\sigma_w^2$, $c[n]$ is the AR(1) noise, $q$ is the AR correlation, and $v[n]$ is white noise with variance $\sigma_v^2$. The power spectrum of the noise is given by

$$S_{xx}(e^{i\omega}) = \sigma_w^2 + \sigma_v^2\left|1 - qe^{-i\omega}\right|^2$$

(2)

The overall noise power due to the AR(1) process is given by

$$ARpow = \frac{\sigma_v^2}{1-q^2}$$

(3)
and can be thought to represent the noise power due to physiological fluctuations.

The white noise variance was estimated by averaging the estimated variance over voxels outside of the brain. The remaining AR(1) parameters were estimated using an Estimate-Maximize algorithm.

**Variability Analysis**

To explore noise variability within individual trials, we computed the average and standard deviation of the noise parameters across cortical gray matter for each trial. The variance of these parameters across gray matter pixels can be thought to come from two sources: 1) the intrinsic variance of the parameters and 2) the variance of the estimation process. If the noise is essentially the same across gray matter, then we would expect most of the variance to come from the estimation process. In order to differentiate between these two sources of variance, we determined the variance of the estimation process through simulation and compared it to the actual variance across gray matter. Noise time series (1024 independent time-series for each simulation) were synthesized over a range of parameter values representative of the average values obtained from the brain estimates, and then analyzed using the methods described above. We then computed the standard deviation of the resulting parameter estimates to obtain an estimate for the standard error of the estimation procedure. We compared this standard error for the estimator to the standard deviation of the experimental noise estimates using an F-ratio test. To explore noise variability across subjects, we computed the average and standard deviation for each parameter across all subjects and trials, again comparing this standard deviation to an empirically derived standard error for the estimation process using an F-ratio test.

For the between trial variability, the question we are addressing is whether or not the noise from a single data set is similar enough to the noise in a data set taken at a later time from the same subject. To this end, we averaged the AR power over cortical gray matter in successive trials from the same subject and examined the percentage change between trials.

**RESULTS AND DISCUSSION**

Empirical measurements of the standard error for AR power values were obtained from simulations. Table 1 shows the parameter values of the synthesized data, the resulting parameter estimates, and the empirically obtained standard error estimates. Synthesized parameter values were chosen to cover a range of realistic values. Overall, the
standard error for the AR power is seen to be roughly 20-25% of its value. The average and standard deviation for gray matter noise parameters in one subject is shown in Table 2, illustrating the large standard deviations in the AR power and $\sigma^2_c$ observed, as well as the variability possible between trials. An F-ratio test comparing the variance in the AR power for trial “B” to the empirically determined variance of the estimator (Table 1, line 1) at those parameter values indicates that the observed variance in the AR power is significantly greater than the estimator variance ($P < 5e^{-4}$). Other trials show a similar difference in observed variance compared to the estimator variance. Figure 2 shows an image of the AR power for a given subject, illustrating a noticeable contrast between gray and white matter.

Table 1: Parameter Estimates from Synthesized Data

<table>
<thead>
<tr>
<th>Synthesized Parameters</th>
<th>Parameter Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma^2_w$</td>
<td>$\sigma^2_c$</td>
</tr>
<tr>
<td>----------------</td>
<td>------------</td>
</tr>
<tr>
<td>26.2</td>
<td>104</td>
</tr>
<tr>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>25</td>
<td>50</td>
</tr>
<tr>
<td>25</td>
<td>100</td>
</tr>
<tr>
<td>25</td>
<td>36.6</td>
</tr>
</tbody>
</table>

Table 2: Average Parameter Estimates From One Subject

<table>
<thead>
<tr>
<th>Trial</th>
<th>ARpow</th>
<th>$\sigma^2_c$</th>
<th>$q$</th>
<th>$\sigma^2_w$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Average</td>
<td>44</td>
<td>27</td>
<td>0.49</td>
</tr>
<tr>
<td></td>
<td>Std. Dev.</td>
<td>±40</td>
<td>±23</td>
<td>±0.3</td>
</tr>
<tr>
<td>B</td>
<td>Average</td>
<td>104</td>
<td>83</td>
<td>0.36</td>
</tr>
<tr>
<td></td>
<td>Std. Dev.</td>
<td>±71</td>
<td>±58</td>
<td>±0.3</td>
</tr>
</tbody>
</table>

The percentage change in AR power between trials is shown in Table 3. In many cases, the amount of AR power changes greatly from one trial to the next (Subjects 1, 2, 4, and 6), while in other subjects this change is not as drastic (subjects 3, 5, and 7).

Averages of the AR noise parameters across all subjects and trials are shown in Table 4. The F-test comparing the across-subject to the estimator variance indicates that the across-subject variance is significantly greater than that of the estimator alone ($P < 5e^{-4}$).
Table 3: % Change in AR power in Gray Matter between trials

<table>
<thead>
<tr>
<th>Subject #</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Change</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>in AR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>power</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>135%</td>
<td>37%</td>
<td>7%</td>
<td>33%</td>
<td>6%</td>
<td>75%</td>
<td>12%</td>
<td></td>
</tr>
</tbody>
</table>

Table 4: Gray Matter Parameter Averages Across All Subjects

<table>
<thead>
<tr>
<th></th>
<th>AR power</th>
<th>$\sigma_c^2$</th>
<th>q</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
<td>36.6</td>
<td>26.9</td>
<td>0.43</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>± 25.0</td>
<td>± 22.1</td>
<td>± 0.1</td>
</tr>
</tbody>
</table>

CONCLUSIONS

The large variance in AR power within a trial, above and beyond estimator variance, indicates that there is wide-ranging spatial variability in the physiological fluctuations present in the brain. The large percentage changes in AR power between trials observed in some subjects suggests that pre-scanning for noise estimates on an individual basis may not be generally reliable. The large variance in AR power across-subjects, again above and beyond that expected from the estimation process, suggests that there is no meaningful global noise estimate. These findings suggest that for a simple, parametric, time-series based noise model such as the one used here, the noise should be estimated locally and concomitantly with the activation response.

REFERENCES


Chapter 4: Implementing A Method for Estimating Physiologically Relevant Activation Parameters and Noise in fMRI Data

INTRODUCTION

The prevailing analytical methods in fMRI are aimed at detecting responses in the brain to experimental treatments. The analysis has been motivated as a detection problem because 1) The brain is known to have a functionally segregated spatial organization, and hence an important scientific and clinical goal is to identify the spatial boundaries of these functional areas, and 2) The heretofore lack of knowledge about the mechanism of the fMRI BOLD response has made it difficult to gain more detailed information about the activation response. However, as our knowledge of the mechanisms for the BOLD response improves, we will be in a position to gain detailed information about underlying physiological mechanisms by estimating parameters from physiologically-inspired models. Work to develop the models and methods for this kind of analysis are therefore a high priority in the field of fMRI.

Two important goals in developing an analytical method of this kind are to 1) use a physiologically-realistic model of the fMRI signal and 2) employ a robust method for estimating the noise. Using a physiologically-realistic model of the fMRI signal is important because it can lend itself naturally to physiological interpretation, and because a physiologically-realistic model will be more successful at capturing the gross features of the activation signal, improving the accuracy of the noise estimates. For example, if we used a model for the fMRI response which failed to capture the "undershoot" dynamics in block-paradigm experiments, the undershoot would appear in the residuals and might bias the noise estimates. Robust estimation of the noise is also essential because it provides the basis for accurate estimates of the variance of the parameter estimates, and thus for statistical inference. This was demonstrated in Chapter 1 from the standpoint of P-value bias in univariate hypothesis tests.

In current fMRI analysis techniques, the fMRI signal is modeled in a Linear, Time-Invariant (LTI) systems framework, where a unimodal Poisson [Friston, et al., 1994] or Gamma [Boynton, et al., 1996; Cohen, 1997] function is chosen as the impulse response (or "hemodynamic response," as it is referred to in the fMRI literature). While such a model can faithfully represent the observed dynamics for sparsely separated single trials [Buckner and Koutstaal, 1998], it will fail to account for the spatially-variable "undershoot" present in many fMRI time series with longer stimulus durations. Furthermore, there may be physiologically interesting differences in the shape or
composition of the fMRI signal across different regions of the brain, as well as delays in onset of the response, that are often unaccounted for in such methods.

As illustrated in Chapter 2, the temporally correlated noise in fMRI data varies greatly between trials and subjects and across different regions of the brain, suggesting the need for local noise estimates obtained concomittantly with the activation estimates. However, because fMRI time series are typically very short, estimating the noise robustly can be difficult. One way of overcoming this is to exploit any spatial smoothness in the noise by local averaging. Because the fMRI BOLD response may have a spatial resolution that is much finer than that of the noise, it is essential to preserve the spatial resolution of the fMRI activation while doing this local averaging on the noise.

In this Chapter we describe a method which 1) utilizes a physiologically-inspired model for the hemodynamic response that both accounts for spatially-varying “undershoot” and delay dynamics and 2) simultaneously estimates the noise using a regularization technique which exploits spatial smoothness of the brain noise without smearing spatial localization of activation.

METHODS

Modeling and Estimation

The signal at time \( t \) and spatial location \( P \) is given by

\[
x_P(t) = m_P + b_P t + s_P(t) + v_P(t) + w_P(t)
\]

\( t = 0 \ldots n-1, \ P = 0 \ldots M-1 \) \hfill (1)

where \( m_P \) is the baseline level, \( b_P \) is a linear drift term, \( s_P(t) \) is the BOLD activation signal, \( w_P(t) \) is the AR(1) noise, and \( v_P(t) \) is the white noise. Under a number of simplifying assumptions, we model the BOLD signal to include both a flow-induced decrease in deoxyhemoglobin, HbR, and blood volume-induced increase in HbR with different time constants [Mandeville, et al., 1998] and a spatially-varying time shift, \( D_P \) between stimulus and activation:

\[
s_P(t) = S_0 \exp(-k_1[HbR](t))V(t) = S_0(1-k_1[HbR](t))V(t)
\]

\[
[HbR](t) = k_2 g(t-D_P) * c(t)
\]

\[
V(t) = V_0 + k_3 h(t-D_P) * c(t)
\]

\[
g(t) = t \exp(-t/15s)u(t), \ h(t) = \exp(-t/12s)u(t)
\]

\hfill (2)
where V(t) is the blood volume, c(t) is the stimulus signal, u(t) is a unit step function, \( S_0 \), \( k_1 \), \( k_2 \), \( k_3 \), and \( V_0 \) are arbitrary constants, and "\(*\)" denotes convolution. Combining terms in (2) and absorbing the constant term into \( m_r \), we obtain three terms:

\[
s(t) = f_{a,P} g(t-D_P) * c(t) + f_{b,P} h(t-D_P) * c(t) + f_{c,P} [g(t-D_P) * c(t)][h(t-D_P) * c(t)]
\]  

(3)

where \( f_{a,P}, f_{b,P}, \) and \( f_{c,P} \) are resultant parameters to be estimated.

The relative contributions of the three terms above can be understood from Figure 1, which shows the contributions of the above terms when a block-paradigm (square-wave) stimulus is given. The first panel shows the contribution of the \( f_a \) term. It resembles an fMRI response, except that the undershoot behavior is absent. When we include the \( f_b \) term, we can obtain undershoot behavior, but the "On" portion of the activation signal is skewed somewhat. When we include the third \( f_c \) term, the "On" portion of the activation signal is no longer skewed, providing a physiologically realistic fMRI signal.

![Figure 1. Components of fMRI Signal Model](image)

To estimate these parameters, we used a maximum-likelihood technique employing a cyclic descent algorithm that iterated between estimating the signal parameters \([m, b, f_a, f_b, f_c]\) and the AR noise parameters \([\sigma_m, q]\).
[Solo, et al., 1997; Solo, et al., 1998]. The white noise parameter was estimated from the background of the image, while the AR parameters were estimated using an Estimate-Maximize algorithm employing regularization by local spatial averaging. Since the cyclic descent algorithm separates the signal estimation from the noise estimation, the local spatial averaging during the noise estimation did not impair the resolution of the activation signal estimates.

Data Analysis

To investigate the efficacy of the regularization technique, a synthetic data set containing spatial variations in AR(1) noise levels and activation levels was created and analyzed with the above technique. Figure 2 provides an illustration of the spatial distributions of the noise and activation signal. The noise varies abruptly across space from regions of high AR power to regions of lower AR power, while the activation varies between regions of activity and inactivity. To compare the regularization technique to other techniques, the data set was analyzed in three different ways: 1) With no regularization, 2) With spatial smoothing prior to analysis, but with no regularization, 3) With regularization.

![Simulated Activation Map](image)

![Simulated Noise Distribution](image)

**Figure 2. Spatial Distribution of Activation Signal and Noise for Synthetic Data Set**

To investigate the efficacy of these methods on real fMRI studies, data were collected from a combined visual and motor experiment using a GE Signa 1.5 T scanner modified by ANMR for echo-planar imaging (EPI).
The data were collected by Kathy O'Craven and Robert Savoy at the MGH-NMR Center (Charlestown, MA). A subject was presented with a 12.8s OFF-ON period of full-field flickering checkerboard and was instructed to finger-tap during the OFF period. The experiment included 8 such periods, imaged at a TR of 800ms, with a single slice chosen to transect the visual and motor areas. These data were analyzed in a manner analogous to the simulated data set from above.

RESULTS AND DISCUSSION

The physiologically-inspired hemodynamic model used here was found to provide a better account for the gross features of the fMRI signal than more commonly employed unimodal LTI kernels. Figure 3 shows the fit of our model, averaged over the primary visual cortex, compared to that using an LTI system with a unimodal gamma function as the impulse response (equivalent to removing the $f_b$ and $f_c$ terms from the model). Both fits were obtained in the same way, without regularization, with the appropriate substitutions for the form of the activation signal $s_r(t)$. The physiologically-inspired model is able to fit the undershoot behavior, while the unimodal gamma function model is not. This suggests that our model, though slightly more complicated than the more common LTI models, may yield important information in studies where quantifying the undershoot and its spatial distribution are important. Furthermore, the residuals from our model will have less of the activation signal left-over, which will improve our noise estimates.
Figure 4 shows the results of the data analysis from the simulated data set. The data processed with regularization on the AR parameters showed a much smoother estimate of the noise (as seen through a map of the AR power) than either of the other two cases, even though the spatial filtering kernels for the regularized and spatially pre-filtered cases were chosen to be equivalent. The contrast is particularly clear in the regions of high AR noise power relative to the white noise. The activation map, on the other hand, shows substantial blurring for the spatially pre-filtered case, particularly at the lower portion of the image where the activation varies rapidly across space, whereas the regularized data show identical spatial resolution to the untreated case. These data demonstrate that the regularization scheme is effective in improving the noise estimates while retaining the native spatial resolution of the activation map.

Figure 5 shows the analogous results for the real data. The activation maps shown in this figure are quantified in terms of the weighted 2-norm of the activation signal, where the weighting is given by the local noise power. An activation metric of this kind ensures that truly “active” regions of the brain can be distinguished from regions that simply have large values of $f_a$, $f_n$, and $f$, due to large amounts of noise. In all cases, both the visual and
motor areas of the brain were correctly identified, in agreement with the experimental paradigm. The noise images given here are displayed as the square-root of the AR power. As in the simulated data case, we see that the noise estimates for the regularized case are smoother than either the spatially pre-filtered data or the untreated data. Furthermore, the noise for the regularized case shows some degree of contrast between cortical gray matter regions (at the outer surface of the brain) and white matter, which coincides with our intuition about hemodynamic origins of the noise. In the untreated and spatially pre-filtered cases, this contrast is less clear. The activation maps for the regularized case have the same high spatial resolution as the untreated case, whereas the activation patterns in the spatially pre-filtered case are blurred.

<table>
<thead>
<tr>
<th></th>
<th>Untreated</th>
<th>Regularized</th>
<th>Pre-filtered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Activation</td>
<td><img src="untreated.png" alt="Image" /></td>
<td><img src="regularized.png" alt="Image" /></td>
<td><img src="pre-filtered.png" alt="Image" /></td>
</tr>
<tr>
<td>AR Noise Power</td>
<td><img src="untreated.png" alt="Image" /></td>
<td><img src="regularized.png" alt="Image" /></td>
<td><img src="pre-filtered.png" alt="Image" /></td>
</tr>
</tbody>
</table>

*Figure 4. Results from Simulated Data*
Conclusions

In this chapter we have described a signal processing method for analyzing fMRI data that incorporates a physiologically realistic model of the hemodynamic model within a framework that can provide robust estimates of the noise concomittantly with the activation estimates. We have demonstrated that the model is able to capture dynamics of the fMRI signal that other, simpler models do not, and have shown that the noise estimation method provides superior noise estimates to other treatements, in a way which does not spoil the spatial resolution of the activation maps.

References

